NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Contents:

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Premeeting briefing

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282)

This premeeting briefing presents:

- the key evidence and views submitted by the company, the consultees and their nominated clinical experts and patient experts and
- the Evidence Review Group (ERG) report.

It highlights key issues for discussion at the first Appraisal Committee meeting and should be read with the full supporting documents for this appraisal.

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies.

This document presents results using the confidential patient access scheme discount for pirfenidone. Nintedanib also has a patient access scheme for which the level of discount is confidential and not known to Roche Products. Accordingly, the Evidence Review Group ran the company's analyses using the patient access scheme discounts for both pirfenidone and nintedanib; the results are in the confidential appendix to this document.

Key issues for consideration

Clinical effectiveness

Natural history of idiopathic pulmonary fibrosis

- Is a decline in lung function temporary or permanent? Can lung function be restored? *(relevant to modelling assumptions)*
- Is the natural history of disease likely to change in the 20 weeks from week 52 to 72? *(relevant to modelling assumptions)*

Generalisability

- Are results from ASCEND and the CAPACITY trials of pirfenidone generalisable to patients seen in clinical practice in England?
  - Fewer comorbidities, fewer concomitant medications and lower mortality risk in trials
  - 25% had percent predicted FVC >80% (compared with 30–50% in UK practice)
  - ASCEND: no UK sites, excluded people with percent predicted FVC >90%
  - *Note that committee C was willing to accept the generalisability issues with the CAPACITY trials in TA282*
- Is the SP3 trial relevant to the decision problem (given the unlicensed dose) and generalisable to clinical practice (given the different baseline characteristics)?
  - *Note that, during committee B's appraisal of nintedanib for idiopathic pulmonary fibrosis, the committee's preference was to include the SP3 trial and exclude the other Japanese trial (SP2) from the evidence informing the model.*
- Given the heterogeneity between the trials, is it more appropriate to report the credible intervals (used in the company model) or predictive intervals (used in the ERG model) from a network meta-analysis?

Effectiveness of pirfenidone: overall survival

- Pooled analysis of ASCEND and the CAPACITY trials showed pirfenidone reduced 1-year mortality by 48% compared with placebo. Robust?
  - Difference between pirfenidone and placebo not statistically significant in individual trials.

  - None of the studies powered to assess effect of pirfenidone on mortality.
  - In the CAPACITY trials, mortality increased between weeks 52 and 72.
- How does the clinical effectiveness of pirfenidone compare with nintedanib?
- Is there robust evidence for the long-term effectiveness of pirfenidone?
  - How robust is the company's comparison between registry and trial data?
  - Does this comparison support the assumption of proportional hazards between pirfenidone and best supportive care over a patient's lifetime? *(relevant to modelling assumptions)*
- Is there a relationship between progression-free survival and overall survival? *(relevant to model structure)*

Effectiveness of pirfenidone in subgroups

- Is pirfenidone effective regardless of percent predicted FVC (≤ or > 80%)?
  - No treatment-by-subgroup interaction test for overall survival and progression-free survival.
  - Only 25% of patients across ASCEND and the CAPACITY trials had a percent predicted FVC of 80% or more.
- Less likely to observe treatment benefit for mortality, progression and exacerbations in people with a percent predicted FVC above 80%?
- Do people who have had a 10% or more decline in percent predicted FVC decline in a 1-year period benefit from continued pirfenidone treatment? *(relevant to modelling assumptions)*

Stopping rules

- In clinical practice, when would it be clinically appropriate to stop treatment with pirfenidone? How would the decision to stop treatment be made? *(relevant to modelling assumptions)*

Cost effectiveness

Model structure and key assumptions

- Was it appropriate to model the condition using a partitioned survival model structured around disease progression?

- Would people continue to have nintedanib or pirfenidone after their disease progresses?

Stopping rules

- Appropriate to include a stopping rule? Based on percent predicted FVC or other outcomes?
- How robust are the company's and the ERG's scenarios including the stopping rule, given that treatment duration and treatment outcomes are disconnected in the partitioned survival model?

Estimation of treatment effect

- Appropriate for company to assume that treatment effect is constant over a patient's lifetime?
  - Would the treatment effect continue in people who continue to take treatment for more than 72 weeks?
  - Would the treatment effect continue after stopping treatment? How long for?
- Appropriate for the company to use 1-year data (week 48 or 52) in the model instead of week 72, given the reduction in treatment benefit in CAPACITY 1 & 2?

Clinical outcomes

- Should overall survival have been modelled using the Weibull (preferred by the company) or Gompertz (preferred by the ERG) distribution?
- Appropriate to assume that people who have experienced an exacerbation have the same risk of recurrent exacerbations as people that have not had an exacerbation?
- Would an exacerbation change the rate of progression (in terms of percent predicted FVC)? The model assumes no relationship between exacerbations and progression.

Costs and utilities

- Were the results of the company's utility mapping exercise plausible?
- Appropriate to assume that routine healthcare costs remain constant in the progressed health state? In the progression-free health state?

- Are end-of-life costs higher when deaths are attributable to idiopathic pulmonary fibrosis (compared with deaths unrelated to the disease)?

Subgroups by disease severity (percent predicted FVC)

- Appropriate to generalise outcomes data from all randomised patients to the subgroups (with the exception of baseline risk of death, disease progression and stopping treatment)?
- Is it relevant to consider the analyses in all randomised patients (which exclude nintedanib as a comparator)? Focus on subgroups?
- ASCEND excluded people with a percent predicted FVC above 90%. Can the model results be generalised to this population?

Equalities

- Would the committee's recommendations affect any protected groups?
  - Does percent predicted FVC disadvantage minority ethnic groups, older people or disabled people?

Innovation

- Any benefits not captured in QALY?
- Is pirfenidone likely to have a clinically meaningful impact on dyspnoea?
  - What is the minimal clinically important difference on the University of San Diego Shortness of Breath Questionnaire?
  - Results for University of San Diego Shortness of Breath Questionnaire were only clinically meaningful using pooled data from ASCEND and the CAPACITY trials.

Patient access scheme (PAS)

- The Department of Health has approved a reduced discount level for pirfenidone [redacted], on the condition that NICE recommends pirfenidone for a similar or greater number of patients than have access under the current guidance in TA282.
  - All analyses in this document use the new, reduced PAS.

1 Remit and decision problems

1.1 The remit from the Department of Health for this appraisal was: to appraise the clinical and cost effectiveness of pirfenidone within its marketing authorisation for treating idiopathic pulmonary fibrosis.

Table 1 Decision problem

	Final scope issued by NICE	Decision problem addressed in the submission
Population	Adults with mild to moderate idiopathic pulmonary fibrosis	As in the scope
Intervention	Pirfenidone	As in the scope
Comparators	• Best supportive care • Nintedanib (only for people with a percent predicted FVC of 50–80%)	As in the scope
Outcomes	• pulmonary function parameters • physical function • exacerbation rate • progression-free survival • mortality • adverse effects of treatment • health-related quality of life	As in the scope
Subgroups	If evidence allows, subgroup analysis by disease severity, defined by FVC (such as above and below 80%) and/or DLCO, will be considered	Subgroup analysis by percent predicted FVC: 50–80% (“moderate”) and >80% (“mild”)
Abbreviations: DLCO, diffusing capacity for carbon monoxide; FVC, forced vital capacity		

1.2 The Evidence Review Group (ERG) advised that the company’s decision problem reflected the scope. The main measure of pulmonary function in the company submission was percent predicted forced vital capacity (FVC) (see section 2.2). The ERG noted that its clinical advisors considered that percent predicted FVC in isolation was not widely used in clinical practice, except to implement the recommendations in TA282, commenting that diffusing capacity for carbon monoxide (DLCO) is

clinically more meaningful. However, the ERG's clinical advisors acknowledged that DLCO is harder to measure and more variable than FVC. The ERG therefore considered that it was reasonable for the company to focus on FVC as the main measure of pulmonary function.

2 The technology and the treatment pathway

2.1 Idiopathic pulmonary fibrosis is a chronic, progressive lung disease in which scarring (fibrosis) occurs. The cause is unknown, but it is thought to be related to an abnormal immune response. Symptoms may include breathlessness and cough. Over time, people can experience a decline in lung function, reduced quality of life, and death. The median survival for people with idiopathic pulmonary fibrosis in the UK from the time of diagnosis is approximately 3 years. People with mild-to-moderate disease live longer than people with severe disease.

2.2 Assessing the severity of idiopathic pulmonary fibrosis usually involves 2 measures of lung function:

- forced vital capacity (FVC): the maximum volume of gas that can be exhaled, from full inhalation, by exhaling as forcefully and rapidly as possible
- diffusion capacity of the lung for carbon monoxide (DLCO): measures how much oxygen travels from the lungs into the blood stream.

FVC and DLCO can be expressed as a percentage of the predicted normal value for a person of the same sex, age, height and ethnicity. DLCO is also corrected for haemoglobin level. Lower percent predicted values indicate worse health. Clinical trials generally define mild-to-moderate idiopathic pulmonary fibrosis as a percent predicted FVC of 50% or more and a percent predicted DLCO of 35% or more. A percent predicted FVC lower than 50%, and a percent predicted DLCO lower than 35%, indicates severe disease.

2.3 The company stated that a decline in percent predicted FVC of more than 10% (absolute value) represents a clinically important difference and indicates a higher mortality risk (see page 185 of the company submission). However, the company suggested that using percent predicted FVC alone to define mild and moderate disease has the potential to underestimate severity for 2 reasons:

- FVC can be elevated in people with emphysema
- the normal range for percent predicted FVC is 90% to 120%, meaning that, of all the people who have an FVC of 80%, some may have lost a third of their baseline lung capacity and others may have lost a tenth.

2.4 The aim of treatment is to manage symptoms and slow progression. NICE clinical guideline 163 on the diagnosis and management of suspected idiopathic pulmonary fibrosis recommends that best supportive care should be offered from diagnosis and be tailored according to disease severity, rate of progression and patient preference. Best supportive care may include symptom relief, managing co-morbidities, withdrawing ineffective or harmful therapies and end of life care. If pharmacological treatment is appropriate, the guideline recommends pirfenidone if the FVC is between 50% and 80% of the predicted value in line with recommendations in NICE technology appraisal guidance 282. The recently published NICE technology appraisal guidance 379 recommends nintedanib as a treatment option for people with a percent predicted FVC of 50–80%. Both technology appraisals recommend stopping treatment if there is evidence of disease progression (an absolute decline in percent predicted FVC of 10% or more within any 12 month period). NICE clinical guideline 163 recommends lung transplantation as an option for people with idiopathic pulmonary fibrosis. The present appraisal, a review of technology appraisal guidance 282, was triggered by the ASCEND study which – in the company's opinion – showed that people with a predicted FVC greater than 80% could benefit from pirfenidone.

2.5 During the original NICE appraisal of pirfenidone (technology appraisal 282), the company agreed a simple discount patient access scheme for pirfenidone. The company has written to the Department of Health to say that it wishes to reduce the level of discount for pirfenidone [redacted]. The Department of Health is content for the reduced discount level to be taken into account in the present appraisal, with the proviso that the change to the discount would be implemented only if NICE issues revised guidance recommending pirfenidone for the same or more patients than have access under the current guidance in TA282. Please see the confidential appendix to this document for further details.

Table 2 Technology and comparator

	Pirfenidone	**Nintedanib**
Marketing authorisation	Indicated in adults for the treatment of mild to moderate idiopathic pulmonary fibrosis	Indicated in adults for the treatment of idiopathic pulmonary fibrosis
Mechanism of action	Immunosuppressant; anti-inflammatory and antifibrotic effects	Tyrosine kinase inhibitor; targets 3 growth factor receptors involved in idiopathic pulmonary fibrosis
Posology and method of administration	Oral, three 267 mg capsules three times daily (total of 2403 mg/day across nine capsules). Start with one 267 mg capsule, three times daily, on days 1–7. Titrate to recommended dose over 14 days. Duration of treatment is not specified in the summary of product characteristics	Oral, 150 mg twice daily. If the 150 mg dose is not tolerated, 100 mg twice daily is recommended. Duration of treatment is not specified in the summary of product characteristics
Cost	£501.92 for a 63-capsule pack [BNF online, March 2016]. This equates to a daily cost of £71.70 (9 capsules/day). A confidential simple discount patient access scheme makes pirfenidone available to the NHS at a lower cost.	£2151.10 for 60 x 150mg capsules [BNF online, March 2016]. This equates to a daily cost of £71.70 (2 capsules /day). A confidential simple discount patient access scheme makes nintedanib available to the NHS at a lower cost.

See the confidential appendix to this document for details of patient access schemes. See the summary of product characteristics for details on adverse reactions and contraindications.

3 Comments from consultees

3.1 Submissions from professional groups highlighted that there are no treatment options for people with a percent predicted FVC above 80% (who represent more than 40% of people with idiopathic pulmonary fibrosis). They highlighted that people with mild disease are less likely to benefit from active treatment because their rates of mortality, progression and exacerbations are low. However, they suggested that there is good clinical evidence for treating mild idiopathic pulmonary fibrosis, noting that ASCEND shows that a benefit of pirfenidone for people with a percent predicted FVC up to 90%. Professional groups suggested starting treatment as early as possible, because preventing or slowing disease progression improves life expectancy.

3.2 The professional groups highlighted that the restrictions on prescribing pirfenidone imposed by current NICE guidance (TA282) create a burden for the NHS. People with a percent predicted FVC above 80%, for whom pirfenidone is not recommended, have regular lung function tests to check whether their disease has progressed sufficiently to allow treatment pirfenidone. This requires paying for pulmonary physiologists to perform the tests. The submission highlighted that a patient's percent predicted FVC may remain above 80% despite their condition worsening, as evidenced by other clinical outcomes such as diffusion capacity of the lung for carbon monoxide or 6-minute walking distance, and yet they will not be eligible for treatment. The professional groups suggested that recommendations for initiating treatment should not be restricted to percent predicted FVC.

3.3 The professional groups commented on the side effects of pirfenidone, noting that about 1 in 5 patients cannot tolerate it.

4 Clinical-effectiveness evidence

Overview of the clinical trials

4.1 The company's clinical evidence came from 4 randomised double-blind trials, comprising 3 multinational phase III studies (CAPACITY 1, CAPACITY 2 and ASCEND) and 1 study in Japan (SP3, a phase III study). All 4 trials compared pirfenidone with placebo in adults with idiopathic pulmonary fibrosis. Lower doses of pirfenidone were investigated in SP3 (1,800 mg/day or 1,200 mg/day) than in the multinational studies (2,403 mg/day); the company explained that this was because the Japanese people have a lower average body weight than European people. The marketing authorisation for pirfenidone in the UK recommends a dosage of 2,403 mg per day. The results of SP3 and the CAPACITY trials were considered during the original NICE technology appraisal of pirfenidone (TA282), but results from the ASCEND trial were not available at that time.

4.2 The primary endpoint in the multinational trials was the change in percent predicted forced vital capacity (FVC) from baseline. Secondary outcomes included progression-free survival, time to acute exacerbation, and time to death. The definition of an acute exacerbation and progression of disease varied across trials (for definitions of acute exacerbation see company submission pages 103–4 and response to clarification question A15; for definitions of progression-free survival see company submission pages 98–9; the company modelled both outcomes in its cost-effectiveness analysis). Health-related quality of life was assessed in the ASCEND and CAPACITY trials using the University of San Diego Shortness of Breath Questionnaire (SOBQ). The CAPACITY trials also included the St. George's Respiratory Questionnaire (SGRQ). The company used the SGRQ to estimate utility values in its cost-effectiveness model.

Table 3 Summary of pirfenidone randomised placebo-controlled trials

Trial name	Inclusion criteria			Trial length	Pulmonary function assessed	Primary endpoint
	Age	% predicted FVC	% predicted DLCO			
ASCEND (n=555)	40–80	50–90%	30–90%	52 wks	Wks 13, 26, 39, 52	Change in % predicted FVC from baseline to week 52
CAPACITY 1 (n=344) and CAPACITY 2 (n=435)	40–80	≥50%	≤90%	72 wks	Every 12 wks	Change in % predicted FVC from baseline to week 72
SP3 (n=275)	20–75	No requirements reported		52 wks	Every 12 wks (VC every 4 wks)	Change in VC at week 52

[a] Unblinded at week 36 (open-label pirfenidone until week 48) because of high incidence of acute exacerbations in the placebo arm

Abbreviations: 6MWD, 6-minute walking distance; DLCO, diffusing capacity for carbon monoxide; FVC, forced vital capacity; SpO2, blood oxygen saturation level; VC, vital capacity (a non-forced measurement, usually greater than FVC); wks, weeks

4.3 The company presented the results of an additional Japanese phase II trial, SP2 (n=107), in its submission. The company excluded SP2 from its network meta-analysis and cost-effectiveness model (see section 4.21) because Committee B previously considered this trial to be an outlier in the NICE technology appraisal of nintedanib (TA379). This document does not include the results of SP2.

4.4 The ASCEND trial was conducted in 9 countries; there were no UK sites. The CAPACITY trials recruited patients from 13 countries including 3 sites in the UK. The company indicated that baseline characteristics were generally similar in the different arms of these 3 trials (see pages 82–4 of the company submission). The baseline characteristics of the Japanese studies differed from that of the multinational studies; for example, SP3 had a higher proportion of male participants and people who smoked, higher percentage predicted DLCO, lower trial corticosteroid use, and smaller proportions having received surgical lung biopsies. There were no

differences between the study arms within SP2 and SP3, and no differences between the 2 Japanese trials. The mean percent predicted FVC at baseline was approximately 68% in ASCEND and 75% in the CAPACITY trials.

4.5 The trials prohibited other treatments for idiopathic pulmonary fibrosis, with a few exceptions. In the CAPACITY trials, short courses of azathioprine, cyclophosphamide, corticosteroids, or acetylcysteine were permitted if the person deteriorated or had an acute exacerbation. In ASCEND, concomitant medications for treating idiopathic pulmonary fibrosis were not permitted unless they were prescribed to treat another condition and there was no clinically acceptable alternative. In SP2 and SP3, corticosteroids were permitted but immunosuppressants were not. Approximately 20–35% of patients in the ASCEND and CAPACITY trials, and 5–10% in the SP3 trial, received corticosteroids.

4.6 People with obstructive airway disease (asthma or chronic obstructive pulmonary disease) or major comorbidities such as a history of unstable or deteriorating cardiac or pulmonary disease (other than idiopathic pulmonary fibrosis) were excluded from the trials.

ERG comments

4.7 The ERG suggested that the clinical trials of pirfenidone might not be generalisable to clinical practice in the UK. The ERG highlighted that:

- trial patients had fewer comorbidities, fewer concomitant medications and a lower mortality risk compared with the population likely to receive pirfenidone in clinical practice
- approximately 25% of people across the CAPACITY trials and ASCEND had mild disease (percent predicted FVC above 80%), compared with around 30–50% in UK clinical practice
- only 1 of the pirfenidone trials, CAPACITY 2, recruited people from UK centres (3 of 110 centres were UK).

4.8 The ERG agreed with the company's judgement that ASCEND and the CAPACITY trials were at a low risk of bias. However the ERG considered that the trials had a moderate risk of reporting bias and "other bias", because of inconsistency between outcomes specified in the trial protocols and those presented in the company submission (see page 54 of the ERG report). The ERG noted that SP3 was at a higher or more unclear risk of bias because of the absence of any published protocols and the inadequacy of the information contained within the published manuscripts.

4.9 The ERG considered that SP3 may not be relevant to the decision problem because it evaluated lower, unlicensed doses of pirfenidone; applied different eligibility criteria; and reported noticeable differences from the other 3 trials in some baseline characteristics.

Clinical trial results

4.10 Table 4 presents the primary outcomes from the 4 key trials of pirfenidone (change in percent predicted FVC, or change in vital capacity volume, from baseline). ASCEND and CAPACITY 2 showed a treatment benefit of pirfenidone. Specifically, the decline in FVC compared to baseline was smaller in patients randomised to pirfenidone than in patients randomised to placebo (44% smaller in ASCEND, 35% smaller in CAPACITY 2). In both trials, this difference between treatment groups was statistically significant. In contrast, CAPACITY 1 showed a smaller treatment benefit of pirfenidone and the difference between treatment groups was not statistically significant. The company suggested that the smaller treatment benefit observed in CAPACITY 1 was because the people randomised to placebo had a smaller decline in percent predicted FVC than expected, because (see response to clarification question A26):

- a higher proportion of people randomised to placebo had borderline obstructive disease, compared with people randomised to pirfenidone; obstructive disease and emphysema are associated with a smaller decline in FVC despite worsening of fibrosis in the lungs

- idiopathic pulmonary fibrosis is a heterogeneous disease and there is a natural variability in decline percent predicted FVC across the population.

The results of the company's composite outcome (10% or more decline in percent predicted FVC or death) are presented on pages 89–90 of the company submission. According to the company, the composite outcome was used in order to impute a FVC measurement for patients who have died (see company response to clarification questions A11 and A13). In ASCEND, pirfenidone was associated with a relative reduction of 47.9% in the proportion of patients who died or had a decline in percent predicted FVC of 10% of more (p=0.000001), compared with placebo.

Table 4 Primary outcomes in randomised placebo-controlled trials of pirfenidone

Study	Time (wks)	Pirfenidone	Placebo	Absolute difference (95% CI)	Relative difference	p value
Absolute change in percent predicted forced vital capacity from baseline, %						
ASCEND	52	−6.17	−10.95	4.78 (95% CI not reported)	43.7%	<0.001
CAPACITY 1	72	−9.0	−9.6	0.6 (−3.5 to4.7)	6.5%	0.501
CAPACITY 2	72	−8.0	−12.4	4.4 (0.7 to 9.1)	35.3%	0.001
Change in vital capacity from baseline, ml						
SP3[a]	52	−90	−160	70	43.8%	0.042
[a] unlicensed dose of pirfenidone: 1800 mg/day Sources: section 4.7 company submission and Noble et al. Lancet 2011; 377(9779): 1760–9						

4.11 Pirfenidone significantly prolonged progression-free survival in ASCEND, CAPACITY 2 and SP3, but not CAPACITY 1 (Table 5).

Table 5 Progression-free survival in randomised placebo-controlled trials of pirfenidone

	Risk of death or disease progression[a]: pirfenidone compared with placebo		
	Time point	Hazard ratio (95% CI)	p value
ASCEND	52 weeks	0.57 (0.43 to 0.77)	0.0001

CAPACITY 1	72 weeks	0.84 (0.58 to 1.22)	0.355
CAPACITY 2	72 weeks	0.64 (0.44 to 0.95)	0.023
SP3	52 weeks	0.45 (0.11 to 0.79)	0.028

[a] Progression-free survival defined as time until 1 of the following events: ≥10% decline in percent predicted FVC, ≥50m decline in 6MWD or death in ASCEND; ≥10% decline in percent predicted FVC, ≥15% decline percent predicted DLCO or death in the CAPACITY trials; ≥10% decline in VC or death in SP3.

Abbreviations: 6MWD, 6-minute walking distance; DLCO, diffusion capacity of the lung for carbon monoxide; FVC, forced vital capacity; VC, vital capacity

Source: section 4.7 company submission

4.12 Neither the ASCEND trial nor the CAPACITY trials were powered to assess the effect of pirfenidone on mortality. In a pre-planned pooled analysis of the 3 trials at 52 weeks, the risk of death was reduced by 48% in people treated with pirfenidone compared with people randomised to placebo (hazard ratio 0.52; 95% confidence interval 0.31 to 0.87; p=0.011) (Table 6). For the results of other secondary outcomes, see pages 98–112 of the company submission.

Table 6 All-cause mortality in pivotal trials of pirfenidone

	Week	Pirfenidone, n (%)	Placebo, n (%)	Hazard ratio (95% CI)	p value
ASCEND	52	11 (4.0)	20 (7.2)	0.55 (0.26 to 1.15)	0.105
CAPACITY 1	52	Not reported (NR)		0.66 (0.24 to 1.84)	NR
	72			0.87 (0.41 to 1.82)	
CAPACITY 2	52			0.37 (0.13 to 1.05)	
	72			0.51 (0.22 to 1.20)	
Pooled: CAPACITY 1 & 2	52	11 (3.2)	22 (6.3)	0.49 (0.24 to 1.01)	0.047
	72	27 (8)	34 (10)	0.77 (0.47 to 1.28)	0.315
Pooled: ASCEND, CAPACITY 1 & 2	52	22 (3.5)	42 (6.7)	0.52 (0.31 to 0.87)	0.011

Source: section 4.7 company submission (table 23) and appendix 9 company submission (figures 5 and 6; reproduced on page 73 of the ERG report)

4.13 There was no significant difference between pirfenidone and placebo in change in health-related quality of life at the end of week 72 of the trials (that is, after 72 weeks in CAPACITY and 52 weeks in ASCEND) (Table 7). Health-related quality of life worsened from baseline in both arms of all 3 trials. However, pooled results across all 3 trials showed that fewer

patients randomised to pirfenidone experienced a 'clinically meaningful' decline in breathlessness (20 points or more on the University of San Diego Shortness of Breath Questionnaire) or death compared with patients randomised to placebo (p=0.0471).

Table 7 Quality of life outcomes from CAPACITY 1, CAPACITY 2 and ASCEND

	Change from baseline in SGRQ score (possible scores range from 0–100)[a]			Change from baseline in SOBQ score (possible scores range from 0–120)[b]		
	Pirfenidone	**Placebo**	**Difference, (p value[c])**	**Pirfenidone**	**Placebo**	**Difference, (p value[c])**
ASCEND, week 52	Not collected			14	17.3	−3.3 (NR)
CAPACITY 1, week 72	7.2	7.3	−0.1 (p=0.766)	11.9	13.9	−2.0 (0.604)
CAPACITY 2, week 72	7.6	9.0	−1.4 (p=0.495)	12.1	15.2	−3.1 (0.509)

Higher scores on the SGRQ and SOBQ indicate worse symptoms (worse quality of life)
[a] All randomised patients (CAPACITY 1: n=335; CAPACITY 2: n=328)
[b] Intent-to-treat population (CAPACITY 1: n=344; CAPACITY 2: n=348; ASCEND: n=555)
[c] p value for the difference between pirfenidone and placebo in change from baseline
Abbreviations: NR, not reported; SGRQ, St. George's Respiratory Questionnaire; SOBQ, University of San Diego Shortness of Breath Questionnaire
Source: section 4.7 company submission (table 32, 35, 36)

4.14 The company's post-hoc subgroup analyses compared treatment effect in people who had a baseline percent predicted FVC of more than 80% (mild disease) with people who had percent predicted FVC of 50–80% (moderate disease), using pooled data from ASCEND and the CAPACITY trials. The company reported that baseline characteristics and demographics were similar across these 2 subgroups and that there was no statistically significant interaction between treatment and disease severity, for the primary outcome (change in percent predicted FVC, see Table 8). Pirfenidone was associated with a statistically significant (p<0.0001) benefit compared with placebo for the primary outcome in both subgroups (mild disease and moderate disease). The company also provided subgroup analyses according to disease severity for overall survival and progression-free survival. The company did not report a

treatment-by-subgroup interaction test so it is unclear if the treatment effect differed between these subgroups (Table 9).

Table 8 Treatment effect of pirfenidone (change in percent predicted FVC from baseline to week 52), according to baseline disease severity

Trial	Percent predicted FVC	Standardised treatment effect[a] (95% CI)	Interaction test, p value
ASCEND	≤80%	0.47 (0.26 to 0.68)	0.78
	>80%	0.52 (0.09 to 0.95)	
CAPACITY 1	≤80%	0.25 (−0.04 to 0.53)	0.20
	>80%	0.58 (0.14 to 1.02)	
CAPACITY 2	≤80%	0.4 (0.11 to 0.69)	0.73
	>80%	0.48 (0.07 to 0.89)	

[a] values greater than 0 indicate a treatment benefit of pirfenidone
CI, confidence interval; FVC, forced vital capacity
Source: company response to clarification question A29

Table 9 Treatment effect of pirfenidone (overall survival and progression-free survival to week 52), according to baseline disease severity

Trial	Percent predicted FVC	Hazard ratio[a] (95% CI)	Treatment effect[b], p value
Overall survival			
ASCEND	≤80%	0.63 (0.29 to 1.34)	0.22
	>80%	<0.01 (0.00 to not evaluable)	0.12
CAPACITY 1	≤80%	0.60 (0.17 to 2.04)	0.41
	>80%	0.77 (0.11 to 5.59)	0.80
CAPACITY 2	≤80%	0.25 (0.08 to 0.76)	0.01
	>80%	Not evaluable	Not evaluable
Progression-free survival			
ASCEND	≤80%	0.56 (0.41 to 0.76)	<0.05
	>80%	0.64 (0.30 to 1.40)	0.26
CAPACITY 1	≤80%	0.84 (0.53 to 1.32)	0.44
	>80%	0.63 (0.29 to 1.41)	0.26
CAPACITY 2	≤80%	0.60 (0.40 to 0.92)	0.02
	>80%	0.40 (0.18 to 0.89)	0.02

[a] hazard ratios below 1 indicate that patients having pirfenidone had a lower risk of the event than patients having placebo
[b] p value indicates significance of the difference between pirfenidone and placebo within the subgroup; p values for treatment-by-subgroup interaction test not reported
CI, confidence interval; FVC, forced vital capacity

Trial	Percent predicted FVC	Hazard ratio[a] (95% CI)	Treatment effect[b], p value
Source: company response to clarification question A31			

4.15 The subgroups in Table 8 and Table 9 were not pre-specified, although the method used was pre-specified. The company also presented the results of an analysis of pre-specified subgroups according to disease severity: percent predicted FVC lower than 70%, between 70% and 80%, and more than 80%. The company reported differences in observed mean effects, using pooled CAPACITY data. The NICE technical team noted that in the mild subgroup (percent predicted FVC over 80%), there was a non-significant treatment benefit in favour of placebo. The decline in lung function in people randomised to placebo was 1.2% smaller than the decline in people randomised to pirfenidone (see figure 16 of the company submission). The treatment-by-subgroup interaction test was not significant (p=0.35), but the ERG advised the technical team that this might be because the company assessed 3 subgroups rather than 2. In response to clarification question A29, the company suggested that the method used to assess the pre-specified subgroups (differences in observed mean effects) was not as robust as the method used to assess the subgroups defined post hoc (that is, a rank analysis of covariance model [ANCOVA]). The company suggested that the pre-specified subgroup analysis "should not be further used for assessment of robustness and consistency of results in subpopulations".

4.16 The company's additional post-hoc analysis focused on the subgroup of people who had a decline in percent predicted FVC of 10% or more during the first 6 months of treatment; this analysis used pooled data from the 3 pivotal trials. In this subgroup, a further 10% decline in percent predicted FVC was observed in fewer people in the pirfenidone arm compared with the placebo arm (p=0.032) (table 37). Based on these results, the company suggested that pirfenidone should not be stopped if percent predicted FVC declines by 10% or more, because continuing pirfenidone

after disease progression might improve health outcomes compared with best supportive care.

Table 10 Outcomes following previous ≥10% decline in percent predicted FVC (pooled data from ASCEND, CAPACITY 1 and CAPACITY 2)

Outcome, n (%)	Pirfenidone (n=24)	Placebo (n=60)	Relative difference	p value
≥10% decline in FVC or death	1 (4.2%)	15 (25.0%)	−83.3%	0.032
Death	0 (0%)	10 (16.7%)	−100%	0.056
>0% and <10% decline in FVC	9 (37.5%)	23 (38.3%)	-2.2%	NR
No further decline in FVC	14 (58.3%)	22 (36.7%)	59.1	0.089
Abbreviations: FVC, forced vital capacity; NR, not reported Source: table 37 (page 116) company submission				

ERG comments

4.17 The ERG noted that efficacy results were not consistent across the pirfenidone trials; in CAPACITY 1 there was no statistically significant difference between pirfenidone and placebo for the primary outcome, or the secondary outcome of progression-free survival.

4.18 The ERG observed that, across all key trials, the treatment effect of pirfenidone was either not statistically significant or did not meet the threshold for a clinically important difference for a number of clinically important and patient-reported outcomes including mortality, diffusing capacity of the lungs for carbon monoxide and health-related quality of life. The ERG acknowledged that when trial data were pooled or included in the network meta-analysis, the differences between pirfenidone and placebo were statistically significant for mortality and the University of California San Diego Shortness of Breath Questionnaire.

4.19 Regarding mortality, the ERG noted that there was a substantial increase in all-cause mortality in the CAPACITY trials between 52 weeks and 72 weeks, with a smaller increase in the placebo group than in the pirfenidone group (see table 14 on page 72 of the ERG report). The

company did not provide an explanation for these relative increases in mortality.

4.20 The ERG observed no significant difference between people with mild disease and people with moderate disease for the primary endpoint. However, the ERG was unclear whether there was a significant difference in progression-free survival and overall survival according to disease severity because the company did not report the results of a treatment-by-subgroup interaction test.

Meta-analyses

4.21 As there are no head-to-head trials comparing pirfenidone with nintedanib, the company did a network meta-analysis. Based on a systematic review, the company included the following randomised trials in its base case network meta-analysis:

- 4 trials of pirfenidone compared with placebo: ASCEND, CAPACITY 1, CAPACITY 2 and SP3.
- 3 trials of nintedanib compared with placebo: INPULSIS 1, INPULSIS 2 and TOMORROW.

The company also included trials of N-acetylcysteine monotherapy and "triple therapy" (N-acetylcysteine, prednisone and azathioprine) in its network meta-analysis. However, the company did not include this evidence in its cost-effectiveness model (and it is not presented in this document) because the final scope did not include these treatments as comparators. The company excluded the SP2 trial from its network meta-analysis and cost-effectiveness model because Committee B previously considered this trial to be an outlier in the NICE technology appraisal of nintedanib (TA379). In addition, the SP2 trial was stopped at 36 weeks so did not provide data at the time point used in the company's base case network meta-analysis (1 year). The company's base case network meta-analysis used a Bayesian random effects model and data on outcomes after approximately 1 year. The company used results from week 52 for all

studies, even though the primary endpoint for the CAPACITY trials was at 72 weeks. For some outcomes in the CAPACITY trials only week 48 data were available (for example, decline in percent predicted FVC and health-related quality of life).

4.22 The outcomes in the network meta-analysis included: change from baseline in percent predicted FVC (the company included the change in vital capacity from SP3 in its analysis of this endpoint), mortality, progression-free survival, acute exacerbations of idiopathic pulmonary fibrosis, health-related quality of life, and stopping treatment.

4.23 The company did not include adverse events in its original analyses. It performed an additional network meta-analysis in response to clarification (question A39) to compare the rates of diarrhoea, rash, serious cardiac adverse events and stopping treatment because of an adverse event.

4.24 The definition of progression-free survival varied between studies, so the company re-analysed the data from CAPACITY 1 and 2 using the definition of progression-free survival from the ASCEND trial for the network meta-analysis. Progression-free survival was defined as the time until 1 of the following events:

- a decline from baseline in percent predicted FVC of 10% or more
- a confirmed decline from baseline in 6-minute walking distance of 50 metres or more
- death.

SP3 and the 3 nintedanib trials also used different definitions of progression-free survival, but the company did not adjust the data from these trials. The definition of an acute exacerbation also varied between studies; the company used the original trial results in its base case and adjusted the data in a sensitivity analysis.

4.25 The results of the company's base-case network meta-analysis are summarised in Table 11 and Table 12. The NICE technical ream noted that the network meta-analysis suggested:

- Both pirfenidone and nintedanib slow the rate of lung function decline, compared with placebo.
  - There is no evidence of a difference between pirfenidone and nintedanib.
- Pirfenidone reduces all-cause mortality and prolongs progression-free survival compared with placebo; these effects are significant.
  - There is a tendency for nintedanib to reduce mortality and increase progression-free survival compared with placebo, but this is not significant.
  - There is a tendency for pirfenidone to reduce mortality and increase progression-free survival compared with nintedanib, but this is not significant.
- With either pirfenidone or nintedanib, more people stop treatment, have diarrhoea, or have rash compared with placebo (although these effects are generally not significant). There is a tendency for more serious cardiac adverse events with pirfenidone than nintedanib (not significant).
- For acute exacerbations, there is no difference between pirfenidone and nintedanib.
  - The company noted that a limitation of the analysis of acute exacerbations is that the studies defined exacerbations differently.

Table 11 Results from the company's base case network meta-analysis (random effects model): continuous outcomes (mean difference in change from baseline)

Outcome	Mean difference in change from baseline (95% CrI)		
	Pirfenidone versus placebo	Nintedanib versus placebo	Pirfenidone versus nintedanib
FVC%pred[a]	3.39 (1.94 to 4.84)	3.33 (2.34 to 4.5)	0.05 (-1.81 to 1.80)
FVC/VC, litres[a]	0.12 (0.04 to 0.20)	0.12 (0.04 to 0.21)	0.00 (-0.12 to 0.11)
6MWD, metres[a]	22.70 (8.82 to 36.31)	6.00 (−28.25 to 40.66)	16.63 (−20.83 to 53.81)
SGRQ[b]	−1.24 (−4.94 to 2.39)	−2.11 (−5.48 to 0.37)	0.88 (−3.45 to 5.94)
UCSD SOBQ[b]	−3.19 (−6.24 to −0.17)	NA	NA

[a] Numbers greater than zero indicate that people having the intervention showed a smaller decline in outcomes than people having the comparator (favourable for the intervention)
[b] Numbers below zero indicate that people having the intervention showed a smaller decline in outcomes than people having the comparator (favourable for the intervention)

Abbreviations: 6MWD, 6-minute walking distance; CrI, credible interval; FVC, forced vital capacity; FVC%pred, percent predicted FVC; NA, data not available for this comparison; SGRQ, St Georges respiratory questionnaire; UCSD SOBQ, University of California San Diego Shortness of Breath Questionnaire; VC, vital capacity

Source: tables 41, 42 and 51–3 company submission (pages 133–5 and 149–52)

Table 12 Results from the company's base case network meta-analysis (random effects model): hazard ratios for survival outcomes and odds ratios for binomial outcomes

	Pirfenidone compared with placebo	Nintedanib compared with placebo	Pirfenidone compared with nintedanib
Hazard ratio[a] (95% CrI)			
All-cause mortality, week 52	0.52 (0.30 to 0.89)	0.71 (0.41 to 1.27)	0.72 (0.32 to 1.58)
PFS[b], week 52	0.63 (0.50 to 0.80)	0.74 (0.51 to 1.08)	0.85 (0.55 to 1.34)
Odds ratio[a] (95% CrI)			
Acute exacerbations	0.62 (0.29 to 1.39)	0.55 (0.26 to 1.09)	1.14 (0.41 to 3.44)
Stopping treatment (all-cause)	1.28 (0.91 to 1.78)	1.42 (1.01 to 2.01)	0.90 (0.55 to 1.44)
Stopping treatment (because of an AE)	1.58 (1.04 to 2.39)	1.52 (1.01 to 2.29)	1.04 (0.58 to 1.85)
Diarrhoea	1.39 (0.94 to 2.11)	7.32 (4.82 to 11.13)	0.19 (0.11 to 0.35)
Rash	3.85 (2.38 to 6.29)	1.29 (0.49 to 3.35)	2.99 (1.03 to 8.88)
Serious cardiac AE	1.36 (0.54 to 3.46)	0.64 (0.17 to 1.49)	2.11 (0.65 to 11.34)

[a] Hazard ratios and odds ratios below 1 indicate that patients having the intervention had a lower risk of the event than patients having the comparator

[b] Data from the CAPACITY trials re-analysed using the ASCEND definition of PFS; PFS data from SP3 and the nintedanib trials were not re-analysed or adjusted to match ASCEND

Abbreviations: AE, adverse event; CrI, credible interval; PFS, progression-free survival

Source: tables 43–50 and 54 company submission (pages 136–47 and 153) and table 16 company response to clarification (page 37)

ERG comments

4.26 Given that idiopathic pulmonary fibrosis is a progressive disease, and fewer negative outcomes are expected with a shorter follow-up, the ERG's key concern with the network meta-analysis was that the company used data from week 48 or 52 rather than the full 72 week data available. The ERG noted that there was a difference in the treatment effect at these 2 time points. For example, the CAPACITY trials reported smaller benefit with pirfenidone at week 72 than at 1 year for change in percent predicted FVC, mortality, and the University of California San Diego Shortness of Breath Questionnaire. Regarding mortality, the ERG considered that using week 52 data would be appropriate if the purpose of the analysis was to estimate treatment effect at the specified time point, and there was reason to believe that treatment effects may change over the extended follow up period. However, because the purpose of the company's analysis was to estimate the population mean survival time, and the company considered it appropriate to extrapolate the treatment effects over the full lifetime of the cost-effectiveness model (see section 5.23), the ERG considered that the company should have used the full evidence base including 72 weeks follow up.

4.27 The company summarised the uncertainty around the results of its network meta-analysis using 95% credible intervals; the ERG highlighted that the NICE Decision Support Unit recommends that the predictive distribution better represents uncertainty about comparative effectiveness when there is heterogeneity between trials. The company provided the predictive intervals in response to clarification (question A34, see

appendix C of the response), but did not use these in its cost-effectiveness analyses.

4.28 The ERG considered that it may not have been appropriate for the company to combine vital capacity data from SP3 with percent predicted FVC data from ASCEND and the CAPACITY trials. The company justified its methods because there is little difference between these 2 endpoints in people without obstructive airway disease. However, the ERG highlighted that although ASCEND and the CAPACITY trials excluded patients with obstructive airway disease, the exclusion criteria for SP3 are not as clear.

4.29 The ERG had other concerns about SP3 and did not consider it should be included in the base-case network meta-analysis or model because it:

- used an unlicensed dose of pirfenidone
- reported noticeable differences from the other 3 trials in some baseline characteristics (see section 4.4)
- had a higher or unclear risk of bias than other trials
- may have overestimated the treatment effect of pirfenidone because the company imputed missing data using last observation carried forward.

The ERG acknowledged that the lower dose of pirfenidone in SP3 was to account for the lower body weight of the Japanese population, but it noted that no reported dose adjustments were made in the INPULSIS trials of nintedanib despite a high proportion of Japanese participants.

4.30 The ERG considered that the relative effectiveness of pirfenidone compared with nintedanib is uncertain because of the heterogeneity between the trial populations. ASCEND excluded people with a percent predicted FVC above 90%, whereas the nintedanib trials included them. This resulted in a clinically meaningful difference in baseline percent predicted FVC between ASCEND and the INPULSIS trials, suggesting that people in ASCEND may have had more advanced disease. The ERG

suggested that this difference in baseline disease severity may have influenced the number of deaths in the trials and impacted the ability to observe a mortality benefit with nintedanib.

4.31 The ERG did an alternative network meta-analysis to inform its exploratory cost-effectiveness analyses, and reported the predictive distribution. The ERG's analysis used a random effects model, excluded the SP3 study and used data up to 72 weeks for all-cause mortality, progression-free survival and acute exacerbations where available (that is, week 72 data from CAPACITY and week 52 data from ASCEND) (Table 13). For all-cause mortality and progression-free survival, the ERG's results were less favourable than the company's results for pirfenidone compared with placebo (Table 12); hazard ratios for pirfenidone were higher and the predictive intervals crossed 1.

Table 13 Results from the ERG's network meta-analysis of ASCEND, CAPACITY 1 and CAPACITY 2

	Pirfenidone compared with placebo	Nintedanib compared with placebo	Pirfenidone compared with nintedanib
Hazard ratio[a] (95% PrI)			
All-cause mortality, week 72	0.63 (0.32 to 1.28)	0.71 (0.36 to 1.37)	0.9 (0.35 to 2.42)
PFS[b], week 72	0.62 (0.35 to 1.10)	0.74 (0.38 to 1.50)	0.84 (0.34 to 2.02)
Acute exacerbations	0.52 (0.12 to 1.89)	0.57 (0.16 to 1.88)	0.92 (0.13 to 5.310
Stopping treatment (all-cause)	1.24 (0.70 to 2.12)	1.43 (0.83 to 2.43)	0.87 (0.39 to 1.81)

[a] Hazard ratios below 1 indicate that patients having the intervention had a lower risk of the event than patients having the comparator

[b] Data from the CAPACITY trials re-analysed using the ASCEND definition of PFS; PFS data from the nintedanib trials were not re-analysed or adjusted to match ASCEND

Abbreviations: PFS, progression-free survival; PrI, predictive interval

Source: figures 28–32 ERG report (pages 143–5)

Adverse effects of treatment

4.32 The company reported that, in the pivotal phase III trials (ASCEND and CAPACITY 1 and 2), the most frequent adverse events with pirfenidone were gastrointestinal or skin-related. These were mild-to-moderately

severe and rarely led to stopping treatment. Adverse events commonly reported in ASCEND (15% or greater in either treatment group) are presented in Table 14. The adverse events leading to stopping treatment in 1% or more people in the pirfenidone arms of the 3 pivotal trials were pneumonia, rash, raised hepatic enzyme levels and decreased weight (in ASCEND), photosensitivity, rash and respiratory failure (in CAPACITY 1) and bladder cancer, nausea and rash (in CAPACITY 2).

Table 14 Adverse events in ≥15% of patients in either treatment group in ASCEND

Adverse event, n (%)	Pirfenidone (n=278)	Placebo (n=277)
Nausea	100 (36)	37 (13.4)
Rash	78 (28.1)	24 (8.7)
Headache	72 (25.9)	64 (23.1)
Cough	70 (25.2)	82 (29.6)
Diarrhoea	62 (22.3)	60 (21.7)
Upper respiratory tract infection	61 (21.9)	56 (20.2)
Fatigue	58 (20.9)	48 (17.3)
Dizziness	49 (17.6)	36 (13)
Dyspepsia	49 (17.6)	17 (6.1)
Anorexia	44 (15.8)	18 (6.5)
Dyspnoea	41 (14.7)	49 (17.7)
Worsening of idiopathic pulmonary fibrosis	26 (9.4)	50 (18.1)

4.33 The company reported that pirfenidone has a different adverse event profile compared with nintedanib. The most frequently reported adverse reactions associated with nintedanib, reported in its summary of product characteristics, are diarrhoea, nausea and vomiting, abdominal pain, decreased appetite, weight loss and elevation of hepatic enzymes. The company noted that pirfenidone is the only idiopathic pulmonary fibrosis treatment without a special warning or precaution for use in patients at risk of cardiovascular disease or bleeding.

Non-randomised and non-controlled evidence

4.34 The company presented the non-randomised and non-controlled evidence for pirfenidone on pages 156–167 of its submission, including:

- The RECAP open-label extension study: patients who completed ASCEND or either of the CAPACITY trials could enrol in RECAP. In addition to pirfenidone, the study permitted concomitant therapy with corticosteroids, azathioprine, cyclophosphamide, and N-acetylcysteine. The study provides 8.8 years of follow-up data for pirfenidone (based on the latest data cut in June 2015, n=1058; the next data cut is planned for June 2016).
- 3 disease registry studies: INOVA (US-based, n=81), Edinburgh (UK-based, n=323) and EuroIPF (based in Europe, n=409). These registries provide 5–15 years of follow-up data for people receiving best supportive care.

The company used the data from RECAP to estimate long-term survival with pirfenidone. The registry studies informed estimates about long-term survival for people having best supportive care.

4.35 The company compared the relative effectiveness of pirfenidone in the clinical trials with data from the registry studies to understand whether the survival benefit observed in clinical trials of pirfenidone might continue long-term. To do this, the company first selected and matched individuals from registries to people enrolled in the ASCEND and CAPACITY trials, to improve comparability (see pages 160–2 of the company submission). It then generated hazard ratios to show the effectiveness of pirfenidone (using outcomes from the pirfenidone arms of clinical trials) compared with best supportive care (using outcomes from the registries). The company compared the new hazard ratios with those from the clinical trials (Table 15). The company noted that the survival results for pirfenidone compared with best supportive care were better in all 3 registries than in the clinical trials. The company acknowledged that there

are limitations to comparing data from clinical trials with real-world evidence, but suggested that the results indicate that the comparative benefit of pirfenidone observed in the phase III clinical trials are likely to continue long-term.

Table 15 Mortality with pirfenidone compared with placebo/best supportive care: results from clinical trials and registries

	Edinburgh registry	INOVA registry	EuroIPF registry	Pooled CAPACITY and ASCEND data (week 52)
Hazard ratio for mortality[a] (95% CrI)	[redacted]	[redacted]	[redacted]	0.52 (0.31 to 0.88)

[a] Hazard ratios below 1 indicate that patients having pirfenidone had a lower risk of dying than patients having placebo/best supportive care (favourable to pirfenidone)

Abbreviations: CrI, credible interval

Source: table 57 company submission (page 162)

ERG comments

4.36 The ERG commented that the company did not report its methods and criteria for identifying and selecting these non-randomised studies. It was unclear whether additional, relevant evidence might have been excluded.

4.37 The ERG considered that the RECAP open-label extension may have overestimated the survival benefit with pirfenidone, because the study excluded people who were considered to be not taking their medication properly (defined by the company as people who had less than 80% of the assigned study treatment during the clinical trial period).

4.38 The ERG advised considerable caution in the interpretation of the company's comparisons between the pirfenidone arm of clinical trials and data from registries, because the analyses are subject to considerable bias. It was concerned that:

- Despite the attempt by the company to adjust registry data to match clinical trial populations, the survival of people from the registries was

shorter than that of people randomised to placebo in the clinical trials (see figure 35 on page 188 of the ERG report). The company did not comment on this discrepancy.

- There were discrepancies in the inclusion criteria applied to the registry data which may bias the estimate in favour of pirfenidone. For example, the company excluded people from the registry studies if they had a percent predicted FVC of 90% or more. However, approximately 8% of people across the CAPACITY trials and ASCEND had a percent predicted FVC of 90% or more. The ERG considered that excluding this population could underestimate the survival for people having best supportive care, based on the potential link between FVC and mortality.

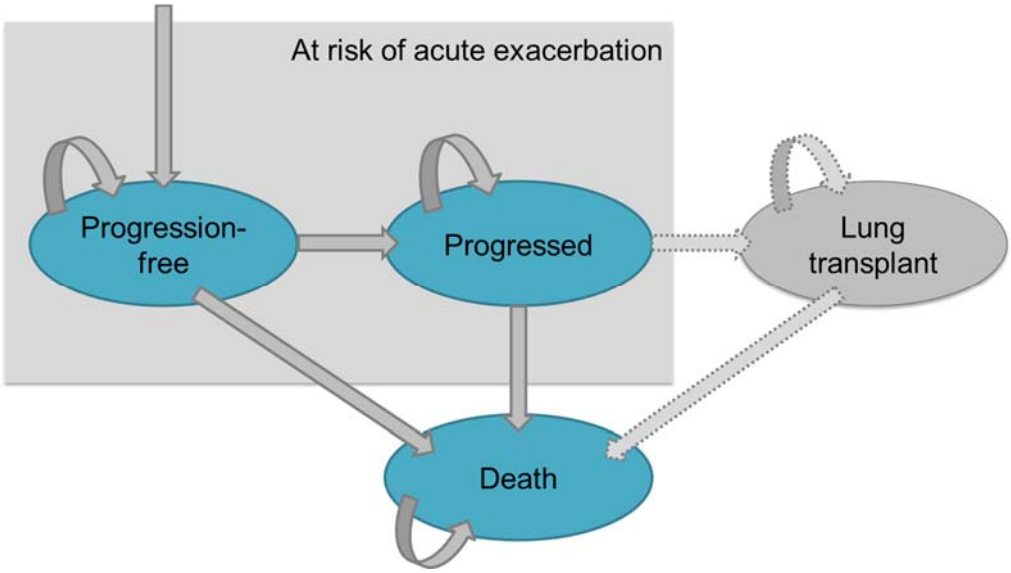
5 Cost-effectiveness evidence

Model structure

5.1 The company provided a partitioned survival model (that is, the proportions of people in each health state were derived from curves of progression-free survival and overall survival; movement between health states is not modelled using transition probabilities). The model assessed the cost effectiveness of pirfenidone compared with nintedanib and best supportive care in adults with mild-to-moderate idiopathic pulmonary fibrosis. The model comprised 3 mutually exclusive health states: progression-free, progressed, and dead. The company included an additional health state for lung transplant in a scenario analysis. The company provided a rationale for its choice of model structure on pages 198–204 of the submission. It modelled people with a percent predicted FVC of 50% or more, based on the population in the pirfenidone clinical trials and marketing authorisation, using a lifetime horizon (34 years) and a 3-month cycle length. The company used an annual discount rate of 3.5%, included a half-cycle correction and calculated costs from the perspective of the NHS and Personal Social Services.

5.2 Patients entered the model in the progression-free health state and either remained in this state or moved into the 'progressed' or 'death' health states. An acute exacerbation did not change the rate of progression; people in the progression-free health state could remain in that health state following an exacerbation. Within the scenario analysis, patients could only enter the lung transplant state from the progressed disease state and if they were under the age of 65, which the company said was in line with eligibility criteria for lung transplantation in clinical practice. Patient characteristics were based on the CAPACITY and ASCEND trials.

Figure 1 Diagram of the company's model (figure 42 on page 204 of the company submission)



5.3 Current NICE guidance on the use of pirfenidone (NICE technology appraisal guidance 282) and nintedanib (NICE technology appraisal guidance 379) recommend that treatment should be stopped if there is evidence of disease progression (a decline in percent predicted FVC of 10% or more within any 12 month period). The company did not apply this 'stopping rule' to the pirfenidone arm of its base-case model, because it considered that people who experience disease progression may continue to benefit from pirfenidone treatment (see section 4.15 and Table 10). The company did apply the stopping rule for nintedanib in the base-case

analysis, to align with current practice, and explored the use of a stopping rule for pirfenidone in a scenario analysis (scenario 31).

ERG comments

5.4 The ERG had a number of major concerns regarding the structure and logic of the company's model. Firstly, the ERG considered that separating idiopathic pulmonary fibrosis into 2 distinct phases (before and after disease progression) is overly simplistic and does not reflect the natural history of the condition. The ERG believes that the company's approach may have overestimated the lifetime gains in quality-adjusted life years (QALYs) with pirfenidone, because the impact of subsequent disease progression for people already in the "progressed" health state is not captured. The ERG investigated the impact of adjusting utility by age, to represent further progression for people in the progressed health state, in its exploratory analyses; it had only a minor impact on the ICER. The ERG also highlighted that, in using a partitioned survival structure, the company assumed that all disease progression is equally detrimental. The ERG suggested that the impact of disease progression on health-related quality of life would differ according to the person's baseline disease severity. For example, someone whose percent predicted FVC decreased from 90% to 80% would have a different decline in quality of life to someone whose percent predicted FVC decreased from 60% to 50%. That is, the model does not capture the impact of different levels of disease severity on quality of life and costs.

5.5 Secondly, the ERG was concerned that using a partitioned-survival model means that overall survival, progression-free survival and stopping treatment are modelled independently of each other. That is, a change in either progression-free survival or time on treatment does not impact overall survival, but does impact treatment costs. The ERG considered this approach to be reasonable when the treatment 'stopping rule' in current NICE guidance is excluded. However, the ERG did not consider the partitioned-survival model, in which treatment duration and treatment

outcomes are disconnected, to be appropriate when the stopping rule is included in the analysis. The ERG also highlighted that the absence of a relationship between these outcomes is not clinically plausible. The ERG advised that, if the company had assumed some dependence between these outcomes, the results may have been more sensitive to the company's choice of parametric curves and its assumption that people remain on treatment regardless of progression status (see sections 5.26 and 5.31).

5.6 Thirdly, the ERG was concerned that the impact of acute exacerbations was underestimated as a result of the model structure, in which acute exacerbations were disconnected from the outcomes of progression and survival, and were associated only with costs and utility reductions. The ERG advised that exacerbations have substantial impact on quality of life, morbidity and mortality and it was therefore concerned that the results of the model were not sensitive to changes in the rate of acute exacerbations.

5.7 Regarding the 'stopping rule' recommended in current NICE guidance for pirfenidone, which the company modelled in a scenario analysis, the ERG's clinical advisors agreed with the company's statement that a prior decline in lung function does not predict a future decline in lung function. The advisors noted that this statement is supported by a retrospective analysis of pulmonary function tests from 734 patients in the UK and US.

Model details

5.8 The company derived the rates of progression and mortality with pirfenidone from patient-level data from the pivotal phase III studies (ASCEND and CAPACITY 1 and 2). Mortality rates were also informed by patient-level data from the open-label extension study of pirfenidone (RECAP, 7-year data cut). For other pirfenidone outcomes, the company used the results of the network meta-analysis (see section 4.21). The company used results from the network meta-analysis to estimate the

relative effectiveness and safety of nintedanib and best supportive care compared with pirfenidone.

5.9 To estimate overall survival for people having pirfenidone, the company fitted a Weibull parametric model to the clinical trial data (see pages 211–4 of the company submission). The company explored alternative survival curves for pirfenidone in scenario analyses (scenarios 12–16; Appendix 21 of company submission).

5.10 To estimate overall survival for people having nintedanib and best supportive care, the company took the hazard ratios from the network meta-analysis and applied them to the curve predicting overall survival with pirfenidone (1.33 for nintedanib and 1.92 for best supportive care, compared with pirfenidone). The company explored alternative methods of estimating overall survival with best supportive care in 3 scenario analyses (scenarios 17–19):

- using trial data up to week 52 followed by applying the hazard ratio
- using real-world data from the UK-based and US-based registries Edinburgh (n=182) and INOVA (n=286).

5.11 The company's model distinguished between deaths that were related or unrelated to idiopathic pulmonary fibrosis, because it considered that disease-related costs increase substantially in the last year of life, attributable to increased routine healthcare use, home care and length of stay in hospital. The company updated its estimates of disease-related deaths in response to clarification (questions B9 and B10): approximately 53% with pirfenidone, 62% with nintedanib and 70% with best supportive care.

5.12 To estimate progression-free survival, the company used the same approach as it had for overall survival. That is, it used a Weibull distribution for the pirfenidone curve, with hazard ratios applied for nintedanib (1.18) and best supportive care (1.59). The definition of

progression-free survival was based on the definition in ASCEND (see section 4.24). The company explored alternative parametric distributions for pirfenidone in scenario analyses (pages 215–7 of company submission) but registry data were not available for progression-free survival.

5.13 To estimate time to stopping treatment for reasons other than death and lung transplant (most commonly because of adverse events), the company again selected a Weibull distribution for pirfenidone and applied a relative risk from the network meta-analysis for nintedanib. As noted in section 5.3, in its base-case analysis the company applied the 'stopping rule' recommended in NICE guidance (that is, if there is evidence of disease progression) for the nintedanib arm but not the pirfenidone arm. The company modelled stopping treatment because of disease progression separately to stopping for other reasons. The company assumed that people received best supportive care after stopping treatment with pirfenidone or nintedanib.

5.14 The baseline risk of an acute exacerbation in the model, applied to people having best supportive care, was 1.46% per model cycle (based on the company submission for the NICE technology appraisal of nintedanib). The company applied odds ratios from the network meta-analysis to calculate the risk of acute exacerbation for people having pirfenidone (0.62) and nintedanib (0.55). Because of lack of evidence on the incidence of recurrent events, the company assumed that people who experienced at least 1 exacerbation had the same risk of another exacerbation as people that had never had an exacerbation.

5.15 The risk of adverse events in the model was informed by a network meta-analysis performed by the company in response to clarification. The company included serious cardiac events, diarrhoea, gastrointestinal perforation, photosensitivity and rash.

5.16 The company assigned utility values to each health state by mapping the St. George's Respiratory Questionnaire data (collected every 12 weeks in the CAPACITY trials) onto the EQ-5D using a mapping algorithm by Freemantle (2015). EQ-5D data were not collected in the pirfenidone trials. In a scenario analysis the company used an alternative mapping algorithm by Starkie (2012). Utility values within each health state remained constant.

Table 16 Utility values applied to health states in the company's base case model

Health state	Utility value	Source
Progression-free	0.847	CAPACITY trial data using St. George's Respiratory Questionnaire, mapped onto EQ-5D using algorithm by Freemantle
Progressed	0.782	
Lung transplant[a]	0.847	Assumed equivalent to progression-free disease
[a] Lung transplant health state included only as a scenario analysis Source: table 88 company submission		

5.17 The company included the following costs in its model: drug treatments, treating adverse events, routine healthcare (including patient monitoring, liver-function tests and oxygen use), lung transplant (scenario analysis only), treating acute exacerbations in hospital, and end of life care. The company calculated costs using the 2014–15 NHS reference costs and the 2015 Personal Social Service Research Unit (PSSRU) Unit Costs of Health & Social Care. Costs within each health state remained constant.

5.18 The company assumed that patients took 7.88 pills of pirfenidone per day on average, based on data from the CAPACITY trials and ASCEND, leading to a daily cost of £62.80 (or a cost per 3-month model cycle of £5,730.62) using the list price. The summary of product characteristics for pirfenidone recommends 9 pills per day. The company assumed the same daily cost for nintedanib as for pirfenidone. It assumed there were no administration costs for pirfenidone and nintedanib because they are oral treatments. The company did not include any costs for best supportive

care. The model inputs are summarised in table 103 of the company submission.

5.19 The company did not explore the sequential use of treatments for idiopathic pulmonary fibrosis in its model (that is, pirfenidone followed by nintedanib or vice versa). The company stated that there is no evidence about the efficacy or safety of sequencing these treatments.

Subgroup analyses

5.20 The company presented results for 3 populations:

- all randomised patients in the pirfenidone trials; the comparator was best supportive care
- the subgroup of people with mild idiopathic pulmonary fibrosis (percent predicted FVC over 80%); the comparator was best supportive care
- the subgroup of people with moderate disease (percent predicted FVC between 50% and 80%); the comparators were best supportive care and nintedanib.

The company's choice of comparators was based on NICE technology appraisal 379, which recommends nintedanib only for people with percent predicted FVC between 50% and 80%. Because the company did not have access to patient-level data for nintedanib, it assumed that the treatment effect for nintedanib was the same regardless of percent predicted FVC. The NICE technical team noted that post-hoc subgroup analyses of the INPULSIS trials, presented by the company in the NICE technology appraisal of nintedanib (TA379), showed that there were no statistically significant differences in the primary or key secondary outcomes by subgroups according to percent predicted FVC.

5.21 The company applied subgroup-specific data for the baseline risk of death, disease progression and stopping treatment. For overall survival and time on treatment, the company fitted a Weibull parametric model to the trial data, using percent predicted FVC as covariates (<50%, 50–80% and >80%). For progression-free survival, the parametric distribution was

fitted to the observed data for each subgroup separately. For other model parameters, the company used the same estimates as those used for all randomised patients.

Changes to company model before first committee meeting

5.22 The company provided updated cost-effectiveness results for the base case model and all sensitivity and scenario analyses, for all 3 populations, in its response to clarification (refer to section B, pages 38–9, and Appendix E, pages 129–66, of the response to clarification). The company also updated its patient access scheme submission. The updated results:

- Included the network meta-analysis of adverse events (see clarification question A39)
- Corrected an error in implementing the stopping rule for nintedanib (see clarification question B8)
- Revised estimates of mortality related to idiopathic pulmonary fibrosis (see clarification questions B9 and B10)
- Corrected an error in the oxygen cost (see clarification question B23).

ERG comments

5.23 The ERG's main concerns about the company's modelling assumptions related to the estimations of treatment effect. The ERG considered that the evidence base did not support the company's assumption that the survival benefit from the pirfenidone trials (based on data up to 52 weeks) is constant over the entire model duration (34 years); changing this assumption of lifetime proportional hazards between pirfenidone and its comparators had the biggest impact on the ICER. The ERG acknowledged that the company's post hoc analyses did not show a significant interaction between the treatment effect and time, but it was concerned that the CAPACITY trials reported a smaller survival benefit for pirfenidone (compared with best supportive care) at week 72 than at week 52. The ERG's plot of overall survival data from ASCEND, RECAP and the CAPACITY trials (see figure 34 on page 175 of the ERG report) also suggested that treatment effect is not maintained. Finally, the ERG

considered that the company's comparison of pirfenidone trial data with registry data (which, according to the company, supported the assumption of long-term proportional hazards compared with best supportive care) was subject to considerable bias (see section 4.38). The ERG's clinical advisors commented that it is possible that if a drug fundamentally changes disease progression over the duration of a clinical trial, then continued treatment may prevent decline over longer time periods. The ERG considered that this supports the possibility of continued effectiveness but does not necessarily support a treatment effect that is constant over the entire model duration. The ERG noted that the assumption of proportional hazards between pirfenidone and best supportive care for progression-free survival was more appropriate.

5.24 Regarding the comparison between pirfenidone and nintedanib, the ERG considered that the company's assumption that pirfenidone was superior to nintedanib, and that this benefit was maintained for the entire model duration, was overly optimistic. The ERG noted that the efficacy of nintedanib and pirfenidone could be similar, and that results of the network meta-analyses were uncertain given the considerable heterogeneity between the populations included in the trials for pirfenidone and nintedanib (see section 4.30).

5.25 The ERG had other concerns about the company's estimation of treatment effect, which were also important drivers of the ICER. For overall survival and progression-free survival, the ERG considered that the company should have:

- used trial data up to week 72 instead of week 52 (see section 4.26); this was less important for progression-free survival, which had a minimal impact on the company's base-case ICER
- excluded the SP3 trial (see section 4.29)
- used the predictive distribution from the network meta-analysis rather than credible intervals (see section 4.27)

The ERG applied these suggestions in its alternative base-case analysis. For time to stopping treatment, the ERG noted that the company's network meta-analysis produced an odds ratio for stopping treatment due to any reason, but the company applied this in the model for people who stopped treatment for reasons other than death and lung transplant. The ERG suggested that this may have introduced bias if the rates of death or lung transplant differed between the trial arms compared in the network meta-analysis.

5.26 Another key driver of the ICER in the company model was the choice of parametric curve for estimating overall survival, progression-free survival and time to stopping treatment. The ERG considered that, although the Weibull curve (selected by the company) fitted the observed data well, the Gompertz distribution may provide a more clinical plausible long-term extrapolation for these outcomes (see section 5.2.3 on pages 177–85 of the ERG report). The ERG did not agree with the company using registry data to justify using the Weibull curve given the ERG's concerns about these data, such as the fact that the survival outcomes from the registries did not match the placebo arm of the trials (see section 4.38). In addition, the ERG highlighted that the probability of death for older people in the model, when using the Weibull distribution, was lower than in the general UK population; the ERG did not consider this to be clinically plausible. The ERG noted that the model was more sensitive to the choice of parametric curve for overall survival than for progression-free survival or time to stopping treatment, but suggested that the latter 2 outcomes may have had more impact if the relationship between all 3 outcomes had been modelled more realistically.

5.27 The ERG evaluated the company's algorithm for mapping the trial-based results from the St. George's Respiratory Questionnaire data onto the EQ-5D, to generate utility values for the model's health states. The ERG considered that the company's method was generally appropriate but noted that, under some circumstances, the algorithm predicted utility values exceeding the maximum possible value of 1. In response to

clarification question B18, the company capped the maximum possible utility value at 1.0, which decreased the utility values for the progression-free health state (from 0.85 to 0.82) and progression health state (from 0.78 to 0.76). The ERG used these new utility values in its exploratory analyses.

5.28 The ERG had a few issues with the company's cost estimates:

- The company calculated dose interruptions and reductions for pirfenidone after the 2-week titration period. The ERG considered that it would have been more appropriate to separate the costs for the first model cycle from those for subsequent cycles. This was amended in the ERG's alternative base case.
- The ERG suggested that the company's assumption of the same daily costs for pirfenidone and nintedanib, based on the dose reductions and interruptions observed in the pirfenidone trials, was likely to be unfavourable to pirfenidone. The ERG used the dose reductions and interruptions observed in the INPULSIS trials of nintedanib in its alternative base case.
- The ERG considered that routine healthcare costs are likely to increase over time within each health state, or at least in the progressed health state, rather than remain constant. However, the ERG was unclear how changing this would affect the ICER.

5.29 The ERG had a number of concerns with the company applying a one-off cost to deaths attributable to idiopathic pulmonary fibrosis (£9,996), but not including costs to deaths unrelated to the disease. The ERG:

- considered that it would have been more appropriate to use the results of the network meta-analysis to estimate the proportion of deaths that were disease-related
- was uncertain whether the company's estimate for end-of-life costs, which was based on the costs associated with people dying from organ

failure estimated in a modelling study, was representative of clinical practice for idiopathic pulmonary fibrosis

- considered that deaths from causes other than idiopathic fibrosis are also likely to be associated with costs, which the company did not include.

In its exploratory analyses, the ERG applied the cost associated with end-of-life care to all deaths in the model, irrespective of the cause.

5.30 The ERG suggested that it was possible for patients to have nintedanib after stopping pirfenidone treatment, and vice versa, noting that the company did not model these options. The company explained that there was insufficient clinical or safety evidence to model treatment sequencing, and the ERG was unclear what the impact on the ICER would be.

5.31 The ERG did not consider that the company's assumption that people whose disease progresses continue to have pirfenidone or nintedanib was supported by the available evidence. However, the ERG suggested that the ICER would not change if the company used different stopping rates according to whether patients had progressed or not, because of the partitioned-survival model structure, in which time on treatment is used to calculate the treatment costs but had no impact on health outcomes.

5.32 Regarding the company's subgroup analyses, the ERG considered it appropriate that treatment effects from all randomised patients were applied to the subgroups, because subgroups were not stratified in the trial and analyses of outcomes were post hoc. However, because it was not possible to rule out a different treatment effect by subgroup, the ERG's exploratory analyses investigate the impact of using the treatment effects by subgroups from the company's post hoc analyses (see section 4.14). The ERG did not consider it appropriate for the company to assume no differences between the subgroups in utility or routine healthcare costs, given that these parameters would be expected to change with disease severity, but was unsure of the effect of this on the ICER.

Company's base-case results and sensitivity analysis

5.33 This document presents results using the confidential patient access scheme discount for pirfenidone. Nintedanib also has a patient access scheme for which the level of discount is confidential and not known to Roche Products. Accordingly, the Evidence Review Group ran the company's analyses using the patient access scheme discounts for both pirfenidone and nintedanib; the results are in the confidential appendix to this document.

5.34 The company's base case deterministic cost-effectiveness analysis showed that pirfenidone was more costly and more effective than best supportive care in all randomised patients, resulting in an incremental cost-effectiveness ratio (ICER) of £21,387 per QALY gained using the patient access scheme discount for pirfenidone (Table 17). The results of the company's probabilistic sensitivity analysis were very similar, producing an ICER for pirfenidone of £20,794 per QALY compared with supportive care. Results from the probabilistic sensitivity analysis indicated that the probability of pirfenidone being cost-effective compared with best supportive care was 45% and 85% at willingness to pay thresholds of £20,000 and £30,000 per QALY gained, respectively. The cost-effectiveness results for the subgroups according to percent predicted FVC were as follows:

- In the mild population (percent predicted FVC above 80%): £24,187 (deterministic) or £23,476 (probabilistic) per QALY gained compared with best supportive care.
- In the moderate population (percent predicted FVC 50–80%): £21,318 (deterministic) or £20,863 (probabilistic) per QALY gained compared with best supportive care. The results compared with nintedanib are presented in the confidential appendix to this document.

Table 17 The company's base case cost-effectiveness results for pirfenidone compared with best supportive care (using the discount in the patient access scheme for pirfenidone)

	Total cost (£)	Total QALYs	Inc. cost (£)	Inc. QALYs	Pairwise ICER vs. BSC (£/QALY)
All randomised patients					
BSC	26,627	3.80			
Pirfenidone	66,638	5.67	40,010	1.87	21,387
Mild idiopathic pulmonary fibrosis (% predicted FVC >80%)					
BSC	31,729	4.82			
Pirfenidone	84,209	6.99	52,480	2.17	24,187
Moderate idiopathic pulmonary fibrosis (% predicted FVC 50–80%)[a]					
BSC	24,868	3.44			
Pirfenidone	61,012	5.14	36,145	1.70	21,318

[a] The results for pirfenidone compared with nintedanib are presented in the confidential appendix to this document

Abbreviations: BSC, best supportive care; FVC, forced vital capacity; ICER, incremental cost effectiveness ratio; Inc., incremental; QALY, quality-adjusted life year

Source: tables 4, 6 and 8 company PAS submission template

5.35 The company performed one-way sensitivity analyses using the 95% confidence interval values of all model parameters. The parameter which had the biggest impact on the ICER was the hazard ratio for mortality (see figures 1–4 of the company PAS submission template).

5.36 The model predicted that pirfenidone was associated with an additional 3.29 years of life compared with best supportive care in all randomised patients; this increased to 4.15 years in the mild population and decreased to 2.87 in the moderate population. Compared with nintedanib, pirfenidone provided an additional 1.61 years of life (moderate population). The company stated that the survival benefit predicted by the model was similar to that observed in clinical trials of pirfenidone and registry studies (Table 18).

Table 18 Percentage of people alive over time, comparison of company model with clinical trials and registry data

Time	Clinical trial result		Registry data		Model result (includes half cycle correction)	
			INOVA	Edinburgh		
	Pirfenidone	BSC	BSC	BSC	Pirfenidone	BSC
1 year	96.4%	93.2%	89.8%	88.0%	95.5%	91.5%
2 years	87.7%	-	74.4%	75.9%	88.8%	79.5%
5 years	70.4%	-	45.6%	43.6%	65.9%	44.9%
7 years	56.8%	-	38.6%	32.2%	51.7%	28.1%

BSC, best supportive care
Source: table 108 (page 258) company submission

Company scenarios

5.38 The company performed 58 scenario analyses for all 3 populations (summarised in table 104 on page 249 of the company submission). The biggest drivers of the ICER were the time horizon, the duration over which the treatment effect remains constant, the parametric distributions for overall survival in people randomised to pirfenidone, and the inclusion of the stopping rule.

- Time horizon: costs were accrued over a relatively short time period, with the long-term benefits of treatment accrued over the longer-term; the majority were captured within 15 years; shortening the time horizon increased the ICER in all 3 populations modelled.
- Overall survival: the company explored fitting different parametric models to the trial data for pirfenidone; the ICER increased when a Gompertz distribution was used, and decreased with all other alternatives (in all 3 populations modelled). When the company used overall survival results from week 72 instead of week 52 of the clinical trials, the ICER increased in all 3 populations.
- Duration of treatment effect: the company explored the impact of removing the treatment effect for pirfenidone at 7 years (which marks the end of available trial data for pirfenidone), compared with a continued treatment effect through to 10 and 14 years; the ICER increased in all scenarios, for all 3 populations modelled.

- Stopping rules: including a stopping rule for pirfenidone (that is, people stop treatment if their percent predicted FVC declines by 10% or more within any 12 month period) reduced the ICER substantially in all 3 populations modelled.

The changes in all other scenarios (including the addition of the lung transplant health state) had a minimal impact on the ICER. The overall range of ICERs compared with best supportive care, using the patient access scheme price for pirfenidone, was £14,847 to £31,540 per QALY gained. The results of the key scenario analyses described above are presented in Table 19. The results of other scenario analyses, and results for the 2 subgroups, are presented in tables 16–18 of the company's PAS submission template (pages 25–40).

Table 19 Cost-effectiveness results from the company scenario analyses with the biggest impact on the ICER, all randomised patients

Scenario	Total cost (£)	Total QALYs	Inc. cost (£)	Inc. QALYs	Pairwise ICER vs. BSC (£/QALY)
Base case	66,638	5.67	40,010	1.87	21,387
Time horizon					
10 years	60,683	4.72	35,737	1.13	31,540
15 years	64,678	5.34	38,338	1.58	24,300
Overall survival					
Gompertz model	64,362	5.20	38,366	1.51	25,360
72 week data	66,638	5.67	37,766	1.44	26,309
Treatment effect					
Up to 7 years	66,638	5.67	37,985	1.48	25,776
Up to 10 years	66,638	5.67	39,218	1.72	22,865
Up to 14 years	66,638	5.67	39,815	1.83	21,731
Stopping rule for PFN	54,360	5.66	27,733	1.86	14,847

Abbreviations: BSC, best supportive care; ICER, incremental cost effectiveness ratio; Inc., incremental; QALY, Quality adjusted life year; OS, overall survival

Source: table 16 company PAS submission template (page 25)

ERG comments

5.39 The ERG considers that it is more appropriate to conduct an economic analysis separately within the mild and moderate subgroups, as the comparators vary by subgroup, rather than considering all randomised patients (which does not consider nintedanib as a comparator).

5.40 As described in section 5.23, the ERG considered that the company's base-case model produced a favourable estimate of the ICER for pirfenidone because the company assumed that treatment benefit of pirfenidone was maintained for the model duration. The ERG noted that in the scenario analyses where the company assumed treatment effects stops after 7, 10 or 14 years, the ICER increased. However, the ERG considered that the estimates of the duration of treatment effect used in the scenario analyses were arbitrary, and in its preferred based case it used an optimistic and pessimistic assumption:

- optimistic: lifetime treatment effect, as in the company base case
- pessimistic: treatment effect stops at 2 years, approximately at the end of the randomised clinical trial evidence.

5.41 The ERG considered that the company's scenario analysis including a stopping rule for pirfenidone probably underestimated the ICER, because treatment duration and treatment outcomes were disconnected in the company's model (see section 5.5). This meant that the life-time costs of treatment were reduced when the stopping rule was applied, but the incremental QALYs were not reduced by the shorter duration of treatment.

5.42 The ERG highlighted discrepancies between the model's prediction of overall survival for people having best supportive care and the observed trial data for patients who were randomised to placebo. The ERG noted that the company did not comment on these discrepancies. In addition, the ERG considered that the company's comparison between the model predictions with registry data for patients receiving best supportive care

was inappropriate because the registry data does not match the trial data for people randomised to placebo (see section 4.38).

ERG exploratory analyses

5.43 The ERG explained that it could not address all of the issues it had identified in the company model without substantially restructuring it. It also noted that changing the model was challenging because the outcomes were disconnected from each other. The ERG was not able to adequately amend the stopping rule within the company's existing model structure and therefore suggested that ICERs from scenario analyses exploring the stopping rule should be interpreted with caution. The ERG considered that the ICERs involving the stopping rule are likely to be favourable to pirfenidone when compared with best supportive care.

5.44 The ERG explored the following changes in 10 individual scenario analyses (presented individually in tables 69 and 71 of the ERG report, and tables 3–8 and of the ERG confidential appendix):

- Estimating treatment effect using the predictive distribution instead of the median hazard ratio from the network meta-analysis, after excluding SP3, based on trial data up to week 72. This scenario required probabilistic, rather than deterministic, analyses to incorporate the predictive distribution.
- Stopping the treatment effect of pirfenidone after 2 years (approximately the end of follow-up of the randomised clinical trial evidence).
- Including a stopping rule for pirfenidone (that is, people stop treatment if their percent predicted FVC declines by 10% or more within any 12 month period).
- Modelling overall survival using the Gompertz distribution (rather than the Weibull).
- Capping utility estimates at 1.0.

- Adjusting utility by age, based on the change in utility values observed in the general population using data from the Health Survey for England (Ara 2010).
- Including costs associated with end of life for all people irrespective of the cause of death.
- Including titration in the first cycle based on data provided by the company at the clarification stage (a different dose intensity is used between the first and subsequent cycles).
- Changing the mean dose of nintedanib, based on that observed in the INPULSIS trials of nintedanib after dose reductions or interruptions (96.4% of the indicated dose)
- Correcting minor programming errors in the company model (described in appendix 4 of the ERG report).

5.45 The ERG's individual exploratory analyses showed that the ICERs were sensitive to 4 changes:

- the duration of extrapolation of the treatment effect of pirfenidone (largest impact)
- including a stopping rule for pirfenidone (large impact but ERG suggest that results should be interpreted with caution)
- using the ERG's preferred estimate of the treatment effect (predictive distribution using week 72 data and excluding SP3) (moderate to large impact)
- using the Gompertz rather than the Weibull distribution for overall survival (moderate impact).

The results of these 4 individual scenarios, for pirfenidone compared with best supportive care, are presented in Table 20 to Table 23. The ERG's results for pirfenidone compared with nintedanib are presented in the confidential appendix to this document.

Table 20 ERG exploratory analysis: treatment effect assumed to stop after 2 years (deterministic analysis)

	Total cost (£)	Total QALYs	Inc. cost (£)	Inc. QALYs	Pairwise ICER vs. BSC (£/QALY)
All randomised patients					
BSC	33,798	5.215			
Pirfenidone	66,638	5.667	32,840	0.452	72,599
Mild idiopathic pulmonary fibrosis (% predicted FVC >80%)					
BSC	40,671	6.606			
Pirfenidone	84,209	6.994	43,539	0.388	112,214
Moderate idiopathic pulmonary fibrosis (% predicted FVC 50–80%)[a]					
BSC	31,180	4.690			
Pirfenidone	61,035	5.138	29,854	0.449	66,503

[a] The results for pirfenidone compared with nintedanib are presented in the confidential appendix to this document

Abbreviations: BSC, best supportive care; FVC, forced vital capacity; ICER, incremental cost effectiveness ratio; Inc., incremental; QALY, quality-adjusted life year

Source: table 76 (page 243) ERG report

Table 21 ERG exploratory analysis: including a stopping rule for pirfenidone (deterministic analysis)

	Total cost (£)	Total QALYs	Inc. cost (£)	Inc. QALYs	Pairwise ICER vs. BSC (£/QALY)
All randomised patients					
BSC	26,627	3.797			
Pirfenidone	54,360	5.664	27,733	1.868	14,847
Mild idiopathic pulmonary fibrosis (% predicted FVC >80%)					
BSC	31,729	4.824			
Pirfenidone	65,740	6.989	34,011	2.165	15,707
Moderate idiopathic pulmonary fibrosis (% predicted FVC 50–80%)[a]					
BSC	24,868	3.443			
Pirfenidone	50,596	5.136	25,728	1.693	15,197

[a] The results for pirfenidone compared with nintedanib are presented in the confidential appendix to this document

Abbreviations: BSC, best supportive care; FVC, forced vital capacity; ICER, incremental cost effectiveness ratio; Inc., incremental; QALY, quality-adjusted life year

Source: table 78 (page 245) ERG report

Table 22 ERG exploratory analysis: ERG's preferred estimate of the treatment effect (probabilistic analysis)

	Total cost (£)	Total QALYs	Inc. cost (£)	Inc. QALYs	Pairwise ICER vs. BSC (£/QALY)
All randomised patients					
BSC	29,694	4.393			
Pirfenidone	66,685	5.672	36,991	1.279	28,922
Mild idiopathic pulmonary fibrosis (% predicted FVC >80%)					
BSC	35,220	5.520			
Pirfenidone	84,133	6.999	48,913	1.480	33,060
Moderate idiopathic pulmonary fibrosis (% predicted FVC 50–80%)[a]					
BSC	27,683	3.995			
Pirfenidone	61,097	5.157	33,414	1.162	28,766

[a] The results for pirfenidone compared with nintedanib are presented in the confidential appendix to this document

Abbreviations: BSC, best supportive care; FVC, forced vital capacity; ICER, incremental cost effectiveness ratio; Inc., incremental; QALY, quality-adjusted life year

Source: table 79 (page 246) ERG report

Table 23 ERG exploratory analysis: Gompertz distribution for overall survival (deterministic analysis)

	Total cost (£)	Total QALYs	Inc. cost (£)	Inc. QALYs	Pairwise ICER vs. BSC (£/QALY)
All randomised patients					
BSC	25,996	3.687			
Pirfenidone	64,362	5.200	38,366	1.513	25,360
Mild idiopathic pulmonary fibrosis (% predicted FVC >80%)					
BSC	30,124	4.520			
Pirfenidone	79,543	6.094	49,420	1.575	31,379
Moderate idiopathic pulmonary fibrosis (% predicted FVC 50–80%)[a]					
BSC	24,430	3.374			
Pirfenidone	59,276	4.776	34,846	1.402	24,855

[a] The results for pirfenidone compared with nintedanib are presented in the confidential appendix to this document

Abbreviations: BSC, best supportive care; FVC, forced vital capacity; ICER, incremental cost effectiveness ratio; Inc., incremental; QALY, quality-adjusted life year

Source: table 77 (page 244) ERG report

5.46 The results of the ERG's probabilistic alternative base case, which combines the changes from all of its exploratory analyses, are presented in Table 24 and Table 25 (for the comparison between pirfenidone and best supportive care; comparisons with nintedanib are presented in the confidential appendix to this document). The ERG presented its alternative base case as a range (most optimistic to most pessimistic scenario) given the uncertainty about the extrapolation of the treatment effect of pirfenidone (assuming either lifetime or 2 years). When the ERG included the stopping rule for pirfenidone, the probabilistic ICERs for pirfenidone compared with best supportive care ranged from:

- In all randomised patients: £27,124–£75,121 per QALY gained
- In the mild population (percent predicted FVC above 80%): £31,722–£113,365 per QALY gained
- In the moderate population (percent predicted FVC 50–80%): £27,432–£70,234 per QALY gained.

The ERG also presented the results of its alternative base case without a treatment stopping rule for either pirfenidone or nintedanib. When the stopping rules were removed, the probabilistic ICERs for pirfenidone compared with best supportive care ranged from:

- In all randomised patients: £39,895–£115,751 per QALY gained
- In the mild population (percent predicted FVC above 80%): £49,921–£186,260 per QALY gained
- In the moderate population (percent predicted FVC 50–80%): £39,166–£104,915 per QALY gained.

Table 24 ERG alternative base case, pessimistic assumption for duration of treatment effect (2 years) including stopping rule for pirfenidone (probabilistic analysis)

	Total cost (£)	Total QALYs	Inc. cost (£)	Inc. QALYs	Pairwise ICER vs. BSC (£/QALY)
All randomised patients					
BSC	34,430	4.610			
Pirfenidone	57,048	4.911	22,618	0.301	75,121
Mild idiopathic pulmonary fibrosis (% predicted FVC >80%)					
BSC	39,063	5.501			
Pirfenidone	66,794	5.745	27,731	0.245	113,365
Moderate idiopathic pulmonary fibrosis (% predicted FVC 50–80%)[a]					
BSC	32,081	4.20			
Pirfenidone	53,249	4.50	21,169	0.30	70,234

[a] The results for pirfenidone compared with nintedanib are presented in the confidential appendix to this document

Abbreviations: BSC, best supportive care; FVC, forced vital capacity; ICER, incremental cost effectiveness ratio; Inc., incremental; QALY, quality-adjusted life year

Source: table 83 (page 250) ERG report

Table 25 ERG alternative base case, optimistic assumption for duration of treatment effect (lifetime) including stopping rule for pirfenidone (probabilistic analysis)

	Total cost (£)	Total QALYs	Inc. cost (£)	Inc. QALYs	Pairwise ICER vs. BSC (£/QALY)
All randomised patients					
BSC	30,947	3.964			
Pirfenidone	57,216	4.932	26,269	0.968	27,124
Mild idiopathic pulmonary fibrosis (% predicted FVC >80%)					
BSC	35,035	4.757			
Pirfenidone	66,796	5.759	31,761	1.001	31,722
Moderate idiopathic pulmonary fibrosis (% predicted FVC 50–80%)[a]					
BSC	29,225	3.64			
Pirfenidone	53,790	4.53	24,565	0.90	27,432

[a] The results for pirfenidone compared with nintedanib are presented in the confidential appendix to this document

Abbreviations: BSC, best supportive care; FVC, forced vital capacity; ICER, incremental cost effectiveness ratio; Inc., incremental; QALY, quality-adjusted life year

Source: table 82 (page 249) ERG report

Innovation

5.47 The company provided justifications for considering pirfenidone to be innovative:

- Pirfenidone was the first treatment with a marketing authorisation for idiopathic pulmonary fibrosis, and so represented a step change in the management of the condition.
- The pre-specified pooled analysis of ASCEND and the CAPACITY trials (n=1247) demonstrated pirfenidone to be the first and only treatment to improve survival for people with idiopathic pulmonary fibrosis; hazard ratio for pirfenidone compared with placebo: 0.52 (95% confidence interval, 0.31 to 0.87; p=0.011). Long term data (RECAP) support the effect on mortality.
- The clinical benefit with pirfenidone is comparable in people with earlier (percent predicted FVC above 80%) and more advanced disease (FVC below 80%).
- Pirfenidone is associated with health-related benefits which cannot be adequately captured in the QALY calculation:
  - Clinically meaningful reduction in dyspnoea (breathlessness), measured on the University of San Diego Shortness of Breath Questionnaire, which has a substantial impact on quality of life.
  - Improved patient choice, based on different adverse event profile.
  - Improved NHS capacity, through reducing inpatient stays attributed to acute exacerbations.
  - Impact on people of a working age (10% of patients are under 60).

6 Equality issues

6.1 No equality issues were raised in the consultee's evidence submissions. During the scoping process, the following issues were noted:

- Estimating a person's percent predicted FVC using expected values from the European Community of Coal and Steel (ECCS) reference tables discriminates against some groups:

  - minority ethnic groups, particularly people of south Asian family origin, where equations for predicting lung function are not adequately developed
  - older people, because the reference tables are derived from populations under the age of 70 whereas the average age of people with idiopathic pulmonary fibrosis is 72
  - disabled people who have difficulty standing straight, because percent predicted FVC is expressed as a percentage of the predicted normal value for a person of the same height
- Using FVC alone to assess disease severity is discriminatory because some people with idiopathic pulmonary fibrosis die when their percent predicted FVC remains above 80%.
- People with percent predicted FVC above 80% have clinically significant fibrosis and should be considered for treatment.

7 Authors

Sophie Laurenson
Technical Lead

Rosie Lovett
Technical Adviser

with input from the Lead Team (Mark Chapman and Sanjay Kinra).

Appendix A: Clinical efficacy section of the draft European public assessment report

http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002154/WC500103073.pdf

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282)

Final scope

Remit/appraisal objective

To appraise the clinical and cost effectiveness of pirfenidone within its marketing authorisation for treating idiopathic pulmonary fibrosis.

Background

Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive lung disease in which scarring (fibrosis) occurs. The cause of IPF is unknown although it is thought to be related to an abnormal immune response. It is a difficult disease to diagnose and requires a multidisciplinary team. Most people with IPF experience symptoms of breathlessness, which may initially be only on exertion. Cough, with or without sputum, is a common symptom. Over time, these symptoms are associated with a decline in lung function, reduced quality of life and death.

The rate of disease progression can vary greatly in people with IPF. The median survival for people with IPF in the UK is approximately 3 years from the time of diagnosis, but is generally longer for people with mild-to-moderate IPF. Prognosis is difficult to estimate at the time of diagnosis and may only become apparent after a period of careful follow-up. Although there are challenges in assessing the severity of the condition, it is widely accepted that severe idiopathic pulmonary fibrosis is defined as forced vital capacity (FVC) less than 50% predicted and a diffusing capacity for carbon monoxide less than 35% predicted. Therefore, mild-to-moderate idiopathic pulmonary fibrosis could be assumed to include a FVC greater than or equal to 50% predicted and a diffusing capacity for carbon monoxide greater than or equal to 35%.

The incidence of IPF is approximately 8 to 9 per 100,000 person-years, which equates to more than 5000 new diagnoses each year in the UK. The incidence is higher in men than women, and increases with age (median age of presentation is 70 years).

The aim of treatment is to manage the symptoms and slow progression. NICE clinical guideline 163 on the diagnosis and management of suspected idiopathic pulmonary fibrosis recommends that best supportive care (including symptom relief, managing co-morbidities, withdrawing therapies suspected to be ineffective or causing harm, and end of life care) should be offered to people from diagnosis and be tailored according to disease severity, rate of progression and the person's preference. If pharmacological treatment is considered appropriate, NICE technology appraisal guidance 282 recommends pirfenidone if a person's forced vital capacity (FVC) is between

50% and 80% of their expected value. Preliminary guidance from an ongoing NICE technology appraisal recommends nintedanib as an option for treating idiopathic pulmonary fibrosis, in people with a percent predicted FVC of 50–80%. Both technology appraisals recommend that treatment should be stopped if there is evidence of disease progression (a decline in per cent predicted FVC of 10% or more within any 12 month period). Lung transplantation is an option if there are no contraindications.

NICE is reviewing the technology appraisal guidance 282 following publication of the ASCEND study, which showed that people with a predicted FVC greater than 80% could potentially benefit from treatment with pirfenidone.

The technology

Pirfenidone (Esbriet, Roche) is an oral immunosuppressant that is thought to have anti-inflammatory and antifibrotic effects. Pirfenidone has a marketing authorisation in the UK for treating mild to moderate IPF in adults.

Intervention(s)	Pirfenidone
Population(s)	Adults with mild to moderate idiopathic pulmonary fibrosis
Comparators	• Best supportive care • Nintedanib (only for people with a percent predicted FVC of 50–80%, subject to ongoing NICE appraisal)
Outcomes	The outcome measures to be considered include: • pulmonary function parameters • physical function • exacerbation rate • progression-free survival • mortality • adverse effects of treatment • health-related quality of life.

Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies should be taken into account.
Other considerations	Guidance will only be issued in accordance with the marketing authorisation. If evidence allows, subgroup analysis by disease severity, defined by FVC (such as above and below or 80% FVC) and/or diffusing capacity for carbon monoxide, will be considered.
Related NICE recommendations and NICE Pathways	**Related Technology Appraisals**: Technology Appraisal No. 282, April 2013, 'Pirfenidone for treating idiopathic pulmonary fibrosis'. Technology Appraisal in preparation, 'Nintedanib for treating idiopathic pulmonary fibrosis [ID752]', Anticipated publication date: January 2016. **Related Guidelines:** Clinical Guideline No.163, July 2013, 'Idiopathic pulmonary fibrosis: The diagnosis and management of suspected idiopathic pulmonary fibrosis'. Review Proposal Date June 2015. **Related Quality Standards:** Quality Standard No. 79, January 2015, 'Idiopathic pulmonary fibrosis'. http://www.nice.org.uk/guidance/qs79 **Related NICE Pathways:** NICE pathway: Idiopathic pulmonary fibrosis, Pathway created June 2013. http://pathways.nice.org.uk/pathways/idiopathic-pulmonary-fibrosis

Related National Policy	National Service Frameworks: Older People Department of Health, November 2013, 'NHS Outcomes Framework 2014-2015'.

APPENDIX C

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Final matrix of consultees and commentators

Consultees	Commentators (no right to submit or appeal)
Company • Roche (pirfenidone) Patient/carer groups • Action For Pulmonary Fibrosis • Afiya Trust • Black Health Agency • British Lung Foundation • Equalities National Council • Muslim Council of Britain • Muslim Health Network • Pulmonary Fibrosis Trust • South Asian Health Foundation • Specialised Healthcare Alliance Professional groups • Association for Respiratory Technology and Physiology • Association of Respiratory Nurse Specialists • British Geriatrics Society • British Thoracic Society • Primary Care Respiratory Society UK • Royal College of General Practitioners • Royal College of Nursing • Royal College of Pathologists • Royal College of Physicians • Royal Pharmaceutical Society • Royal Society of Medicine – Intellectual Disabilities Forum • UK Clinical Pharmacy Association Others • Department of Health • NHS England • NHS Lincolnshire East CCG • NHS Greenwich CCG • Welsh Government	General • Allied Health Professionals Federation • Board of Community Health Councils in Wales • British National Formulary • Care Quality Commission • Department of Health, Social Services and Public Safety for Northern Ireland • Healthcare Improvement Scotland • Medicines and Healthcare Products Regulatory Agency • National Association of Primary Care • National Pharmacy Association • NHS Alliance • NHS Commercial Medicines Unit • NHS Confederation • Scottish Medicines Consortium Comparator Company(ies) • Boehringer Ingelheim (nintedanib) Relevant research groups • Breathing Matters • British Association for Lung Research • Cochrane Airways Group • MRC Clinical Trials Unit • National Institute for Health Research Evidence Review Group • Evidence Review Group TBC • National Institute for Health Research Health Technology Assessment Programme Associated Guideline Groups • National Clinical Guideline Centre Associated Public Health Groups

	• Public Health England • Public Health Wales

NICE is committed to promoting equality, eliminating unlawful discrimination and fostering good relations between people who share a protected characteristic and those who do not. Please let us know if we have missed any important organisations from the lists in the matrix, and which organisations we should include that have a particular focus on relevant equality issues.

PTO FOR DEFINITIONS OF CONSULTEES AND COMMENTATORS

Definitions:

Consultees

Organisations that accept an invitation to participate in the appraisal; the manufacturer(s) or sponsor(s) of the technology; national professional organisations; national patient organisations; the Department of Health and the Welsh Government and relevant NHS organisations in England.

The manufacturer/sponsor of the technology are invited to prepare a submission dossier, can respond to consultations, nominate clinical specialists and has the right to appeal against the Final Appraisal Determination (FAD).
All non-manufacturer/sponsor consultees are invited to prepare a submission dossier respond to consultations on the draft scope, the Assessment Report and the Appraisal Consultation Document. They can nominate clinical specialists and/or patient experts and have the right to appeal against the Final Appraisal Determination (FAD).

Commentators

Organisations that engage in the appraisal process but are not asked to prepare a submission dossier. Commentators are able to respond to consultations and they receive the FAD for information only, without right of appeal. These organisations are: manufacturers of comparator technologies; Healthcare Improvement Scotland; the relevant National Collaborating Centre (a group commissioned by the Institute to develop clinical guidelines); other related research groups where appropriate (for example, the Medical Research Council [MRC], National Cancer Research Institute); other groups (for example, the NHS Confederation, NHS Alliance and NHS Commercial Medicines Unit, and the British National Formulary.
All non-manufacturers/sponsors commentator organisations can nominate clinical specialists and patient experts to present their personal views to the Appraisal Committee.

Assessment group

An independent academic group (commissioned by the National Institute for Health Research (NIHR) Health Technology Assessment Programme (HTA Programme) to assist in the appraisal) prepares an Assessment Report on the health technology (a review of the clinical and cost effectiveness of the technology(ies)) based on a systematic review of the manufacturer/sponsor and non-manufacturer/sponsor submission dossier to the Institute.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Pirfenidone for the treatment of adults with idiopathic pulmonary fibrosis

Company evidence submission

Roche Product Limited

February 2016

File name	Version	Contains confidential information	Date
Pirfenidone IPF, Roche evidence submission (final – 3 March update)	**Final**	**Yes**	**3 March 2016**

Table of contents

List of tables

List of figures

List of abbreviations

6MWD	6 minute walking distance
ADR	Actual dose received
AE	Adverse events
AIC	Akaike Information Criterion
ALAT	Latin American Thoracic Society
ALKP	Alkaline phosphatase
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
AR	Adverse reaction
AST	Aspartate aminotransferase
ATS	American Thoracic Society
BI	Boehringer Ingelheim
BIC	Bayesian Information Criterion
BLF	British Lung Foundation
BMI	Body mass index
BSC	Best supportive care
BTS	British Thoracic Society
CADTH	Canadian Agency for Drugs and Technologies in Health
CC	Complication and comorbidity
CDEC	Canadian Drug Expert Committee
CDR	Common Drug Reviews
CDSR	Cochrane Database of Systematic Reviews
CE	Conformité Européenne
CENTRAL	Cochrane Central Register of Controlled Trials
CF	Cystic Fibrosis
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Disease
CPRD	Clinical Practice Research Datalink
CRP	C reactive Protein
CSR	Clinical study reports
CV	Cardiovascular
DARE	Database of Abstracts of Reviews of Effectiveness
DLCO	Diffusing capacity of the lungs for carbon monoxide
DOH	Department of Health
DSMB	Data and Safety Monitoring Board
DSU	Decision Support Unit
EMA	European medicine Agency
EMC	Electronic Medicines Compendium
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
ERS	European Respiratory Society
EU	European Union
FAS	Full analysis set
FDA	Food and Drug Administration
FE	Fixed Effects
FEV1	Forced expiratory volume in 1 second
FVC	Forced vital capacity
GAP	Gender age physiology predictive tool for mortality in IPF
GEE	Generalised estimating equation
GGT	Gamma-glutamyl transpeptidase
GI	Gastrointestinal
GP	General Practitioner
GRE	Generalised random effects
HR	Hazard ratio
HRCT	High-resolution computed tomography
HRQoL	Health related quality of life
HS	Health State
HTA	Health Technology Assessment.
ICER	Incremental cost-effectiveness ratio
ICU	Intensive Care unit
ILD	Interstitial lung disease
IPF	Idiopathic pulmonary fibrosis
ITT	Intention to Treat
JAGS	Just another Gibbs sampler
JRS	Japanese Respiratory Society
KM	Kaplan-Meier
LDH	Lactate Dehydrogenase
LOCF	Last observation carried forward
LY	Life years
LYG	life years gained
MAA	Marketing Authorization Application
MCID	Minimal clinically important difference
MEDRA	Medical Dictionary for Regulatory Activities
MIMS	Monthly Index of Medical

	Specialities
MMRM	Mixed effect Model Repeat Measurement
MRC	Medical research council
MRU	Medical resource use
MRU	Medical resource use
NAC	N-acetylcysteine
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NPD	Number of pills per day
NRH	Non Respiratory Hospitilsations
NTB	Nintedanib
OS	Overall survival
OWSA	One-way sensitivity analysis
PaO_2	Partial pressure arterial oxygen
PAS	Patient access scheme
PBS	Pharmaceutical Benefits Scheme
PEY	Person exposure years
PFN	Pirfenidone
PFS	Progressive free survival
PFT	Pulmonary Function Tests
PICOS	Patient Population or Problem, Intervention, Comparison, Outcomes, and Setting
PRAC	Pharmacovigilance Risk Assessment Committee
PRISMA	Providing Innovative Service Models and Assessment
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Service Research Unit
QALY	Quality-adjusted life year
QOL	Quality of life
QTc	Corrected QT interval
RCT	Randomised Control Trial
RDD	Recommended daily dose
RE	Random effects
RH	Respiratory Hospitilisations
RMSE	Root mean square error
RU	Resource use
SD	Standard deviation
SE	Standard error
SF-36	Short Form (36) Health Survey
SGRQ	St Georges respiratory questionnaire
SMC	Scottish Medicines Consortium
SMPC	Summary of product characteristics
SOBQ	Shortness of Breath Questionnaire
SOC	Standard of care
SpO_2	Blood oxygen saturation level
SSD	Sum of squared differences
STA	Single Technology Appraisal
TE	Treatment-emergent
TGF-β	Transforming growth factor-beta
TID	Ter in die (Three times daily)
TLC	Total lung capacity
TNF-α	Tumour necrosis factor-alpha
TORCH	Towards a Revolution in COPD Health
TSD	Technical Support Document
TTD	Time to treatment discontinuation
TTO	Time trade off
UCSD	University of California San Diego
SOBQ	Shortness of Breath Questionnaire
UIP	Usual interstitial pneumonia
ULN	Upper limit of Normal
VC	Vital capacity
WHO	World Health Organisation

1 Executive summary

Idiopathic pulmonary fibrosis (IPF) is characterised by fibrosis of the lung interstitium and is a debilitating and progressive disease of unknown cause. The disease leads to an irreversible and continuing decrease in lung volume, and progressive worsening of dyspnoea (i.e. shortness of breath) and lung function. (Zibrak and Price, 2014, Oldham and Noth, 2014). The clinical course of IPF is unpredictable and features periods of relative stability with a slow decline in lung function that may be interspersed with episodes of stepwise deterioration in symptoms and acute episodes of rapid respiratory deterioration.

Patients with IPF ultimately die from respiratory failure or a complicating comorbidity. The prognosis for IPF is extremely poor with a median survival time in the UK of only 3 years from the time of diagnosis (Navaratnam 2011). This is a rate which exceeds that of many cancers, and emphasises the severity of IPF and its impact on patients (Vancheri, 2010).

There is currently no cure for IPF. The treatment goal is to slow disease progression and prolong survival, whilst managing the patient's symptoms. As the irreversible scarring of lung tissue occurs during the early stages of IPF, early intervention with effective treatments which delay disease progression should be an important goal for the management of the condition.

Pirfenidone (Esbriet®) was the first licensed treatment for IPF, with EU marketing authorisation granted in February 2011. This medicine has an "orphan designation" which means that it is used to treat life-threatening or chronically debilitating conditions that affect no more than five in 10,000 people in the European Union. It is indicated in adults for the treatment of mild to moderate IPF. Whilst the exact mechanism of action of pirfenidone has not been fully established, existing data suggest that pirfenidone exerts both antifibrotic and anti-inflammatory properties (EMC, 2015a)

In clinical practice, there are no accepted thresholds of percent predicted forced vital capacity (FVC) used to define the disease severity of a patient with IPF, although

there is a general acceptance that an FVC <50% predicted and diffusing capacity of carbon monoxide (DLco) <35% predicted defines severe disease(NICE 2015e). The ATS/ERS/JRS/ALAT[1] guidelines also do not propose a formal staging system for classification of disease severity (Raghu 2015).

NICE TA282 currently recommends pirfenidone in adult patients with a predicted forced vital capacity (FVC) between 50% and 80%. NICE also recommend that treatment should be stopped if a patient's FVC falls by ≥10% in any 12 month period (NICE 2013a).

Since this original submission in 2011 (presented to NICE by InterMune, the Marketing Authorisation Holder of pirfenidone in the UK at the time), nintedanib has also received EU approval, with a similar indication. NICE has recently recommended nintedanib in line with the recommendations for pirfenidone in TA282 (NICE 2016). In patients with a preserved FVC, above this 80% threshold, there are no licensed or NICE-approved therapies except best supportive care (BSC). Such patients face a significant unmet need.

As IPF is a progressive and deteriorating disease, it is clear that preventing irreversible loss of lung function is vital to allow these patients to remain functioning as long as possible. Early treatment with pirfenidone has demonstrated similar efficacy to treatment in patients with more severely reduced FVC. Compared to the current restrictions around its use in clinical practice, pirfenidone provides a valuable treatment option for patients who currently have to reach a significant level of morbidity to be eligible to start a treatment which has been shown to slow that rate of deterioration and prevent future morbidity and mortality.

Furthermore, the nature of IPF means that a prior decline in lung function does not predict a future decline, and periods of stability can only be identified in retrospect. Application of a stopping rule such as that described in TA282 is therefore complicated, since progression with treatment does not always constitute treatment failure.

[1] ATS: American Thoracic Society; ERS: European Respiratory Society; JRS: Japanese Respiratory Society; ALAT: Latin American Thoracic Society

This resubmission will present data from three pivotal randomised, placebo-controlled, double-blind, phase III trials in patients with IPF (ASCEND, CAPACITY-2 and CAPACITY 1), along with two supportive studies in Japanese patients (SP2 & SP3). Data from ASCEND, along with the pre-specified pooled analysis of the pivotal trials, were not available at the time of the original submission and add significant new information to the evidence base for this review, as highlighted by the consultees and commentators to this appraisal during the earlier scoping phase (NICE 2015e). This important study was also accepted as evidence by the EMA and included in the SmPC in 2014 (EMC 2015a).

Most significantly, the findings constitute the first time a treatment in IPF has demonstrated a mortality benefit for patients. (King 2014 Suppl). Further analyses found that the magnitude of clinical benefit offered through pirfenidone treatment was comparable in both patients with earlier (FVC ≥80%) and more advanced IPF (FVC <80%). This finding supports the prompt use of pirfenidone in IPF patients after diagnosis, regardless of disease severity (Albera, 2015). The use of pirfenidone in patients with mild IPF has been recommended by national agencies in a number of countries, including Canada, Sweden and Switzerland.

Clinical effectiveness

Despite the classification of pirfenidone as an orphan treatment in the management of IPF, its treatment benefit is supported by a wealth of clinical data from 3 large randomised controlled trials, with almost 9 years of follow-up to date.

The three pivotal trials (ASCEND, CAPACITY 1 & 2) provide evidence for the safety and efficacy of pirfenidone based on over 1300 IPF patients (King 2014; Noble 2011). Further evidence from the two Japanese studies (SP2 and SP3) provides supportive data which are consistent to the findings from the pivotal trials (Azuma 2005; Taniguchi 2010). Overall, the use of pirfenidone in mild to moderate IPF patients slows the rate of decline in FVC and improves survival. Pirfenidone also reduced the decline in 6-minute walking distance (6MWD), thus demonstrating an important functional outcome and reinforcing why pirfenidone is a valuable treatment option for patients with IPF.

A pooled analysis of ASCEND, CAPACITY 1 & 2 at Week 52 showed treatment with pirfenidone reduced the proportion of patients with disease progression or death by 43.8%, and increased the proportion of patients with no decline in percent predicted FVC by 59.3%, compared with placebo ($p<0.0001$) (Noble 2014a).

Treatment-emergent IPF-related mortality was lower in the pirfenidone group than the placebo group in the pooled analysis of CAPACITY 1 & 2 (HR: 0.48 [95% CI: 0.24, 0.95] p=0.03). There was no significant difference in overall survival between pirfenidone and placebo in these trials (which were not powered to assess OS) (Noble 2011).

In ASCEND, pirfenidone significantly reduced the change in percent predicted FVC or death compared to placebo at Week 52 ($p<0.001$) (King 2014). A similar effect was seen in the absolute change in percent predicted FVC in CAPACITY 2 ($p=0.001$), but did not reach statistical significance in CAPACITY 1 ($p=0.501$) (Noble 2011).

The overall findings provide strong and consistent evidence for a clinically meaningful benefit of pirfenidone on exercise tolerance and overall survival in IPF patients.

Recent analyses also raise the possibility that pirfenidone might have salutary effects beyond just the slowing in the rate of loss of lung function: a post-hoc analysis from the pooled CAPACITY and ASCEND datasets has been shown that those patients on pirfenidone with a ≥ 10% reduction in FVC have significantly less mortality than those on placebo with a similar change (Nathan 2015a, Nathan 2016).

Substantial long-term exposure to pirfenidone (up to 8.8 years), together with post-marketing experience, leads to a well described adverse-event profile. This experience shows pirfenidone is well-tolerated, with a manageable side effect profile. Safety results from all five RCTs demonstrate gastrointestinal and skin-related events were more common in the pirfenidone group compared to placebo. However the most common adverse events have typically been mild to moderate in intensity, generally occur within the first six months of treatment, and infrequently led to drug

discontinuation. There was no significant signal of an adverse effect on the cardiovascular system.

The treatment effect of pirfenidone in slowing disease progression, symptom control, and prolonging survival have been confirmed in real-world studies (Harari and Caminati, 2015).

Indirect treatment comparison vs. nintedanib

As there is no head-to-head comparison between pirfenidone and nintedanib, a network meta-analysis was performed to allow comparison between these two treatment options (along with BSC). The results of our network meta-analysis (NMA) show that pirfenidone is a more effective treatment compared to placebo in terms of all-cause mortality, IPF-related mortality, progression free survival (PFS), FVC outcomes, 6MWD, and shortness of breath questionnaire (SOBQ). Pirfenidone has more data on more outcomes available than nintedanib to support its all-rounded efficacy profile.

Base-case cost-effectiveness results

In the base-case analysis, the comparison of pirfenidone to BSC in patients with mild-to-moderate IPF (the intention-to-treat [ITT] population) led to additional costs of [redacted] over the patient's lifetime (at list price). Treatment was associated with a life-year and quality-adjusted life year (QALY) gain of 3.29 and 1.87, respectively. This leads to an incremental cost-effectiveness ratio (ICER) of [redacted] per QALY gained. When compared to nintedanib in patients with moderate IPF, treatment with pirfenidone is estimated to incur an additional cost of [redacted] and generate a QALY gain of 0.92, leading to an ICER of [redacted] (at list prices for both treatments).

Expert advisory Panel

This submission has been developed with input from an Expert Panel of Interstitial Lung Disease (ILD) consultants;

- Helen Parfrey, Papworth Hospital NHS Foundation Trust
- Nazia Chaudry, University Hospital of South Manchester NHS Foundation Trust
- Felix Woodhead, University Hospitals Of Leicester NHS Trust

- Toby Maher, Royal Brompton and Harefield NHS Foundation Trust
- Michael Gibbons, Royal Devon and Exeter NHS Foundation Trust

1.1 Statement of decision problem

Table 1. The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with mild to moderate IPF	Same as final scope issued by NICE	N/A
Intervention	• Pirfenidone	Same as final scope issued by NICE	N/A
Comparator(s)	• Best supportive care • Nintedanib (only for people with a percent predicted FVC of 50-80%, subject to ongoing NICE appraisal)	Same as final scope issued by NICE	N/A
Outcomes	Outcome measures to be considered include: • Pulmonary function parameters • Physical function • Exacerbation rate • PFS • Mortality • Adverse effects of treatment • Health-related quality of life	Same as final scope issued by NICE	N/A
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	Same as final scope issued by NICE	N/A

	The availability of any patient access schemes for the intervention or comparator technologies should be taken into account.		
Subgroups to be considered	If evidence allows, subgroup analysis by disease severity, defined by FVC (such as above and below or 80% FVC) and/or diffusing capacity for carbon monoxide, will be considered	Same as final scope issued by NICE. Subgroup analysis by FVC and DLco status at baseline was investigated, but the available data only allowed FVC to be assessed and reported in this submission.	N/A
Special considerations	Guidance will only be issued in accordance with the marketing authorisation	Same as final scope issued by NICE	N/A

1.2 Description of the technology being appraised

Table 2. Technology being appraised

UK approved name and brand name	Pirfenidone (Esbriet®)
Marketing authorisation	Pirfenidone was granted EU marketing authorisation on the 28th February 2011 by the European Commission (EU/1/11/667/001; EU/1/11/667/002; EU/1/11/667/003) (EMA, 2015a; EMC, 2015a). Pirfenidone was designated orphan status on 16th November 2004 by the European Commission (EMA EU/3/04/241) (EMA, 2015a).
Indications and any restriction(s) as described in the summary of product characteristics	Pirfenidone is indicated in adults for the treatment of mild to moderate IPF (EMC, 2015a)
Method of administration and dosage	Pirfenidone is available in 267mg capsules, and is administered orally. When initiating treatment, the dose should be titrated over a 14-day period as follows (EMC, 2015a): • **Days 1 to 7:** one capsule, three times a day (801 mg/day) • **Days 8 to 14:** two capsules, three times a day (1602 mg/day) • **Day 15 onward:** three capsules, three times a day (2403 mg/day) The recommended daily dose of pirfenidone is three 267 mg capsules three times a day with food (total dose: 2403 mg/day). Guidance on dosing after a treatment interruption can be found in the SMPC. Dose adjustments are allowed for the management of AEs, and are specified in further detail in the SMPC.

1.3 Summary of the clinical effectiveness analysis

ASCEND (PIPF-016) (King 2014)

ASCEND was a multinational, randomised, double-blind, placebo-controlled, Phase III trial designed to evaluate the efficacy and safety of pirfenidone 2403 mg/day compared with placebo in patients with IPF (N=555). Patients were treated for 52 weeks. The primary endpoint was the change in FVC or death at Week 52 from baseline.

- **Primary endpoint was met:** At Week 52, there was a relative reduction of 47.9% in the proportion of patients with a ≥10% decline in % predicted FVC or death in the pirfenidone group as compared with placebo ($p<0.001$).

- **Relevant secondary endpoints:** At Week 52, pirfenidone reduced the relative risk of death or disease progression by 43% compared with placebo (HR: 0.57 [95% CI: 0.43-0.77] p<0.001). Treatment with pirfenidone also reduced the decline in 6MWD from baseline compared with placebo (p=0.04). A decrease of ≥50m in 6MWD or death occurred in 25.9% (n=72) of patients in the pirfenidone group compared with 35.7% (n=99) of patients in the placebo group; relative reduction of 27.5% in the pirfenidone group (p=0.04).
- **Key safety findings:** Gastrointestinal (GI) and skin-related adverse events (AEs) were more common in the pirfenidone group than in the placebo group, but were generally mild to moderate and reversible. Elevations in alanine or aspartate aminotransferase levels (3x upper limit of normal [ULN]) occurred in 2.9% of patients receiving pirfenidone, vs 0.7% of patients receiving placebo. All aminotransferase elevations were reversible without clinically significant consequences. There were no significant or serious reports of cardiovascular adverse events in the pirfenidone trials.

CAPACITY 1 and 2 (PIPF-006 and PIPF-004) (Noble 2011)

The CAPACITY programme consisted of two multinational, randomised, double-blind, placebo-controlled, Phase III trials designed to evaluate the efficacy and safety of pirfenidone in patients with IPF compared with placebo. In CAPACITY 2 (study 004), 435 patients were randomised in a 2:1:2 ratio to pirfenidone 2403 mg/day (n=174), 1197 mg/day (n=87), or placebo (n=174). In CAPACITY 1 (Study 006), 344 patients were randomly assigned in a 1:1 ratio to pirfenidone 2403 mg/day (n=171) or placebo (n=173). Patients were treated for a minimum of 72 weeks. The primary endpoint was change in % predicted FVC from baseline to Week 72 in both studies.

- **CAPACITY 2 (PIPF-004) met its primary endpoint:** At Week 72, pirfenidone 2403 mg/day significantly reduced the decline in FVC compared with placebo (-8.0% vs. -12.4%; p=0.001). Mean change in % FVC in the pirfenidone 1197 mg/day group were intermediate to the pirfenidone 2403 mg/day and placebo groups.
- **CAPACITY 1 (PIPF-006) did not meet its primary endpoint:** There was no significant difference between pirfenidone 2403 mg/day and placebo in the

mean change in % predicted FVC at Week 72 (-9.0% vs. -9.6%; p=0.501). However, the study provided supportive data on treatment effect of pirfenidone in patients with IPF. Significant treatment effect was evident at every time point from Week 12 to Week 48, and in repeated-measures analysis of % predicted FVC change over all assessment time points (p=0.007). This difference in FVC outcomes between the two studies may have been partly attributable to baseline imbalances. For example, numerically more patients in CAPACITY 1 had been diagnosed with IPF ≥1 year, and more patients receiving placebo had higher incidence of obstructive airway disease. Analysis of pooled CAPACITY 1 & 2 data showed a significant pirfenidone treatment effect for this endpoint as described in the next paragraph.

- **Primary endpoint analysis of the pooled CAPACITY 1 & 2 population showed statistical difference:** There was a significant difference between pirfenidone 2403 mg/day and placebo in the mean change in % predicted FVC at Week 72 across the pooled patient populations (pre-specified analysis: -8.5% vs. -11%; p=0.005).
- **Relevant secondary endpoints from pooled CAPACITY 1 & 2 populations:** In the pooled CAPACITY 1 & 2 population, the HR for treatment-emergent IPF-related mortality favoured pirfenidone over placebo and was statistically significant (HR: 0.48 [95% CI: 0.24-0.95]; p=0.03). Treatment-emergent was defined as occurrence within 28 days of study drug treatment. HRs for treatment-emergent all-cause mortality, overall all-cause mortality and IPF-related mortality also favoured pirfenidone but were not statistically significant. Exploratory analysis of pooled data from both studies showed a statistically significant treatment effect with pirfenidone for the following endpoints: PFS time (HR: 0.74 [95% CI: 0.57-0.96] p=0.025); proportion of patients with a decline in % predicted FVC of ≥10% (p=0.003), and mean change in 6MWD (p=0.0009). The treatment effect was evident by Week 12 and persisted throughout the duration of the study.
- **Key safety findings from pooled CAPACITY 1 & 2 populations:** The most frequently reported AEs in the pooled pirfenidone group were GI disorders,

skin disorders, and dizziness; these were generally mild to moderate in severity. More patients in the pooled pirfenidone group had elevations in alanine and aspartate aminotransferases (3x ULN) compared with placebo (4% [n=14] vs.<1% [n=2]); all cases reversible without clinical consequences.

RECAP (PIPF-0012) (Costabel 2014, Fisher 2015, Kreuter 2014, Roche 2016a)
RECAP (PIPF-012) is an open-label extension of the ASCEND and CAPACITY trials. The study was designed to assess the long-term safety of pirfenidone 2403 mg/day in patients with IPF who received ≥80% of scheduled doses, and completed the Week 72 final study visit in CAPACITY 1 or CAPACITY 2. Patients in the ASCEND study were also eligible to roll-over into RECAP, although no published data analysis including ASCEND is available to date (Kreuter 2014, Roche 2016a).

As on-going open-label extension of three clinical trials, the RECAP study is not yet complete. The most recent datacut was performed in June 2015, with analyses based on summary data from this datacut presented at a recent congress (Fisher 2015). This includes survival data for pirfenidone, with patient data available through to 8.8 years.

SP3 (Japanese Phase III study) (Taniguchi 2010)
SP3 was a randomised, double-blind, placebo-controlled, Phase III trial designed to evaluate the efficacy and safety of pirfenidone vs. placebo in Japanese patients with IPF (N=275). Patients were randomised in a 2:1:2 ratio to receive pirfenidone 1800 mg/day (n=108), 1200 mg/day (n=55), or placebo (n=104).

The Japanese population tend to have lower body weight than their European counterparts, and therefore the doses used were lower than that of the multinational studies ASCEND, CAPACITY 1 & 2.

Patients were treated for 52 weeks. The primary endpoint was the change in vital capacity (VC) at Week 52.

- **Primary endpoint was met:** Pirfenidone was associated with improvements in VC vs. placebo (-0.09 L vs. -0.16 L, respectively; p=0.0416).

- **Relevant secondary endpoints:** Pirfenidone improved PFS time compared with placebo (p=0.028). Disease progression was defined as 'death and/or ≥10% decline in VC from baseline'.
- **Key safety findings:** Photosensitivity occurred in significantly more patients with pirfenidone at both doses vs. placebo (p<0.01). Anorexia and elevated gamma-glutamyl transpeptidase (GGT) levels were significantly more common in the pirfenidone 1800 mg/day group vs placebo (p<0.01) (NB: Generally, the higher the GGT level the greater the damage to the liver). Despite this, pirfenidone was generally well-tolerated in IPF patients.

SP2 (Japanese Phase II study) (Azuma 2005)

SP2 was a randomised, double-blind, placebo-controlled, Phase II trial designed to evaluate the efficacy and safety of pirfenidone vs. placebo in Japanese patients with IPF. One hundred and nine patients were randomly assigned in a 2:1 ratio to receive pirfenidone 1800 mg/day (n=73) or placebo (n=36).The primary endpoint was the change in the lowest blood oxygen saturation level (SpO2) reached during a 6-minute walk test (6MWT) from baseline to 9 months.

- **Primary endpoint was met:** At 9 months, pirfenidone was not associated with a significant change in the lowest SpO_2 during 6MWT vs. placebo (0.47% vs. 0.94%, respectively; p=0.0722).
- **Relevant secondary endpoints:** At 9 months, there was a statistically significant difference in decline of VC between pirfenidone and placebo (-0.03 L vs. -0.13 L; p=0.0366). Acute exacerbations of IPF were reported in 14% (5/35) of the placebo group and not at all in the pirfenidone group (p=0.0031).
- **Key safety findings:** Photosensitivity, stomach discomfort, anorexia, nausea, heartburn, and fatigue were significantly more frequent in the pirfenidone group compared with placebo (p<0.05). Elevated gamma-guanosine triphosphate levels was more common in pirfenidone-treated patients than placebo. Most adverse events disappeared with decrease of pirfenidone dose or by temporarily withholding the medication. Skin photosensitivity was the most common major adverse event that resulted in discontinuation or dose reduction of pirfenidone.

Network meta-analysis of pirfenidone vs. standard of care in patients with IPF

There are no head-to-head trials comparing pirfenidone with nintedanib in patients with IPF, therefore a network meta-analysis (NMA) was conducted in order to provide an indirect treatment comparison. This NMA was conducted to support pricing and reimbursement submissions across all markets, and so also included N-acetylcysteine (NAC) and triple therapy (combination treatment with prednisolone, azathioprine and NAC) as comparator treatments.

Outcomes analysed in the NMA included: change in % predicted FVC, FVC in litres, 6MWD, Shortness of Breath Questionnaire (SOBQ), St. George's Respiratory Questionnaire (SGRQ) total score, mortality, PFS, acute exacerbations of IPF, discontinuations of the study and of treatment.

For each outcome, the base case analysis included all Phase II and III trials of pirfenidone, nintedanib and NAC (King 2014; Noble 2011; Taniguchi 2010; Richeldi 2014; Martinez 2014; Richeldi 2011; Raghu 2012). SP2 was excluded from the analyses as it was previously considered to be an outlier by NICE in the nintedanib appraisal (NICE, 2015c).

The NMA results provide evidence that pirfenidone is a more effective treatment compared to placebo in terms of all-cause mortality, IPF-related mortality, PFS, FVC, 6MWD, and SOBQ.

Across the clinical trial programmes, the NMA identified data available on more outcomes for pirfenidone than that available for nintedanib, which led to some restrictions within the comparisons made. Data on all outcomes were available for the comparison of pirfenidone and placebo, and on all outcomes for the comparison for both pirfenidone and nintedanib, with the exception of 6MWD, SOBQ, and PFS.

Strengths and limitations of the evidence

The evidence was identified by conducting extensive literature searches in a range of databases, supplemented by manual searches of the websites of key regulatory bodies. No limits were placed on language.

Pivotal RCTs of pirfenidone vs. placebo in patients with IPF

Overall, all five trials (ASCEND, CAPACITY 1, CAPACITY 2, SP3 and SP2) were considered to have a low risk of bias. All five studies were well-designed as RCTs, reducing the risk of selection bias. All five trials used appropriate allocation concealment methods.

Baseline characteristics were similar between intervention and control groups in all five trials. IPF patients enrolled into the ASCEND and CAPACITY programme were generally comparable to those seen in UK clinical practice (age, gender, disease severity). As SP2 and SP3 were conducted in Japan, and Japanese people tend to have lower body weight than people in the UK, there may be limitations on the generalisability of these data.

There were no unexpected imbalances in drop-outs between treatment and control groups. The results of all outcomes measured were described in the study publications. All outcomes reported on the clinical trial registries for the studies have been reported.

NMA of pirfenidone vs. standard of care in patients with IPF

An assessment of similarity of the studies eligible for inclusion in NMA was undertaken, as well as a full assessment of risk of bias for each trial identified for inclusion in the base case network. Due to safety concerns, the triple therapy arm (prednisone, azathioprine, and NAC) of the PANTHER trial was terminated early after a mean follow up of 32 weeks compared to the planned duration of 60 weeks. This treatment arm was excluded from the principal analysis but did feature in a sensitivity analysis.

A limitation of the NMA is that the number of studies in the networks is low, with 10 studies included in the base case analysis, which leads to uncertainty in the estimates. Furthermore, within the network all active treatments are compared directly to placebo but not to each other. Hence comparisons amongst active treatments will be more uncertain than comparisons between placebo and active treatments.

A potential concern was how the key outcomes were measured. The trials differed in their definitions of PFS and acute exacerbations. They also differed in how they handled missing data and statistical analyses. Not all trials explicitly reported loss to follow-up and discontinuation separately making it impossible to fully assess the impact of missing data.

Due to the limited number of studies contributing to each network, a pragmatic approach was adopted, whereby trials were included regardless of minor differences in outcome definitions, timing of assessment and analysis methods. It was assumed that the differences in definitions and methods did not influence the relative treatment effects. Where multiple sets of results were available for a single trial, we used the results from the method that was most consistently reported across trials.

1.4 Summary of the cost-effectiveness analysis

Economic analysis was undertaken to assess the cost-effectiveness of pirfenidone for the treatment of mild to moderate IPF. The analysis was conducted using a three-health-state area under the curve model, considering the proportion of patients in “Pre-progression”, “Post-progression” and “Death” over the model time horizon. Progression has been defined in line with the ASCEND clinical trial (King 2014):

- Confirmed ≥10% absolute decline in percent predicted FVC
- Confirmed ≥50m decline in 6MWD

Such a definition also accounts for views provided by consultees and commentators during the scoping process for this re-review, in that capturing the impact of a treatment on function beyond FVC-alone was an important consideration (NICE 2015e).

The 3 health-state structure, similar to that used in oncology models, was chosen based on prior use by Loveman (Loveman 2014, Loveman 2015). It also allowed for reduced complexity compared with the simulation model used within the previous pirfenidone submission, which was deemed to add considerable complexity with no added benefit in regards to the precision of cost-effectiveness estimates (InterMune 2011). This model structure is consistent with increasing evidence indicating that the

fundamental hallmarks of cancer biology are comparable to those of IPF (Albera 2011). The model structure also considers acute exacerbations (dramatic, singular events that are often fatal) as an important feature.

Resource use and quality of life data are based upon the progression, acute exacerbation, and adverse events profiles for each treatment. The inclusion of lung transplant within the model is tested via a sensitivity analysis. Quality of life data was derived from the CAPACITY trials, using a published mapping algorithm to adjust outcomes from the SGRQ to EQ-5D utility values (Freemantle 2015, Starkie 2011). Resource use was based upon advice on UK clinical practice in IPF from a panel of UK clinicians, stratified by treatment type and progression status; costing of these used NHS Reference Costs, as agreed during the nintedanib appraisal (Boehringer Ingelheim 2015, NICE 2016).

In line with the NICE scope, the model compares pirfenidone to BSC and nintedanib (NICE 2015e). NICE has recently issued guidance that nintedanib is recommended only for the treatment of patients with moderate IPF, and therefore the only comparator chosen for ITT and mild patient populations was BSC (NICE 2015a, NICE 2016). Within the model base case, the NICE-recommended stopping rule – based upon progression status – is only applied to nintedanib. This is based upon clinician feedback that the current stopping rule for pirfenidone may deny treatment to patients who could derive a morbidity/mortality benefit, as there is no information to indicate that these benefits are limited to patients whose lung function declines at a slower rate (NICE 2015e).

The analysis was primarily based upon the results of the pivotal Phase III clinical trials of pirfenidone (CAPACITY and ASCEND), along with newly-available long-term follow up data from the RECAP extension study (Costabel 2014, Fisher 2015, King 2014, Kreuter 2014, Noble 2011). Comparison to nintedanib was conducted via the NMA, including all relevant Phase III and Phase II clinical trials (barring SP2 which was identified as an outlier during the NICE appraisal for nintedanib) (NICE 2016, Roche 2016).

The economic analysis was conducted in line with the NICE reference case, using annual discount rates of 3.5% for both costs and QALYs. A 34-year model time

horizon was utilised in order to ascertain the expected costs and outcomes for patients over their lifetime. The model was constructed considering the perspective of the National Health Service (NHS) and Personal Social Services (PSS).

Sensitivity analysis was conducted to establish the most influential parameters on model results: assumptions regarding the long-term overall survival profile and duration of treatment effect for pirfenidone and comparators were found to have the largest effect. The main area of uncertainty is survival with comparator treatments, due to the lack of equivalently robust long-term follow-up data for comparator therapies. [redacted] (see Section 4.11; Roche 2016a).

Pirfenidone is associated with substantial benefit compared with BSC in the ITT population, with similar levels of cost-effectiveness demonstrated for the mild and moderate subgroups. Pirfenidone provides benefits to patients both pre and post progression (0.44 QALYs pre-progression; 1.44 post-progression vs BSC in the ITT population), supporting the assertion that the economic-based stopping rule defined in previous IPF submissions does not have a sound clinical basis. Pirfenidone is also cost-effective when compared to nintedanib in patients with moderate IPF.

Despite the classification of pirfenidone as an orphan treatment in the management of IPF, the treatment benefit estimated by the model is supported by a wealth of clinical data from 3 large randomised controlled trials, with almost 9 years of follow-up to date.

Base-case model results (assessed at the list prices for pirfenidone and nintedanib) are shown for the ITT, mild and moderate populations; in Table 3, Table 4 and Table 5, respectively.

Table 3: Base-case results – ITT population, list price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	5.38	3.81				
PFN	[redacted]	8.67	5.68	[redacted]	3.29	1.87	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years.

Table 4: Base-case results – Mild population, list price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	7.11	4.84				
PFN	[redacted]	11.26	7.00	[redacted]	4.15	2.17	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years.

Table 5: Base-case results – Moderate population, list prices

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)	Versus baseline (QALYs)
BSC	22,475	4.80	3.45					
NTB	62,639	6.06	4.23	40,164	1.26	0.78	51,611	51,611
PFN	[redacted]	7.67	5.15	[redacted]	1.61	0.92	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years.

2 The technology

On 29 September 2014, following a merger agreement on 24 August 2014,Roche completed the purchase of InterMune. As part of this acquisition, all rights for pirfenidone were transferred to Roche. This submission is based on the study reports, datasets and analyses which have been made available to Roche since September 2014.

2.1 Description of the technology

UK approved name: pirfenidone
Brand name: Esbriet®
Therapeutic class: Immunosuppressant

Whilst the mechanism of action of pirfenidone has not been fully established, studies in animal models and cell cultures suggest that the drug affects the production of two proteins, transforming growth factor-beta (TGF-β), which is involved in cell growth, and tumour necrosis factor-alpha (TNF-α), which plays an important role in inflammation. Existing data suggest that pirfenidone exerts both antifibrotic and anti-inflammatory properties. (EMC, 2015a).

2.2 Marketing authorisation/CE marking and health technology assessment

Pirfenidone was granted EU marketing authorisation on the 28th February 2011 by the European Commission (EU/1/11/667/001; EU/1/11/667/002; EU/1/11/667/003) (EMA, 2015a; EMC, 2015a). Prior to this, pirfenidone was designated as an 'orphan medicine' (a medicine used in rare diseases) on 16th November 2004 by the European Commission (EMA EU/3/04/241) (EMA, 2015a).

In the UK, pirfenidone is indicated for the treatment of adults with mild to moderate IPF (EMC, 2015a).

Pirfenidone is contraindicated in patients (EMC, 2015a):

- With hypersensitivity to the active substance, or to any of the excipients;

- With history of angioedema with pirfenidone;
- Who are using fluvoxamine;
- With severe hepatic impairment/end stage liver disease, and in patients with severe renal impairment/end stage renal disease requiring dialysis

The SmPC for pirfenidone is included as Appendix 1, and the assessment report by the EMA is included as Appendix 2 (EMA, 2011).

The Committee for Medicinal Products for Human Use (CHMP) discussed in the assessment report both non-clinical and clinical aspects, as well as the benefit-risk balance of pirfenidone.

Based on the CHMP review of data on quality, safety and efficacy, the CHMP considered by consensus that the risk-benefit balance of pirfenidone for the treatment of adults with mild to moderate IPF was favourable and therefore recommended the granting of the marketing authorisation. No special conditions were attached to the marketing authorisation, see Appendix 2.

Pirfenidone received marketing authorisation in the EU in 2011. Countries outside of the EU that also have pirfenidone approval are listed in Table 6.

Table 6. List of countries where pirfenidone is licensed outside of EU

Country	Year	Brand
Argentina (Martindale, 2015)	Not reported	Fibridoner®
Japan (FDA, 2014a)	2008	Pirespa®
India (FDA, 2014a)	2010	Pirfenex®
Canada (Health Canada, 2012)	2012	Esbriet®
China (FDA, 2014a)	2013	Etuary®
USA (FDA, 2014b)	2014	Esbriet®
Mexico(FDA, 2014a)	2014	KitosCell LP®

This is a re-review of pirfenidone, considering the recommendations initially made by NICE in TA282 (NICE, 2013a). The initial review of pirfenidone in IPF by NICE concluded:

"*1.1 Pirfenidone is recommended as an option for treating idiopathic pulmonary fibrosis only if:*

- *the person has a forced vital capacity (FVC) between 50% and 80% predicted and*
- *the manufacturer provides pirfenidone with the discount agreed in the patient access scheme.*

1.2 Treatment with pirfenidone that is recommended according to 1.1 should be discontinued if there is evidence of disease progression (a decline in per cent predicted FVC of 10% or more within any 12 month period).

1.3 People currently receiving pirfenidone that is not recommended according to 1.1 should have the option to continue treatment until they and their clinician consider it appropriate to stop."

The review has been triggered by the availability of updated clinical data from the ASCEND study (King 2014). This was a large Phase III clinical trial initiated to further investigate and quantify the efficacy benefit of pirfenidone, and support a request for Marketing Authorisation to the FDA. A pre-specified pooled analysis of ASCEND and CAPACITY 1 & 2 identified a 48% reduction in all-cause mortality (HR: 0.52 [95% CI: 0.31–0.87] p=0.011) and a 68% reduction in treatment-emergent IPF-related mortality (HR: 0.32 [95% CI: 0.14-0.76] p=0.006) with pirfenidone over placebo. These findings constitute the first time a treatment in IPF has demonstrated a mortality benefit, and provides further evidence to support the continued usage of pirfenidone (King 2014 Suppl).

Further analyses found that the magnitude of clinical benefit offered through pirfenidone treatment was comparable in both patients with earlier (FVC ≥80%) and more advanced IPF (FVC <80%). This finding supports the prompt use of pirfenidone in IPF patients after diagnosis, regardless of disease severity (Albera, 2015).

There are also data from European registries which suggest that pirfenidone may also be effective in patients who are more severe than those randomised in clinical trials (Harari 2015).

Based on these updated clinical data, this submission presents the clinical- and cost-effectiveness of pirfenidone when used for the treatment of patients with mild-to-moderate IPF, in line with the product's SmPC.

In August 2013, the SMC recommended pirfenidone for restricted use in the treatment of IPF. The restriction limited use to patients with a FVC of ≤80% predicted, and subject to PAS or a price that is equivalent or lower (SMC, 2013). No stopping rule was imposed by the SMC.

2.3 Administration and costs of the technology

Table 7. Costs of the technology being appraised

	Cost	Source
Pharmaceutical formulation	Hard capsule. Each capsule contains 267 mg pirfenidone. Two piece capsules with a white opaque body and white opaque cap imprinted with “PFD 267 mg” in brown ink and containing a white to pale yellow powder.	EMC, 2015a
Acquisition cost (excluding VAT)*	List prices: • Starter pack: £501.92 (63 caps); • 28 day pack: £2,007.70 (252 caps); • 30 day pack: £2,151.10 (270 caps)	BNF, 2016
Method of administration	Oral	EMC, 2015a
Doses	Upon initiating treatment, the dose should be titrated to the recommended daily dose of nine capsules per day over a 14-day period as follows: • Days 1 to 7: one capsule, three times a day (801 mg/day) • Days 8 to 14: two capsules, three times a day (1,602 mg/day) • Day 15 onward: three capsules, three times a day (2,403 mg/day) The recommended daily dose of pirfenidone for patients with IPF is three 267 mg capsules three times a day (total of 2,403 mg/day).	EMC, 2015a
Dosing frequency	Three times a day	EMC, 2015a
Average length of a course of treatment	Pirfenidone is administered until progression to severe IPF (i.e. % predicted FVC <50%) or unacceptable toxicity.	EMC, 2015a
Average cost of a course of treatment	£2,151.10 every 30 days	BNF, 2016
Anticipated average interval between courses of treatments	Continuous treatment until progression to severe IPF or unacceptable toxicity.	EMC, 2015a
Anticipated number of repeat courses of treatments	Continuous treatment until progression to severe IPF or unacceptable toxicity.	EMC, 2015a
Dose adjustments	If treatment is interrupted for 14 consecutive days or more, the initial titration regimen should be repeated; if treatment is interrupted for less than 14 consecutive days, the dose can be resumed at the previous daily dose without titration. To manage adverse events dose may be reduced to 1-2 capsules (267 mg – 534 mg) 2-3 times/day with re-	EMC, 2015a

	escalation to the recommended daily dose as tolerated.	
Anticipated care setting	Diagnosis and initiation of treatment occurs in a ILD specialist centre with ongoing management as an outpatient	EMC, 2015a
* Indicate whether this acquisition cost is list price or includes an approved patient access scheme. When the marketing authorisation or anticipated marketing authorisation recommends the intervention in combination with other treatments, the acquisition cost of each intervention should be presented.		

In line with recommendation 1.1 from NICE TA282 (NICE, 2013a), a patient access scheme is currently in place for pirfenidone in England and Wales.

2.4 Changes in service provision and management

Pirfenidone is already used in current practice to treat patients with IPF, therefore, there is no requirement for additional tests, investigations, infrastructure or patient monitoring associated with this appraisal, and we do not anticipate any change to current clinical practice. Furthermore, pirfenidone is a monotherapy and there is no requirement for additional therapeutic intervention during treatment.

As pirfenidone is currently licensed for the treatment of mild to moderate IPF, we anticipate minimal impact on resource use.

Pirfenidone is an oral medication which should be initiated and monitored by a specialist physicians experienced in the diagnosis and treatment of IPF. There are 20 pulmonary fibrosis specialist centres in England which are currently authorised to prescribe pirfenidone (BLF, 2014).

Patients requiring pirfenidone will receive supplies direct from the tertiary care provider, or the tertiary care provider may arrange for supplies via Homecare. This service, provided by Roche products, is available to patients at no cost to the NHS and represents a considerable cost saving for the NHS. The savings to the NHS cover the cost of courier fees, pharmacy technician costs to support the service and the additional cost of supply due to VAT charges. At present approximately 60% of all pirfenidone prescriptions are delivered through the Roche provided homecare service

Liver function monitoring is needed at regular intervals, and would be carried out by either the specialist centre or under a shared care agreement with the patient's GP.

2.5 Innovation

Pirfenidone was the first treatment licensed for the management of IPF. As such, it represented a significant step change in the management of this orphan disease at the time of regulatory approval. IPF is ultimately a fatal disease, with mortality rates that exceed those of many cancers (Vancheri 2010).

Following the original NICE appraisal (NICE TA282), which resulted in the recommendation of pirfenidone in moderate IPF patients, the ASCEND study has been published. The pooled clinical data across the ASCEND, CAPACITY 1 & 2 studies captures 1247 patients, of whom 623 received study treatment with pirfenidone (the SP2 and SP3 clinical trials increase these totals to 1623 and 859, respectively).

The pre-specified pooled analysis of the ASCEND and CAPACITY 1 & 2 studies (N=1247) demonstrated pirfenidone to be the first and only treatment to significantly improve survival for patients with IPF; pirfenidone compared with placebo showed a 48% reduction in all-cause mortality (HR: 0.52 [95% CI: 0.31–0.87] p=0.011) and a 68% reduction in treatment-emergent IPF-related mortality (HR: 0.32 [95% CI: 0.14-0.76] p=0.006) (King 2014 Suppl)

Further analyses found that the magnitude of clinical benefit offered through pirfenidone treatment was comparable in both patients with earlier (FVC ≥80%) and more advanced IPF (FVC <80%). This finding supports the prompt use of pirfenidone in IPF patients after diagnosis, regardless of disease severity (Albera 2015).

In comparison to other treatments for IPF – specifically, nintedanib – long-term data are available which demonstrate the sustained treatment effect on mortality with pirfenidone: RECAP (PIPF-0012), the open-label extension of the ASCEND and CAPACITY trials, has recently evaluated data through to 8.8 years (see Section 4.11 Costabel 2014, Kreuter 2014, Fisher 2015, Roche 2016a). Findings are supportive of thise from the clinical trial programme: pirfenidone has a vital role in preventing early morbidity and deaths in IPF (Fisher 2015).

Whilst the economic model supporting this appraisal is based on these clinical data, some of the health-related benefits which have been identified in the trial programme cannot be adequately captured in the QALY calculation. These benefits include:

- QALYs do not take into account a reduction in decline of IPF related symptoms and the physical and social consequential benefits to the patient, family and carer's in reducing such decline. One example would include dyspnoea as measured by SOBQ score. This health related outcome measure is not included in the QALY calculation, despite patients reporting it having a significant impact on their daily life in many studies/reports (FDA 2015, Swigris 2005). Pooled data from ASCEND, CAPACITY 1 & 2 showed pirfenidone treatment reduced the proportion of patients who experienced a greater than 20 point increase in dyspnoea (as measured by SOBQ score) or death (p=0.0471) (Noble 2014a) meeting a clear unmet clinical need. Given the final guidance from NICE on the use of nintendanib in IPF (NICE, 2016), pirfenidone is the only licensed treatment in IPF which could potentially be used in patients with mild IPF. IPF is a chronic, progressive, and fatal lung disease that is characterised by irreversible loss of lung function. Early treatment to delay progression should, therefore, be an important goal for the management of the condition. We note that clinical opinion heard during the nintedanib appraisal strongly advocates for earlier access to treatments (NICE, 2016), but this is not reflected in the prevailing guidance from NICE, which is ultimately at the detriment of patients.
- improving patient choice, based on the different adverse-event profiles of treatments for IPF. For example, nintedanib has special warnings and precautions for use in patients with CV risk (EMC, 2015b). EMA has recently requested BI to update the nintedanib SmPC with a warning on the risk of haemorrhage and epistaxis and new data on mild/moderate hepatic impairment (EMA, 2015b). Pirfenidone is the only approved IPF treatment without a special warning or precaution for use in patients at risk of cardiovascular disease or bleeding (EMC, 2015a).
- improvement in NHS capacity, through a reduction in bed days for acute exacerbations. At 52 weeks, the odds ratio for acute exacerbation with

pirfenidone compared with placebo (based on data from ASCEND, CAPACITY 1 & 2; see Section 4.10) was 0.62. Based on the severity of exacerbation events in IPF, this is also likely to include a reduction in ICU use.

- impact of treatment on population of a working age . Whilst the majority of patients are older, data from the UK BTS registry show that 10% of recorded patients are aged under 60 and a further 34% are under 70,(BTS 2014) This is consistent with data from USA which demonstrated between 22.5% to 28.4% of IPF patients were aged under 65 depending on the exact definition of IPF used.(Raghu 2006)

Based on these landmark findings from the clinical trials, we consider pirfenidone to continue to be an innovative treatment, with a demonstrated significant impact on patients' lives – particularly those with FVC ≥80% predicted. Pirfenidone remains the only treatment with evidence of prolonging patient survival, delaying disease progression and improving patient symptoms.

3 Health condition and position of the technology in the treatment pathway

3.1 Disease overview

Provide a brief overview of the disease or condition for which the technology is being used. Include details of the underlying course of the disease.

ILD is a family of diseases that causes progressive scarring of the lung tissue through inflammation and fibrosis. IPF is the most common form of ILD, and is a chronic debilitating and fatal disease with poor prognosis (Zibrak and Price, 2014). The cause of IPF is unknown although it is thought to be related to an abnormal immune response to an unknown cause.

Most people with IPF experience symptoms of breathlessness and cough (Oldham and Noth, 2014). Over time, these symptoms are associated with a decline in lung and physical function, reduced overall survival, and a devastating effect on a patients' quality of life. IPF and its associated complications (such as hospitalisation for acute exacerbations) place a significant economic burden on the NHS. The financial burden of IPF on NHS bed days in the UK alone has been estimated at £16.2 million annually. Without intervention, this is expected to grow to over £20 million by 2020 (Navaratnam 2013).

The natural history of IPF is highly variable amongst individuals. Fortunate patients will experience a slow decline in respiratory function after diagnosis. Some patients will experience relatively stable periods with bouts of acute exacerbations, which require hospitalisation for respiratory failure. Others will experience rapid deterioration and progression leading to death (Ley 2011).

There is currently no cure for IPF. Due to the progressive and irreversible nature of IPF, the goals of treatment are to reduce the rate of disease progression in order to prolong survival, while optimally balancing benefit with safety and tolerability.

3.2 Effects of the disease on patients, carers and society

Describe the effects of the disease or condition on patients, carers and society.

The burden of IPF on patients is significant and goes beyond respiratory limitations and survival. Prior to death, patients with IPF suffer from dyspnoea, sometimes presenting with a debilitating cough, which can negatively impact their quality of life (QoL) (Swigris 2005). Dyspnoea is a strong driver in health-related quality of life impairment (Nishiyama 2005). In addition to the obvious effect on physical health, energy level, and respiratory symptoms, IPF patients are less likely to be independent (Swigris 2005). Disability increases with the severity of the disease and can make patients feel sad and fearful of losing their independence. After IPF diagnosis, patients are less likely to participate in social activities with their families and friends, which can further affect their mental and spiritual well-being (Swigris 2005).

IPF patients also have an increased risk of comorbidities, including pulmonary hypertension, emphysema, pulmonary embolism, chronic bronchitis and pulmonary infection which can worsen their QoL and survival (Collard 2012). As the incidence of IPF is increasing in the UK, patients who do not receive effective treatment for their IPF or comorbidities may have a negative impact on healthcare resources in the UK. IPF usually first develops in adults aged 50 or above and is thought to be more common in men (Navaratnam 2011). Societal costs may also be impacted, as the disability brought on by IPF can cause patients to lose their job, causing uncertainties around a family's financial security (Swigris 2005).

3.3 Clinical pathway of care

The clinical pathway of care for patients with IPF in England and Wales is presented in Figure 1. Best supportive care (BSC) is considered from the point of diagnosis, and tailored to disease severity, rate of progression, and patient preference.

Patients who have received BSC and/or pharmacological treatment should be referred for lung transplantation assessment if they wish to explore lung transplantation and there are no absolute contraindications. For those patients who do not receive a lung transplant the only option is use of pharmacological agents:

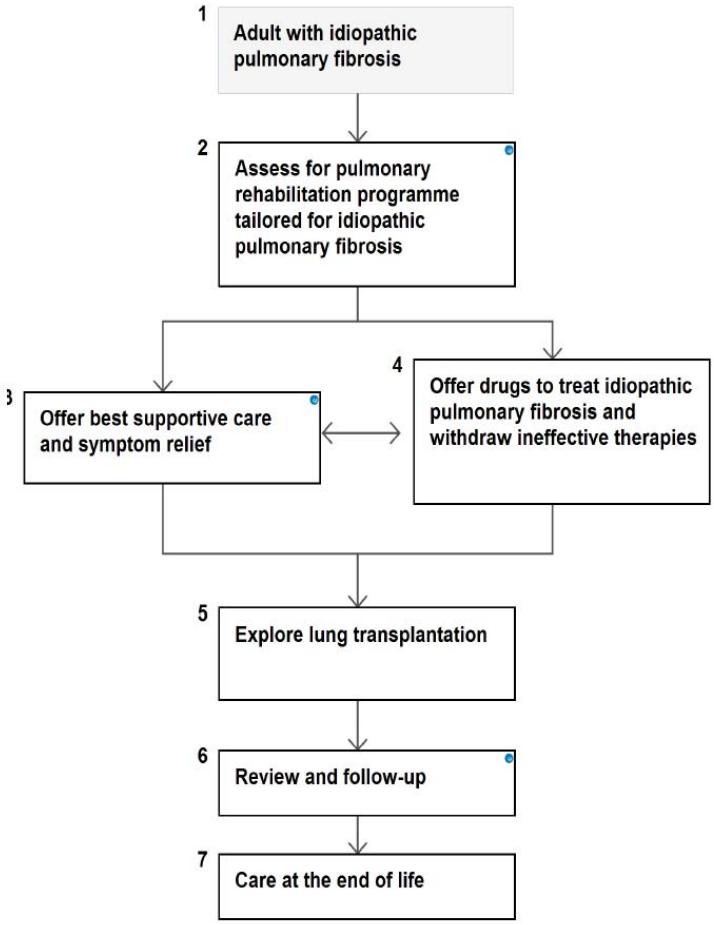
- Pirfenidone is licensed for the treatment of mild to moderate IPF, and is currently recommended by NICE in adults with predicted FVC between 50-80% under a patient access scheme. Treatment should be discontinued if there is evidence of disease progression (a decline in predicted FVC of ≥10%) within any 12 month period (NICE, 2013a).
- Nintedanib has been recommended by NICE in final guidance (TA379) under the same criteria as pirfenidone in TA282 (NICE, 2016).
- N-acetylcysteine (NAC) can also be used for managing IPF, but its benefits are uncertain. As part of the scoping process for TA282 in October 2011, it was agreed that NAC was not an appropriate comparator: it is not licensed for use in IPF, alone or in combination with other treatments (NICE, 2011). It was also not considered an appropriate comparator in this appraisal, or that of nintedanib (NICE, 2015c).

Patients who are not eligible to receive treatment with pirfenidone or nintedanib are limited to BSC only (including symptom relief, managing co-morbidities, withdrawing therapies suspected to be ineffective or causing harm, and end of life care) (NICE, 2013b).

This submission assesses the use of pirfenidone in line with its marketing authorisation; i.e. in both mild and moderate patients:

- For mild patients there is no active treatment option available in NHS clinical practice – patients must wait for their disease to worsen before receiving treatment
- For moderate patients treatment with pirfenidone currently forms the standard of care

Figure 1. NICE pathway for managing IPF



3.4 Life expectancy of people with the disease in England

IPF is ultimately fatal with mortality rates that exceed those of many cancers (Vancheri 2010).

Prognosis is difficult to predict as the rate of progression can vary greatly. In the UK, the median survival for people with IPF is 3 years from the time of diagnosis (Navaratnam 2011), and only 20% of people with IPF survive for more than 5 years (Kim 2006). It is estimated that there are 15 000 people living with IPF, and each year there are 5000 new cases diagnosed, and 5000 deaths due to IPF (Navaratnam, 2011).

3.5 Guidance related to the condition

Provide details of any relevant NICE guidance, pathways or commissioning guides related to the condition for which the technology is being used. Specify whether any subgroups were explicitly addressed.

- NICE TA282, April 2013. Pirfenidone for treating idiopathic pulmonary fibrosis (NICE, 2013a)
- NICE TA379, January 2016. Nintedanib for treating idiopathic pulmonary fibrosis (NICE, 2016)
- NICE CG13, July 2013. Idiopatpathic pulmonary fibrosis: the diagnosis and management of suspected idiopathic pulmonary fibrosis (NICE, 2013b)
- NICE QS79, January 2015. Idiopathic pulmonary fibrosis (NICE, 2015b)
- NICE pathway, June 2013. Idiopathic pulmonary fibrosis (NICE, 2015a)

3.6 Other clinical guidelines

Provide details of other clinical guidelines (for example, UK guidance from the royal societies or European guidance) and national policies.

- **An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline** (Raghu 2015): The strong clinical evidence in support of pirfenidone has been recognised in the latest update to the international consensus clinical practice guidelines for IPF. The guidelines now specifically recommend pirfenidone as a viable treatment option in IPF, acknowledging its FVC and mortality benefits: "*This recommendation puts a high value on the potential benefit of pirfenidone on patient-important outcomes such as disease progression as measured by rate of FVC decline and mortality and a lower value on potentially significant adverse effects and the cost of treatment.*".
- **BTS Guideline: Interstitial lung disease guideline: the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society** (Wells 2008a): The therapeutic benefit of pirfenidone in IPF has been recognised in this guideline, but it has

not been updated to reflect the safety and efficacy evidence from ASCEND, CAPACITY 1 & 2 clinical trials.

3.7 Issues relating to current clinical practice

Describe any issues relating to current clinical practice, including any variations or uncertainty about established practice.

Unmet need for patients with percent predicted FVC ≥80%: As IPF is a chronic and irreversible disease, with permanent damage caused to the lungs through scarring even at early stages of the disease, early treatment is warranted. Patients who are not eligible for (or cannot tolerate) pirfenidone or nintedanib are limited to only receive BSC, including all patients with FVC ≥80% predicted.

Lack of an agreed staging system in IPF - potential for misclassification using percent predicted FVC: In clinical practice, there are no accepted thresholds of percent predicted FVC used to define the disease severity of a patient with IPF, although there is a general acceptance that an FVC <50% predicted and DLco <35% predicted defines severe disease. The ATS/ERS/JRS/ALAT guidelines also do not propose a formal staging system for classification of disease severity (Raghu 2015). Other staging systems include the GAP index (Ley 2012), which includes age and gender as predictors of mortality, along with % predicted DLco. The composite physiology index adds forced expiratory volume in 1 second (FEV1) to FVC and DLco predicted values (Wells 2003). There is not, however, an agreed staging system used in UK clinical practice.

FVC values may not be fully reflective of the severity of the disease. Comments submitted during the scoping stage of this appraisal, along with evidence heard from clinicians during the NICE appraisal of nintedanib, emphasise that percent predicted FVC can be hard to interpret in the presence of comorbidities, specifically it may be elevated in the presence of emphysema thus masking significant lung disease. These patients may benefit from treatment with pirfenidone but are not eligible under the current NICE guidance.

Pirfenidone treatment is licensed for mild to moderate IPF patients in the UK. Pooled analysis of the ASCEND, and CAPACITY 1 & 2 studies demonstrate the magnitude

of treatment effect of pirfenidone on reducing the decline in FVC was comparable in both patients with earlier (FVC ≥80%) and more advanced IPF (FVC <80%). This finding supports the prompt use of pirfenidone in IPF patients with early disease after diagnosis, regardless of percent predicted FVC (Albera, 2015).

Lack of clinical basis for a stopping rule: There is no cure for IPF. NICE currently recommends the use of pirfenidone (or nintedanib) in adult patients with a predicted FVC between 50% and 80%, and recommends stopping treatment if patient's FVC falls by ≥10% in any 12 month period (NICE, 2013a; NICE, 2016).

The unpredictable and heterogeneous clinical course of IPF means that a prior decline in lung function does not predict a future decline: periods of stability can only be identified in retrospect. This implies that application of a stopping rule such as that described in TA282 is complicated, since progression with treatment does not always constitute treatment failure.

Analysis of outcomes from the three pivotal trials for pirfenidone showed that following a FVC decline ≥10% with pirfenidone, continued pirfenidone treatment significantly reduced the risk of death and increased disease stabilisation compared to placebo patients who continued treatment with placebo following a FVC decline ≥10%, see section 4.8 for further details (Nathan 2015a, Nathan 2016). These findings are suggestive of a salutary treatment effect with pirfenidone, beyond just the slowing in the rate of loss of lung function. The effects of pirfenidone on physical function (as measured by 6MWD) could also play a role in longer-term survival.

In addition to this effect on mortality beyond FVC alone, these finding also make clear that there should be no empirical rule for stopping treatment in a real-life setting, and pirfenidone should be continued in case of significant FVC progression.

Such a view is supported by comments submitted during the scoping stage of this appraisal. Stakeholders commented that the use of such a rule: "*may deny treatment to patients who may derive a morbidity/mortality benefit as there is no information to indicate that these benefits are limited to patients whose lung function declines at a slower rate*" (NICE, 2015e)

3.8 Equality

Roche Products Ltd. does not believe that the use of pirfenidone will be associated with any equality issues.

4 Clinical effectiveness

- A systematic literature review identified 10 RCTs of interest which enabled an indirect comparison on outcomes relevant to the decision problem
- Despite the classification of pirfenidone as an orphan treatment in the management of IPF, its treatment benefit is supported by a wealth of clinical data from 5 large randomised controlled trials (n=1716), with almost 9 years of follow-up to date.
- The primary studies (ASCEND, CAPACITY 1 & 2) are well-designed RCTs ranging from 9 months to 72 weeks. These are complemented by long term data capture from the RECAP study, providing evidence of treatment benefit up to 8.8 years.
- The median survival for a patient diagnosed with IPF is 3 years (Navaratnam 2011), and prolonging survival is important goal in order to allow patients to gain extra quality time with their families.
- Pirfenidone is the first and only treatment with a proven mortality benefit in IPF. The pre-specified pooled analysis of ASCEND, CAPACITY 1 & 2 at 52 weeks demonstrated pirfenidone to significantly reduce the mortality rate compared to placebo (HR=0.52; 95% CI: 0.31-0.87; p=0.011) (King 2014; Noble 2011).
- Lung function is an established predictor of outcome and the primary endpoint in most studies of IPF. In the ASCEND study, pirfenidone significantly reduced the change in percent predicted FVC or death compared to placebo at Week 52 ($p<0.001$). A similar effect was seen in the absolute change in percent predicted FVC in CAPACITY 2 (p=0.001), but did not reach statistical significant in CAPACITY 1 (p=0.501)
- In a sub group analysis from the pooled data from CAPACITY 1&2, those patients diagnosed more than a year before randomisation experienced a significantly greater treatment effect supporting
- Overall, safety results from the clinical study programme showed that pirfenidone was well-tolerated with a manageable side effect profile. Gastrointestinal and skin-related events were more common in the pirfenidone group compared to placebo, but rarely led to treatment discontinuation

4.1 Identification and selection of relevant studies

Search objective

The objective of this systematic review is to assess the efficacy and safety of pirfenidone and its comparators for the treatment of adult patients with mild to moderate IPF as per the decision problem. A search strategy was developed using PICOS elements to identify relevant studies for the technology. The literature search was conducted to identify RCTs of pirfenidone for the treatment of IPF, or any comparator studies in IPF to be used in a NMA. For the purpose of the systematic review of pirfenidone in IPF, only pirfenidone studies will be summarised in section 4.1, with results for comparator treatments described in section 4.10 of this submission.

The systematic literature review was conducted according to the NICE guide to the methods of technology appraisal 2013 and therefore adhered to the Centre for Reviews and Dissemination guidance for undertaking systematic reviews in health care.

A detailed search strategy for the systematic literature review can be found in Appendix 3. The strategy was designed to search for three concepts, structured as follows:

(IPF AND RCTs) OR pirfenidone

For the population and intervention concepts, the strategy took a sensitive search approach (both in structure and choice of terms). Given the reasonably low search retrieval numbers and the range and extent of interventions relevant to the indirect and mixed treatment comparisons, it was decided to not include specific terms for these interventions. The approach taken therefore maximises sensitivity – aiming to identify all RCTs in IPF, including interventions of interest. For the same reason (and again in the context of low retrieval numbers) it was decided to search on pirfenidone as an additional stand-alone concept.

Systematic searches were conducted in October 2011 to inform the InterMune NICE STA submission. The searches reported in this review were update searches to identify any literature published since October 2011. The searches were limited to the following records:

- Those added to the databases from January 2011 onwards;
- Those updated or indexed from January 2011 onwards;
- Those published from January 2011 onwards.

These dates reflect a conservative search approach to ensure records were not missed. The searches in the database and congress proceedings were conducted in April 2015.

The following sources were searched:

- Cochrane Central Register of Controlled Trials (CENTRAL)
- Cochrane Database of Systematic Reviews (CDSR)
- Database of Abstracts of Reviews of Effectiveness (DARE)
- Embase 1974 to 2013 December 19
- Health Technology Assessment database
- Ovid MEDLINE® in-process and other non-indexed citations
- Ovid MEDLINE® 1946 to present
- European Respiratory Society congress abstracts
- British Thoracic Society congress abstracts
- American Thoracic Society conference abstracts
- World Association for Sarcoidosis and Other Granulomatous Disorders conference abstracts

The reference list of relevant studies, papers and review articles were hand-searched for potentially relevant additional studies that may have been missed in the database searches.

Study selection for RCTs

The selection of the relevant RCTs was based on the eligibility criteria (Table 8) and was conducted in two phases as detailed below:

Phase 1:

The title and abstracts of the search results were assessed and categorised based on the eligibility criteria. Irrelevant records (such as animal studies, editorials, case reports and studies of conditions and interventions outside of the systematic review scope) were removed and the number of records removed was recorded. This process was undertaken by a single experienced information specialist.

Phase 2:

Full-text records were assessed in detail to select those addressing the review eligibility criteria. This assessment was undertaken by two independent researchers, with disagreements discussed and a third researcher involved when required. Where there was uncertainty about the relevance of a record based on the abstract it was included.

Electronic or paper copies of potentially relevant full papers meeting the systematic review inclusion criteria were obtained.

Table 8. Eligibility criteria used in the search strategy for RCTs

Clinical effectiveness	Inclusion criteria	Exclusion criteria
Population	Adults (aged 18 or older) with suspected or diagnosed IPF	Studies of children and young people <18 years Studies of people with a diagnosis of pulmonary fibrosis as a complication of either of the following: • Connective tissue disorders • A known exogenous agent (for example, drug induced disease or asbestosis)
Intervention	Pirfenidone	Any studies not containing pirfenidone
Comparators	Any comparator; • Best supportive care* (placebo) • Nintedanib	N/A
Outcomes	**Pulmonary function parameters** • Lung capacity (VC/FVC) • Categorical declines in FVC • Gas transfer (carbon monoxide diffusing capacity [DLco]) **Physical function** • Physical functioning (6MWD) **Exacerbation rate** • Hospitalisations • Acute exacerbations **Progression-free survival** **Mortality** • All-cause mortality • IPF-related mortality **AEs of treatment** **HRQoL** • St George's Respiratory Questionnaire (SGRQ) • University of California, San Diego Shortness of Breath Questionnaire (SOBQ) • EuroQoL five dimensions questionnaire (EQ-5D)	• Anticoagulation for the treatment of pulmonary hypertension • Treatment of lung cancer • Lung transplantation other than timing and referral
Study design	• Studies in humans • Phase II or III RCTs • Studies published as abstracts, conference presentations or press releases were eligible if adequate data were provided • Systematic reviews of RCTs**	• Cross-over RCTs
Language	No language limits	No language limits

*Best supportive care is defined as information and support, symptom relief, management of comorbidities, withdrawal of therapies suspected to be ineffective or causing harm, end-of-life care, oxygen therapy and/or pulmonary rehabilitation

**Systematic reviews were eligible for inclusion as a source of references to primary studies

Results

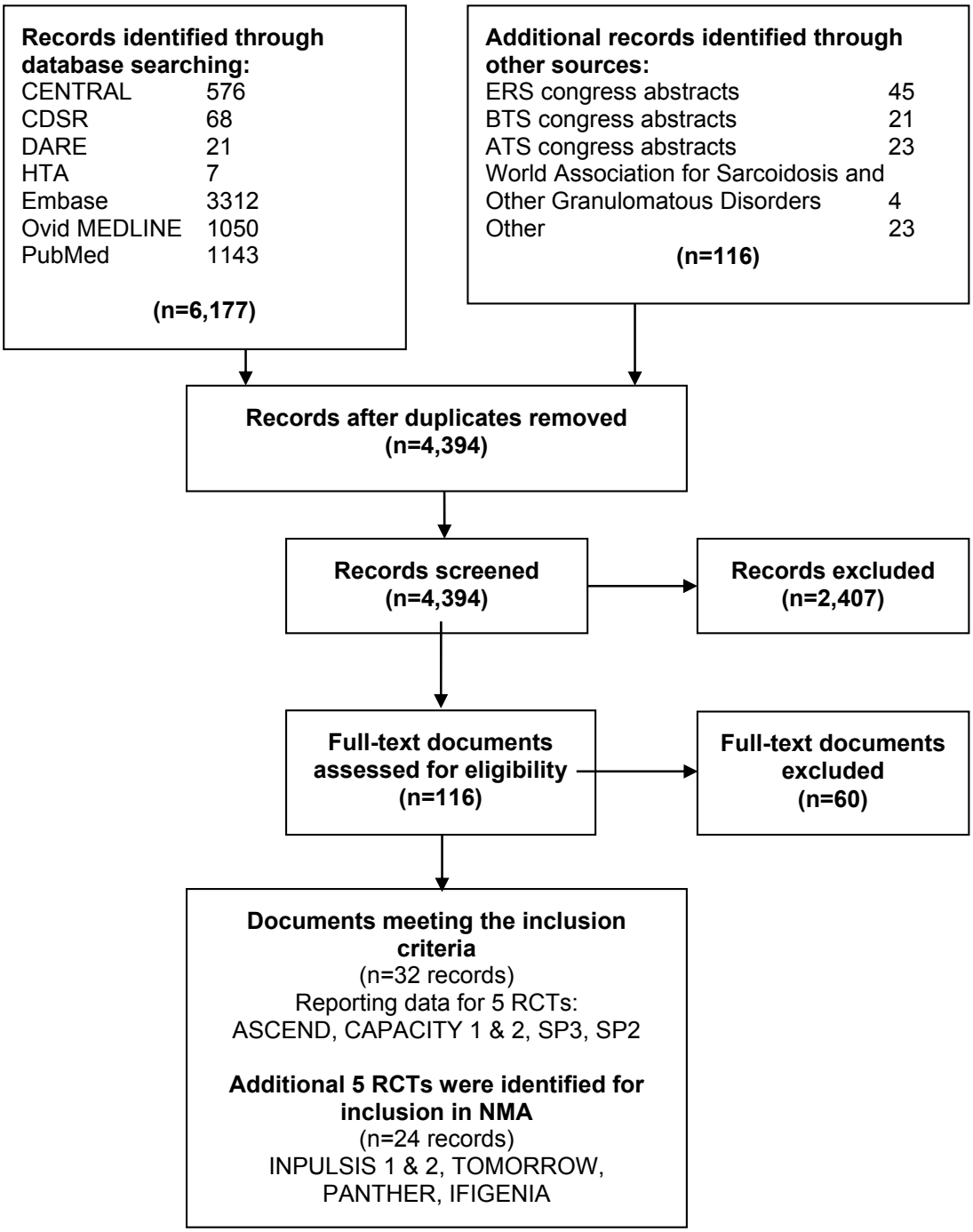
The search identified 6295 records from databases, registries, congress proceedings and reference lists of relevant articles (including 23 records from the original InterMune submission). Following the removal of duplicated records, the title and abstracts of 4394 records were screened against the eligibility criteria and 2407 records were further excluded. 116 full text documents were assessed against the eligibility criteria and as a result, 32 documents were identified which reported data on five RCTs of pirfenidone in IPF:

- PIPF-016 (ASCEND) 3 documents
- PIPF-006 and PIPF-004 (CAPACITY 1 and 2) 18 documents
- SP3 (Japanese Phase III study) 9 documents
- SP2 (Japanese Phase II study) 2 documents

A flow diagram of the numbers of studies included and excluded at each stage of selection process is provided in Figure 2. The systematic literature review identified 10 RCTs of interest which enabled an indirect comparison on outcomes relevant to this submission. RCTs of pirfenidone in IPF are summarised in sections 4.5 to 4.7, whilst RCTs of comparators in IPF can be found in section 4.10.

RECAP is an open-label extension to the ASCEND and CAPACITY trials, and will be described in section 4.11.

Figure 2. Flow diagram showing the study identification process



Records were divided into two groups: primary and associated references. Primary references presented original evidence and thus selected for data extraction. Associated references presenting additional data not already reported in the primary manuscript will be summarised within this submission.

Table 9. List of primary manuscripts and associated references for each trial

Study Name	References
ASCEND	**Primary manuscript**
	King TE et al. N Engl J Med 2014; 370 (22): 2083-2092
	Associated references
	Manuscript
	Lederer DJ et al. Chest 2015; 148(1): 196-201
	Congress abstract
	King TE et al. Am J Respir Crit Care. 2014: A6602.
CAPACITY 2 (PIPF-004) and CAPACITY 1 (PIPF-006)	**Primary manuscript**
	Noble PW et al. Lancet 2011; 377 (9779): 1760-9.
	Associated references
	Manuscript
	King TE et al. Am J Respir Crit Care. 2014;189(7):825-31.
	Valeyre D et al. Respirology. 2014;19(5):740-47.
	Kreutzkamp B. Krankenhauspharmazie. 2014;35(9):2142-43
	Congress abstracts
	Costabel U et al. Am J Respir Crit Care. 2010;181 (Meeting Abstracts 1)
	King TE et al. Am J Respir Crit Care. 2011;183 (Meeting Abstracts 1):[A5302].
	Noble PW et al. Am J Respir Crit Care. 2010;181 (Meeting Abstracts 1):[A1257].
	Sahn SA et al. Am J Respir Crit Care. 2011;183 (Meeting Abstracts 1):[A3810].
	Sahn SA et al. Am J Respir Crit Care. 2010;181 (Meeting Abstracts 1):[A6025].
	Valeyre D et al. Am J Respir Crit Care. 2010;181 (Meeting Abstracts 1):[A6026].
	Costabel U et al. ERS Annual Congress, Barcelona, Spain, September 18-22; 2010. p. [388].
	Noble P et al. ATS International Conference, May 15-20, San Diego; 2009. p. A1129 [Poster #216].
	Noble PW et al. ATS International Conference, May 15-20, 2009 San Diego; 2009. p. [C98].
	Roskell R et al. ERS International Congress 2014. Munich: European Respiratory Society; 2014. p. [A1905].
	Albera C et al. ERS International Congress. Barcelona; 2010. p. [A389].
	Du Bois R et al. ERS International Congress. Vienna; 2009. p. [A2823].
	Valeyre D et al. ERS International Congress. Barcelona; 2010. p. [A391].
	Karimi-Shah BA et al. N Engl J Med. 2015;372(13):1189-91.
SP3	**Primary manuscript**
	Taniguchi H et al. European Respiratory Journal. 2010; 35 (4): 821–829.
	Associated references
	Manuscript
	Azuma A et al. Respiratory Research. 2011;12:[143].
	Congress abstracts
	Ebina M et al. Am J Respir Crit Care. 2010;181 (Meeting Abstracts 1):[A3988].
	Taniguchi H et al. Resp Res. 2011;12:[93].
	Ogura T et al. ATS annual conference, May 16-21, 2008, Toronto; 2008. p. [A768].

	Baba T. Asian Pacific Society of Respirology 14th Congress. Seoul; 2009. p. [A375].
	Ebina M et al. ERS International Congress. Vienna; 2009. p. [P666].
	Taguchi Y et al. ERS International Congress. Barcelona; 2010. p. [A390].
	Taniguchi H et al. ERS International Congress. Amsterdam; 2011. p. [A649].
SP2	**Primary manuscript**
	Azuma A et al. Am J Respir Crit Care. 2005; 171 (9): 1040-1047
	Associated references
	Congress abstract
	Azuma, A et al. Am J Respir Crit Care, 2002. Abstract A729.

The clinical effectiveness summary of the pirfenidone trials is based on data from the primary manuscripts and clinical study reports (CSRs). Congress abstracts will be summarised if the outcome has not already been reported in the primary manuscript.

We are also aware that the following publication has been recently published, which has not been captured in the literature search:

- Noble PW, Albera C, Bradford WZ. Pirfenidone for idiopathic pulmonary fibrosis: analysis of pooled data from three multinational Phase 3 trials. European Respiratory Journal 2015; doi: 10.1183/13993003.00026-2015
- Albera C. Pirfenidone is efficacious in Patients With Idiopathic Pulmonary Fibrosis (IPF) and Mild Restrictive Disease: Benefit of Early Intervention. Oral presentation presented at ATS annual congress 2015; 15-20 May; Denver, USA

The outcomes for each study included in this pooled analysis have been reported in the primary manuscripts of each study (which were used for the NMA). The exclusion of this publication therefore does not affect the NMA results. A list of excluded studies can be found in Appendix 4.

4.2 List of relevant randomised controlled trials

Five studies were identified that met the inclusion criteria, providing evidence of pirfenidone in IPF; ASCEND (Phase III), CAPACITY 1 & 2 (Phase III), SP3 (Phase III), and SP2 (Phase II). All studies compared pirfenidone to placebo. There are no

head-to-head studies comparing pirfenidone to nintedanib. A summary of the identified RCTs are listed in Table 10.

All the relevant studies which have been identified from the literature search and have met the selection criteria have been included in this report.

Table 10. List of relevant RCTs

Trial number (Acronym) No. of patients	Patient population & enrolment criteria			Intervention	Comparator	Primary study ref.
	Country	FVC (or VC); %-predicted	DLco; %-predicted			
PIPF-016 (ASCEND) n=555	Multi-national	50-90%	30-90%	Pirfenidone (p.o.) 2403 mg/day	Placebo	King 2014
PIPF-004 (CAPACITY 2) n=435	Multi-national	≥50%	FVC or Hgb-corrected DLco ≤90%	Pirfenidone (p.o.) 2403 mg/day or 1197 mg/day	Placebo	Noble 2011
PIPF-006 (CAPACITY 1) n=344	Multi-national	≥50%	FVC or Hgb-corrected DLco ≤90%	Pirfenidone (p.o.) 2403 mg/day	Placebo	Noble 2011
SP3* n=275	Japan	Not reported	Not reported	Pirfenidone (p.o.) 1800 mg/day or 1200mg/day	Placebo	Taniguchi 2010
SP2* n=107	Japan	Not reported	Not reported	Pirfenidone (p.o.) 1800mg/day	Placebo	Azuma 2005

* SP3 and SP2 enrolled patients with IPF who met the following arterial oxygen saturation measured by pulse oximetry (Sp,O2) criteria:
  - 1) oxygen desaturation of ≥5% difference between resting Sp,O2 and the lowest Sp,O2 during a 6-min steady-state exercise test (6MET); and
  - 2) the lowest Sp,O2 during the 6MET of ≥85% while breathing air.

Full inclusion exclusion criteria are described in Appendix 5

4.3 Summary of methodology of the relevant randomised controlled trials

ASCEND (PIPF-016) (King 2014; InterMune, 2014)

Trial design: ASCEND was a Phase III, double-blind, randomised, placebo-controlled, 52-week study designed to provide additional evidence of the effect of pirfenidone on disease progression in IPF patients. Patients were randomly assigned (1:1 ratio) to receive either oral pirfenidone 2403 mg/day or placebo using permuted block randomisation, without stratification on any variable.

Participants: Eligible participants were all adults aged between 40-80 years whom received a centrally confirmed clinical and radiographic diagnosis of IPF. Key criteria for enrolment include 50-90% of the predicted FVC, 30-90% of the predicted DLco, FEV_1:FVC ratio of ≥0.80, and a 6MWD of ≥150 m. A comprehensive list of the inclusion and exclusion criteria can be found in Appendix 5.

Study settings: The multicentre study was conducted in 127 investigational sites in nine countries (US, Australia, Peru, Brazil, Israel, Mexico, New Zealand, Croatia, and Singapore).

Interventions: In the ASCEND (PIPF-016) study, patients received pirfenidone or placebo equivalent orally in three equally divided doses with food at the same time each day, with the dose gradually increased over 14 days to the full dose. The study drug was titrated according to the following schedule:

- **Days 1–7:** One capsule TID
- **Days 8–14:** Two capsules TID
- **Day 15 and continuing:** Three capsules TID (full dose)

For those assigned to the pirfenidone study group, each capsule contained pirfenidone 267 mg. Pirfenidone and placebo capsules were supplied as opaque, hard, white gelatin capsules that were visually indistinguishable, with identical packaging and labeling.

Patients remained on a stable maintenance dose for the duration of the study unless the dose was reduced to manage an AE. It was the responsibility of the investigator to monitor patients as frequently as clinically indicated for toxicities and to manage the patient accordingly. The ultimate decision regarding study treatment interruption, restart, and dose modification was the responsibility of the investigator.

Outcomes: The primary outcome was the change in FVC or death at Week 52. Secondary outcomes were the 6MWD, PFS, dyspnoea, and death from any cause or from IPF. Physical examination and clinical laboratory assessments were performed at baseline and at Weeks 2, 4, 8, 13, 26, 39, and 52. Pulmonary function, exercise tolerance, and dyspnoea were assessed at baseline and at weeks 13, 26, 39, and 52. Central reviewers of Biomedical Systems, who were blinded to the study drug assignments, evaluated all FVC results for adequacy and repeatability, according to the ATS criteria.

CAPACITY 2 (PIPF-004) and CAPACITY 1 (PIPF-006) (Noble 2011)

Trial design: CAPACITY 1 & 2 were concurrent studies designed to confirm the effect of pirfenidone on reduction of decline in lung function in IPF patients. Both were Phase III, double-blind, randomised, placebo-controlled, multicentre studies evaluating the safety and efficacy of pirfenidone in IPF patients.

In CAPACITY 2, participants were randomly assigned (2:1:2 ratio) to receive oral pirfenidone 2403 mg/day, pirfenidone 1197 mg/day, or placebo.

In CAPACITY 1, participants were randomly assigned (1:1 ratio) to receive oral pirfenidone 2403 mg/day or placebo.

Participants: Eligible patients were 40-80 years with a diagnosis of IPF in the previous 48 months, with no evidence of improvement in measures of disease severity over the preceding year. Key criteria for enrolment include percent predicted FVC ≥50%, predicted DLco ≥35%, either FVC or Hgb-corrected DLco ≤90%, 6MWD ≥150m. Exclusion criteria include patients with obstructive airway disease, connective tissue disease, alternative explanation for interstitial lung disease, and being on a waiting list for lung transplant. A comprehensive list of the inclusion and exclusion criteria can be found in Appendix 5.

Study settings: The studies were conducted in 110 investigational sites in thirteen countries (US, Canada, Italy, Germany, France, Spain, UK, Australia, Belgium, Poland, Ireland, Mexico, and Switzerland).

Interventions: Patients received the study drug (pirfenidone or placebo equivalent) orally in three equally divided doses with food at the same time each day, with the dose gradually increased to the full dose over 15 days. The study drug was titrated according to the following schedule:

- **Days 1–7:** One capsule TID
- **Days 8–14:** Two capsules TID
- **Day 15 and continuing:** Three capsules TID (full dose)

Pirfenidone was supplied as 267 mg capsules for the 2403 mg/day dose, or as 133 mg capsules for the 1197 mg/day dose. Pirfenidone and placebo capsules were supplied as opaque, hard, white gelatin capsules that were visually indistinguishable. Patients who followed a different dose-escalation schedule in the initial 2-week period were not considered as protocol deviators.

Patients remained on a stable maintenance dose for the duration of the study unless the dose was reduced to manage an AE. The dose of study treatment was to be modified at the investigator's discretion in the event of adverse effects or intolerability after discussion with the medical monitor.

Outcomes: The primary outcome in both studies was change in percentage of predicted FVC from baseline to Week 72. Secondary outcomes in both studies were categorical FVC (5-level scale), PFS, worsening in IPF, dyspnoea (using the University of California San Diego Shortness of Breath Questionnaire), 6MWD, worst peripheral oxygen saturation during the 6MWT, and percentage predicted DLco. In PIPF-006 (CAPACITY 1), fibrosis by use of high resolution computed tomography (HRCT) was an additional secondary outcome. Physical examinations and clinical laboratory assessments were performed at weeks 2, 4, 6, and 12, and then every 12 weeks. Pulmonary function, exercise tolerance, and dyspnoea were assessed every 12 weeks. Patients continued assessments until study completion, even after

permanent treatment discontinuation, and all such assessments were included in the ITT analyses.

RECAP: Patients who completed ASCEND or either of the CAPACITY studies were offered the opportunity to enrol in RECAP, an open-label extension study evaluating the effect of long-term treatment with pirfenidone, see section 4.11 (Costabel 2014, Kreuter 2014).

SP3 (Taniguchi 2010)

Trial design: SP3 was a Shionogi-sponsored Phase III, double-blind, randomised, placebo-controlled, multicentre study evaluating the efficacy and safety of pirfenidone in IPF patients in 73 centres in Japan. Participants were randomly assigned (2:1:2 ratio) to oral pirfenidone 1800 mg/day, pirfenidone 1200 mg/day, or placebo using a modified minimisation method, which included some random allocation based on biased coin design to balance baseline SpO_2.

Participants: All participants were aged between 20-75 years and must have received a confirmed diagnosis according to the ATS/ERS consensus statement (ATS/ERS, 2000). Details of the inclusion and exclusion criteria for the trial are included in Appendix 5.

Study settings: The study was conducted in 73 centres in Japan.

Interventions: In the SP3 study, participants received the tablet-formulation of the study drug (pirfenidone and/or placebo equivalent) orally in three divided doses, with the dose gradually increased over four weeks to the full dose. All participants received the same counts of tablets. As described in Section 1.3, the Japanese 1800 mg/day dose is considered to be equivalent to UK doses as the dose by weight would be similar for all studies, and has previously been considered by NICE (NICE, 2013a section 3.24).

Table 11. Treatment regimen titration in the SP3 study

	Pirfenidone 1800 mg/day arm	Pirfenidone 1200 mg/day arm	Placebo arm
Days 1-14	1x pirfenidone 200mg tablet TID	1x pirfenidone 200mg tablet TID	1x placebo tablet TID
Days 14-28	2x pirfenidone 200mg tablets TID	1x pirfenidone 200mg tablet TID and 1x placebo tablet TID	2x placebo tablets TID
Day 29 and continuing	3x pirfenidone 200mg tablets TID	2x pirfenidone 200mg tablets TID and 1x placebo tablet TID	3x placebo tablets TID

The appearance of pirfenidone and placebo tablets was manufactured as similar as possible with respect to physical characteristics.

Outcomes: The primary outcome was the change in VC from baseline to Week 52. Secondary outcomes were PFS time and change in the lowest SpO_2 during the 6MWT. Tertiary outcomes were pulmonary function tests (arterial oxygen tension, alveolar-arterial oxygen tension difference at rest, total lung capacity, and DLco), acute exacerbation, serum levels of markers of interstitial pneumonias, and subjective/objective symptoms (cough, presence/absence of sputum and dyspnoea in daily living assessed with Hugh-Jones classification).

VC was measured every four weeks, and the lowest SpO_2 during 6MWT and other pulmonary function tests were assessed every 12 weeks.

The original primary endpoint was change in lowest SpO_2 during the 6MWT over 52 weeks. However, a decision was made to revise the primary endpoint due to recommendation by the independent DSMB due to evolved knowledge of assessment with objective measurements in IPF, as well as the lack of validation in the 6MWT study, and difficulties in reproducibility of the SpO_2 measurements during the 6MWT. Change in lowest SpO_2 during 6MWT was changed to a secondary endpoint.

SP2 (Azuma 2005)

Trial design: SP2 was a Phase II, multicentre, randomised, double-blind, placebo-controlled, prospective study in IPF patients in 25 centres in Japan. Participants were

randomly assigned (2:1 ratio) to receive oral pirfenidone 1800 mg/day or placebo using a modified permuted-block randomisation method with block sizes of six. The planned study duration was 48 weeks, but based on a planned interim analysis at Week 24, the Efficacy and Safety Assessment Committee recommended stopping the study due to high incidence of acute exacerbations in the placebo arm. The study was unblinded at Week 36, and all patients were given the option to be treated with open-label pirfenidone through to Week 48.

Participants: All participants were aged between 20-75 years, and must have received a confirmed diagnosis of IPF according to the ATS/ERS consensus statement (ATS, 2000). Details of the inclusion and exclusion criteria for the trial are included in Appendix 5.

Study settings: The study was conducted in 25 sites in Japan.

Interventions: Participants received the study drug (pirfenidone or placebo equivalent) in three divided doses, with the dose gradually increased over a week to the full dose. The study drug was titrated using the following schedule:

- **Days 1 and 2:** One tablet TID
- **Days 3 and 4:** Two tablets TID
- **Days 5 and continuing:** Three tablets TID (full dose)

For those assigned to receive pirfenidone treatment, each tablet was pirfenidone 200mg. The maximum pirfenidone dose (1800 mg/day) was maintained in patients tolerating it throughout the study unless the dose was reduced to manage an AE. The dose of study treatment was to be modified according to a prespecified regimen utilising the Standards for Classification of Serious Adverse Drug Reactions.

Outcomes: The primary endpoint was the change in the lowest SpO2 during 6MWT. Secondary endpoints were changes in resting pulmonary function tests while breathing air (VC, TLC, DLco, PaO2), disease progression patterns assessed via HCRT, acute exacerbations, change in serum markers (pneumocyte damage), and change in HRQoL (Chronic Respiratory Disease Questionnaire Score and Hugh-Jones Classification Score).

Lowest SpO2 during 6MWT was measured at baseline and then at 3-monthly follow-up visits. In addition, the difference in SpO2 area between baseline and follow-up test at 6 or 9 months was also determined. Pulmonary function tests were performed at baseline and 3-monthly thereafter. All HRCT scans were performed at baseline and 6-month intervals. Serum KL-6 and surfactant protein-D levels were measured to assess changes in blood levels during the study. The Chronic Respiratory Disease Questionnaire Score and Hugh-Jones Classification Score were used to assess patient's perceived QoL during the study.

4.3.2 Provide a comparative summary of the methodology of the RCTs in a table. A suggested table format is presented below.

Table 12. Comparative summary of trial methodology for the relevant RCTs

Trial number (Acronym) Reference	PIPF-0016 (ASCEND) King 2014	PIPF-004 (CAPACITY 2) Noble 2011	PIPF-006 (CAPACITY 1) Noble 2011	SP3 Taniguchi 2010	SP2 Azuma 2006
Location	127 sites in Australia, Brazil, Croatia, Israel, Mexico, New Zealand, Peru, Singapore, and the US.	110 centres in Australia, Europe, and North America.(including 3 sites in the UK)		73 centres in Japan.	25 centres in Japan.
Trial design	Phase III, randomised, double-blind, placebo-controlled trial.	Phase III, randomised, double-blinded, placebo-controlled trial.		Phase III, randomised, double-blind, placebo-controlled trial.	Phase II, randomised, double-blind, placebo-controlled trial
Eligibility criteria for participants	Patients aged 40–80 years with a diagnosis of IPF in accordance with the International consensus statement [ATS, 2000] at least 6 months and no more than 48 months before randomisation, confirmed by central review. Percent predicted FVC ≥50% and ≤90% at screening were included. The subjects had to meet all of the inclusion criteria and none of the exclusion criteria.	Patients aged 40–80 years with a diagnosis of IPF in accordance with the International consensus statement [ATS, 2000] in the previous 48 months with no evidence of improvement in measures of disease severity over the preceding year. Percent predicted FVC ≥50% at Screening and Day 1 (before randomisation) were included. The subjects had to meet all of the inclusion criteria and none of the exclusion criteria.		Patients aged 20 -75 years, diagnosed with IPF in accordance with the International consensus statement [ATS/ERS, 2000]. The subjects had to meet all of the inclusion criteria and none of the exclusion criteria.	Patients aged 20 -75 years, diagnosed with IPF in accordance with the International consensus statement [ATS/ERS, 2000]. The subjects had to meet all of the inclusion criteria and none of the exclusion criteria.
Trial drugs, permitted and disallowed concomitant medication	Patients received pirfenidone 2403 mg/day or placebo. Concomitant treatment with any investigational drug for the treatment of IPF was prohibited. However, concomitant medications	Patients received pirfenidone 2403 mg/day, pirfenidone 1197 mg/day, or placebo. Concomitant treatments	Patients received pirfenidone 2403 mg/day or placebo. Concomitant treatments for IPF were prohibited, with exceptions of short	Pirfenidone 1800 mg daily, 1,200 mg daily or placebo. Concomitant corticosteroid ≤10mg/day (as the	Pirfenidone 1800 mg daily or placebo. Concomitant prednisone ≤10mg/day was allowed. The following immunosuppressants or

	used for another indication were permitted if there was no clinically acceptable alternative.	for IPF were prohibited, with exceptions of short courses of azathioprine, cyclophosphamide, corticosteroids, or acetylcysteine for protocol-defined acute exacerbation of IPF, acute respiratory decompensation, or progression of disease.	courses of azathioprine, cyclophosphamide, corticosteroids, or acetylcysteine for protocol-defined acute exacerbation of IPF, acute respiratory decompensation, or progression of disease.	prednisone equivalent) was allowed. However, concomitant immunosuppressants or other investigational drugs for IPF were not allowed.	other anti-inflammatory/ antifibrotic drugs were not allowed: cyclophosphamide, azathioprine, methotrexate, d-penicallimine, cochicine, erythromycin, IFNs, N-acetylcysteine, cyclosporine, tacrolimus and other investigational drugs for IPF.
Primacy outcomes	Change in percent predicted FVC or death at Week 52.	Change in percent predicted FVC from baseline to Week 72.		Change in VC from baseline to Week 52 (originally was the change in lowest SpO_2 during the 6MWT).	Change in the lowest SpO_2 reached during the 6-MWT.
Secondary outcomes	Change from baseline to Week 52 in the 6-minute walk distance and progression-free survival (defined as the time to the first occurrence of any one of the following: a confirmed decrease of 10 percentage points or more in the percentage of the predicted FVC, a confirmed decrease of 50 m or more in the 6-minute walk distance, or death); change in dyspnoea measured with the use of the University of California San Diego Shortness of Breath Questionnaire (UCSD SOBQ); the rate of death from any cause; and	Categorical FVC, progression-free survival (time to confirmed ≥10% decline in percentage predicted FVC, ≥15% decline in percentage predicted DLco or death), worsening IPF(time to acute exacerbation, death, lung transplantation, or admission to hospital for respiratory problems), dyspnoea (SOBQ), 6MWD, worst peripheral oxygen saturation (SpO_2) during the 6MWT, percentage predicted DLco, and fibrosis by use of HRCT.		Progression free survival time, change in the lowest SpO_2 during the 6MWT	Changes in resting PFTs while breathing air (VC, TLC, DLco, PaO_2), disease progression patterns (HRCT), acute exacerbation episode (IPF), change in serum markers (pneumocyte damage), change in HRQoL

	the rate of death from IPF during the period from baseline to 28 days after the last dose of the study drug.			

4.4 Statistical analysis and definition of study groups in the relevant randomised controlled trials

The primary outcome of ASCEND compared change in percentage predicted FVC between the 2403 mg/day and placebo groups (King 2014). It was estimated that 250 patients per group would provide at least 90% power to detect a difference in normalised ranks of 0.08 (± 0.27). For change in 6MWD, 250 patients per group were expected to provide approximately 75% power to detect a difference in normalised ranks of 0.068.

The statistical analysis plan for ASCEND also included a pooled analysis of rates of death from any cause and IPF-related death from the ASCEND and two CAPACITY trials. This analysis was performed for the purpose of increasing the statistical power and deriving a more stable estimate of the treatment effect.

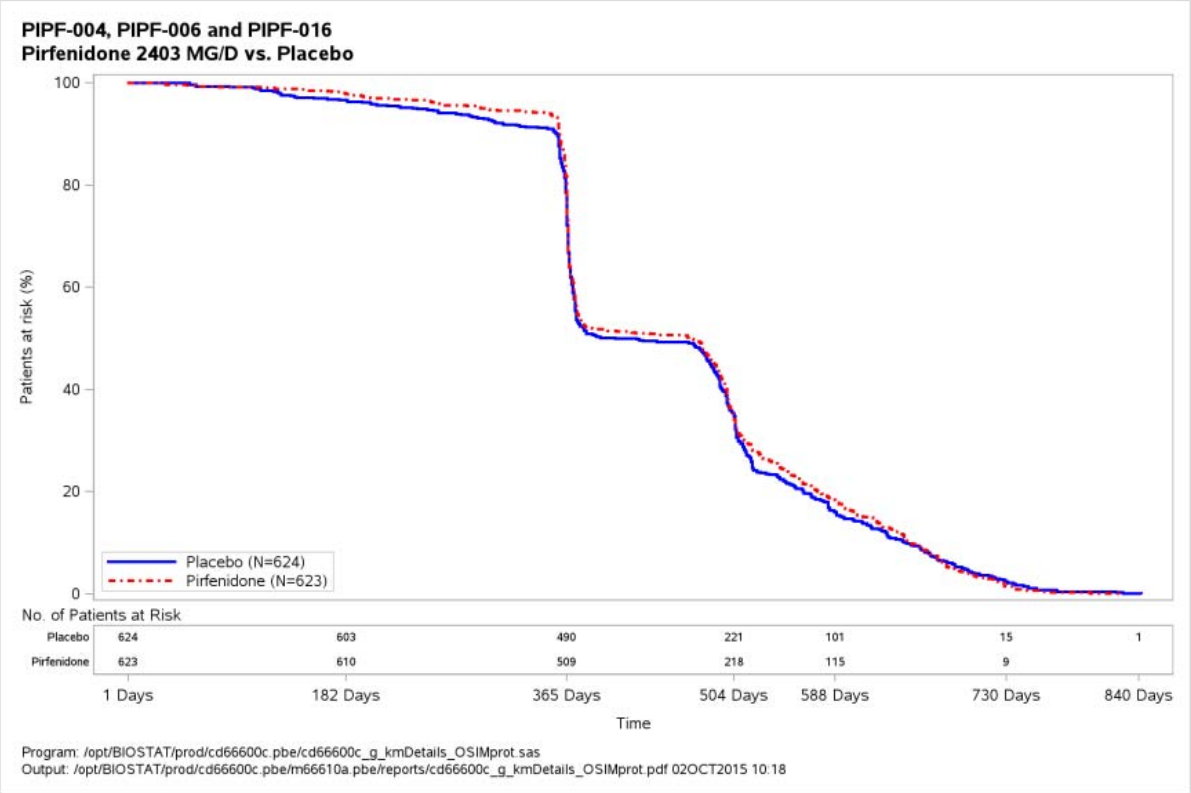
ASCEND, CAPACITY 2 and CAPACITY 1 were substantially similar in study design, eligibility criteria, patient population, intervention and comparator. This renders pooling of the data appropriate in order to provide a single, stable and robust estimate of the treatment effect of pirfenidone. The appropriateness and usefulness of pooling the data from the three pivotal trials has been recognised and accepted as valid by the EMA CHMP: "*…the new pooling of the 52 week efficacy data have further demonstrated the efficacy of pirfenidone in IPF, especially with regards (to the) rate of FVC decline, 6MWT performance and mortality data. The inclusion of PIPF-016 (ASCEND) is warranted within the SmPC as the data further enhances the body of evidence to aid the healthcare professional in prescribing the medicine. The pooled survival data is also useful for the same reason."*

The independent statistical assessor in the EMA evaluation further commented: "*The methods used in the pooled analysis are analogous to those in the individual studies and are appropriate as discussed in the original MAA and above*." (EMA 2014).

The pooled analysis for overall survival was pre-specified to be conducted at 52 weeks, as all patients from the three studies contributing to the analysis were to be followed up until at least 52 weeks. As described in Figure 3 below, the number of patients at risk beyond weeks 52 and weeks 72 falls dramatically. Data from

ASCEND is removed from the information after week 52, and patients became eligible to receive pirfenidone after placebo as part of the RECAP study from week 72 onwards. A 52 week timepoint also matches the available data for nintedanib for this outcome (Section 4.10).

Figure 3 Proportion of patients still at risk by days after randomisation



The CAPACITY 1 & 2 studies' primary outcome compared change in percent predicted FVC between the 2403 mg/day and placebo groups (Noble 2011). It was estimated that 160 patients per group would provide 97% power to detect a 50% relative reduction in the rate of FVC change from baseline to Week 72 at a 0.05 significance level. This assumed an absolute percent predicted FVC change between baseline and Week 72 of 5.5% in the placebo arm and 2.75% in pirfenidone arm with a standard deviation of 6.0%.

The power and sample size calculations in SP2 and SP3 were based on simulation studies with the use of the lowest SpO2 achieved during a 6MWT after one year (Azuma 2005, Taniguchi 2010). The trials recruited 90 patients (SP2) and 250 patients (SP3). These numbers were expected to provide 80% power to detect

assumed differences of the mean changes in the lowest SpO2 from baseline to Week 52 between the two groups at a significance level of 0.025 for the SP2 trial and 0.1 (two-sided) in SP3. The SP3 trial noted that although the primary endpoint was altered from the lowest SpO2 to VC after the study was started, the power calculated on the basis of the change in VC turned out to be the same (maintained at ~ 80%) and therefore the planned sample size was not altered.

Details of the statistical analysis and definition of study groups in the relevant RCTs have been summarised in following table using items 7a, 7b, 12a and 12b of the CONSORT checklist.

Table 13. Summary of statistical analyses in the RCTs

Study (Reference)	Hypothesis	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
ASCEND (King 2014)	Treatment with pirfenidone would reduce the deterioration in lung function in patients with idiopathic pulmonary fibrosis.	The test statistic for the primary efficacy analysis was a ranked analysis of covariance (ANCOVA), with the average standardised rank change in the percentage of the predicted FVC as the outcome variable and the standardised rank baseline value as a covariate. The key secondary end points were analysed with the use of the Hochberg procedure for multiple comparisons. For time-to-event analyses, pirfenidone was compared with placebo using a log-rank test; hazard ratios were based on the Cox proportional-hazards model.	For the primary efficacy comparison of change in %FVC between the 2403 mg/day and placebo groups, 250 patients per group would provide at least 90% power to detect a difference in normalised ranks of 0.08 with a standard deviation of 0.27. For change in 6MWD, 250 patients per group will provide approximately 75% power to detect a difference in normalised ranks of 0.068 with a standard deviation of 0.285	The authors used the intent-to-treat (ITT) population in the efficacy analysis, which consisted of all patients who signed the informed consent form and were randomised, The safety analysis population included all patients who signed informed consent and received any amount of study drug. For the ranked ANCOVA analyses, missing values owing to death were assigned the worst ranks, with early deaths ranked worse than later deaths. In the analyses of mean change, missing values owing to death were assigned the worst possible outcome (e.g. FVC=0). Missing values with reasons other than death were imputed as the average value for the three patients with the smallest sum of squared differences at each visit.

Study (Reference)	Hypothesis	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
CAPACITY 1 & 2 (Noble 2011)	Treatment with pirfenidone would reduce the deterioration in lung function in patients with IPF. The null hypothesis for the primary efficacy outcome variable was that there is no difference in the absolute change in percent predicted FVC from baseline to Week 72 between the pirfenidone 2403 mg/day group and the placebo group. Similarly, in all other hypothesis tests of secondary and exploratory outcome variables, the null hypothesis was that there is no treatment difference between the pirfenidone 2403 mg/day group and the placebo group.	The primary efficacy analysis was by use of a rank analysis of covariance (ANCOVA) model, stratified by region, with standardised rank change in FVC as the outcome and standardised rank baseline percentage predicted FVC as a covariate, evaluated against a final adjusted two-tailed p value of 0·0498. Magnitude of treatment effect was estimated by use of differences in treatment group means and categorical change in FVC. To assess treatment effect over the full study, a repeated-measures analysis with averaging of percentage predicted FVC change over all assessment time points was pre-specified. In the efficacy analyses, pirfenidone 2403 mg/day was compared with placebo in the intention to treat (ITT) population. The group assigned to pirfenidone 1197 mg/day in study PIPF-004 was summarised descriptively. Analyses of pooled data were pre-specified to derive precise estimates of magnitude of treatment effect.	The sample size and treatment duration provided approximately 97% power to detect a 50% reduction in the rate of FVC progression after 72 weeks of treatment with pirfenidone compared to placebo and also increased the power on the various secondary endpoints. For the primary efficacy comparison of change in percent predicted FVC between the 2403 mg/day and placebo groups, 160 patients per group were expected to provide 97% power to detect a 50% relative reduction in the rate of FVC change from Baseline to Week 72 at 0.05 significance level. This assumed an absolute percent predicted FVC change between Baseline and Week 72 of 5.5% in the placebo arm and 2.75% in pirfenidone arm with a standard deviation of 6.0%.	In all analyses, missing values at Week 72 due to death, discontinuation, or other reasons were imputed to ensure that all patients were included. In the rank ANCOVA, patients with missing data due to death were ranked worse than those who remained alive. Patients who died were ranked according to the number of days from randomisation until death, with the shortest time until death as the worst rank. To estimate mean change in percent predicted FVC from Baseline at Week 72, missing Week 72 FVC values for patients who died before Week 72 were replaced with a percent predicted FVC of zero. Missing values at Week 72 due to reasons other than death (e.g., early withdrawal from the study, lung transplantation) were imputed with the average value from three patients with the smallest sum of squared differences at each visit with data that were not missing. A data monitoring committee reviewed safety and efficacy data and undertook two interim analyses of all-cause mortality in the pooled dataset against a conservative stopping boundary of p=0·0001. For time-to-event analyses, appropriate censoring methodology wasused.

Study (Reference)	Hypothesis	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
SP3 (Taniguchi 2010)	Treatment with pirfenidone would reduce the deterioration in lung function in patients with IPF.	Analyses of the change in VC and the lowest SpO_2 from baseline were performed with ANCOVA using the respective baseline measurements as covariates. Analyses of the change in other pulmonary function tests and the serum levels of the markers of interstitial pneumonias were performed with the least significant difference method based on one-way ANOVA. The cumulative PFS rates were estimated using the Kaplan–Meier method and compared using the log-rank test. Incidences were compared with Fisher's exact test.	The planned sample size was 250 in total: 100, 50 and 100 patients in the high-dose, low-dose and placebo groups, respectively. The sample sizes of 100 for the high-dose and placebo groups were determined on the basis of simulations that would provide a statistical power of 0.8 to detect assumed differences of the mean changes in the lowest SpO_2 from baseline to Week 52 between the two groups at a significance level in this study of 0.1 (two-sided). Although the primary endpoint was altered from the lowest SpO_2 to VC after the study was started, the power calculated on the basis of the change in VC turned out to be the same (maintained at ~ 0.8) and, thus, the planned sample size was not altered.	As the low-dose group was intended to assess benefit–risk profiles of pirfenidone treatment at a tapered dose, the sample size of the low-dose group was obtained by halving the sample size of the high-dose and placebo groups. Multiplicity problems were not taken into account because the main analysis was the comparison between the high-dose and placebo groups. The principle of the last observation carried forward (LOCF) was adopted to impute missing values if patient data were available for ≥4 Weeks after the baseline. The mixed model approach using the available repeated measures of changes in VC was performed as a sensitivity analysis.

Study (Reference)	Hypothesis	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
SP2 (Azuma 2005)	Treatment with pirfenidone would reduce the deterioration in lung function in patients with IPF.	Analysis of change from baseline was performed with the Welch's t test. Categorical variables were analysed with the Wilcoxon's test. Analyses of incidences were performed with Fisher's exact test.	The pre-specified sample size was 90 patients based on a simulation study with the use of the lowest SpO_2 achieved during a 6MWT after 1-year duration of the study. This minimum number of patients provided statistical power greater than 0.8 to detect assumed efficacy at the significance level of 0.025.	For missing values, the principle of last observation carry forward (LOCF) was adopted. Immediately after initiating the trial, a decision to conduct a pre-specified analysis (i.e., before breaking the code) in the subset of patients who were able to complete the 6MWT without the SpO_2 reaching less than 80% at baseline was made. Based primarily on important 6-month trends in a secondary endpoint, the DSMB recommended early termination of the trial on ethical grounds. Due to the length of time needed to collect, analyse, report a minimum of 6 months of data for DSMB review, all patients in the trial completed a minimum of 9 months on their assigned treatment arm. While both 6 and 9 month results are presented, the latest complete follow-up exam (9 months) most closely match the planned length of follow-up and are considered of primary importance.

4.5 *Participant flow in the relevant randomised controlled trials*

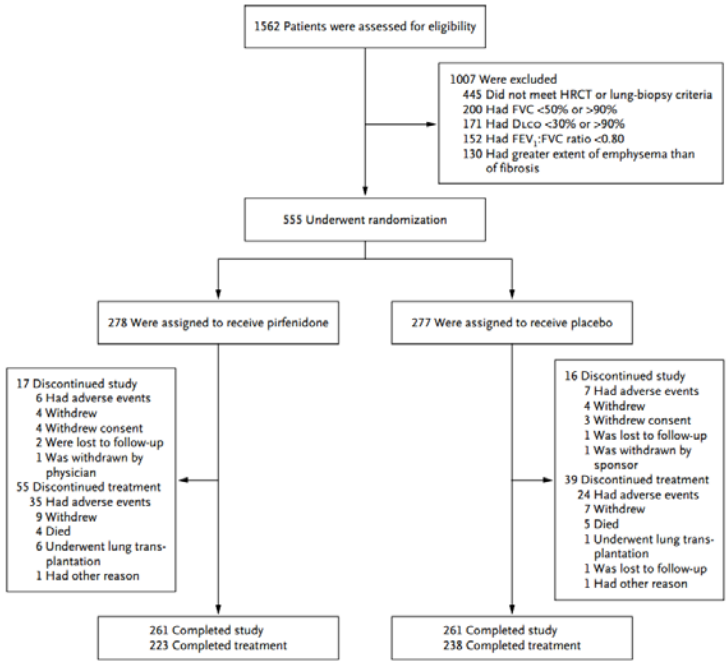
ASCEND study (King 2014)

A total of 1562 patients were screened, and 555 patients with IPF were randomly assigned in a 1:1 ratio to receive pirfenidone 2403 mg/day or placebo. The most common reason for exclusion from the study was that the HRCT or lung biopsy criteria were not met.

At 52 weeks, 55 patients discontinued treatment in the pirfenidone group compared to 39 in the placebo group. The most common reason for discontinuation of treatment and study were adverse events. The numbers of participants who withdrew from the study do not include participants who died or underwent lung transplantation.

The CONSORT diagram shows the flow of participants through each stage of the trial is shown in Figure 4.

Figure 4. Participant flow in ASCEND

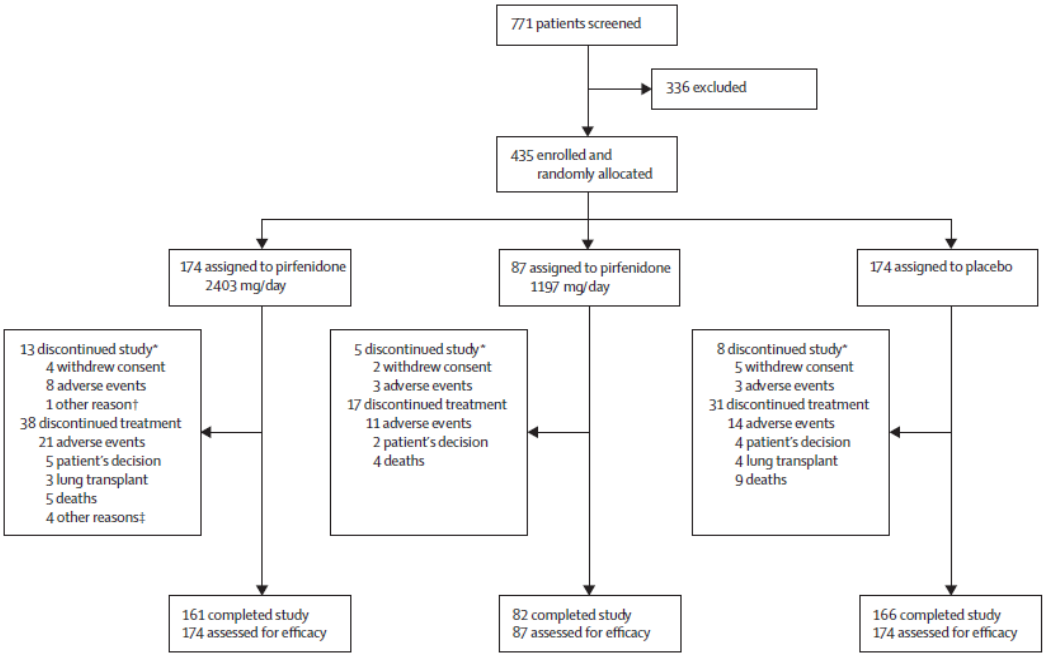


CAPACITY 2 study (Noble 2011)

A total of 771 patients were screened, and 435 patients with IPF were randomly assigned in a 2:1:2 ratio to receive pirfenidone 2403 mg/day, pirfenidone 1197 mg/day or placebo.

The most common reason for discontinuation of treatment was adverse events. The most common reason for discontinuation of study was consent withdrawal. The CONSORT diagram shows the flow of participants through each stage of the trial is shown in Figure 5.

Figure 5. Participant flow in the CAPACITY 2 study



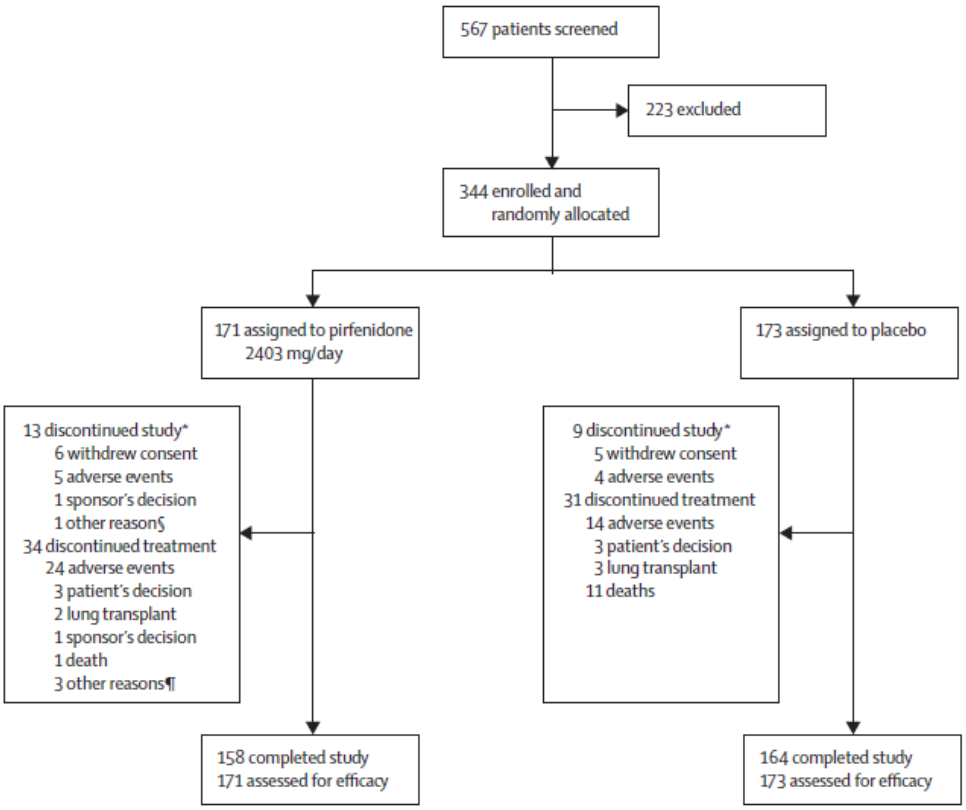
*Does not include death or lung transplantation
†Discontinued study because of deportation
‡Includes unknown interaction with chemotherapy (n=1), deportation (n=1), non-adherence to assigned treatment regimen (n=1), and spontaneous discontinuation of study drug (n=1)

CAPACITY 1 study (Noble 2011)

A total of 567 patients were screened, and 344 patients with IPF were randomly assigned in a 1:1 ratio to receive pirfenidone 2403 mg/day or placebo.

The most common reason for discontinuation of treatment was adverse events. The most common reason for discontinuation of study was consent withdrawal. The CONSORT diagram shows the flow of participants through each stage of the trial is shown in Figure 6.

Figure 6. Participant flow in the CAPACITY 1 study



*Does not include death or lung transplantation
§Discontinued study due to placement on lung transplantation schedule
¶Includes placement on lung transplantation schedule (n=1), prolonged QTc interval that was subsequently ascertained to be present at baseline (n=1), and unknown (n=1)

We would like to highlight that the data extracted from CAPACITY 1 and CAPACITY 2 to support the the meta-analyses and NMAs (Sections 4.9 and 4.10) are different to those presented in the nintedanib NICE manufacturer submission (Boehringer Ingelheim 2015). The data used within our analyses are presented in Table 14, along with the rationale for the difference.

Table 14. Extracted data used to input into the NMA for all-cause discontinuation of treatment

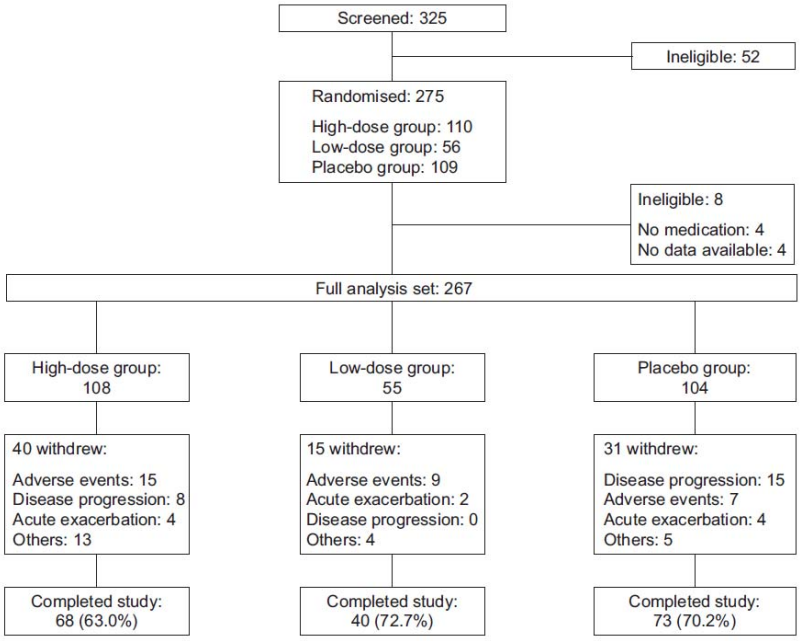
Study (source)	Treatment	Time point	Events	Patients at risk
CAPACITY 1 (Data on file[1])	PFN	48 weeks	18	171
	PBO		18	173
CAPACITY 2 (Data on file[1])	PFN	48 weeks	22	174
	PBO		21	174
Explanation of differences to BI submission: • CAPACITY 1 & 2 were treated as individual studies in our NMA. • Assessments were conducted every 12 weeks in CAPACITY 1 & 2 and therefore 48 weeks was considered most appropriate data cut to use to compare with 52 week data from other trials				
[1] Roche 2016a				

SP3 study (Taniguchi 2010)

A total of 325 patients were screened, and 275 patients with IPF were randomly assigned in a 2:1:2 ratio to receive pirfenidone 1800 mg/day, pirfenidone 1200 mg/day or placebo. Eight patients had no medication, or had no data available and therefore were excluded from the full analysis.

The most common reason for withdrawal in the pirfenidone-treated arms was adverse events. The most common reason for withdrawal in the placebo arm was disease progression. The CONSORT diagram shows the flow of participants through each stage of the trial is shown in Figure 7.

Figure 7. Participant flow in the SP3 study

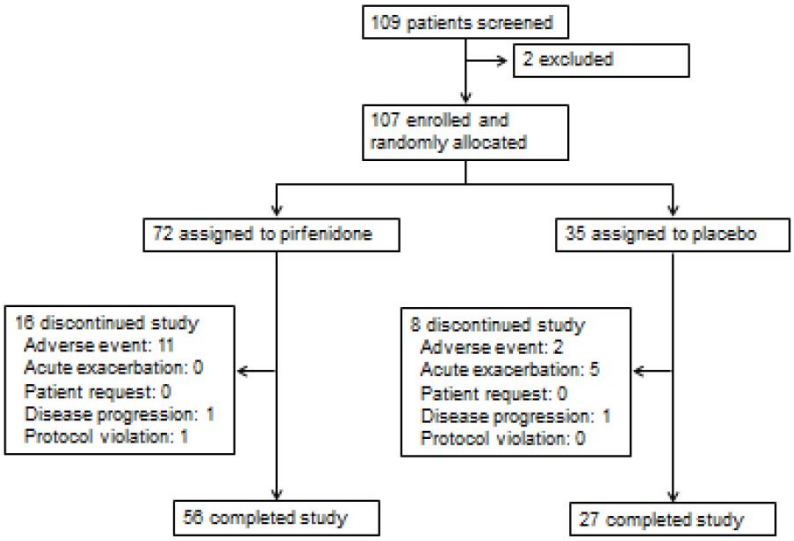


SP2 study (Azuma 2005)

A total of 109 patients were screened, and 107 patients with IPF were randomly assigned in a 2:1 ratio to receive pirfenidone 1800 mg/day or placebo.

The most common reason for withdrawal in the pirfenidone-treated arm was adverse event. The most common reason for withdrawal in the placebo arm was acute exacerbation. The CONSORT diagram shows the flow of participants through each stage of the trial is shown in Figure 8.

Figure 8. Participant flow in the SP2 trial



The key inclusion exclusion criteria across the five RCTs are summarised in the table below. The resulting baseline characteristics are then described and tabulated.

Table 15: Key Inclusion/Exclusion Criteria for RCTs

	ASCEND	CAPACITY2	CAPACITY1	SP3	SP2
Key inclusion criteria **Lung function parameters**	• FVC (% predicted value) 50-90% • DLco 30-90% • 6MWT ≥150 m	• FVC (% predicted value) ≥50% • DLco ≥35% • FVC or DLco ≤90%	• FVC (% predicted value) ≥ 50% • DLco ≥35% • FVC or DLco ≤90% • - 6MWT ≥150 m	• O2 desaturation of 5% between resting SpO2 and min SpO2 during 6 min exercise test (6MET) • SpO2 >85% during 6MET (air).	• Adequate oxygenation at rest (PaO2 70 mm Hg) and SpO2 ≤ 90% during exertion
Key inclusion criteria **IPF Diagnosis**	• Confident clinical and radiographic diagnosis of IPF, confirmed centrally with diagnosis of IPF >6 months but <48 months. • No improvement of IPF in preceding year.	• Confident clinical and radiographic diagnosis of IPF, confirmed locally (diagnosis previous 48 months) • No improvement of IPF in preceding year	• Confident clinical and radiographic diagnosis of IPF, confirmed locally (diagnosis previous 48 months) • No improvement of IPF in preceding year	• Confident clinical and radiographic diagnosis of IPF (as per ATS/ERS guideline concensus) • No decrease in symptoms during the preceding 6 months	• Confident clinical and radiographic diagnosis of IPF (as per guideline concensus)
Key exclusion criteria **Patient factors**	• Abnormal lab parameters • Obstructive airway disease • History of unstable /deteriorating cardiac or pulmonary disease • History of severe hepatic impairment/ end-stage liver disease/end-stage renal disease requiring dialysis	• Abnormal lab parameters • Obstructive airway disease, • History of unstable /deteriorating cardiac or pulmonary disease • History of severe hepatic impairment/ end-stage liver disease/end-stage renal disease requiring dialysis	• Abnormal lab parameters • Obstructive airway disease, • History of unstable /deteriorating cardiac or pulmonary disease • History of severe hepatic impairment/ end-stage liver disease/end-stage renal disease requiring dialysis	• Coexisting pulmonary hypertension, asthma, Tb, sarcoid, bronchiectasis or respiratory infection; • Comorbid conditions including malignancy, severe hepatic, renal, DM or cardiac disease;	• Coexisting pulmonary hypertension, asthma, Tb, sarcoid, bronchiectasis or respiratory infection; • Comorbid conditions including malignancy, severe hepatic, renal, DM or cardiac disease;

Patients who participated in the ASCEND, and CAPACITY 1 & 2 studies were recruited from specialist centres around the world, including 3 in the UK.

The demographic characteristics of the patients were well balanced across the treatment groups within all of the relevant RCTs (see Table 16). Comparison across the studies shows that the age of the patients was similar and overlapping in ASCEND, CAPACITY 1 & 2. In the Japanese studies, SP2 and SP3, the age of patients were also similar across the treatment groups, although slightly younger than those in the international studies.

The proportion of male patients was larger in ASCEND (78.4%) compared with CAPACITY 2 (70.7%) and CAPACITY 1 (71.8%) and there was a greater representation of non-white patients (8.8% vs. 3.4% and 1.2% respectively). In terms of these demographic findings, while there is a clear overlap of the study populations, given the modest differences with respect to age, sex and race, the data from the three studies can be regarded as inclusive of the overall range of patients likely to present with mild to moderate IPF in routine clinical practice. The Japanese studies, SP2 and SP3, reported the highest percentage of males (89% and 79%, respectively).

Percentage predicted FVC was reported at baseline by the CAPACITY trials and the ASCEND trial. For ASCEND, patients at higher risk of disease progression were enroled (percent predicted FVC 50-90% and percent predicted DLco 30-90% at Screening). For the CAPACITY studies, patients with percent predicted FVC ≥50% at Screening and Day 1 (before randomisation) and FVC or Hgb-corrected DLco ≤90% of predicted value at the Screening were included. There were no significant differences between treatment arms within the trials, however the ASCEND trial reported a percentage predicted FVC approximately 7-8% lower than the CAPACITY trials.

The SP2 and SP3 trials reported percentage predicted total lung capacity and vital capacity. There were no differences between the study arms within trials and no differences between the two trials.

Table 16. Characteristics of participants in the studies across treatment groups

Baseline characteristic	PIPF-016 (ASCEND) (King 2014; InterMune 2014)		PIPF-004 (CAPACITY 2) (Noble 2011)			PIPF-006 (CAPACITY 1) (Noble 2011)	
	Pirfenidone 2403 mg/day (n=278)	Placebo (n=277)	Pirfenidone 2403 mg/day (n=174)	Pirfenidone 1197mg/day (n=87)	Placebo (n=174)	Pirfenidone 2403 mg/day (n=171)	Placebo (n=173)
Age, mean years ± SD	68.4 ± 6.7	67.8 ± 7.3	65.7 ± 8.2	68.0 ± 7.6	66.3 ± 7.5	66.8 ± 7.9	67.0 ± 7.8
Male, n (%)	222 (79.9)	213 (76.9)	118 (68)	65 (75)	128 (74)	123 (72)	124 (72)
Percentage of predicted FVC, mean % ± SD	67.8 ± 11.2	68.6 ± 10.9	74.5 ± 14.5	76.4 ± 14.4	76.2 ± 15.5	74.9 ± 13.2	73.1 ± 14.2
Percentage of predicted DLco, mean % ± SD	43.7 ± 10.5	44.2 ± 12.5	46.4 ± 9.5	47.2 ± 8.2	46.1 ± 10.2	47.8 ± 9.8	47.4 ± 9.2
Dyspnoea score, mean ± SD	34.0 ± 21.9	36.6 ± 21.7	NR	NR	NR	NR	NR
Mean 6MWD, m ± SD	415.0 ± 98.5	420.7 ± 98.1	411.1 ± 91.8	417.5 ± 112.8	410.0 ± 90.0	378.0 ± 82.2	399.1 ± 89.7
Supplemental O_2 use, n (%)	78 (28.1)	76 (27.4)	29 (16.7)	15 (17)	25 (14)	48 (28)	49 (28)
HRCT definite IPF, n (%)	266 (95.7)	262 (94.6)	159 (91)	83 (95)	164 (94)	149 (87)	158 (91)
Surgical lung biopsy, n (%)	86 (30.9)	79 (28.5)	86 (49)	32 (37)	85 (49)	94 (55)	94 (54)
Time since IPF diagnosis, years ± SD	1.7 ± 1.1	1.7 ± 1.1	1.3 ± 0.96	1.4 ± 1.16	1.4 ± 1.12	1.2 ± 1.09	1.1 ± 1.04
Former smoker, n (%)	184 (66.2)	169 (61.0)	110 (63)	57 (66)	114 (66)	112 (66)	101 (58)
Pre-enrolment corticosteroid use, n (%)	6 (2.2)	2 (0.7)	14 (8.0)	10 (11.5)	9 (5.2)	22 (12.9)	17 (10.0)
Concomitant corticosteroid use, n (%)	82 (29.5)	101 (36.5)	38 (21.8)	24 (27.6)	52 (29.9)	42 (24.6)	50 (29.0)

Baseline characteristic	SP3 (Taniguchi 2010)			SP2 (Azuma 2005)	
	Pirfenidone 1800 mg/day (n=108)	Pirfenidone 1200 mg/day (n=55)	Placebo (n=104)	Pirfenidone 1800 mg/day (n=72)	Placebo (n=35)
Age, mean years ± SD	65.4 ± 6.2	63.9 ± 7.5	64.7 ± 7.3	64.0 ± 7.1	64.3 ± 7.6
Male, n (%)	85 (78.7)	47 (85.5)	81 (77.9)	62 (86.0)	33 (94.0)
Percentage of predicted VC, mean % ± SD	77.3 ± 16.8	76.2 ± 18.7	79.1 ± 17.4	81.6 ± 20.3	78.4 ± 17.2
Percentage of predicted TLC, mean % ± SD	73.2 ± 16.5	72.4 ± 15.6	75.2 ± 15.7	78.5 ± 17.9	73.9 ± 16.4
Percentage of predicted DLco, mean % ± SD	52.1 ± 16.8	53.6 ± 19.1	55.2 ± 18.2	57.6 ± 17.2	57.7 ± 13.8
Lowest SpO_2 during 6MWT, mean % ± SD	89.0 ± 2.3	88.8 ± 2.4	89.0 ± 2.0	87.1 ± 3.9	87.1 ± 4.2
Desaturation <88% during 6MWT, n (%)	34 (31.5)	19 (34.5)	24 (23.1)	NR	NR
Mean $P(A\text{-}a)O_2$ ± SD	18.4 ± 11.3	16.9 ± 9.6	17.4 ± 9.7	NR	NR
Percentage of predicted SpO_2, mean % ± SD	89.0 ± 2.3	88.8 ± 2.4	89.0 ± 2.0	NR	NR
Mean PaO_2 at rest, mmHg ± SD	79.8 ± 10.2	81.6 ± 8.4	81.0 ± 9.5	80.3 ± 7.7	82.0 ± 17.6
Mean VC, mL ± SD	2400.8 ± 638.4	2437 ± 684.8	2472.3 ± 698.9	NR	NR
Surgical lung biopsy, n (%)	26 (24.1)	16 (29.1)	28 (26.9)	15 (21.0)	8 (23.0)
IPF diagnosis, n (%) ≤1 year 1-3 years >3 years	 38 (35.2) 29 (26.9) 41 (38.0)	 20 (36.4) 13 (23.6) 22 (40.0)	 41 (39.4) 25 (24.0) 38 (36.5)	 20 (28.0) 17 (24.0) 35 (49.0)	 6 (17.0) 10 (29.0) 19 (54.0)
Former smoker, n (%)	81 (75.0)	33 (60.0)	70 (67.3)	57 (79.0)	30 (86.0)
Pre-enrolment corticosteroid use, n (%)	9 (8.3)	6 (10.9)	6 (5.8)	10 (14.0)	5 (14.0)
Concomitant corticosteroid use, n (%)	8 (7.4)	6 (10.9)	5 (4.8)	NR	NR

4.6 Quality assessment of the relevant randomised controlled trials

Four of the five trials (ASCEND, CAPACITY 1 & 2, and SP2) used a computer generated randomisation list (King 2014; Noble 2011; Azuma 2005). In SP3 (Taniguchi 2010), patients were allocated to treatment groups using a modified minimisation method, including some random allocation based on biased coin design to balance baseline SpO_2. In SP3, the original primary outcome was altered to VC after the study had started and it is not clear whether VC outcomes would have been affected considering the attempt to balance groups for SpO_2 at randomisation.

All five trials used appropriate allocation concealment methods. In the ASCEND, CAPACITY 1 & 2 studies (King 2014; Noble 2011) used centralised interactive voice response systems to conceal allocation. The ASCEND trial also supplied pirfenidone and placebo in capsules and packaging that were visually indistinguishable. The SP2 trial employed a third party to assign the study drug; no further details were reported (Azuma 2005). The SP3 trial reported using identical packaging of pirfenidone and placebo (Taniguchi 2010).

Four of the five trials (ASCEND, CAPACITY 1 & 2, SP2) explicitly reported that all personnel involved in the study were masked to treatment group assignment (King 2014; Noble 2011; Azuma 2005). The SP3 trial was reported to be double-blinded, but it is not clear who was blinded. Our access to source data on this trial is limited by the historical ownership of data, as decribed in Section 2 (Taniguchi 2010).

There were no unexpected imbalances in drop-outs between treatment and control groups. The CAPACITY 1 and 2 trials reported that compared with placebo patients, a higher proportion of patients in the pirfenidone 2403 mg/day group permanently discontinued treatment because of adverse events: 14.0% versus 8.1% (CAPACITY 1) and 12.1% versus 8.0% (CAPACITY 2). However, a lower proportion of patients in the pirfenidone 2403 mg/day group died while on study treatment: 0.6% versus 6.4% (CAPACITY 1) and 2.9% versus 5.2% (CAPACITY 2). None of these differences were statistically significant. No differences were reported in the rates of drop outs for the other trials (Noble 2011).

There is no evidence to suggest that authors of any of the five trials measured more outcomes than they reported. The results of all outcomes measured were described in the study publications.

The CAPACITY and ASCEND trials used intention to treat analysis (ITT). The SP2 trial reported results for both ITT and per protocol analyses. It was unclear whether SP3 used ITT principles; for some outcomes it appears that ITT was used and for other outcomes, the full number of randomised participants was not used. Two analysis sets were pre-specified; the full analysis set (FAS) and the per-protocol set (PPS). The full analysis set excluded from the efficacy analysis those patients who were deemed ineligible; four patients who did not take any of the study medication and four with no available data. The per protocol set excluded patients who were excluded from the FAS and also excluded patients who did not meet all inclusion and exclusion criteria, did not take study drug, did not comply with the study drug titration schedule or otherwise were noncompliant with the treatment regimen.

Overall, all five trials were considered to have a low risk of bias.

Table 17. Quality assessment results for parallel group RCTs

	ASCEND	CAPACITY 1 & CAPACITY 2	SP3*	SP2
Was randomisation carried out appropriately?	Yes	Yes	Unclear	Yes
Was the concealment of the treatment allocation adequate?	Yes	Yes	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes	Yes
Were the care providers, participants, and outcome assessors blind to treatment allocation?	Yes	Yes	Unclear	Yes
Were there any unexpected imbalances in drop-outs between groups?	No	No	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	No	No
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes	Yes	Yes
Risk of bias of the study	Low risk	Low risk	Low risk	Low risk
* During the course of this multi-year trial, the academic community's views on appropriate primary endpoints in IPF evolved, and the data safety and monitoring board recommended a change of the primary endpoint to VC after a discussion of blinded interim comparative data.				

4.6.4 The complete quality assessment for each RCT should be included in an appendix

A complete quality assessment for each RCT is included in Appendix 6.

4.7 Clinical effectiveness results of the relevant randomised controlled trials

- Overall, the use of pirfenidone in mild to moderate IPF patients slows the rate of decline in FVC and improves survival. Clinical evidence shows pirfenidone also reduces the decline in 6MWD, it is therefore a valuable treatment option for patients with IPF
- In the pre-specified pooled analysis of ASCEND, CAPACITY 1 & 2 at 52 weeks, patients treated with pirfenidone had an approximate (and statistically significant) 50% reduction in the risk of death compared to patients receiving placebo (HR=0.52; 95% CI: 0.31-0.87; p=0.011)
- Treatment-emergent IPF-related mortality was significantly lower in the pirfenidone group than the placebo group in the pooled analysis of CAPACITY 1 & 2 (HR: 0.48 [95% CI: 0.24, 0.95] p=0.03).
- A pooled analysis of ASCEND, CAPACITY 1 & 2 studies at 52 weeks showed patients treated with pirfenidone had significantly higher rates of progression free survival (HR: 0.62; 95% CI: 0.51–0.76; p<0.0001)
- The benefit of treatment with pirfenidone on percent predicted FVC was apparent from the three pivotal Phase III trials. In the ASCEND study, pirfenidone significantly reduced the change in percent predicted FVC or death compared to placebo at Week 52 (p<0.001). A similar effect was seen in the absolute change in percent predicted FVC in CAPACITY 2 (p=0.001), but did not reach statistical significant in CAPACITY 1 (p=0.501)
- Across the studies, there were a low number of exacerbations, and inconsistent definitions were used to define events. As such, results for this outcome should be interpreted with caution. ASCEND demonstrated a clear numerical difference in numbers of exacerbations between the study arms at 52 weeks, with 40 events being reported in the placebo arm and only 24 in those patients treated with pirfenidone.
- Results of two earlier Japanese studies (SP2 and SP3) provide supportive evidence on the effectiveness of pirfenidone in IPF, and are consistent with

those of the large phase III trials (ASCEND, CAPACITY 1 & 2).

- When considering patients with earlier (FVC ≥ 80% predicted) vs. later (FVC <80% predicted) disease, the pooled analysis demonstrated significantly improved outcomes in 6MWD and quality of life (UCSD SOBQ) for those with less severe disease at 12 months. There was also a numerically lower risk of FVC decline ≥10% or death in those with FVC ≥ 80% predicted, although this was not statistically significant.(p= 0.2403)
- In the pooled analysis of ASCEND and CAPACITY 1 & 2 a consistent treatment effect was seen across all demographic subgroups

This section presents the results of each clinical trial, grouped by outcome of interest. A tabulated summary of each trial can be found in Appendix 7.

Extracted data used in the meta-analyses and NMAs (Sections 4.9 and 4.10) are summarised at the end of each outcome category. Any differences from the values reported in the nintedanib NICE manufacturer submission are highlighted within the relevant tables, along with a rationale for the difference (Boehringer Ingelheim 2015).

Lung Function

FVC categorical decline of ≥10% percent predicted

A decline in percentage predicted FVC of ≥10% is a decrement that has long been recognised as both clinically significant and highly predictive of mortality (Collard 2003; Flaherty 2003; Latsi 2003; Zappala 2010).

The ASCEND study met its primary endpoint: at Week 52, there was a relative reduction of 47.9% in the proportion of patients who experienced a decline in FVC by ≥10% or death (p=0.000001) (King 2014). CAPACITY 2 showed a statistically significant difference for this outcome at Week 72 (absolute difference: 14.4 [95% CI: 7.4, 21.3] p=0.001), but not in CAPACITY 1 (absolute difference: 3.8 [95% CI: -2.7, 10.2] p=0.440). However, in CAPACITY 1, a statistically significant treatment effect

was evident at every time point to Week 48 (p=0.007) (Noble 2011). This is discussed further in Section 4.13.

In all three trials, the benefit of treatment with pirfenidone 2403 mg/day was apparent early and was persistent, with a progressively increasing difference in decline in percent predicted FVC in favour of pirfenidone 2403 mg/day over the study period.

Table 18. Categorical analysis of change from baseline in percent predicted FVC or death for the relevant RCTs

Study (Ref)	Time point	Treatment group	Decline ≥10% FVC or death, n (%)	No decline* in FVC, n (%)	p-value†
ASCEND (King 2014)	52 weeks	PFN 2403 mg/day (N=278)	46 (16.5)	63 (22.7)	p=0.000001
		PBO (N=277)	88 (31.8)	27 (9.7)	
CAPACITY 2 (Noble 2011;)	72 weeks	PFN 2403 mg/day (N=174)	35 (20.1)	42 (24.1)	p=0.001
		PBO (N=174)	60 (34.5)	24 (13.8)	
CAPACITY 1 (Noble 2011;)	72 weeks	PFN 2403 mg/day (N=171)	39 (22.8)	44 (25.8)	p=0.440
		PBO (N=173)	46 (26.6)	38 (22.0)	
Pooled CAPACITY 1 & 2 (Noble 2011;)	72 weeks	PFN 2403 mg/day (N=345)	74 (21)	86 (24.9)	p=0.003
		PBO (N=347)	106 (31)	62 (17.9)	

PFN- pirfenidone; PBO- placebo
*Change in predicted FVC ≥10%
†Rank ANCOVA (pirfenidone 2403 mg/day vs placebo)

A pre-specified pooled analysis of ASCEND, CAPACITY 1 & 2 at Week 52 showed treatment with pirfenidone reduced the proportion of patients with a ≥10% decline in percent predicted FVC or death by 43.8%, and increased the proportion of patients with no decline in percent predicted FVC by 59.3%, compared with placebo (p<0.0001) (Noble 2014a). Please see section 4.4 for a discussion on the pooling of these trials.

Data used in the meta-analyses and NMAs (Sections 4.9 and 4.10) are provided in Table 19.

Table 19. Extracted data used to input into NMA for categorical decline ≥10% FVC

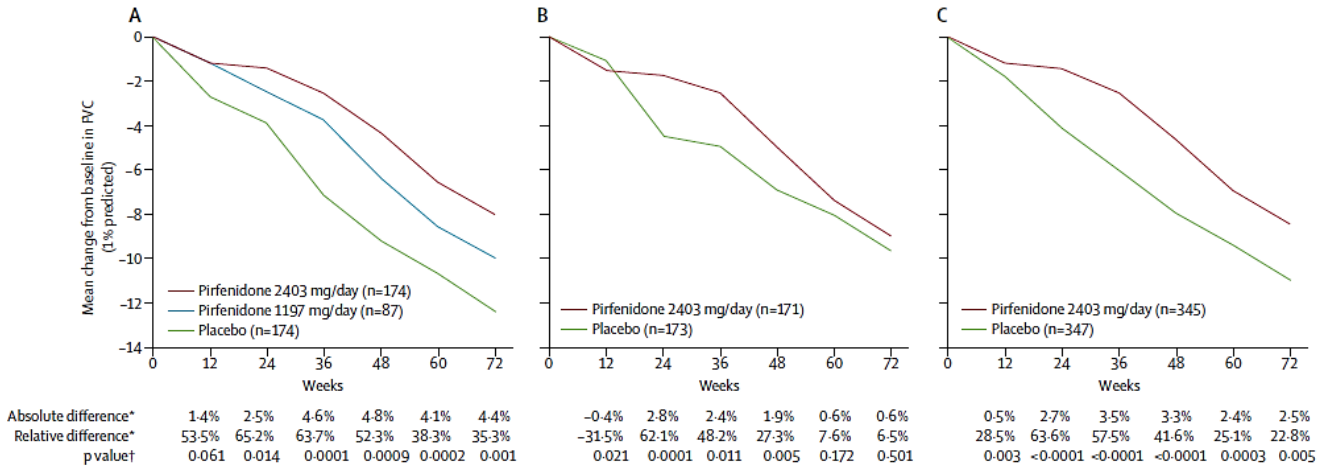
Study (source)	Treatment	Time point	Events	Patients at risk
CAPACITY 2 (Data on file[1])	PFN 2403 mg/day	48 weeks	28	174
	PBO		49	174
CAPACITY 1 (Data on file[1])	PFN 2403 mg/day	48 weeks	34	171
	PBO		38	173
ASCEND (Data on file[2])	PFN 2403 mg/day	52 weeks	62	278
	PBO		100	277
Explanation for differences to the BI submission: • CAPACITY 1 & 2 assessments were performed every 12 weeks, and therefore 48 week data was taken as it is the closest data to 52 weeks for the NMA • ASCEND data was re-calculated as per the INPULSIS definition (where deaths are counted as non-responders). Patients with no FVC values at 52 weeks were assumed to be non-responders.				
[1] Roche 2016a Patients with no FVC values at 52 weeks were assumed to be non-responders. [2] Roche 2016a				

Change in percent predicted FVC/VC

The primary efficacy outcome in the CAPACITY 1 & 2 studies was the change in percent predicted FVC from baseline to Week 72 (CAPACITY 1 & 2) (Noble 2011). In SP3, the primary efficacy outcome variable was the change from baseline in VC in the pirfenidone 1800 mg/day group compared with the placebo group at 52 weeks (Taniguchi 2010).

CAPACITY 2 reached its primary endpoint; pirfenidone 2403 mg/day resulted in a significant absolute difference in change in percent predicted FVC compared with placebo at Week 72 (absolute difference 4.4%; relative difference 35.3%;CI 0·7 to 9·1 p=0.001). Outcomes in the pirfenidone 1197 mg/day group were intermediate to the pirfenidone 2403 mg/day and placebo groups. At Week 72, the absolute difference in percent predicted FVC did not reach statistical significance in CAPACITY 1 (absolute difference: 0.6%; relative difference: 6.5%;CI -3·5 to 4·7 p=0.501), see Figure 9 (Noble 2011). However, a significant treatment effect of pirfenidone 2403 mg/day was evident at every time point from Week 12 until Week 48, and in the repeated measures analysis of percentage predicted FVC change over all assessment time points (p=0.007) (Noble 2011).

Figure 9. Mean change from baseline in % predicted FVC in the CAPACITY 2 (A), CAPACITY 1 (B), and in the pooled population (C)



Absolute difference*	1·4%	2·5%	4·6%	4·8%	4·1%	4·4%
Relative difference*	53·5%	65·2%	63·7%	52·3%	38·3%	35·3%
p value†	0·061	0·014	0·0001	0·0009	0·0002	0·001
Absolute difference*	−0·4%	2·8%	2·4%	1·9%	0·6%	0·6%
Relative difference*	−31·5%	62·1%	48·2%	27·3%	7·6%	6·5%
p value†	0·021	0·0001	0·011	0·005	0·172	0·501
Absolute difference*	0·5%	2·7%	3·5%	3·3%	2·4%	2·5%
Relative difference*	28·5%	63·6%	57·5%	41·6%	25·1%	22·8%
p value†	0·003	<0·0001	<0·0001	<0·0001	0·0003	0·005

*Pirfenidone 2403 mg/day versus placebo
†Rank ANCOVA (pirfenidone 2403 mg/day vs placebo). 95% CIs were only calculated for absolute differences for the Week 72 time point in CAPACITY 2 (95% CI: 0.7-9.1) and CAPACITY 1 (95% CI: -3.5-4.7)

In a pooled analysis of the CAPACITY 1 & 2 trials, there was a significant treatment effect on percent predicted FVC with a mean change of -8.5% in the pirfenidone 2403 mg/day arm and -11% in the placebo arm (absolute difference: 2.5%; relative difference: 22.8%; p=0.005, rank ANCOVA) at Week 72 (Noble 2011).

In SP3, an analysis of the mean decline from baseline in percent predicted VC at Week 52 showed a significant treatment effect of pirfenidone 1800 mg/day compared with placebo, respectively: -2.91% ± 0.77 compared with -5.13% ± 0.78 (p=0.044, ANCOVA) (FDA, 2010).

The ASCEND manuscript did not report the change in % predicted FVC, but this was analysed to inform the NMA. These data were consistent with the categorical findings, and showed a statisticaly significant mean difference of 4.78% less decline at 52 weeks with pirfenidone compared to placebo. Data used in the meta-analyses and NMAs (Sections 4.9 and 4.10) are provided in Table 20.

Table 20. Extracted data used to input into the NMA for change in baseline in % predicted FVC

Study (source)	Treatment	Time point	Mean change from baseline	SE	Mean difference from PBO	p-value
ASCEND* (Data on file[1])	PFN 2403 mg/day (n=278)	52 weeks	-6.17	0.875	4.781	<0.001
	PBO (n=277)		-10.95	0.877		
* The ASCEND manuscript did not report the change in % predicted FVC but this was analysed to inform the NMA.						
[1] Roche 2016a						

Change in FVC/VC (ml)

Change in VC or FVC were available for all five RCTs. CAPACITY 1, CAPACITY 2 and ASCEND measured FVC, while SP2 and SP3 measured VC. The decision of whether to use VC or FVC as the end point in the trials was dictated by the guidelines at the time.

For SP2 and SP3, the ATS international consensus statement published in 2000 recommended measurement of VC (ATS ERS, 2000). Given that there is little difference between VC and FVC in subjects without obstructive pathology (Chhabra, 1998), and IPF patients have a restrictive pathology, it is appropriate that VC and FVC are treated as comparable endpoints.

All trials except CAPACITY 1 showed a statistically significant difference in change in FVC/VC with pirfenidone treatment compared with placebo.

Results across all trials are presented in Table 21.

The CAPACITY trials assessed this outcome at 48 weeks, with these data being used in the NMA to allow comparison of studies across a similar time point (see Section 4.10). Data extracted to inform the meta-analysis and NMA re reported in Table 22.

Table 21: Mean change from baseline in FVC/VC of the relevant RCTs

Study (Ref)	Time point	Treatment group	Mean decline in FVC/VC	Difference, p-value†
ASCEND (King 2014)	52 weeks	PFN 2403 mg/day (N=278)	FVC: 235 ml	Absolute difference: 193 ml Relative difference: 45.1% p<0.001
		PBO (N=277)	FVC: 428 ml	
CAPACITY 2	72 weeks	PFN 2403 mg/day (N=174)	FVC: 318 ml	Absolute difference:157 ml Relative difference: 33% p-value=0.004
		PBO (N=174)	FVC: 475 ml	
CAPACITY 1	72 weeks	PFN 2403 mg/day (N=171)	FVC: 379 ml	Absolute difference: -5 ml Relative difference: -1.4% p-value=0.508
		PBO (N=173)	FVC: 373 ml	
SP3 (Taniguchi 2010)	52 weeks	PFN 1800 mg/day (N=108)	VC: 90 ml	PFN 1800 mg/day vs. PBO: Absolute difference: 70 ml Relative difference: NR p=0.042
		PFN 1200 mg/day (N=55)	VC: 80 ml	
		PBO (N=104)	VC: 160 ml	
SP2 (Azuma 2005)	9 months	PFN 1800 mg/day (N=72)	VC: 30 ml	Absolute difference: 100 ml Relative difference: NR p=0.037
		PBO (N=35)	VC: 130 ml	

PFN- pirfenidone; PBO- placebo; NR- not reported
†Rank ANCOVA: ASCEND, CAPACITY 1 & 2 (pirfenidone 2403 mg/day vs placebo); SP2 and SP3 (pirfenidone 1800 mg/day vs. placebo)

Table 22. Extracted data used to input into NMA for change in FVC in L

Trial	Intervention	Time point	Change from baseline	
			Mean change	SD
CAPACITY 2 (Data on file[1])	PFN 2403 mg/day (n=174)	48 weeks	-0.1808	0.5276
	PBO (n=174)		-0.3498	0.67495
CAPACITY 1 (Data on file[2])	PFN 2403 mg/day (n=171)	48 weeks	-0.2196	0.72228
	PBO (n=173)		-0.2739	0.64423

[1] Roche 2016a
[2] Roche 2016a

Mortality

All-cause and treatment-emergent (TE) IPF-related mortality

CAPACITY 1 & 2 and ASCEND studies reported hazard ratios and proportion of deaths while SP2 and SP3 only reported the proportion of deaths for all-cause mortality and TE IPF-related mortality. TE IPF-related mortality was defined as deaths occurring between randomisation and within 28 days of last dose of study drug.

In the three large Phase III studies (ASCEND and CAPACITY 1 & 2), there were relatively few deaths (<10%) in either treatment group at the end of the study. Details of the all-cause mortality and TE IPF-related mortality at the common time point of 52 weeks are presented in Table 23.

None of the studies were powered to assess the effect of pirfenidone on mortality, and a pre-specified pooled analysis of the three trials was requested by the FDA to increase statistical power (King 2014a).

Table 23. Mortality rates in the ASCEND, CAPACITY 1 & 2 studies and pooled populations at Week 52

Patients, n (%)	Pirfenidone 2403 mg/day	Placebo	HR (95% CI)*	p-value**
ASCEND	**n=278**	**n=277**		
All-cause mortality	11 (4.0)	20 (7.2)	0.55 (0.26, 1.15)	0.105
TE IPF-related mortality	3 (1.1)	7 (2.5)	0.44 (0.11, 1.72)	0.226
CAPACITY 1 & 2 studies†	**n=345**	**n=347**		
All-cause mortality	11 (3.2)	22 (6.3)	0.49 (0.24-1.01)	0.047
TE IPF-related mortality	4 (1.2)	15 (4.3)	0.27 (0.09–0.81)	0.012
Pooled data for ASCEND, CAPACITY 1 & 2	**n=623**	**n=624**		
All-cause mortality	22 (3.5)	42 (6.7)	0.52 (0.31–0.87)	0.011
TE IPF-related mortality	7 (1.1)	22 (3.5)	0.32 (0.14–0.76)	0.006

†Data in the CAPACITY 1 & 2 studies were censored at one year
*Cox proportional hazards model
**Log-rank test
Abbreviations: TE- treatment-emergent; TE IPF-related mortality defined as deaths occurring between randomisation and within 28 days of last dose of study drug.

In the pooled analysis of CAPACITY 1 & 2 at 72 weeks, there were fewer overall deaths and TE IPF-related deaths in the pirfenidone groups compared to the placebo groups. Overall, there was a 23% reduction in all-cause mortality vs placebo among patients treated with pirfenidone 2403 mg/day (HR=0.77; 95% CI: 0.47-1.28; p=0.315) at Week 72. For TE IPF-related mortality, the HR also favoured pirfenidone at Week 72 (HR=0.48; 95%CI: 0.24-0.95; p=0.03) (Noble 2011).

In SP3, eleven patients died during the study in the pirfenidone 1800 mg/day (n=3), pirfenidone 1200 mg/day (n=4), and placebo arms (n=4) (Taniguchi 2010). In SP2, there were no patient deaths reported in the pirfenidone 1800 mg/day treatment arm during the study period. There was one death in the placebo arm (Azuma 2005).

Data informing the meta-analysis and NMA at the pre-sepcified common time point of 52 weeks are presented in Table 24 (see Section 4.4 and Figure 3). Data are also extracted for 72 weeks to inform a sensitivity analysis descried in Section 4.10.

Table 24. Extracted data used to input into the NMA for all-cause mortality

Study (Source)	Treatment	Time point	Events	Patients at risk
CAPACITY 1 (Data on file[1])	PFN	52 weeks	6	171
	PBO		9	173
CAPACITY 2 (Data on file[2])	PFN	52 weeks	5	174
	PBO		13	174
SP3 (Data on file[3])	PFN	52 weeks	3	110
	PBO		6	109
CAPACITY 1 (Data on file[4])	PFN	72 weeks	13	171
	PBO		15	173
CAPACITY 2 (Data on file[5])	PFN	72 weeks	8	174
	PBO		15	174

Explanation for differences to the BI submission:

- We handled CAPACITY 1 & 2 as individual studies.
- For the NMA at 52 weeks, we used internal CAPACITY 1 & 2 data at 52 weeks. BI used data from Noble 2011 which reports all-cause mortality at 72 weeks.
- For the NMA at 72 weeks, we used internal CAPACITY 1 & 2 data with mortality censored at 72 weeks. The results published in Noble 2011 included some deaths beyond 72 weeks.
- SP3 overall mortality data, censored at 52 weeks was available internally. BI used data from Taniguchi 2010 which only presents treatment-emergent mortality.

[1] Roche 2016a
[2] Roche 2016a
[3] Roche 2016a
[4] Roche 2016a
[5] Roche 2016a

IPF-related mortality data are consistent with the all-cause mortality data. There were fewer deaths due to IPF or any cause in the pirfenidone-treated arms compared with placebo. Data informing the meta analysis and NMA are presented in Table 25.

Table 25. Extracted data used to input into the NMA for IPF-related mortality

Study (Source)	Treatment	Time point	Events	Patients at risk
CAPACITY 1 (Data on file[1])	Pirfenidone	52 weeks	4	171
	Placebo		7	173
CAPACITY 2 (Data on file[2])	Pirfenidone	52 weeks	2	174
	Placebo		10	174
CAPACITY 1 (Data on file[3])	Pirfenidone	72 weeks	9	171
	Placebo		13	173
CAPACITY 2 (Data on file[4])	Pirfenidone	72 weeks	4	174
	Placebo		11	174
ASCEND (Data on file[5])	Pirfenidone	52 weeks	4	278
	Placebo		11	277

Explanation for differences to the BI submission:

- We handled CAPACITY 1 & 2 as individual studies.
- Treatment-emergent IPF-related mortality were reported the primary manuscripts for ASCEND and CAPACITY 1 & 2 (measured randomisation to 28 days after the last dose of the study drug). For our NMA, data was extracted from internal data and measured IPF-related mortality from randomisation to 52 weeks and 72 weeks, regardless of adherence

[1] Roche 2016a
[2] Roche 2016a
[3] Roche 2016a
[4] Roche 2016a
[5] Roche 2016a

Progression-Free Survival

Since there is increasing evidence indicating that the fundamental hallmarks of cancer biology are comparable to those of IPF, progression-free survival (PFS) – which is usually employed in oncology studies – could be an appropriate endpoint for IPF studies, using predictive endpoints such as categorical changes in FVC and distance walked in 6MWT (Albera 2011).

Four trials reported data for PFS. The definitions of PFS varied across the trials:

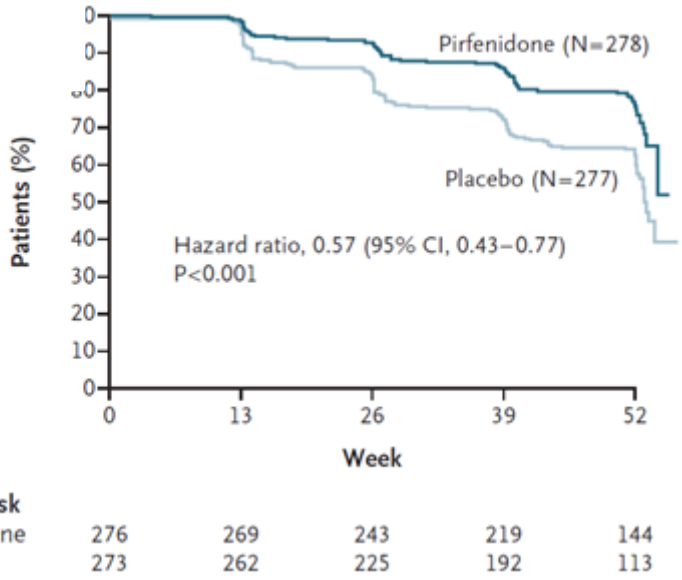
- The ASCEND trial defined PFS as a confirmed ≥10% decline from baseline in %FVC, confirmed ≥50 m decline from baseline in 6MWD, or death;
- The CAPACITY 1 & 2 trials defined PFS as confirmed ≥10% decline in % predicted FVC, ≥15% decline in % predicted DLco or death. In a post-hoc

analysis, the ASCEND definition of PFS was applied to the CAPACITY trials at Week 52 and Week 72;

- The SP3 trial defined PFS as VC decline of ≥10% or death.

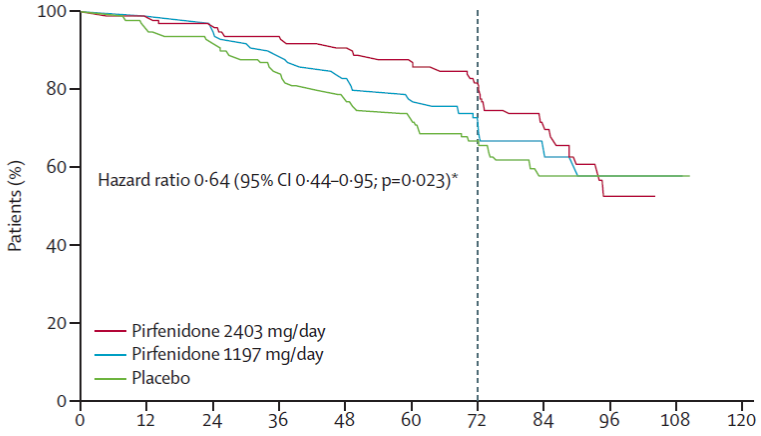
In ASCEND, treatment with pirfenidone reduced the risk of death or disease progression by 43% compared with placebo (HR 0.57; 95% CI, 0.43–0.77 p=0.0001, log-rank test), see Figure 10 (King 2014). For each component of the composite endpoint, fewer patients in the pirfenidone group than in the placebo group had a qualifying event, including death (3.6% vs. 5.1%), a confirmed absolute decrease of ≥10% predicted FVC (6.5% vs. 17.7%), and a confirmed decrease of 50 m or more in the 6-minute walk distance (16.5% vs. 19.5%) (King 2014).

Figure 10. Kaplan–Meier estimates for PFS in ASCEND (all randomised patients)



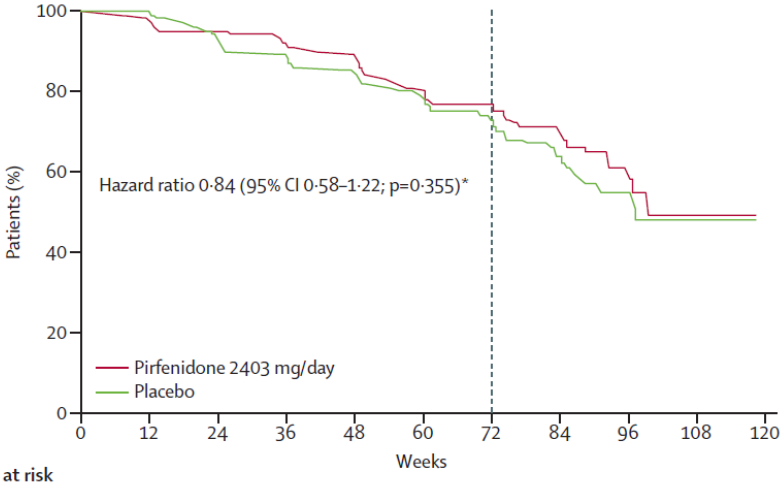
In CAPACITY 2, pirfenidone 2403 mg/day significantly prolonged PFS with a 36% reduction in the risk of death or disease progression (HR:0.64 [95% CI: 0.44, 0.95] p=0.023), see Figure 11 (Noble 2011). However, no significant difference was noted in CAPACITY 1 in PFS for the pirfenidone 2403 mg/day group compared with placebo (HR: 0.84 [95% CI: 0.58, 1.22] p=0.355), see Figure 12 (Noble 2011). In the pooled population from CAPACITY 1 & 2, pirfenidone 2403 mg/day significantly prolonged the PFS by 26% (HR=0.74; 95% CI: 0.57 to 0.96; p=0.025), see Figure 13 (Noble 2011).

Figure 11. Kaplan-Meier estimates for PFS in CAPACITY 2 (all randomised patients)



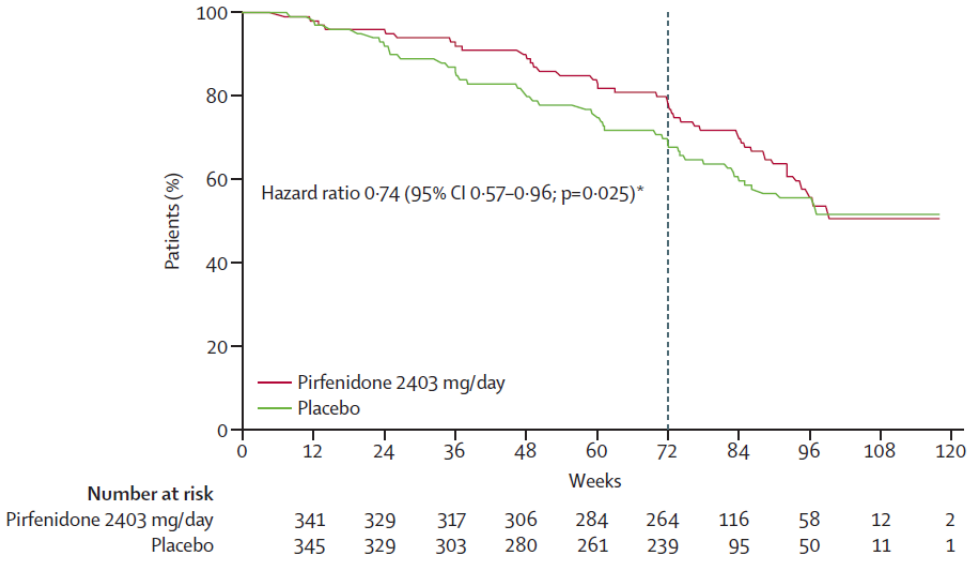
Number at risk									
Pirfenidone 2403 mg/day	171	167	160	157	148	138	55	23	5
Pirfenidone 1197 mg/day	87	86	79	74	68	64	27	11	5
Placebo	173	162	150	136	126	116	44	21	4

Figure 12. Kaplan-Meier estimates for PFS in CAPACITY 1 (all randomised patients)



Number at risk										
Pirfenidone 2403 mg/day	170	162	157	149	136	126	61	35	7	2
Placebo	172	167	153	144	135	123	51	29	7	1

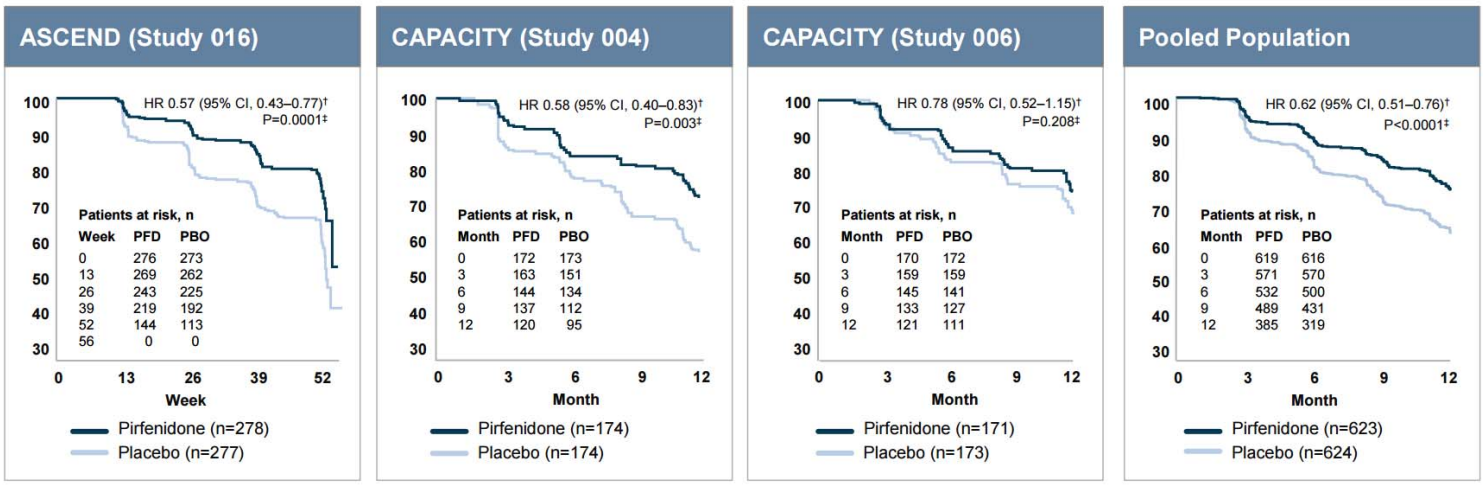
Figure 13. Kaplan-Meier estimates for PFS in the CAPACITY 1 & 2 pooled population (all randomised patients)



As described above, an exploratory analysis of PFS in the pooled CAPACITY 1 & 2 population was conducted using a definition for disease progression which was in line with that used in ASCEND: time to the first occurrence of death, confirmed ≥10% decline in percent predicted FVC, or confirmed ≥50 m decrement in 6MWD (Sahn 2011). All patients who were randomised to receive pirfenidone or placebo were included in this pooled exploratory analysis of PFS. Incorporation of the 6MWD criterion for disease progression resulted in a 29% relative reduction in the risk of death or disease progression in patients receiving pirfenidone compared with placebo (HR: 0.71 [95% CI: 0.57, 0.88]).

A pooled analysis of ASCEND, CAPACITY 1 & 2 studies at Week 52 showed a significant treatment benefit for pirfenidone over placebo for PFS (HR: 0.62; 95% CI: 0.51–0.76; p<0.0001), see Figure 14 (Noble 2014a).

Figure 14. Progression free survival at Week 52 in ASCEND, CAPACITY studies, and in the pooled population



† Cox proportional hazards model
‡ Log-rank test

In SP3, PFS was a pre-specified secondary endpoint. Results showed pirfenidone 1800 mg/day reduced the risk of death or disease progression by 55% compared to placebo (HR 0.45; 95% CI 0.11–0.79; p=0.028, log-rank test). There was evidence of a strong treatment effect that began at approximately Day 70 and was sustained throughout the 52-week study period. (FDA, 2010; Taniguchi 2010).

Table 26. Extracted data for input into the NMA for PFS at 52 & 72 weeks

Study	Time point	HR (95% CI)	p-value
SP3 (Lederer 2014)	52 weeks	0.65 (0.48, 0.88)	NR
CAPACITY 1 (Data on file[1])	52 weeks	0.78 (0.52, 1.15)	0.2084
CAPACITY 2 (Data on file[2])	52 weeks	0.58 (0.40, 0.83)	0.0028
CAPACITY 1 (Data on file[3])	72 weeks	0.75 (0.54, 1.06)	0.1006
CAPACITY 2 (Data on file[4])	72 weeks	0.57 (0.41, 0.80)	0.0008
Explanation on differences to BI's NMA: • SP3 data was taken from a congress poster Lederer 2014 • CAPACITY 1 & 2 were handled as individual studies			
[1] Roche 2016a [2] Roche 2016a			

Acute Exacerbations

The criteria for acute exacerbations varied across the trials. In CAPACITY 1 and 2, acute exacerbation required all of the following within a 4-week interval:

- Worsening of PaO2 (≥8 mm Hg drop from the most recent value)
- Clinically significant worsening of dyspnoea
- New, superimposed ground-glass opacities on HRCT in one or more lobes
- All other cardiac, thromboembolic, aspiration, infectious processes ruled out

For ASCEND, acute exacerbations were identified via a post-hoc analysis of adverse events based on the MedDRA lower level term “acute exacerbation of IPF”.

In SP3, the diagnostic criterion of acute exacerbation of IPF was manifestation of the following within a month:

- Increase in dyspnoea
- New ground-glass opacities appear on HRCT in addition to previous honeycomb lesion
- Oxygen partial pressure in resting arterial blood (PaO2) is lower by more than 10 Torr than previous one
- Exclusion of obvious causes, such as infection, pneumothorax, cancer, pulmonary embolism or congestive heart failure.
- The serum levels of CRP, LDH are usually elevated as well as serum markers of interstitial pneumonias, such as KL-6, Sp-A or Sp-D

In SP2, the definition of acute exacerbation of IPF was manifestation of all of the following:

- Worsening, otherwise unexplained clinical features within 1 month
- Progression of dyspnoea over a few days to less than 5 weeks
- New radiographic/HRCT parenchymal abnormalities without pneumothorax or pleural effusion (e.g., new, superimposed ground-glass opacities)

- A decrease in the PaO2 by 10 mm Hg or more
- Exclusion of apparent infection based on absence of Aspergillus and pneumococcus antibodies in blood, urine for Legionella pneumophila, and sputum cultures.

In SP3, the incidence of acute exacerbation during the study or within 28 days after the termination of the study was 5.6% (n=6), 5.5% (n=3) and 4.8% (n=5) in the pirfenidone 1800 mg/day, pirfenidone 1200 mg/day and placebo groups, respectively. No significant differences were seen among the three groups (Taniguchi 2010). A stepwise multivariate analysis revealed that decline in VC ≥10% within 6 months (HR, 3.951, p=0.012) was a significant risk factor for acute exacerbations (Taniguchi 2011b).

In SP2, the incidence of acute exacerbation of IPF was 14% (n=5) in the placebo group and was none in the pirfenidone group during the 9 months (p=0.0031) (Azuma 2005). All five patients with acute exacerbation required hospitalisation for supportive care that included high-dose corticosteroid therapy and oxygen supplementation. One of the five patients in the placebo group died after the onset of acute exacerbation.

The publications for ASCEND and the CAPACITY studies did not report the incidence of acute exacerbations. These data have been extracted from the CSRs for use in the meta-analysis and NMA, and are reported in Table 27.

Table 27. Extracted data for acute exacerbations for ASCEND, CAPACITY 1 & 2 used to input into NMA

Trial	Intervention	Time point	n
CAPACITY 2 (Data on file[1])	**Pirfenidone 2403 mg/day n=174**	52 weeks	0
	Placebo n=174		3
CAPACITY 1 (Data on file[1])	**Pirfenidone 2403 mg/day n=171**	52 weeks	1
	Placebo n=173		0
ASCEND (Data on file[2])	**Pirfenidone 2403 mg/day n=278**	52 weeks	24
	Placebo n=277		40
Explanations for differences to BI's NMA: • For CAPACITY 1 & 2, acute exacerbations were not reported in the primary manuscript Noble 2011. Data at 52 weeks were available as data on file and were handled as separate studies. • For ASCEND, acute exacerbations were not reported in the primary manuscript King 2014. Acute exacerbations at 52 weeks were available as data on file.			
[2] Roche 2016a.			

Hospitalisations

Details of hospitalisations were available for the CAPACITY trials and SP2. In the pooled CAPACITY 1 & 2 population, the number of patients with at least one hospitalisation for respiratory causes (14.8% pirfenidone 2403 mg/day arm vs. 15% placebo arm) and non-respiratory causes (20.9% pirfenidone 2403 mg/day vs. 16.1% placebo arm) was similar across treatment arms (Noble 2011), see Table 28.

Whilst the incidence of hospitalisation was similar, the duration of these hospital stays was consistently numerically longer in the placebo arms. In SP2, five patients in the pirfenidone treatment arm were hospitalised due to exacerbations (Azuma 2005).

Table 28. Post-hoc analysis of hospitalisation observed in CAPACITY 1 & 2

Study arm	CAPACITY 2		CAPACITY 1		Pooled	
	Pirfenidone 2403 mg/day n=174	Placebo n=174	Pirfenidone 2403 mg/day n=171	Placebo n=173	Pirfenidone 2403 mg/day n=345	Placebo n=347
Respiratory Hospitalisations (RH)						
Number of patients with at least 1 RH	29 (16.7%)	29 (16.7%)	22 (12.9%)	23 (16.7%)	51 (14.8%)	52 (15.0%)
Number of RH	34	40	31	37	65	77
Mean length of RH (days)	7.6	12.1	8.5	17.3	8.0	14.6
Total number of days in hospital	259	484	264	640	522	1124
Average number of NRH days per patient	1.5	2.8	1.5	3.7	1.5	3.2
Non-respiratory hospitalisations (NRH)						
Number of patients with at least 1 NRH	35 (20.1%)	31 (17.8%)	37 (21.6%)	25 (14.5%)	72 (20.9%)	56 (16.1%)
Number of NRH	38	42	48	31	86	73
Mean length of NRH (days)	7.2	16.0	10.1	20.8	8.8	8.0
Total number of days in hospital	274	672	485	645	758	1317
Average number of NRH days per patient	1.6	3.9	2.8	3.7	2.2	3.8

Physical Functioning

Categorical analysis of change from baseline in 6MWD

A decrement of ≥50 metres in 6MWD represents an appropriate and clinically relevant threshold for a categorical assessment of response to therapy as it has been associated with an increased risk of mortality (du Bois 2011). Categorical analysis of 6MWD data was carried out post-hoc in the CAPACITY 1 & 2 studies, but was pre-specified as a secondary endpoint in ASCEND. Results are summarised in Table 29.

Table 29. Proportion of patients with a mean decline ≥50 m in 6MWD or death from baseline the relevant RCTs (ITT population)

Study (Ref)	Time point	Treatment group	Mean decline ≥50 m in 6MWD or death, n (%)	Difference, p-value
ASCEND (King 2014 Suppl)	52 weeks	PFN 2403 mg/day (N=278)	72 (25.9)	Absolute difference: 9.8% Relative reduction: 27.5% p=0.04*
		PBO (N=277)	99 (35.7)	
CAPACITY 2 (FDA, 2010)	72 weeks	PFN 2403 mg/day (N=170)	62 (36.5)	p=0.049**
		PBO (N=170)	80 (47.1)	
CAPACITY 1 (FDA, 2010)	72 weeks	PFN 2403 mg/day (N=169)	56 (33.1)	p=0.10**
		PBO (N=168)	79 (47.0)	
Pooled CAPACITY 1 & 2 (FDA, 2010)	72 weeks	PFN 2403 mg/day (N=339)	118 (34.8)	Absolute difference: 12.2% Relative risk: 26% p=0.001**
		PBO (N=338)	159 (47.0)	

PFN- pirfenidone; PBO- placebo
*Rank ANCOVA (pirfenidone 2403 mg/day vs placebo)
**CMH test

A pooled analysis of ASCEND, CAPACITY 1 & 2 studies at comparable time points of Week 48 to 52 showed treatment with pirfenidone significantly improved 6MWD compared with placebo (p=0.0004) (Nathan 2014).

Mean change in 6MWD from baseline

The reliability and validity of 6MWD as a responsive measure of disease status and a valid endpoint for clinical trials has been demonstrated in a recent study, where the minimally clinical important difference (MCID) was estimated at 24-45 meters (du Bois 2011).

Three trials reported data in relation to change in 6MWD (ASCEND, CAPACITY 1 & 2). The overall findings of the three Phase III studies provide strong and consistent evidence for a clinically meaningful benefit of pirfenidone on the exercise tolerance of patients with IPF, see Table 30.

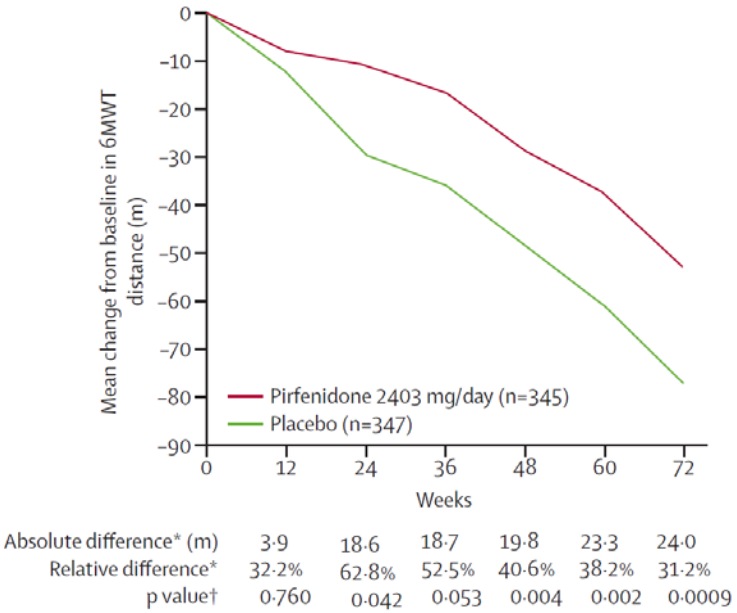
Table 30. Mean change from baseline in 6MWD in the relevant RCTs

Study (Ref)	Time point	Treatment group	Mean decline, metres	Difference, p-value†
ASCEND (King 2014 Suppl; InterMune 2014)	52 weeks	PFN 2403 mg/day (N=278)	33.5 m	Absolute difference: 26.7 m Relative reduction: 44.2% p=0.036
		PBO (N=277)	60.2 m	
CAPACITY 1 (Noble 2011)	72 weeks	PFN 2403 mg/day (N=174)	45.1 m	Absolute difference: 31.8 Relative difference: NR p<0.001
		PBO (N=174)	76.9 m	
CAPACITY 2 (Noble 2011)	72 weeks	PFN 2403 mg/day (N=171)	60.4 m	Absolute difference: 16.4 m Relative difference: NR p=0.171
		PBO (N=173)	76.8 m	
Pooled CAPACITY 1 & 2 (Noble 2011)	72 weeks	PFN 2403 mg/day (N=345)	52.8 m	Absolute difference: 24 m Relative difference: 31.2% p=0.0009
		PBO (N=347)	76.8 m	

PFN- pirfenidone; PBO- placebo
†Rank ANCOVA (pirfenidone 2403 mg/day vs placebo)

In the pooled analysis of the CAPACITY trials, the mean decline from baseline for 6MWD was markedly reduced in the pirfenidone group compared to the placebo group at Week 72 (-52.8m and -76.8m, respectively; absolute difference of 24 m; $p < 0.001$) (Noble 2011), see Figure 15.

Figure 15. Mean change from baseline in 6MWD in CAPACITY 1 & 2 pooled population



*Pirfenidone 2403 mg/day vs placebo
†Rank ANCOVA (pirfenidone 2403 mg/day vs placebo)
174

Table 31. Extracted data used to input into NMA for 6MWD

Study (source)	Treatment	Time point	Mean change from baseline	SD	Mean difference from PBO	p-value
CAPACITY 1 (Data on file)	PFN 2403 mg/day (n=169)	48 weeks	-23.5	114.7	21.5	0.023
	PBO (n=168)		-44.9	105.74		
CAPACITY 2 (Data on file)	PFN 2403 mg/day (n=170)	48 weeks	-34.50	100.02	18.0	0.068
	PBO (n=170)		-52.60	121.01		
ASCEND	PFN 2403 mg/day (n=278)	52 weeks	-33.6	95.73	26.7	0.036
	PBO (n=277)		-60.2	122.56		

Change in Health Related Quality of Life

St. George's Respiratory Questionnaire

The validity of the SGRQ to capture QoL and changes over time in patients with IPF has been shown in a number of studies, with several showing a strong correlation between physical impairment and disease severity, clinical symptoms, and functional disability in patients with IPF (Lee 2014).

Only the CAPACITY trials reported data for this outcome. At Week 72, there was no evidence of a treatment effect in either trial.

Table 32. Summary of change in health status from baseline to Week 72 in CAPACITY 1 & 2 (all randomised patients)

	Change from baseline to Week 72 (mean ± SD)		p-value*
	Pirfenidone	Placebo	
CAPACITY 2	**(n=163)**	**(n=165)**	
SGRQ	7.6 ± 18.89	9.0 ± 18.86	0.495
CAPACITY 1	**(n=166)**	**(n=169)**	
SGRQ	7.2 ± 16.85	7.3 ± 20.37	0.766
*Rank ANCOVA stratified by geographic region (USA and rest of world). Missing data due to a patient's death were ranked as worse than any non-death and according to time until death			

Table 33. Extracted data used to input into NMA for change in SGRQ score from baseline

Trial	Treatment	Time point	Change from baseline		Diff vs. placebo	
			Mean change	SD	Mean diff	p value
CAPACITY 2 (Data on file[1])	PFN 2403 mg/day (n=163)	48 weeks	4.4	14.94	-2.5	0.525
	PBO (n=165)		6.8	16.96		
CAPACITY 1 (Data on file[2])	PFN 2403 mg/day (n=166)	48 weeks	4	13.87	-0.1	0.6738
	Placebo (n=169)		4.1	16.36		
[1] Roche 2016a [2] Roche 2016a						

University of San Diego Shortness of Breath Questionnaire

It has been reported that the SOBQ can be used to formulate clinically relevant inferences about IPF patients (Gries 2013; Swigris 2014). This total score in this questionnaire increases with increased dysponea, and an increment of 20 points is

considered a clinically relevant threshold based on estimates of the minimal important difference for the USCD SOBQ that range from 5-11 (Swigris 2012).

In ASCEND, the proportion of patients with ≥20 point increase in shortness of breath as measured by SOBQ was smaller in patients receiving pirfenidone compared with placebo (p=0.1577), see Table 34 (InterMune, 2014). In CAPACITY 1 & 2, no significant differences were observed between the pirfenidone and placebo groups for the change from baseline to Week 72, see Table 35. Pooled data from the three studies showed pirfenidone treatment reduced the proportion of patients who experienced a ≥20 point increase or death (p=0.0471) (Noble 2014a).

Table 34. Categorical outcomes for UCSD SOBQ in ASCEND at Week 52† (all randomised)

Outcomes, n (%)	Pirfenidone 2403 mg/day (n=278)	Placebo (n=277)	p-value*
Worsening score ≥20 points or death	81 (29.1)	100 (36.1)	0.1577
Worsening score <20 to 0 points	124 (44.6)	115 (41.5)	
No worsening (score change <0 points)	73 (26.3)	62 (22.4)	

†Missing data due to reasons other than death were imputed using the sum of squared differences (SSD) method and included in the ≥20 points category
*p-value by rank ANCOVA

Table 35. Mean change in UCSD SOBQ score from baseline for the relevant RCTs (ITT population)

Study (Ref)	Time point	Treatment group	Mean change in dyspnoea score	p-value*
CAPACITY 2 (FDA, 2010)	72 weeks	PFN 2403 mg/day (N=174)	12.1	p=0.509
		PBO (N=174)	15.2	
CAPACITY 1 (FDA, 2010)	72 weeks	PFN 2403 mg/day (N=171)	11.9	p=0.604
		PBO (N=173)	13.9	

PFN- pirfenidone; PBO- placebo
*Rank ANCOVA (pirfenidone 2403 mg/day vs placebo)

Data on the change in SOBQ score from baseline were extracted from ASCEND and CAPACITY studies at a common timepoint, and are presented in Table 36.

Table 36. Extracted data used to input into NMA for change in SOBQ score from baseline

Trial	Treatment	Time point	N	Change from baseline	
				Mean change	SD
CAPACITY 2 (Data on file[1])	PFN 2403 mg/day (n=171)	48 weeks	171	8.3	21.17
	PBO (n=169)		169	12.1	23.6
CAPACITY 1 (Data on file[2])	PFN 2403 mg/day (n=168)	48 weeks	168	7.6	18.78
	PBO (n=171)		171	10.2	24.22
ASCEND (Data on file[3])	PFN 2403 mg/day (n=278)	52 weeks	278	14	23.68
	PBO (n=277)		277	17.3	24.49

[1] Roche 2016a
[2] Roche 2016a
[3] Roche 2016a

Gas Transfer

Measurement of the carbon monoxide diffusing capacity of the lungs (DLco)

Four of the trials (CAPACITY 1 & 2, SP3, SP2) reported data in relation to the change from baseline in DLco. The CAPACITY trials reported the change in % predicted DLco while SP2 and SP3 reported the mean decline (mL/min/mmHG).

There was a reduced mean decline from baseline in percent predicted Hgb-corrected DLco that tended to favour pirfenidone 2403 mg/day in CAPACITY 2 (mean change of -7.9% and -9.9%, for pirfenidone and placebo respectively, p=0.145) that was not observed in CAPACITY 1 (mean change of -9.8% and -9.2%, respectively, p = 0.996) at Week 72 (Noble 2011). A published, pooled analyses indicated that there is no evidence of a treatment effect for this outcome (p=0.301) (Noble 2011).

In both the SP2 and SP3 trials, there was no statistical difference in mean decline of DLco between pirfenidone 1800 mg/day and placebo (Taniguchi 2010; Azuma 2005).

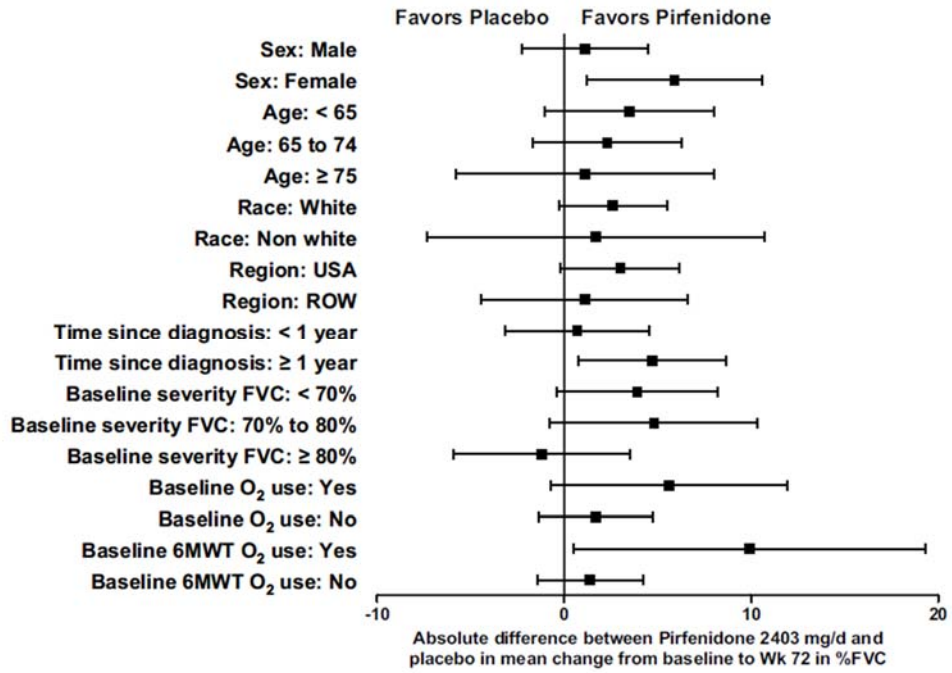
4.8 Subgroup analysis

CAPACITY 1 & 2

Pre-specified analyses of study outcomes from CAPACITY 1 & 2, performed in subgroups defined by baseline patient characteristics were generally consistent with the overall population findings (FDA, 2010).

Pooled CAPACITY 1 & 2 data were also used for analysis of the primary efficacy outcome variable in subpopulations defined by baseline patient characteristics. There was no evidence of interaction between treatment and the following subgroups: sex (p=0.263), age (p=0.864), race (p=0.807), geographic region (p=0.359), and baseline IPF severity (p=0.352), see Figure 16. There was evidence of an interaction between treatment and time from IPF diagnosis to randomisation, with those patients diagnosed more than a year before randomisation experiencing a significantly greater treatment effect. Both subgroups still favoured pirfenidone over placebo (absolute difference in percent predicted FVC at 72 week 0.7 vs 4.7 p=0.021 (FDA 2010)).

Figure 16: Subgroup analyses of difference between pirfenidone and placebo in mean change from baseline to Week 72 in % predicted FVC (pooled CAPACITY studies)



Overall, a treatment effect was seen in all subgroups and no difference in treatment effect was found across all demographic and baseline disease subgroups.

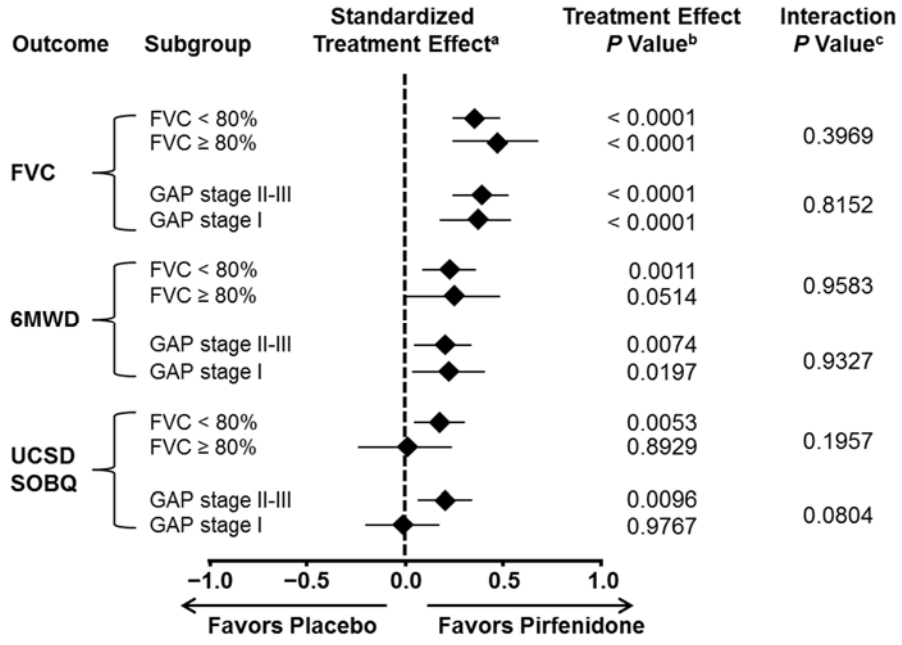
Pooled subgroup analysis of ASCEND, CAPACITY 1 & 2 (post-hoc)

A post-hoc analysis of pooled data from ASCEND, CAPACITY 1 & 2 was conducted to examine the effects of pirfenidone on patients stratified by earlier vs. more advanced disease severity (Albera 2015). Populations were stratified by baseline FVC ≥80% (pirfenidone, n=146; placebo, n=170) and baseline FVC <80% (pirfenidone, n=477; placebo n=454). Baseline characteristics and demographics were similar across groups. Efficacy outcomes of interest include absolute ≥10% FVC decline, ≥50m 6MWD decline, and ≥20-point worsening of dyspnoea as measured by UCSD SOBQ. Treatment-by-subgroup interaction was tested based on rank ANCOVA model. Missing FVC, 6MWD and UCSD SOBQ values were imputed by using the sum of squared differences method. Factors in the model include study, geographic region, treatment group, subgroups, and treatment-by-subgroup interaction. A proportional hazards model estimated the HR between subgroups.

Results demonstrated that treatment with pirfenidone leads to a consistent effect in the risk of FVC decline ≥ 10% or death regardless of baseline FVC level, with no significant difference found between those patients with baseline FVC ≥ 80% predicted vs, those with FVC < 80% predicted (

Figure **17**).

Figure 17. Treatment effect of pirfenidone by baseline disease severity from pooled data of ASCEND, CAPACITY 1 & 2



A separate post-hoc analysis was conducted to evaluate the outcomes of patients who experienced a ≥10% decline in % FVC during the first 6 months of treatment in ASCEND, CAPACITY 1 & 2 studies (N=1247) (Nathan 2015a, Nathan 2016). Significantly fewer patients receiving pirfenidone experienced a subsequent ≥10% decline in percent predicted FVC or death compared with placebo.

These findings suggest a potential benefit to continued treatment with pirfenidone despite an initial decline in FVC, which is not consistent with the stopping rule currently recommended in TA282 (NICE 2013a).

Table 37. Outcomes following previous ≥10% decline in FVC

Outcome, n (%)	Pirfenidone (n=24)		Placebo (n=60)		Relative Difference*	P value
≥10% decline in FVC or death	1	(4.2)	15	(25.0)	−83.3%	0.032
Death	0	(0)	10	(16.7)	−100%	0.056
>0% to <10% decline in FVC	9	(37.5)	23	(38.3)	-2.2%	ND
No further decline in FVC	14	(58.3)	22	(36.7)	59.1	0.089
*Relative difference calculated using the following formula: 100 × [pirfenidone − placebo]/[placebo]						

These results were supported by an additional post-hoc analysis of pooled ASCEND, CAPACITY 1 & 2 data, which evaluated the effect of pirfenidone on patient subgroups (age [< 65 years, 65-74 years, ≥75 years], and smoker status [current/former smoker, never smoked] and baseline disease status (e.g. time since diagnosis, supplemental oxygen use, % predicted FVC [<65, 65 to ≤80, >80), % predicted DLco [<40, 40 to 50, ≥50], and 6MWD) (Noble 2014b). Overall, a treatment effect was seen in all subgroups and no difference in treatment effect was found across all demographic and baseline disease subgroups.

SP3

An exploratory subgroup analysis was conducted in SP3 (Azuma 2011).

Patients were stratified by baseline % predicted VC, PaO2, and the lowest SpO2 during the 6MWT. IPF patients with baseline % predicted VC ≥ 70% and SpO2 < 90% were most likely to benefit from pirfenidone when evaluated using changes in VC, and cough and dyspnoea symptoms (Azuma 2011). Stratified analysis from other exploratory examination showed similar results (Taniguchi 2010; Ebina 2010). Pirfenidone-treated patients with % predicted VC ≥ 70% at baseline, had significantly improved clinical outcomes by reducing decline in VC, and elongated the PFS duration at Week 52 (Taniguchi 2010; Ebina 2010). These results demonstrated pirfenidone exerted more pronounced effects in IPF patients with mild impairment.

RECAP

Patients who entered the long-term extension study, RECAP, were evaluated in two groups: those with a % predicted FVC <50% and those at >50%. Mean FVC and

DLco values were [redacted] and [redacted] in patients with FVC <50% predicted, compared to [redacted] and [redacted], in the FVC ≥50% predicted cohort (Roche 2016a).

Both groups declined similarly over 180 weeks, with an annual rate of decline of [redacted]. Long-term treatment with pirfenidone had a similar rate of decline in patients with baseline FVC<50% compared to patient with more preserved lung function (Figure 18).

This suggests that the treatment benefit from pirfenidone is consistent across FVC levels, supporting early intervatntion and continued treatment.

Figure 18: Mean (SD) of percent predicted FVC over time (RECAP) by baseline FVC category (< 50% predicted or ≥ 50% predicted)

[redacted]

4.9 Meta-analysis

- The methods and results of the meta-analyses conducted to inform the NMA are fully described in Appendix 9.
- Meta-analysis was used compare pirfenidone with placebo in patients with mild to moderate IPF.
- The outcomes evaluated include: change from baseline in lung capacity; all-cause and IPF-related mortality; progression free survival; acute exacerbation; changes from baseline in 6MWD and USCD SOBQ; discontinuation of treatment and treatment related serious adverse events.
- To ensure the internal and external validity of the analyses, each meta-analysis summarises an outcome for a specific time point. The principal meta-analyses summarise each outcome at one year. Data at this time point was available for SP3, CAPACITY 1 and 2 and ASCEND. For some outcomes, sensitivity analyses also include CAPACITY data up to 72 weeks.
- SP2 was excluded from the analyses as it was previously considered to be an outlier by NICE in the nintedanib appraisal (NICE 2015d). Furthermore, SP2 did not provide any data at one year (SP2 was stopped at 36 weeks).
- Random effects models were specified as the principal analysis.
- The meta-analyses suggest that pirfenidone slows the rate of decline in FVC and improves survival. For percent predicted FVC, on average, over 52 weeks, patients receiving pirfenidone decline by 3.4 percentage points less than patients receiving placebo (95% CI: 1.9% to 4.9%). The meta-analysis of change from baseline in FVC (measured in litres) also shows that pirfenidone slows the decline in lung function.
- Compared to placebo, at 52 weeks, pirfenidone reduces all-cause mortality [HR: 0.52, 95% CI: 0.31 to 0.88] and IPF-related mortality [HR: 0.37 (0.18, 0.76)]. Pirfenidone also improves progression free survival [HR 0.63 (0.53, 0.74)].
- The meta-analyses also suggest that pirfenidone reduces the decline in 6MWD and USCD SOBQ.

- For acute exacerbation, the random effects meta-analysis suggests no evidence of a difference between pirfenidone and placebo. A critical limitation of this analysis is that different studies used different definitions of acute exacerbation (see Section 4.7).
- There was no evidence that treatment with pirfenidone affects the rate of all-cause discontinuation of treatment or the rate of treatment related serious adverse events.

4.10 Indirect and mixed treatment comparisons

Summary of network meta-analyses

- There are no head-to-head trials comparing pirfenidone with nintedanib in patients with IPF, therefore a NMA was conducted. The NMA was conducted to support pricing and reimbursement submissions across all markets, and so also includes NAC and triple therapy.
- The outcomes evaluated include: change in % predicted FVC, FVC in litres, 6MWD, SOBQ, SGRQ total score, mortality, PFS, acute exacerbations of IPF and discontinuation of treatment. The NMA was not conducted for change from baseline in DLco due to the lack of comparator trial data.
- To ensure the internal and external validity of the analyses, each meta-analysis summarises an outcome for a specific time point. The principal NMA summarise each outcome at one year (sensitivity analyses include CAPACITY data up to 72 weeks)
- The base case network included all Phase II and III trials of pirfenidone, nintedanib, NAC and triple therapy with the exception of SP2 (King 2014; Noble 2011; Taniguchi 2010; Richeldi 2014; Martinez 2014; Richeldi 2011; Raghu 2012).
- Extracted (unpublished) TOMORROW and INPULSIS data reported in the nintedanib submission have been included in our NMA where appropriate [Boehringer Ingelheim 2015]
- The NMA was conducted using standard Bayesian approaches, following the guidelines set out by the NICE Technical Support Documents on evidence synthesis [NICE DSU 2011a]. Random effects models were specified as the principal analysis.
- For the changes from baseline in percent predicted FVC/VC and FVC/VC (litres), the NMA suggests that both pirfenidone and nintedanib are superior to placebo.
- The NMA suggests that pirfenidone reduces all-cause and IPF-related mortality, compared to placebo. Pirfenidone also increases PFS. There is no

evidence that nintedanib reduces mortality or improves PFS.

- For acute exacerbation and change from baseline in SGRQ there is no evidence of a difference between any of the treatments.
- For all-cause discontinuation of treatment, there is no evidence of a difference between pirfenidone and placebo. Nintedanib increases the odds of all-cause discontinuation of treatment, relative to placebo.

Search strategy

In absence of head-to-head trials of pirfenidone with other IPF treatments, a systematic literature review and NMA was performed to estimate the relative effectiveness of pirfenidone to other available treatments (Roche 2016). A search strategy was developed to identify any RCT of IPF so this would identify pirfenidone and comparator studies for the treatment of IPF in line with the decision problem. Information sources and literature search for study selection has been described in section 4.1.2 and full details of the search strategies, databases and resources used to identify studies are provided in Appendix 3.

Study selection

The NMA was conducted to support pricing and reimbursement submissions across all markets, and included placebo, pirfenidone, nintedanib, NAC and triple therapy. NAC and triple therapy as treatments of interest.

The process of study selection for abstract and full-text review was the same as detailed in section 4.1. However, the PICOS criteria were amended to capture RCTs of pirfenidone and other comparators (including nintedanib), see Table 38.

Table 38. Criteria used in the trial selection process for NMA

Clinical effectiveness	Inclusion criteria	Exclusion criteria
Population	Adults (aged 18 or older) with suspected or diagnosed IPF	Studies of children and young people (younger than 18) Studies of people with a diagnosis of pulmonary fibrosis as a complication of either of the following: • Connective tissue disorders (systemic lupus erythematosus, rheumatoid arthritis, scleroderma, polymyositis, and dermatomyositis) A known exogenous agent (for example, drug induced disease or asbestosis)
Intervention	• Pirfenidone • Double therapy (with prednisone and azathioprine) • N-acetylcysteine (NAC) • Nintedanib • Triple therapy (Prednisone and azathioprine and NAC)	Interventions or combinations of interventions other than those listed
Comparators	• Placebo or best supportive care • Double therapy (with Prednisone and azathioprine) • N-acetylcysteine (NAC) • Nintedanib • Pirfenidone • Triple therapy (Prednisone and azathioprine and NAC)	Comparators or combinations of comparators other than those listed
Outcomes	• Lung capacity (VC/FVC) • Gas transfer (DLco) • Physical functioning (6MWD) • Progression-free survival (PFS) • Adverse effects of treatment • Health related quality of life measured using SGRQ, SOBQ,	• Anticoagulation for the treatment of pulmonary hypertension Treatment of lung cancer • Lung transplantation other than timing and referral

	dyspnoea score or EQ-5D • Hospitalisations • Acute exacerbations • Mortality (All cause or IPF-related) • Categorical declines in FVC (0%, 5% and 10%) • Discontinuation • Compliance	
Trial design	• Studies in humans, published as Phase II or III RCTs • Studies published as abstracts, conference presentations or press releases were eligible if adequate data were provided • SRs were eligible for inclusion as a source of references to primary studies	• Cross-over RCTs
Language restrictions	No language limits	No language limits

Results

A total of 4394 records were identified (after de-duplication) and assessed for relevance. Subsequent titles and abstract review was performed, and 116 full text documents were retrieved and assessed against the eligibility criteria. Subsequently, 56 documents reporting results of 10 RCTs were identified. A study flow diagram for the NMA systematic literature review is presented in Figure 2.

We are aware of further outcomes reported from the TOMORROW and INPULSIS studies, which were presented in the nintedanib NICE manufacturer submission (Richeldi 2011, Richeldi 2014); specifically:

- Acute exacerbation data from the TOMORROW study (Richeldi 2011)
- 6MWD data from the TOMORROW study
- PFS data from the pooled INPULSIS trials (Richeldi 2014)

These data became available after our systematic literature review was conducted in April 2015. Where appropriate, these have been included in the NMA.

Of the ten trials identified, two of these trials (SP2 and IFIGENIA) were excluded from the NMA for reasons presented in section 4.10.5.

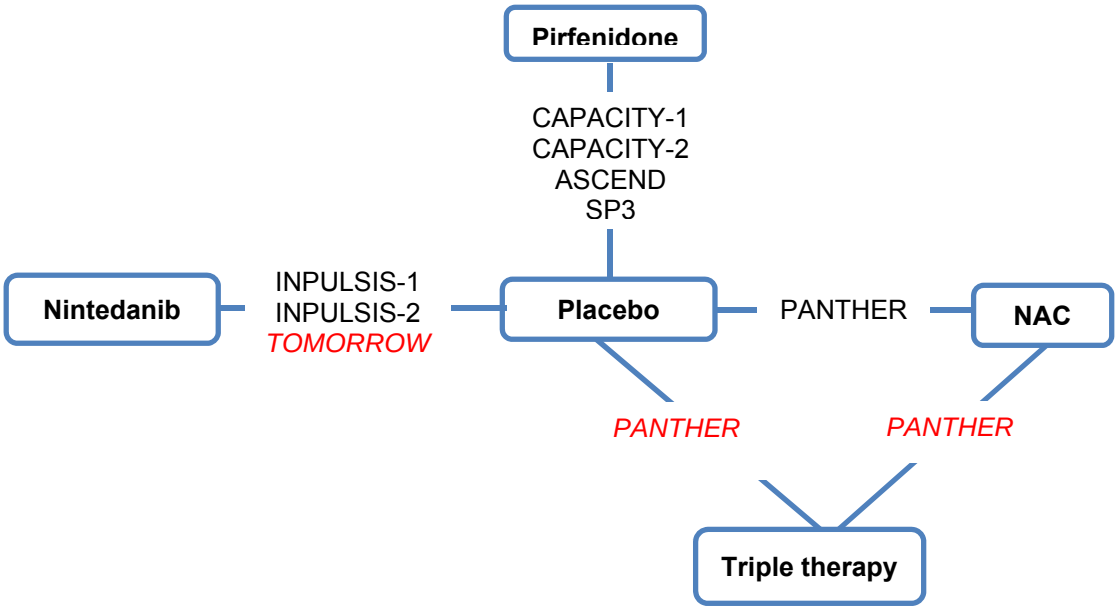
The trials included in the base case network are summarised in Table 39. As a sensitivity analysis, NMA was also applied to a restricted network. The restricted network is limited to phase III trials, and excludes triple therapy. Trials included in the restricted network are indicated in Table 39.

Table 39. Summary of the trials used to carry out the indirect or mixed treatment comparison

Trial (Reference)	Included in restricted network?	Placebo	PFN	NAC (oral)	Triple therapy	NTB
ASCEND (King 2014)	Yes	Yes	Yes			
CAPACITY 1 (Noble 2011)	Yes	Yes	Yes			
CAPACITY 2 (Noble 2011)	Yes	Yes	Yes			
SP3 (Taniguchi 2010)	Yes	Yes	Yes			
INPULSIS-1 (Richeldi 2014)	Yes	Yes				Yes
INPULSIS-2 (Richeldi 2014)	Yes	Yes				Yes
PANTHER* (Raghu 2012; Martinez 2014)	Yes*	Yes		Yes		
TOMORROW (Richeldi 2011)		Yes				Yes

* Only the placebo and NAC arms of PANTHER are included in the restricted network

Figure 19. Network diagram including all trials for NMA



All trials are included in the base-case network
Red italicised text indicates trials excluded from the restricted network

The literature search and NMA described above were developed to support submissions of pirfenidone to all national agencies: as such, some comparators of interest included in the searches are beyond the scope of this appraisal. Given the IFIGENIA trial compares double and triple therapy, inclusion of this trial would not contribute any information to the comparisons of interest in this appraisal and so will not feature in any of the NMA.

Consistent with the meta-analysis (Section 4.9) and conclusions of the ERG and NICE Committee on review of the nintedanib manufacturer submission, SP2 was excluded from all (extended) networks as it is considered an outlier, and did not provide any data at one year (SP2 was stopped at 36 weeks) [NICE 2015f, NICE 2016].

Methods and outcomes of included studies

Time points

A key assumption of NMA is that studies should be similar in terms of any characteristics that impact the treatment effect. The principal NMAs in this submission summarise each outcome at approximately one year (48 weeks to 60 weeks). We make the assumption that treatment effects will be similar across these time points. The NMA presented in the nintedanib NICE manufacturer submission

BI's included data from 36 to 72 weeks [Boehringer Ingelheim 2015]. It is difficult to justify whether treatment effects will be stable over this longer time period.

Outcomes included in studies

Outcome measures were chosen to reflect the decision problem. The NMA was conducted for the following outcomes:

Survival outcomes

- IPF-related mortality
- All-cause mortality
- PFS

Survival outcomes were analysed on the log-hazard scale (leading to HRs):

- For the principal analysis, we assumed that proportional hazards is an acceptable assumption up to 52 weeks, but not beyond. For the CAPACITY studies we used the estimates of the HR reported at 48 weeks/12months;
- As a sensitivity analysis, we assumed that proportional hazards holds up to Week 72. For the CAPACITY studies we used the estimates of the HR at Week 72.

We evaluated overall survival (all-cause mortality) on the hazard ratios scale, whereas the nintedanib NICE manufacturer submission only evaluated this outcome using the odds ratio scale. Hazard ratios provide a more precise measure of survival because they account for the time of death, whereas odds ratios only account for the number of deaths.

Continuous outcomes

- Percent predicted FVC
- FVC litres
- 6MWD
- SGRQ total score
- UCSD SOBQ

Continuous outcomes were analysed according to change from baseline at 12 months. Where data at 12 months/52 weeks was not available, data from 48 to 60 weeks was used instead.

Binomial outcomes:

- Decline of ≥10% in FVC percent predicted
- Discontinuation of the treatment for any reason
- Acute exacerbation of IPF

Binomial outcomes were analysed as the proportion of patients with the event at 12 months (leading to odds ratios). Where data at 12 months/52 weeks was not available, data from 48 to 60 weeks was used instead.

Populations included in the studies

Populations included in the trials are in line with the licensed indications and the scope. All patients had mild to moderate impairment in pulmonary function at baseline.

Full details of the patient characteristics in each trial can be found in Appendix 10.

Although the eligible age varied slightly across studies, the average ages of the actual populations were similar (62-69 years). In all studies over 70% of the population in each trial was male and in seven trials the majority of patients were white. There was a high Japanese contingent in the INPULSIS trials compared to the other trials assessed. In all trials, IPF diagnosis was made in accordance with applicable consensus guidelines. This was conducted by central review in the ASCEND trial.

Whilst the SP3 study used a lower dose of pirfenidone than that licensed in the UK, this reflects the difference in mean weights in the North American and European population compared to the Japanese population.

All patients had mild to moderate impairment in pulmonary function at baseline however measures of function were reported inconsistently across trials at baseline. Approximately half of the patients in the CAPACITY trials had a diagnosis for less

than 1 year (Noble 2011). The majority of patients in the remaining trials had been diagnosed for 1 year or more.

All studies (with the exception of the TOMORROW trial) reported patients' smoking history. The majority of patients had smoked in the past but were not smokers when joining the trial, and this was consistent across the trials.

Although there were some differences between the baseline populations in the included trials, there were no major concerns regarding the inclusion of any of these trials in the network.

Further detail on the baseline characteristics, methods, outcomes, and results for each of the comparator studies can be found in Appendix 10. Further detail on the baseline characteristics, methods, outcomes, and results for the pirfenidone studies can be found in section 4.7.

Extracted data for the NMA can be found in Appendix 11.

Risk of bias

A summary of the quality assessment for each RCT included in the NMA is summarised below (Table 40). For the full quality assessments for each study, please refer to Appendix 10.

The majority of studies were of good quality, with low risk of bias. All studies were Phase II or III RCTs. Although all trials reported randomisation, some did not report details of the randomisation process so it was unclear whether this process was adequate in all trials. For the purpose of these analyses, it is assumed that randomisation process was adequate for all. Planned treatment duration varied from 52 weeks (ASCEND) to 72 weeks (CAPACITY 1 & 2). Planned analysis in PANTHER was 60 weeks, however, the triple therapy arm was terminated after a mean follow-up of 32 weeks. The implications of this will be explored in each for each outcome in turn.

Table 40. Quality assessment summary of RCTs included for NMA

	Was randomisation carried out appropriately?	Was the concealment of treatment allocation adequate?	Groups similar at baseline in terms of prognostic factors?	Were the care providers, participants, and outcome assessors blind to treatment allocation?	Unexpected imbalances in drop-outs between groups?	Authors measured more outcomes than they reported?	Did the analysis include an ITT basis?	Risk of bias of the study [2]
CAPACITY 1 & CAPACITY 2	Yes	Yes	Yes	Yes	No	No	Yes	Low risk
SP3	Unclear	Yes	Yes	Unclear	No	No	Yes	Low risk
ASCEND	Yes	Yes	Yes	Yes	No	No	Yes	Low risk
PANTHER	Yes	Yes	Yes	Unclear	Yes (at interim analysis)	No	No	Some risk of bias
TOMORROW	Unclear	Yes	Yes	Yes	No	No	No	Low risk
INPLUSIS-1 & INPULSIS-2	Yes	Yes	Yes	Yes	No	No	No	Low risk

All the trials reported some degree of discontinuation and loss to follow up. The trials used different methods for handling missing data. In the CAPACITY trials, a conservative approach was applied whereby missing values as a result of death were assigned the worst rank in ANCOVA analyses, and worst possible outcome in mean change analyses (eg, FVC=0) and categorical analyses. Other missing data were imputed with the average value from three patients with the smallest sum of squared differences at each visit with data that were not missing.

For SP3 and the analysis of secondary endpoints in the TOMORROW trial, the last-observation-carried-forward (LOCF) approach was used when data for the entire 52-week period were not available.

In the INPULSIS trials the statistical model used for the primary analysis allowed for missing data, assuming that they were missing at random; missing data were not imputed for the primary analysis.

A pragmatic approach was taken for the analysis, i.e. including all trials in the NMA, regardless of discontinuation, loss to follow-up and how missing data was handled. Whilst some methods for handling missing data may produce biased results, strict criteria on the handling of missing data, could lead to the exclusion of most trials from the network.

Method of analysis

The NMA was conducted using standard Bayesian approaches, following the guidelines set out by the NICE Technical Support Documents on evidence synthesis (full details of the statistical methodology are provided in Appendix 12 (NICE DSU 2011a).

Choice of Model (random effects models versus fixed effect models)

For the network meta-analyses, random effects models were specified as the principal analysis.

Random effects models are considered more appropriate than fixed effect models because there is heterogeneity in the way some outcomes are measured and in the way missing data is handled (for further details please see the feasibility assessment). Fixed effect models will also be fitted for comparison as sensitivity analyses.

With respect to the network meta-analyses, a key limitation of the random effects models is that there are few studies to estimate the between-study variance. This lack of information may lead to a high level of uncertainty in the results. Despite this limitation, we consider the assumption of no heterogeneity made in the fixed effect model to be unrealistic. As such, we favour the random effects over the fixed effect model.

Assessment of Heterogeneity

Statistical heterogeneity was assessed for each outcome. For each pairwise comparison that was informed by at least two trials an ordinary meta-analysis was performed.

The results were displayed in forest plots. The forest plots provide an assessment of heterogeneity including the I-squared statistic, between-study variance (tau-squared) and the p-value of the heterogeneity statistic Q.

Reporting of results

For each outcome, summaries of the difference between treatments are reported. For continuous outcomes (e.g. FVC, 6MWD) this report presents the mean difference in the change from baseline; for binary outcomes (e.g. adverse events, discontinuation) this report presents the odds ratio, and for survival outcomes (e.g. PFS, all-cause mortality) this report presents the hazard ratio. We have provided posterior medians as point estimates, accompanied by 95% credible intervals (CrI) (the Bayesian analogue of confidence intervals). The estimates and their 95% credible intervals are presented in matrix tables and forest plots.

Results of the NMA

This section presents the results of the random effects models for the base case network (phrased: base case network (random effects model). A summary of the results for all outcomes is provided in Table 55. Forest plots for the statistical assessments of heterogeneity are provided in Appendix 13. Three sensitivity analyses were conducted: base case network (fixed effect model), restricted network (random effects model) and restricted network (fixed effect model). The results of the sensitivity analyses are provided in Appendix 11.

Lung Capacity

- For the changes from baseline in percent predicted FVC/VC and FVC/VC (litres), the NMA suggests that both pirfenidone and nintedanib are superior to placebo. There is no evidence of a statistically significant difference between pirfenidone and nintedanib.
- For decline in percent predicted FVC of 10% or more, the NMA suggests that pirfenidone is superior to placebo. There is no evidence of a difference between nintedanib and placebo. Nor is there evidence of a difference between pirfenidone and placebo.

All studies, with the exception of SP3, used FVC to measure lung capacity. In SP3, lung capacity was measured by VC. As noted in Section 4.9, FVC and VC were considered to be identical in patients with restrictive lung disease such as IPF and therefore were combined for the NMA, similar to the nintedanib NICE manufacturer submission which was accepted as a valid approach by the ERG and NICE Committee [Boehringer Ingelheim 2015, NICE 2015f, NICE 2016].

Change in Percent Predicted FVC/VC

For the base case network, change in percentage of predicted FVC/VC was reported for SP3, CAPACITY 1 and 2, ASCEND, TOMORROW, INPULSIS 1 & 2, and PANTHER (NAC versus placebo arm only).

A difference was noted between the studies in terms of how missing data was analysed. In the key publications, SP3, CAPACITY 1 and 2, ASCEND and TOMORROW presented results based on the imputation of missing values. For SP3 and TOMORROW, missing values were imputed using the last observation carried forward (LOCF) approach (Taniguchi, Richeldi 2011). For CAPACITY 1 and 2 and ASCEND, missing values as a result of death were assigned the worst possible outcome and missing data due to reasons other than death were imputed using the smallest sum of squared differences method (Noble 2011). The other studies (INPULSIS I and II and PANTHER) assumed that data was missing at random and did not impute missing values.

Assessment of statistical heterogeneity

For the comparison of placebo with pirfenidone, the I-squared statistic is 24.4%, indicating low heterogeneity. For the comparison of nintedanib with placebo the I-squared statistic was 0%, indicating minimal heterogeneity.

Base case network (random effects model)

The results of the random effects (RE) model are provided in Figure 20and Table 41 below. The results suggest that both pirfenidone and nintedanib are superior to placebo. There is no evidence of a difference between pirfenidone and nintedanib. The sensitivity analysis is consistent with the base case analysis and can be found in Appendix 14..

Figure 20. Forest plot of the mean difference in change from baseline in percentage of predicted FVC/VC (%) (base case network, RE model)

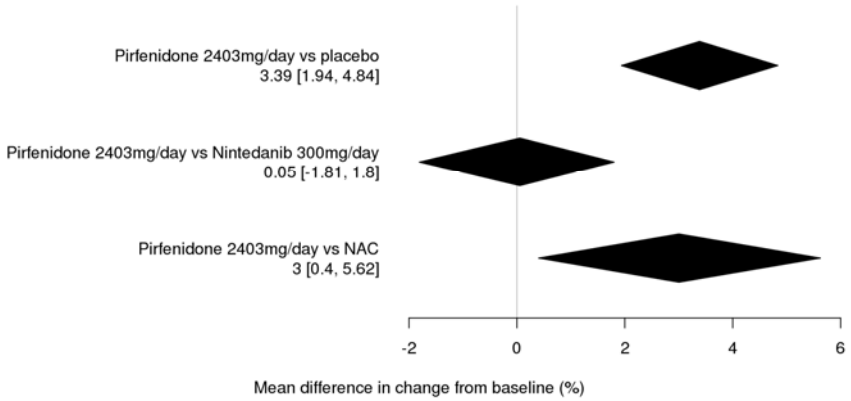


Table 41. Estimates and 95% credible intervals for the mean difference in change from baseline in percentage of predicted FVC/VC (%) (base case network, RE model)

	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		3.39 (1.94, 4.84)	3.33 (2.34, 4.5)	0.38 (-1.79, 2.54)
Pirfenidone 2403mg/day (*)	-3.39 (-4.84, -1.94)		-0.05 (-1.80, 1.81)	-3.00 (-5.62, -0.40)
Nintedanib 300mg/day	-3.33 (-4.50, -2.34)	0.05 (-1.81, 1.80)		-2.95 (-5.44, -0.61)
NAC	-0.38 (-2.54, 1.79)	3 (0.40, 5.62)	2.95 (0.61, 5.44)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Mean differences are calculated as column treatment minus row treatment, with a positive result being favourable to the column treatment.

Change from Baseline in FVC/VC (L)

Change in VC/FVC (L) was reported for all trials. As per percentage of predicted FVC/VC, a difference was noted between the studies in terms of how missing data was analysed.

Assessment of statistical heterogeneity

For the comparison of pirfenidone with placebo, the I-squared statistic was 50%, indicating moderate heterogeneity; for the comparison of nintedanib with placebo, the I-squared statistic was 0%, indicating minimal heterogeneity.

Base case network (random effects model)

For the base case, the results of the random effects model are provided in Figure 21 and Table 42. According to the results, pirfenidone and nintedanib are superior to placebo. There is no conclusive evidence of an effect of pirfenidone compared to NAC, nintedanib, and triple therapy. The results of the sensitivity analysis are in line with the base case analysis and indicate efficacy of pirfenidone and nintedanib in slowing the decline of lung volume over time (see Appendix 14). The sensitivity analyses support that there is no evidence of a difference in effect between pirfenidone and nintedanib.

Figure 21. Forest plot of the mean difference in change from baseline in FVC/VC (L) (base case network, RE model)

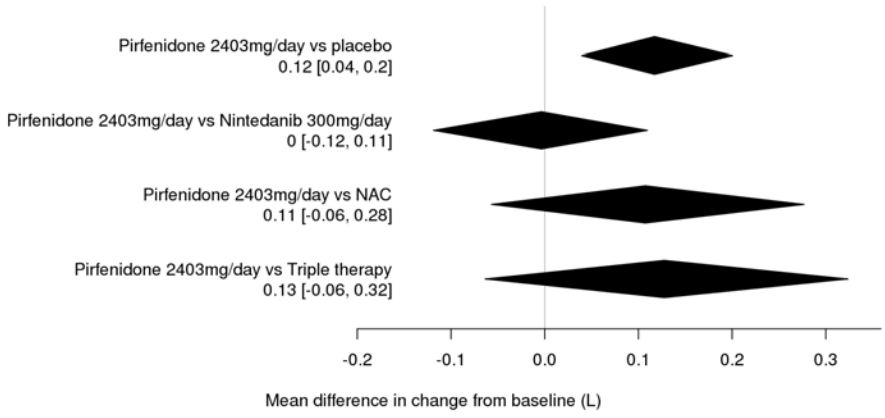


Table 42. Estimates and 95% credible intervals for the mean difference in change from baseline in FVC/VC (L) (base case network, RE model)

	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.12 (0.04, 0.20)	0.12 (0.04, 0.21)	0.01 (-0.14, 0.16)	-0.01 (-0.19, 0.17)
Pirfenidone 2403mg/day (*)	-0.12 (-0.20, -0.04)		0.00 (-0.11, 0.12)	-0.11 (-0.28, 0.06)	-0.13 (-0.32, 0.06)
Nintedanib 300mg/day	-0.12 (-0.21, -0.04)	0.00 (-0.12, 0.11)		-0.11 (-0.28, 0.05)	-0.13 (-0.33, 0.06)
NAC	-0.01 (-0.16, 0.14)	0.11 (-0.06, 0.28)	0.11 (-0.05, 0.28)		-0.02 (-0.25, 0.21)
Triple therapy	0.01 (-0.17, 0.19)	0.13 (-0.06, 0.32)	0.13 (-0.06, 0.33)	0.02 (-0.21, 0.25)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Mean differences are calculated as column treatment minus row treatment, a positive result is favourable to the column treatment.

Categorical decline of ≥10% in percent predicted FVC

The NMA includes data at Week 48 for CAPACITY 1 & 2, and at Week 52 for ASCEND, INPUSIS 1 & 2, and PANTHER. In the CAPACITY studies, FVC assessments were conducted every 12 weeks, therefore it was considered most appropriate to include the 48 week CAPACITY data in this NMA. Results for PANTHER, CAPACITY and ASCEND were re-calculated as per the INPULSIS definition, with missing values of FVC for any reason (including death) assumed to be non-responders.

Results for TOMORROW were only available from BI's submission. The analysis presented in BI's submission was not clearly defined, however the clarification document (page 9) suggests that this analysis might be measuring 10% declines at any point up to 52 weeks. Results for other studies reflect categorical declines at exactly 52 weeks. Hence the measure of categorical decline for TOMORROW was not sufficiently similar to the measures used in the other studies, and TOMORROW was therefore excluded from the NMAs.

Assessment of statistical heterogeneity

For the comparison between placebo and pirfenidone the I-squared statistic was 40.8%, indicating low to moderate heterogeneity; for the comparison between nintedanib and placebo the I-squared statistic was 40%, again indicating low to moderate heterogeneity.

Base case network (random effects model)

The random effects results suggest that pirfenidone is superior to placebo. In all, there is no evidence of a difference between pirfenidone and nintedanib, and pirfenidone and NAC, see Figure 22.

For the base case network, no phase II or triple therapy data was available. Hence, for this outcome, the base case network is equivalent to the restricted network. For the base case network, the fixed effect model suggests that nintedanib is superior to placebo. All other conclusions remain the same.

Figure 22. Forest plot of odds ratios for FVC categorical decline of ≥10% percent predicted (base case network, RE model)

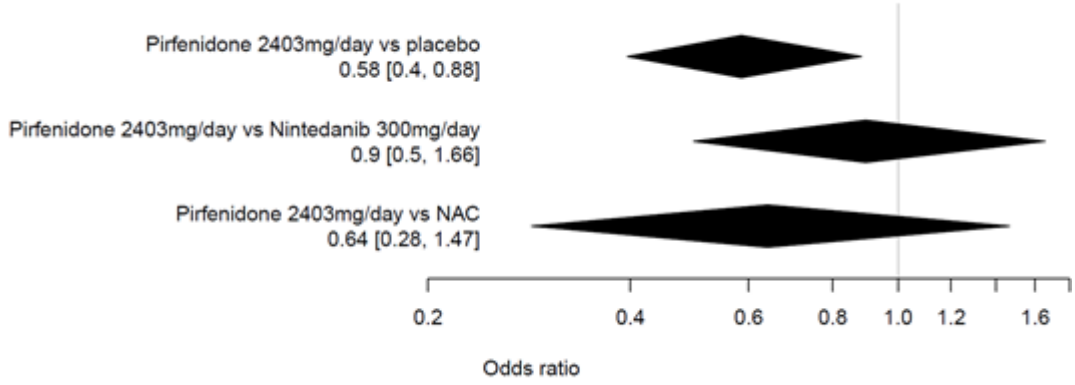


Table 43. Odds ratios and 95% credible intervals for FVC categorical decline of ≥10% percent predicted (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.58 (0.40, 0.88)	0.65 (0.42, 1.02)	0.91 (0.45, 1.87)
Pirfenidone 2403mg/day	1.71 (1.13, 2.53)		1.12 (0.60, 2.01)	1.56 (0.68, 3.51)
Nintedanib 300mg/day	1.53 (0.98, 2.40)	0.90 (0.50, 1.66)		1.40 (0.60, 3.27)
NAC	1.09 (0.53, 2.23)	0.64 (0.28, 1.47)	0.71 (0.31, 1.65)	

Odds ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.

All-cause mortality

- The principal all-cause mortality NMA uses data up to 52 weeks. A sensitivity

analysis uses data up to 72 weeks.

- The NMAs included HR data where available, however proportions could also be included in the analyses.
- The results of the principal NMA suggest that pirfenidone is superior to placebo. There is no evidence of a difference between nintedanib and placebo.

Results for all-cause mortality were available for each of the trials (CAPACITY 1 & 2, ASCEND, pooled INPULSIS, PANTHER [triple therapy and NAC], SP3, TOMORROW). For CAPACITY 1 and 2, results were available at both at Week 52 and 72.

For the CAPACITY and INPULSIS trials two definitions of survival were available; 'overall survival' and 'treatment emergent survival'. Overall survival is used in this analysis in line with an intention-to-treat (ITT) policy for analysis.

Assessment of statistical heterogeneity (HR at 52 weeks)

For the base case analysis, the comparison between placebo and pirfenidone was informed by three HRs (ASCEND, CAPACITY 1 & 2) (for SP3, only the numbers of deaths were available). The I-squared statistic was 0%, indicating minimal heterogeneity. For nintedanib, only a pooled HR was available for the INPULSIS trials and a proportion from TOMORROW and therefore statistical heterogeneity was not assessed.

Base case network (random effects model) (HR at 52 weeks)

The random effects model suggests that pirfenidone 2403mg/day is superior to placebo (HR: 0.52 [95% CrI: 0.30, 0.89] and triple therapy (HR: 0.06 [95% CrI: 0.01, 0.5]), see Figure 23 and Table 44.

Figure 23. Forest plot of hazard ratios for all-cause mortality at 52 weeks (base case network, RE model)

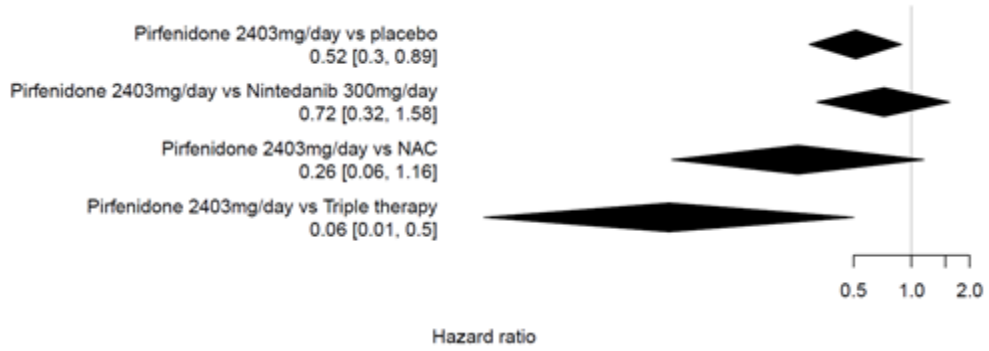


Table 44. Hazard ratios and 95% credible intervals for all-cause mortality at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.52 (0.30, 0.89)	0.71 (0.41, 1.27)	2.00 (0.49, 8.05)	9.25 (1.09, 78.38)
Pirfenidone 2403mg/day (*)	1.94 (1.13, 3.39)		1.39 (0.63, 3.10)	3.88 (0.86, 17.49)	17.99 (1.99, 163.84)
Nintedanib 300mg/day	1.40 (0.79, 2.46)	0.72 (0.32, 1.58)		2.79 (0.61, 12.62)	12.94 (1.41, 118.03)
NAC	0.50 (0.12, 2.03)	0.26 (0.06, 1.16)	0.36 (0.08, 1.63)		4.64 (0.36, 59.53)
Triple therapy	0.11 (0.01, 0.92)	0.06 (0.01, 0.50)	0.08 (0.01, 0.71)	0.22 (0.02, 2.75)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.

Assessment of statistical heterogeneity (HR at 72 weeks)

For the base case analysis, the comparison between placebo and pirfenidone was informed by three HRs (ASCEND, CAPACITY 1 & 2) (for SP3, only the numbers of deaths were available). The I-squared statistic was 0%, indicating minimal heterogeneity. For nintedanib, only a pooled HR was available for the INPULSIS trials and a proportion from TOMORROW and therefore statistical heterogeneity was not assessed.

Base case network (random effects model) (HR at 72 weeks)

For the base case network, the RE model shows a trend in favour of pirfenidone, with evidence of a difference between pirfenidone and triple therapy (HR: 0.07 [95%

CrI: 0.01, 0.6]). Overall, there is no evidence of a difference between placebo, pirfenidone, nintedanib and NAC, see Figure 24 and Table 45.

Figure 24. Forest plot of hazard ratios and 95% credible intervals for all-cause mortality at 72 weeks (base case network, RE model)

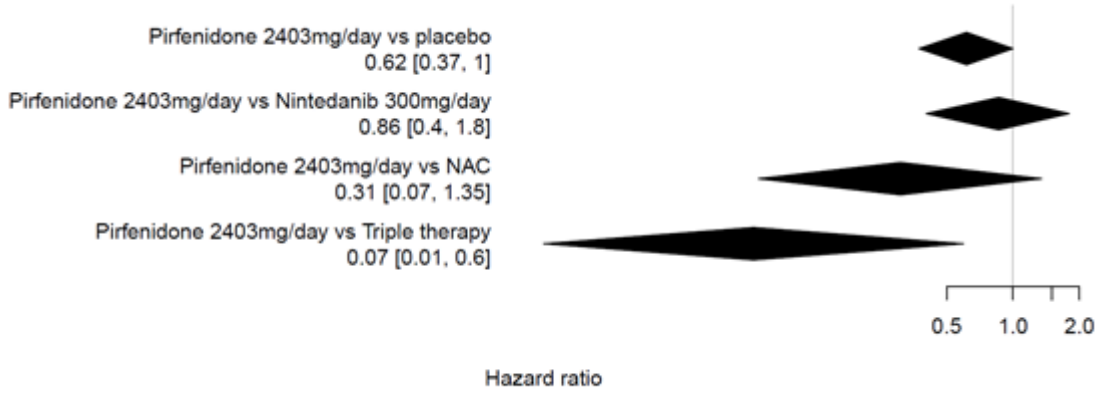


Table 45. Hazard ratios for all-cause mortality at 72 (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.62 (0.37, 1.00)	0.71 (0.41, 1.26)	1.99 (0.49, 8.05)	9.24 (1.09, 78.43)
Pirfenidone 2403mg/day (*)	1.62 (1.00, 2.67)		1.16 (0.55, 2.48)	3.24 (0.74, 14.30)	15.03 (1.67, 134.79)
Nintedanib 300mg/day	1.40 (0.79, 2.46)	0.86 (0.40, 1.80)		2.79 (0.62, 12.61)	12.91 (1.41, 118.09)
NAC	0.50 (0.12, 2.03)	0.31 (0.07, 1.35)	0.36 (0.08, 1.62)		4.63 (0.36, 59.44)
Triple therapy	0.11 (0.01, 0.92)	0.07 (0.01, 0.60)	0.08 (0.01, 0.71)	0.22 (0.02, 2.77)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.

IPF-related mortality

IPF-Related Mortality

- The principal all-cause mortality NMA uses data up to 52 weeks. A sensitivity analysis uses data up to 72 weeks.
- The NMAs included HR data where available, however proportions could also

be included in the analyses.

- The result of the principal NMA suggests that pirfenidone is superior to placebo. There is no evidence of a difference between nintedanib and placebo.

Results for IPF-related mortality were available for CAPACITY 1 and 2, ASCEND, INPULSIS, PANTHER and TOMORROW. For CAPACITY 1 and 2, results were available at both at Week 52 and 72.

The CAPACITY trials and ASCEND reported data for 'IPF-related mortality' and 'IPF-related treatment emergent deaths'. 'IPF-related mortality' is used in this analysis in line with an intention-to-treat (ITT) policy for analysis. In BI's submission, TE IPF-related mortality data from ASCEND and the CAPACITY trials were used.

INPULSIS, PANTHER and TOMORROW reported deaths from a respiratory cause. It is assumed that for the patients included in the studies, deaths from a respiratory cause will be IPF-related. These outcomes will be combined in the network meta-analyses.

The comparison between placebo and NAC is only informed by proportion data from PANTHER. Likewise, the comparison between placebo and triple therapy is also only informed by proportion data from PANTHER. Hence the HR estimates involving NAC and triple therapy are very uncertain.

Assessment of statistical heterogeneity (HR at 52 weeks)

The comparison of pirfenidone with placebo was informed by three HRs (ASCEND, CAPACITY 1 & 2). The I-squared statistic was 0%, indicating minimal heterogeneity. The comparison between nintedanib and placebo is informed by a pooled HR from the INPULSIS trials and a HR from TOMORROW. The I-squared statistic was 37.8%, indicating moderate heterogeneity.

Base case network (random effects model) (HR at 52 weeks)

For the base case analysis, the results of the random effects model are provided in Figure 25 and Table 46. The analysis suggests that pirfenidone is superior to

placebo and to triple therapy. There is no evidence of a difference between pirfenidone and nintedanib, or pirfenidone and NAC.

Figure 25. Forest plot of hazard ratios for IPF-related mortality at 52 weeks (base case network, RE model)

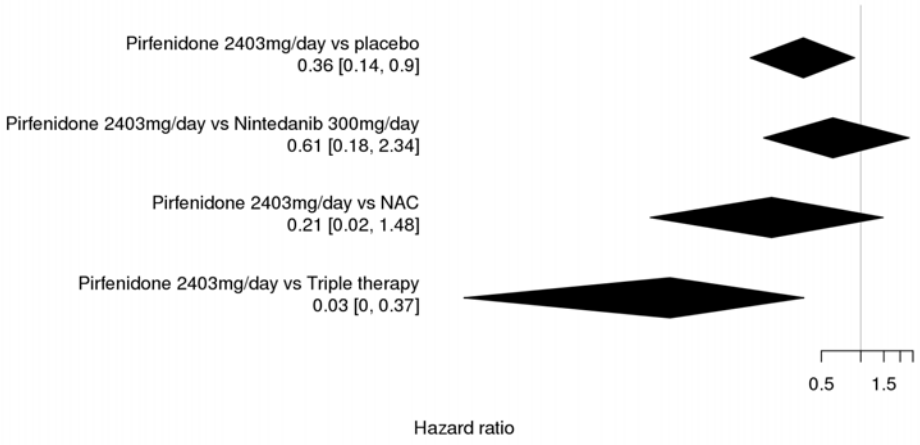


Table 46. Hazard ratios and 95% credible intervals for IPF-related mortality at 52 weeks (base case network, RE model)

	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.36 (0.14, 0.90)	0.60 (0.22, 1.33)	1.74 (0.30, 11.89)	10.26 (1.20, 342.72)
Pirfenidone 2403mg/day (*)	2.75 (1.11, 6.98)		1.63 (0.43, 5.52)	4.80 (0.68, 40.52)	28.65 (2.72, 1072.32)
Nintedanib 300mg/day	1.67 (0.75, 4.61)	0.61 (0.18, 2.34)		2.94 (0.44, 25.62)	17.62 (1.76, 677.29)
NAC	0.57 (0.08, 3.29)	0.21 (0.02, 1.48)	0.34 (0.04, 2.28)		6.11 (0.32, 290.44)
Triple therapy	0.10 (0.00, 0.84)	0.03 (0.00, 0.37)	0.06 (0.00, 0.57)	0.16 (0.00, 3.09)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.

Assessment of statistical heterogeneity (HR at 72 weeks)

For the base case analysis, the comparison between placebo and pirfenidone was informed by three HRs (ASCEND, CAPACITY 1 & 2). The I-squared statistic was 0%, indicating minimal heterogeneity. The comparison between nintedanib and placebo is informed by a pooled HR from the INPULSIS trials and a HR from TOMORROW. The I-squared statistic was 37.8%, indicating moderate heterogeneity.

Base case network (random effects model) (HR at 72 weeks)For the base case analysis, the results of the random effects model are provided in Figure 26 and Table 47. The analysis suggests that pirfenidone is superior to triple therapy. Despite a marked trend in favour of pirfenidone in the comparison with placebo, there is no conclusive evidence of a difference between pirfenidone and placebo, nintedanib, and NAC.

Figure 26. Forest plot of hazard ratios for IPF-related mortality at Week 72 (base case network, RE model)

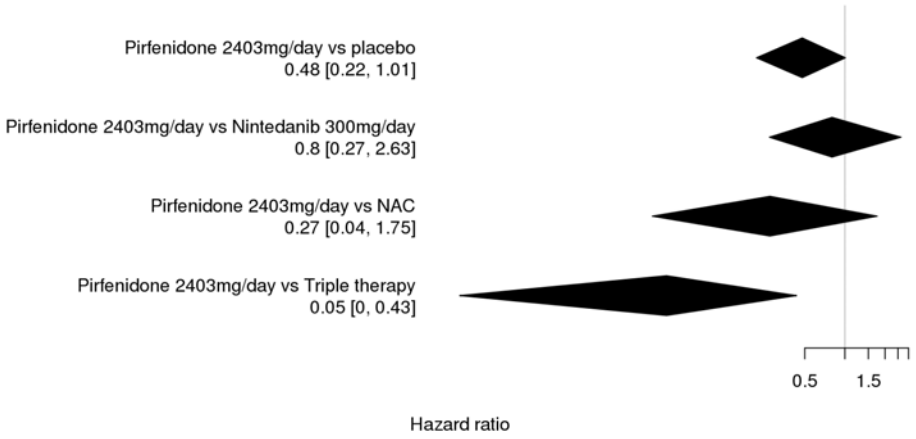


Table 47. Hazard ratios and 95% credible intervals for IPF-related mortality at Week 72 (base case network, RE model)

	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.48 (0.22, 1.01)	0.60 (0.23, 1.28)	1.74 (0.32, 11.24)	10.27 (1.25, 338.65)
Pirfenidone 2403mg/day (*)	2.09 (0.99, 4.60)		1.25 (0.38, 3.68)	3.65 (0.57, 27.79)	21.77 (2.31, 773.43)
Nintedanib 300mg/day	1.67 (0.78, 4.30)	0.80 (0.27, 2.63)		2.93 (0.46, 23.78)	17.54 (1.85, 652.10)
NAC	0.57 (0.09, 3.14)	0.27 (0.04, 1.75)	0.34 (0.04, 2.16)		6.12 (0.35, 279.63)
Triple therapy	0.10 (0.00, 0.80)	0.05 (0.00, 0.43)	0.06 (0.00, 0.54)	0.16 (0.00, 2.85)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.

Progression-free survival

- The principal all-cause mortality NMA uses data up to 52 weeks. A sensitivity analysis uses data up to 72 weeks.
- The NMAs included HR data where available, however proportions could also be included in the analyses.
- The results of the principal NMA suggest that pirfenidone is superior to placebo. There is no evidence of a difference between nintedanib and placebo.

Results for PFS were available for CAPACITY 1 and 2, ASCEND, SP3 and PANTHER. However, the definition of PFS varied between studies.To maintain similarity as far as possible, for CAPACITY 1 and 2, the PFS estimate based on the definition used in the ASCEND trial was included in the analysis. For the other definitions, it is assumed that they will lead to similar hazard ratios and odds ratios between a given pair of treatments, and thus that it is appropriate to combine them in an NMA. We believe this to be a reasonable assumption because in a comparison done between the CAPACITY and ASCEND trials, the replacement of DLco by 6MWD led to an increase in qualifying events without changing the HR estimate.

TOMORROW only reported the proportion of patients who progressed, rather than the proportion of patients who either progressed or died. It was unclear how many patients progressed before they died and therefore PFS cannot be calculated.

Assessment of statistical heterogeneity (HR at 52 weeks)

For the base case analysis, the comparison between pirfenidone and placebo was informed by four trials (ASCEND, CAPACITY 1 & 2, SP3). The I-squared statistic was 0%, indicating minimal heterogeneity. Statistical heterogeneity was not assessed for nintedanib versus placebo because the only data available was a pooled HR from the INPULSIS studies.

Base case network (random effects model) (HR at 52 weeks)

Overall, the results of the random effects models suggest that pirfenidone is superior to placebo. Despite a trend in favour of pirfenidone, there is no conclusive evidence

of a difference between pirfenidone and nintedanib, pirfenidone and NAC, or pirfenidone and triple therapy, see Figure 27 and Table 48.

Figure 27. Forest plot of hazard ratios for PFS at 52 (base case network, RE model)

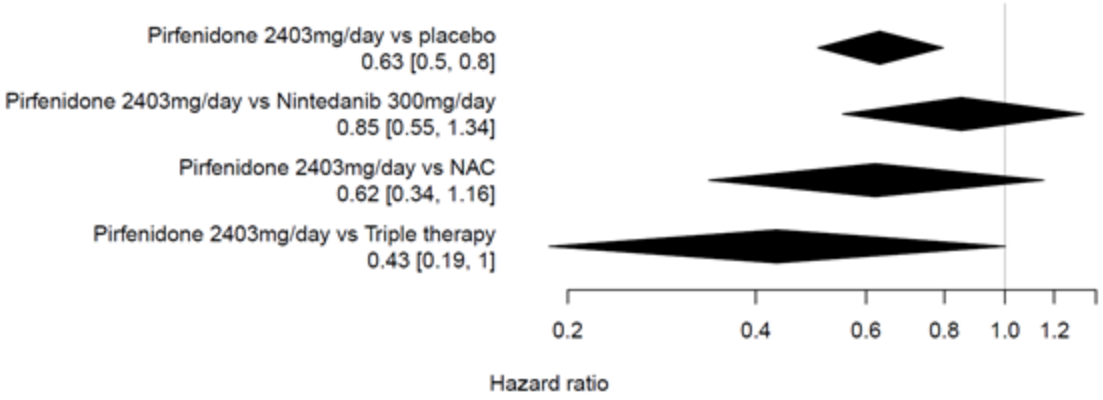


Table 48. Hazard ratios and 95% credible intervals for PFS at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.63 (0.50, 0.80)	0.74 (0.51, 1.08)	1.02 (0.57, 1.80)	1.46 (0.65, 3.28)
Pirfenidone 2403mg/day (*)	1.59 (1.25, 1.99)		1.17 (0.75, 1.82)	1.61 (0.87, 2.98)	2.32 (1.00, 5.35)
Nintedanib 300mg/day	1.35 (0.92, 1.98)	0.85 (0.55, 1.34)		1.37 (0.69, 2.72)	1.98 (0.81, 4.80)
NAC	0.98 (0.56, 1.75)	0.62 (0.34, 1.16)	0.73 (0.37, 1.44)		1.44 (0.53, 3.87)
Triple therapy	0.68 (0.31, 1.54)	0.43 (0.19, 1.00)	0.51 (0.21, 1.24)	0.70 (0.26, 1.87)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.

Assessment of statistical heterogeneity (HR at 72 weeks)

For the base case network, the comparison between placebo and pirfenidone was informed by four HRs (ASCEND, CAPACITY 1 & 2, SP3). The I-squared statistic was 0%, indicating minimal heterogeneity. Statistical heterogeneity was not assessed for nintedanib versus placebo because the only data available was a pooled HR from the INPULSIS studies.

Base case network (random effects model) (HR at 52 weeks)

The results suggest that pirfenidone is superior to placebo. Though trending in favour of pirfenidone, there is no conclusive evidence of a difference between

pirfenidone and nintedanib, pirfenidone and NAC, or pirfenidone and triple therapy, see Figure 28 and Table 49.

Figure 28. Forest plot of hazard ratios for PFS at 72 weeks (base case network, RE model)

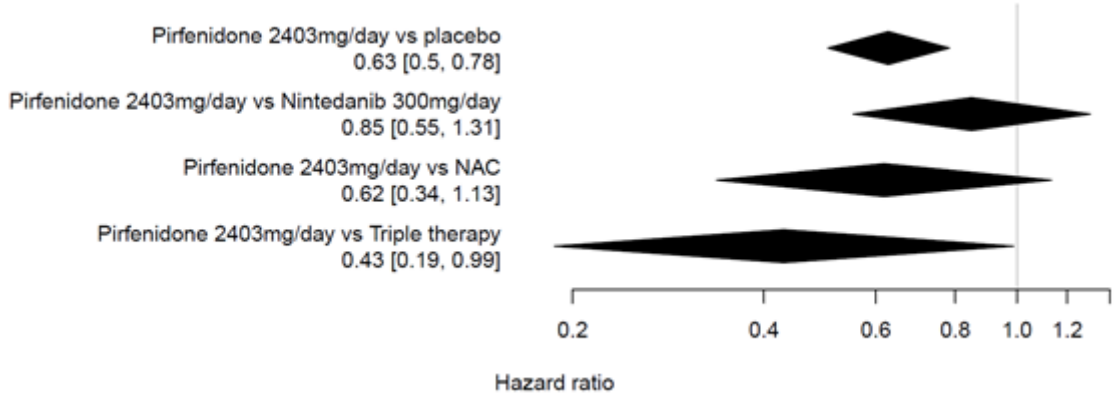


Table 49. Hazard ratios and 95% credible intervals for PFS at 72 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.63 (0.50, 0.78)	0.74 (0.51, 1.07)	1.02 (0.58, 1.79)	1.46 (0.65, 3.26)
Pirfenidone 2403mg/day (*)	1.59 (1.28, 1.98)		1.18 (0.77, 1.81)	1.62 (0.88, 2.97)	2.33 (1.01, 5.34)
Nintedanib 300mg/day	1.35 (0.93, 1.96)	0.85 (0.55, 1.31)		1.37 (0.70, 2.69)	1.97 (0.81, 4.77)
NAC	0.98 (0.56, 1.73)	0.62 (0.34, 1.13)	0.73 (0.37, 1.43)		1.44 (0.54, 3.85)
Triple therapy	0.68 (0.31, 1.53)	0.43 (0.19, 0.99)	0.51 (0.21, 1.23)	0.70 (0.26, 1.86)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.

Acute exacerbations

- The NMA suggests that there is no difference in acute exacerbations between any of the treatments.
- A critical limitation of this NMA is that the outcome was defined differently across the trials (see Section 4.7).

Eight trials (CAPACITY 1 & 2, ASCEND, SP3, INPULSIS 1 & 2, TOMORROW, and PANTHER) reported data for acute exacerbations which informed the NMA, however, the outcome was defined differently across the trials and was not collected systematically

To mitigate the differences in definitions, we reanalysed our IPD to match BI's definition, adjusted for different base case by meta-regression, and corrected actual data based on the baseline prevalence of AEs as an additional sensitivity analysis.

Assessment of statistical heterogeneity

For the comparison between placebo and pirfenidone, the I-squared statistic was 2.5%, indicating minimal heterogeneity. For the comparison between nintedanib and placebo the I-squared statistic was 64.3%, indicating high heterogeneity.

Base case network (random effects model)

For the base case network, the uncertainty in the comparison between pirfenidone and nintedanib was very large; OR: 1.14 (95% CrI: 0.41, 3.44). The odds ratio estimate of pirfenidone compared with placebo was 0.62 (95% CrI: 0.29, 1.39), and with NAC was 0.63 (95% CrI: 0.08, 5.33), see Figure 29 and Table 50.

Figure 29. Forest plot of odds ratios for the probability of acute exacerbations (base case network, RE model)

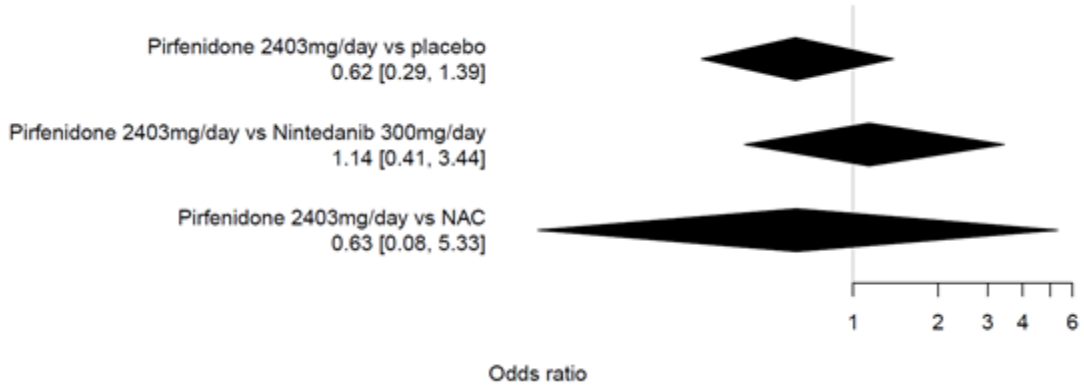


Table 50. Odds ratio estimates and 95% credible intervals for the probability of acute exacerbations (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.62 (0.29, 1.39)	0.55 (0.26, 1.09)	0.99 (0.14, 7.21)
Pirfenidone 2403mg/day (*)	1.60 (0.72, 3.45)		0.88 (0.29, 2.43)	1.59 (0.19, 13.16)
Nintedanib 300mg/day	1.82 (0.92, 3.82)	1.14 (0.41, 3.44)		1.82 (0.22, 15.16)
NAC	1.01 (0.14, 7.35)	0.63 (0.08, 5.33)	0.55 (0.07, 4.49)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Odds ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment (lower probability of acute exacerbations).

Physical Functioning

Change in 6MWD

- For change from baseline in 6MWD, the NMA suggests that pirfenidone is superior to placebo. There is no evidence of a difference between nintedanib and placebo, and there is no evidence of a difference between pirfenidone and placebo.

Five trials reported 6MWD (CAPACITY 1 and 2, ASCEND, TOMORROW and PANTHER).

A difference was noted between the NAC and pirfenidone studies in terms of follow-up schedules and methods used to handle missing data. For CAPACITY 1 and 2 and

ASCEND missing data was imputed using the smallest sum of squared differences (SSD) approach. For PANTHER missing data was not imputed, moreover a linear slope approach was used to model the decline of distance walked over time.

Assessment for statistical heterogeneity

For the base case analysis, the comparison between placebo and pirfenidone was informed by three trials (ASCEND, CAPACITY 1 & 2). The I-squared statistic was 0%, indicating minimal heterogeneity. For nintedanib, 6MWD data was only available from TOMORROW.

Base case network (random effects model)

For the base case analysis, the results of the random effects model are provided in Figure 30and Table 51. The results suggest that pirfenidone is superior to placebo. On average, over 52 weeks, patients receiving pirfenidone decline by 22.7 metres less than patients receiving placebo (95% CrI: 8.8, 36.3). There is no evidence of any differences between pirfenidone, nintedanib, NAC and triple therapy.

Figure 30. Forest plot of the mean difference in change from baseline in 6MWD (base case network, RE model)

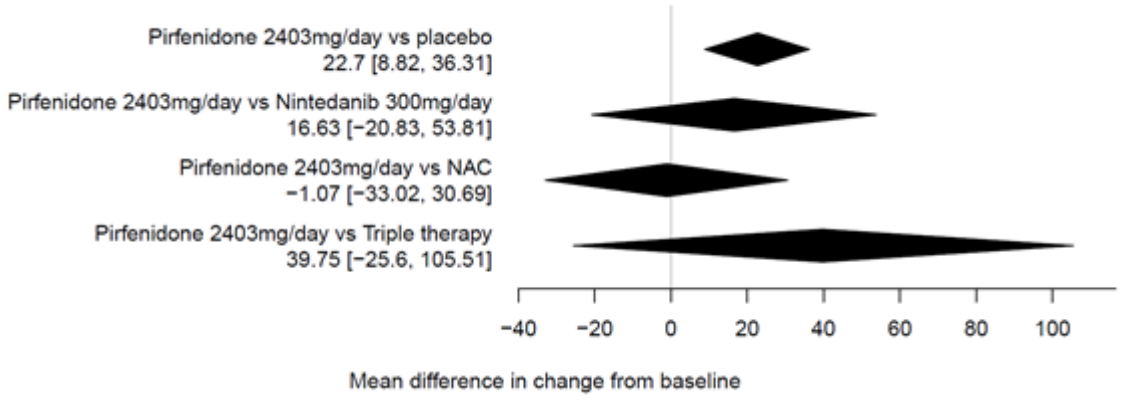


Table 51. Estimates and 95% credible intervals for mean difference in 6MWD (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		22.70 (8.82, 36.31)	6.00 (-28.25, 40.66)	23.80 (-4.79, 52.38)	-17.06 (-81.23, 46.80)
Pirfenidone 2403mg/day	-22.70 (-36.31, -8.82)		-16.63 (-53.81, 20.83)	1.07 (-30.69, 33.02)	-39.75 (-105.51, 25.60)
Nintedanib 300mg/day	-6.00 (-40.66, 28.25)	16.63 (-20.83, 53.81)		17.75 (-27.21, 62.24)	-23.11 (-95.79, 50.06)
NAC	-23.80 (-52.38, 4.79)	-1.07 (-33.02, 30.69)	-17.75 (-62.24, 27.21)		-40.92 (-110.70, 29.00)
Triple therapy	17.06 (-46.80, 81.23)	39.75 (-25.60, 105.51)	23.11 (-50.06, 95.79)	40.92 (-29.00, 110.70)	

Mean differences are calculated as column treatment minus row treatment, a positive result is favorable to the column treatment.

St. George Respiratory Questionnaire

- Based on the results of the NMA for change from baseline in SGRQ total score, there is no evidence of any differences between placebo, pirfenidone, nintedanib, NAC and triple therapy.

For the base case analysis, change from baseline in SGRQ total score was reported for CAPACITY 1 and 2, TOMORROW, INPULSIS 1 and 2 and PANTHER.

As per the other outcomes, a difference was noted between the studies in terms of follow-up schedules and methods used to handle missing data. For CAPACITY 1 and 2 missing data was imputed using the SSD approach. For TOMORROW missing data was imputed using an LOCF approach. For INPULSIS 1 and 2 and PANTHER missing data was not imputed.

Assessment of statistical heterogeneity

For the base case analysis, the comparison between placebo and pirfenidone was informed by two trials (CAPACITY 1 & 2). The I-squared statistic was 0%, indicating minimal heterogeneity. The comparison between nintedanib and placebo was

informed by three trials (INPULSIS 1 & 2, TOMORROW). The I-squared statistic was 64.7% indicating moderately high heterogeneity.

Base case network (random effects model)

For the base case analysis, the results of the random effects model are provided in Figure 31 and Table 52. There is no evidence of a difference between pirfenidone and placebo, nintedanib, NAC and triple therapy.

Figure 31. Forest plot of the mean difference in change from baseline in SGRQ score (base case network, RE model)

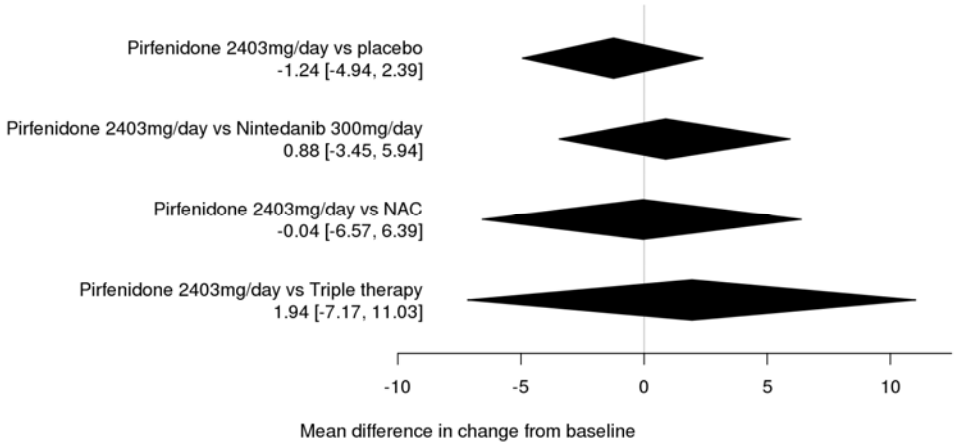


Table 52. Estimates and 95% credible intervals for the mean difference in SGRQ score (base case network, RE model)

	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		-1.24 (-4.94, 2.39)	-2.11 (-5.48, 0.37)	-1.19 (-6.52, 4.17)	-3.19 (-11.54, 5.15)
Pirfenidone 2403mg/day	1.24 (-2.39, 4.94)		-0.88 (-5.94, 3.45)	0.04 (-6.39, 6.57)	-1.94 (-11.03, 7.17)
Nintedanib 300mg/day	2.11 (-0.37, 5.48)	0.88 (-3.45, 5.94)		0.93 (-4.80, 7.41)	-1.03 (-9.69, 7.95)
NAC	1.19 (-4.17, 6.52)	-0.04 (-6.57, 6.39)	-0.93 (-7.41, 4.80)		-1.98 (-11.87, 7.90)
Triple therapy	3.19 (-5.15, 11.54)	1.94 (-7.17, 11.03)	1.03 (-7.95, 9.69)	1.98 (-7.90, 11.87)	

Mean differences are calculated as column treatment minus row treatment, a positive result is favourable to the row treatment.

UCSD Shortness of Breath Questionnaire

- The NMA for change from baseline in UCSD SOBQ suggests that

pirfenidone is superior to placebo. UCSD SOBQ data was not available for nintedanib.

CAPACITY 1 and 2 and ASCEND report change from baseline in the UCSD SOBQ total score for Week 48/52. The PANTHER trial (both triple therapy versus placebo and NAC versus placebo) reports change from baseline data at Week 60. UCSD SOBQ data was not reported for nintedanib.

As per percentage of predicted VC/FVC, a difference was noted between the studies in terms of the methods used to handle missing data. For each of CAPACITY 1 and 2 and ASCEND, missing data were imputed using the SSD method. For the PANTHER trial, a linear MMRM model was fitted to the change in SOBQ score data (Raghu 2012).

Assessment of statistical heterogeneity

For the comparison of pirfenidone with placebo, the I-squared statistic was 0%, indicating minimal heterogeneity. No UCSD SOBQ data is available for nintedanib.

Base case network (random effects model)

For the base case analysis, the results of the random effects model are provided in Figure 32 and Table 53. The results suggest that pirfenidone is superior to placebo. There is no evidence of a difference between pirfenidone and NAC, or between pirfenidone and triple therapy.

Figure 32. Forest plot of the mean difference in change from baseline in UCSD SOBQ total score (base case network, RE model)

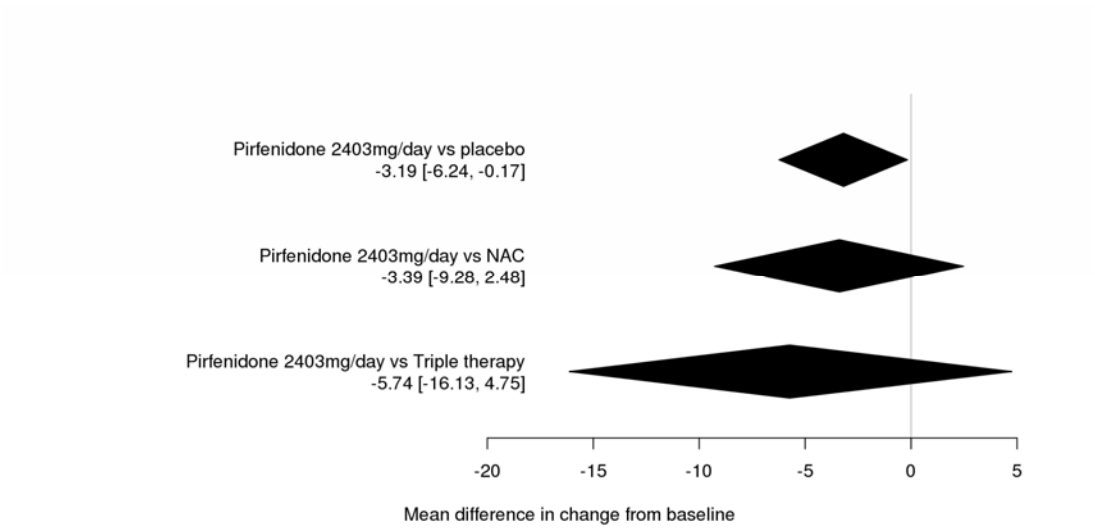


Table 53. Estimates and 95% credible intervals for the mean difference in change from baseline in UCSD SOBQ total score (base case network, RE model)

	Placebo	Pirfenidone 2403mg/day	NAC	Triple therapy
Placebo		-3.19 (-6.24, -0.17)	0.19 (-4.84, 5.25)	2.55 (-7.45, 12.46)
Pirfenidone 2403mg/day	3.19 (0.17, 6.24)		3.39 (-2.48, 9.28)	5.74 (-4.75, 16.13)
NAC	-0.19 (-5.25, 4.84)	-3.39 (-9.28, 2.48)		2.34 (-8.85, 13.53)
Triple therapy	-2.55 (-12.46, 7.45)	-5.74 (-16.13, 4.75)	-2.34 (-13.53, 8.85)	

* Mean differences are calculated as column treatment minus row treatment, a positive result is favourable to the row treatment.

All-cause Discontinuation of Treatment

- The NMA for all-cause discontinuation of treatment found no evidence of a difference between pirfenidone and placebo.
- The analysis suggests that, compared with placebo, nintedanib increases the odds of all-cause discontinuation of treatment.

All eight trials reported data on treatment discontinuations (SP3, CAPACITY 1 and 2, ASCEND, TOMORROW, INPULSIS 1 and 2 and PANTHER [NAC vs. placebo]).

For INPULSIS all-cause discontinuation of study was measured by the count of patients who “did not complete the planned observation time”. We assumed that this count included patients who did not complete the planned observation time due to death or lung transplantation.

Assessment of statistical heterogeneity

For both the comparison of pirfenidone with placebo, and the comparison of nintedanib with placebo, the I-squared statistic was 0%, indicating minimal heterogeneity.

Base case network (random effects model)

For all-cause treatment discontinuation rate, the base case random effects analysis suggests that there is no evidence of a difference between pirfenidone 2403mg/day

and placebo (OR:1.28, 95% CrI: (0.91, 1.78)), see Figure 33 and Table 54. The analysis suggests that, compared with placebo, nintedanib increases the odds of all-cause discontinuation of treatment (OR: 1.42, 95% CrI: (1.01, 2.01)). The odds ratio estimates suggest similar all-cause treatment discontinuation rates for pirfenidone, nintedanib and NAC.

Figure 33. Forest plot of odds ratios for the probability of all-cause discontinuation of treatment (base case network, RE model)

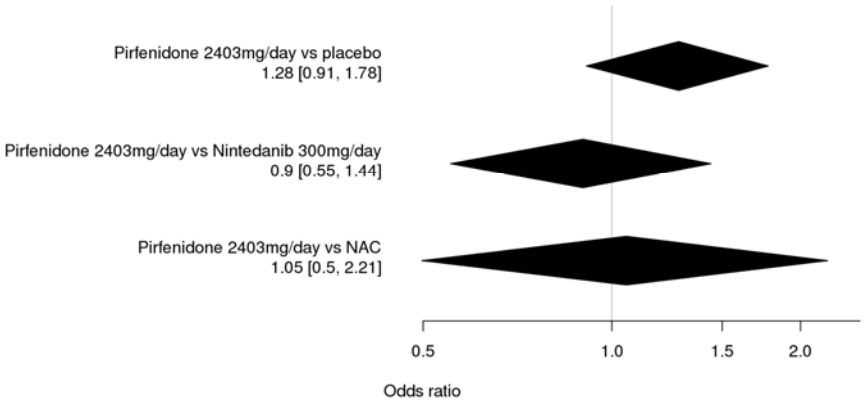


Table 54. Odds ratios and 95% credible intervals for the probability of all-cause discontinuation of treatment (base case network, RE model)

	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		1.28 (0.91, 1.78)	1.42 (1.01, 2.01)	1.21 (0.62, 2.36)
Pirfenidone 2403mg/day (*)	0.78 (0.56, 1.10)		1.11 (0.69, 1.81)	0.95 (0.45, 2.01)
Nintedanib 300mg/day	0.7 (0.50, 0.99)	0.90 (0.55, 1.44)		0.85 (0.40, 1.80)
NAC	0.83 (0.42, 1.60)	1.05 (0.50, 2.21)	1.17 (0.56, 2.48)	

* For SP3, pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. Odds ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment (lower probability of treatment discontinuation).

Limitations of the NMA

Due to the limited number of studies contributing to each network, a pragmatic approach was adopted, whereby trials were included regardless of minor differences in outcome definitions, timing of assessment and analysis methods. It was assumed that the differences in definitions and methods did not influence the relative

treatment effects. Where multiple sets of results were available for a single trial, we used the results from the method that was most consistently reported across trials.

The low number of studies in the networks also leads to uncertainty in the estimates, although sensitivity analyses were performed to assess variability in results across scenarios, and these reported comparable results to the base case results presented above.

Summary of NMA results

Overall, the NMA results provide evidence that pirfenidone is a more effective treatment compared to placebo in terms of all-cause mortality, IPF-related mortality, progression free survival (PFS), as well as the FVC outcomes, the physical functional outcome 6 minute walking test distance, and the health-related quality of life outcome UCSD Shortness of Breath Questionnaire (SOBQ). Despite consistent trends in favour of pirfenidone over nintedanib on mortality outcomes, there is no conclusive evidence from the NMA of a difference between pirfenidone and nintedanib for any outcomes considered in this report. This result is not unexpected, given the uncertainty in between pirfenidone and active treatments.

Sensitivity analyses conducted support the base case finding reported above: results of fixed and random effects models were consistent, and there was also consistency of results across the base case and restricted networks assessed.

The NMA results are summarised in Table 55, and reflect the data used to inform the cost-effectiveness model described in Section 5.

Table 55. Summary of main NMA results compared to pirfenidone (base case network, RE model)

Outcome	PFN better than placebo	PFN better than NAC	PFN better than NTB	*PFN better than triple therapy*
Lung Capacity				
Change in Percent Predicted FVC/VC	++	++	0	*NA*
Change in FVC/VC (Litres)	++	+	0	+
FVC decline of ≥10% Percent Predicted	++	+	0	*NA*
Physical Functioning				
Change in 6MWD	++	0	+	+
Change in Health-Related Quality of Life				
SGRQ	+	0	0	-
UCSD SOBQ	++	+	NA	+
Time to event outcomes				
PFS HR at 52 wks	++	+	+	++
PFS HR at 72 wks	++	+	+	++
All-Cause Mortality HR at 52 wks	++	+	+	++
All-Cause Mortality HR at 72 wks	+	+	0	++
IPF-Related Mortality HR at 52 wks	+	+	+	++
IPF-Related Mortality HR at 72 wks	+	+	+	++
Other				
Acute Exacerbations	+	+	0	*NA*
All-cause Discontinuation of Treatment	-	0	0	*NA*

NA : data not available for this comparison

+: Pirfenidone better than comparator; ++: Pirfenidone better than comparator (CrI do not cross 1 for hazard or odds ratios, or 0 for other); - : Comparator better than pirfenidone

--: Comparator better than pirfenidone (CrI do not cross 1 for hazard or odds ratios, or 0 for other); 0: no evidence of a difference or trend

4.11 Non-randomised and non-controlled evidence

- Non-RCT evidence supporting this submission is available through the RECAP study, an open-label extension of the ASCEND and CAPACITY trials, and international IPF registries
- RECAP (PIPF-012) is an on-going open-label extension of the ASCEND and CAPACITY trials designed to assess the long-term safety of pirfenidone. Data for overall survival and time-on-treatment are presented which describe the use of pirfenidone through to 8.8 years (latest datacut: June 2015)
- Registry evidence was gathered to provide information on long-term survival for patients receiving BSC
- Patient level data was available from 3 registries with patient follow-up for overall survival of between 5 and 15 years
- To improve comparability a two-step process was conducted to sub-set data to match CAPACITY and ASCEND inclusion / exclusion criteria, followed by propensity scoring (trimming and reweighting) to adjust for remaining differences in patient characteristics
- The final propensity score model was used to estimate the comparative effectiveness of pirfenidone to BSC using the real-world data
- The comparative effectiveness estimated across the 3 registries was comparable to that observed in the pooled ASCEND / CAPACITY data. Results are presented in order of comparability of datasets following trimming with the ASCEND / CAPACITY trial population: INOVA: HR 0.47 (95% CI: 0.38, 0.61); EuroIPF: HR 0.34 (95% CI: 0.18; 0.63), and; Edinburgh: HR 0.29 (95% CI: 0.22; 0.40)
- Supportive evidence from alternative registries indicated similar median survival for BSC (median 3.4 – 4.4 years across sources)
- Whilst there are limitations to comparing data from a Phase III trial to real-world evidence this analysis indicates that the comparative benefit observed from pirfenidone in the Phase III RCTs is likely to extend to the long-term

RECAP study (Costabel, 2014)

RECAP (PIPF-012) is an open-label extension of the ASCEND and CAPACITY trials. The study was designed to assess the long-term safety of pirfenidone 2403 mg/day in patients with IPF who received ≥80% of scheduled doses, and completed the Week 72 final study visit in CAPACITY 1 or CAPACITY 2 (Costabel 2014). Patients in the ASCEND study were also eligible to roll-over into RECAP, although no published data analysis including ASCEND is available to date (Kreuter 2014, Roche 2016a).

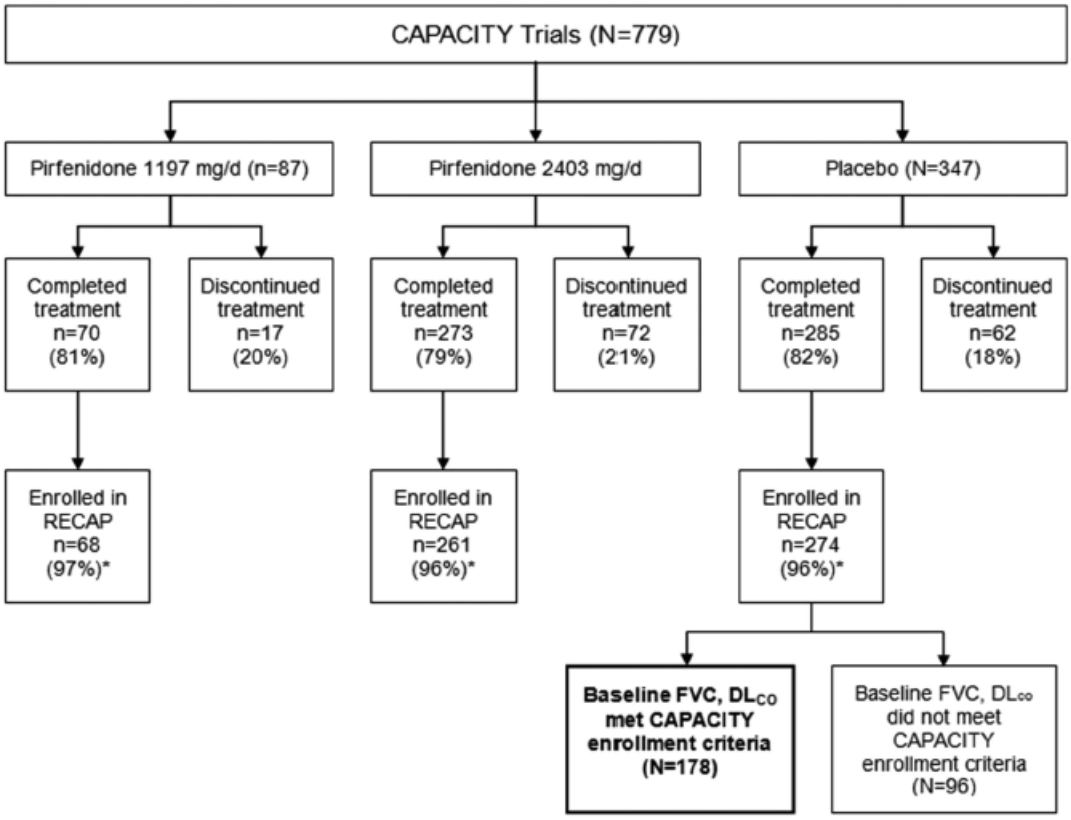
Study design: An overview of the study design and key eligibility criteria are shown in Table 56.

Table 56. Summary of RECAP study design

	RECAP (PIFP-012) (Costebel, 2014; Kreuter 2014)
Study design	Open-label, uncontrolled, Phase III extension study in which eligible patients receive treatment with pirfenidone 2403 mg/day
Intervention	Eligible patients received pirfenidone 2403 mg/day Concomitant therapy with corticosteroids, azathioprine, cyclophosphamide, and/or NAC were permitted if judged appropriate by investigator
Population	IPF patients that completed ASCEND, or CAPACITY 1 & 2 studies
Objectives	**Primary objective:** To examine the long-term safety and tolerability of pirfenidone in patients with IPF who were previously randomised to the placebo group in either CAPACITY 1 or 2 studies (later adjusted to allow enrolment from the ASCEND trial, Kreuter 2014) **Secondary objective:** To obtain additional efficacy data for pirfenidone 2403 mg/day in patients with IPF
Inclusion/Exclusion criteria	**Inclusion criteria:** • Completes the ASCEND or CAPACITY studies final visit • In the opinion of the principal investigator has been generally compliant (received ≥80% of scheduled doses) with study requirements during the qualifying study, or must be considered eligible to enrol in RECAP by the InterMune medical monitor • Is able to provide informed consent and comply with the requirements of the study **Exclusion criteria:** • Pregnant or lactating women • In the opinion of the PI, is not a suitable candidate for study participation • Known hypersensitivity to any of the components of the study drug • Participates in another interventional clinical trial between the end of participation in ASCEND or either CAPACITY studies and time of enrolment in RECAP • Receives concomitant medications defined in the protocol • Permanently discontinues study drug during the ASCEND or CAPACITY studies for any reason

Study population: At the time the interim data-cut, April 2010, a total of 178 patients were recruited for enrolment from 88 sites in North America, Europe and Australia. To facilitate comparison with CAPACITY outcomes, analyses were based on patients newly treated with pirfenidone in RECAP who had baseline FVC and DLco values that met ASCEND or CAPACITY entry criteria (Section 4.3) (Costabel 2014).

Figure 34. RECAP trial profile**



*Percentage of CAPACITY completers who elected to enroll in RECAP

** Since 2014, patients completing the ASCEND study have been eligible to enter into the RECAP study

Results: Safety results of the interim analysis can be found in section 4.12.

As an on-going open-label extension of three clinical trials, the RECAP study is not yet complete. The most recent datacut was performed in June 2015, with analyses based on summary data from this datacut presented at a recent congress (Fisher 2015). This includes survival data for patients continuing treatment with pirfenidone, following receipt of pirfenidone in CAPACITY/ASCEND, with patient data available through to 8.8 years (Figure 35) (Roche 2016a). Time on treatment data for patients

continuing treatment with pirfenidone from the latest datacut of RECAP are also presented in Figure 36 (Roche 2016a). The next datacut is planned in June 2016.

Figure 35. RECAP KM estimates for OS: patients continuing on pirfenidone 2403mg/d (data cut: June 2015)

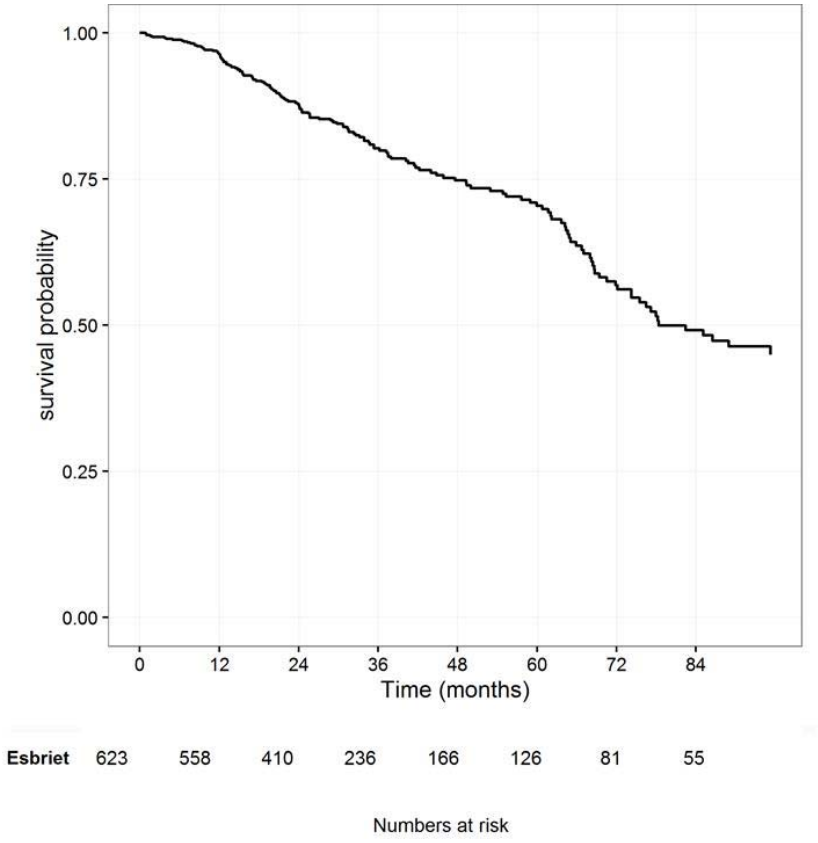
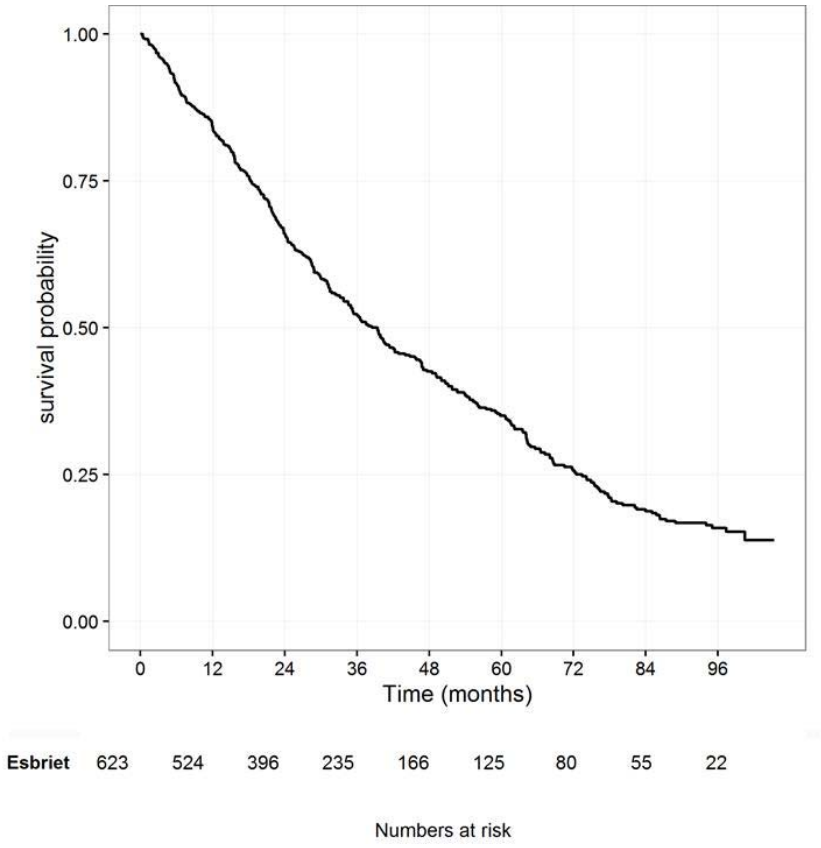


Figure 36. RECAP KM estimates for time on treatment: patients continuing on pirfenidone 2403mg/d (data cut: June 2015)



IPF registry data

Given the lack of long-term data available for BSC from the Phase III RCTs, evidence was gathered to provide similar longevity of information for outcomes expected for patients receiving BSC in real-world practice to the information available for pirfenidone. The authors of a recent review of observational studies performed for IPF treatments concludes: "*the profile of these patients seems to be quite similar all over the world, as does their clinical management*", which gives reassurance on the appropriateness of using these registries (Harari 2015).

The holders of various registries reporting outcomes for patients with IPF in real-world practice were contacted, resulting in the availability of patient level information from three registries collecting information on patients with IPF:

- Edinburgh registry
- INOVA registry

- EuroIPF registry

Patient level data were available for detailed baseline characteristics for all three of these registries (Table 58). To improve the comparability of the patients between the trial information available for pirfenidone and the data available from the registries, a two-stage process was conducted:

1. Apply inclusion/exclusion criteria from the pirfenidone RCTs to select similar patients

2. Define a propensity score model that calculates the probability of being included in a clinical trial based on baseline characteristics and exclude patients with unusual profiles based upon propensity-score based trimming

The propensity score is the probability of assignment to one cohort (e.g. a clinical trial population or the Edinburgh registry), conditional on observed baseline characteristics. A key advantage of propensity score adjustment is that a large number of covariates can be adjusted for, even when studying infrequent outcomes.

A separate logistic model to derive the propensity score for each patient from ASCEND/CAPACITY and each registry included in the analysis was derived. Cohort assignment (i.e. patient is included in clinical trial or not) was regressed on observed baseline characteristics. The estimated propensity score was the predicted probability of cohort assignment in the trial derived from the fitted regression model.

A logistic model was derived using backward stepwise regression with default threshold parameters (0.05). The kernel density distributions for each of the logistic models prior to trimming for each of the three registries are shown in Appendix 16.

The distribution of propensity scores across cohorts derived from this first logistic model was inspected, and trimming was applied in order to restrict the analysis to observations within a propensity score range that was common to all cohorts—that is, by excluding patients in the non-overlapping parts of the propensity score distribution. The cut point for trimming was the lower 2.5th percentile in the trial group. After trimming (i.e. exclusion of all patients with a propensity score below the lower 2.5th percentile), the same logistic model was fitted to the data to derive the

new propensity scores. The Kaplan–Meier curves for each of the patient registries post trimming are shown in Appendix 16.

The kernel density distributions for each of the logistic models post trimming for each of three registries are also shown in Appendix 16. It is clear that the INOVA and EuroIPF registries provide the most comparable patient sample to the patients in the pirfenidone Phase III RCTs.

The final propensity score model was used to estimate the comparative effectiveness of pirfenidone to BSC using the real-world data (Table 57). The comparative effectiveness estimated across the 3 registries was comparable / better to the comparative effectiveness observed in the pooled ASCEND / CAPACITY data. Results were similar comparing the pooled hazard ratio vs BSC from ASCEND / CAPACITY and INOVA which represented the study with the largest sample size and most similar patient characteristics post trimming (HR 0.52 vs [redacted]; Roche 2016a). Whilst there are limitations to comparing data from a Phase III trial to real-world evidence this analysis indicates that the comparative benefit observed from pirfenidone in the Phase III RCTs is likely to extend to the long-term.

Table 57. Overall survival comparison: pirfenidone versus registry data

Outcome	Edinburgh registry	INOVA registry	EuroIPF registry	Pooled CAPACITY and ASCEND data
Hazard ratio for pirfenidone vs BSC (post trimming unadjusted data)	[redacted]	[redacted]	[redacted]	0.64 (0.41;0.99) at 72 weeks
Hazard ratio for pirfenidone vs BSC (post trimming data using propensity score model to adjust for remaining imbalances)	[redacted]	[redacted]	[redacted]	0.52 (0.31; 0.88) at 52 weeks
Key: BSC, best supportive care.				

Table 58. Summary of available registries for best supportive care, registries with patient level demographic data

	Edinburgh	INOVA	EuroIPF
Geographic Region	UK	USA	Europe
Dates of registry information	1 January 2001 – 30 May 2014	November 1996 - June 2015	2008 - 2011
Patient population	• Incident IPF cases with a definite or possible UIP pattern on HRCT based on the 2011 ATS/ERS diagnostic guidelines for IPF • Event time available • Patients diagnosed up to 48 months prior to data collection date	Confirmed as incident IPF cases based on the 2011 ATS/ERS/JRS/ALAT diagnostic guidelines for IPF.	Verified diagnosis of IPF
n	323	815	409
Follow-up	Patients were followed from index date (date of IPF diagnosis) to date of death or May 30, 2014. Vital status was ascertained on May 30, 2014. Patients were censored on May 30, 2014, if their death could not be confirmed. None of the patients seen at this center underwent lung transplantation during the follow-up period, so this was not included as a censoring criterion for this cohort.	Patients were followed from index date (date of IPF diagnosis) to date of death or date of last visit. Date of last vital status is provided in the dataset. Patients were censored on their date of last visit, if their death could not be confirmed. If patients had a transplant, it was indicated in the dataset, but no dates were provided for treatment or transplant.	Patients were followed from index date (date of inclusion in registry) to date of death or date of last visit. Date of last visit and vital status check was provided. Patients were censored on date of last visit, if their death could not be confirmed
Treatments received during follow-up	BSC only	BSC only	BSC only
Inclusion/exclusion criteria applied to match ASCEND/CAPACITY	• FVC/VC<90% and/or DLco<90% • FVC/VC>50% • DLco >30% • FEV1/FVC>0.7 • Age 40 - 80 • Gender known	• FVC/VC<90% and/or DLco <90% • FVC/VC>50% • DLco >30% • FEV1/FVC>0.7 • Age 40 - 80 • Gender known	• FVC/VC<90% and/or DLco <90% • FVC/VC>50% • DLco >30% • FEV1/FVC>0.7 • Age 40 - 80 • Gender known

	Edinburgh	INOVA	EuroIPF
		• Event time available	• Event time available
Number of patients following application of ASCEND/ CAPACITY inclusion/ exclusion criteria	182	286	115
Parameters included in the propensity score model	• Age • Sex • Baseline %predicted FVC • Baseline %predicted DLco • First order interaction terms	• Age • Sex • Baseline %predicted FVC • Baseline %predicted DLco • Baseline FEV/FVC • First order interaction terms	• Age • Sex • Baseline %predicted FVC • Baseline %predicted DLco • Baseline FEV/FVC • Baseline smoking status • First order interaction terms
Number of patients remaining after trimming	125	254	89
Age, mean years ± SD	69.4 ± 7.6	66.2 ± 7.9	66.3 ± 8.4
Male (%)	72%	80%	85%
FVC ± SD	81.2 ± 12.4	70.9 ± 12.8	75.4 ± 14.3
DLco **± SD**	51.6 ± 11.8	46.5 ± 11.1	46.0 ± 10.6
FEV1/FVC ± SD	0.83 ± 0.07	0.83 ± 0.06	0.83 ± 0.07
Propensity score model	logOdds(Trial=1) = Age + Sex + DLco + FVC + Age* DLco + Age*FVC	logOdds(Trial=1) = Age + Sex + DLco + FVC +FEV/FVC + Age* DLco + Sex*FEV/FVC	logOdds(Trial=1) = Age + Sex + DLco + FVC +FEV/FVC + Smoke + Age*FVC + Age*Sex + Age*FEV/FVC + Sex*FVC + Sex*Smoke + DLco *Smoke

Key: DLco, carbon monoxide diffusing capacity of the lungs; FEV, forced expiratory volume; FVC, forced vital capacity; HRCT, high-resolution computed tomography; IPF, idiopathic pulmonary fibrosis; SD, standard deviation; UIP, usual interstitial pneumonia.

In addition to the registries where patient level data were available, three additional sources of supportive information were identified:

1. CPRD data (n=4,527) were obtained from 2000 to 2012 (inclusive), before pirfenidone was available in the UK (Roskell 2014). Patients were selected based on the following criteria:

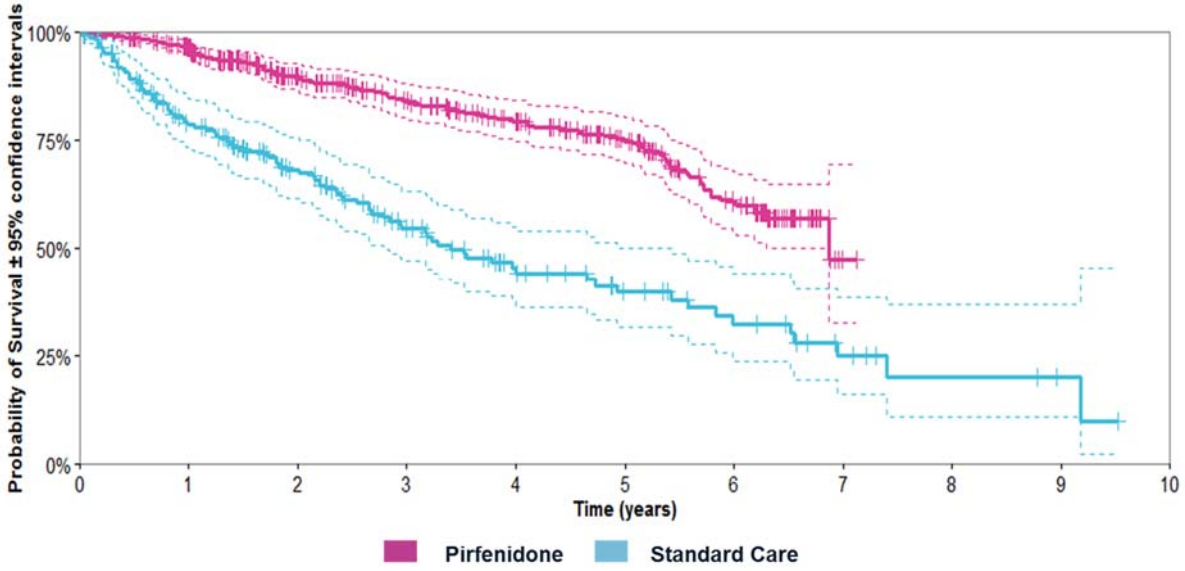
    - A clinical or referral event record for IPF as defined by Read (general practices coding system in the UK) as specified in Navaratnam (Navaratnam 2011).
    - No clinical or referral codes for connective tissue disease, extrinsic allergic alveolitis, sarcoidosis, pneumoconiosis, or asbestosis at any time in the patient record
    - IPF events whilst alive and registered at an up-to-standard general practice
    - At least 1 year of registration prior to the index date (date of IPF record)

    To improve the similarity between the CAPACITY and CPRD cohorts, the following restrictions were applied to the CPRD data:

    - Survival times were adjusted using random-sampling of diagnosis to randomisation collected in the CAPACITY studies (n=2,888)
    - Patients with an FVC<50% were excluded, this was determined based on data within 1 month of the patient's index date (n=193)

    Full propensity scoring was not possible as only FVC data were available for patients within the CPRD dataset. Standard care patients were followed up to 9.53 years; a median survival of 3.41 years was observed (95% CI: 2.67, 4.93).

Figure 37. Comparison of standard care to pirfenidone using CPRD data



2. Strand et al. (2015) report overall survival for patients prospectively enrolled from the National Jewish Health Institutional Review Board-approved ILD database for patients between January 1, 1985 and January 1, 2011 diagnosed with IPF according to consensus guidelines. Median survival was 4.4 years (95% CI: 4.1-5.2) for IPF.

3. Kondoh et al. (2010) retrospectively studied patients diagnosed with IPF based on ATS/ERS criteria (Kondoh 2010). Median survival was 3.7 years. A stepwise multivariate Cox regression model demonstrated the prognostic significance of FVC progression (10% decline in FVC at 6 months), acute exacerbations, BMI and disease severity measured via the modified MRC scale.

Table 59 provides a summary of the characteristics of the patients contained within the three additional supportive registries and the patients in the CAPACITY / ASCEND trials. Patients within the Strand registry appear most similar to those in CAPACITY / ASCEND.

Table 59. Summary of available registries for best supportive care, registries without patient level demographic data

	CPRD	Strand	Kondoh	CAPACITY / ASCEND
Geographic Region	UK	USA	Japan	Global
Data collection dates	2000 - 2012	Jan 1985 – Jan 2011	Jan 2000 - Dec 2005	
Patient population	ICD10 codes: H563.00 H563.11 H563.12 H563100 H563z00	Subgroup diagnosed with IPF according to consensus guidelines including ATS/ERS	Patients diagnosed with IPF based on ATS/ERS criteria	Diagnosis of IPF in accordance with the ATS international consensus statement
n	193 in FVC reported and ≥50 subgroup	321	74	623 on high dose pirfenidone arms
Age, mean years ± SD	73.5 ± 9.2	66.1 ± 9.1	64.1 ± 7.4	67.2 ± 7.6
Male (%)	68%	75%	82%	74%
FVC ± SD	79.3 ± 15.7	71.4 ± 17.4	77.0 ± 19.2	67.8 ± 11.2
DLco ± SD	NR	52.3 ± 18.7	59.3 ± 18.7	47.1 ± 9.7

Key: DLco, carbon monoxide diffusing capacity of the lungs; FEV, forced expiratory volume; FVC, forced vital capacity; HRCT, high-resolution computed tomography; IPF, idiopathic pulmonary fibrosis; SD, standard deviation; UIP, usual interstitial pneumonia.

4.12 Adverse reactions

- Overall, safety results from the pirfenidone clinical study programme showed that the treatment was well-tolerated with a manageable side effect profile.
- Gastrointestinal and skin-related events were more common in the pirfenidone group compared to placebo, but rarely led to treatment discontinuation. There are also a number of reported patients with serum transaminase elevations but these tend to be reversible without long term sequelae.The findings of an analysis of all avalible data (3160 person exposure years; median duration of exposure of 1.7 years) were consistent with prior observations (Lancaster 2016)
- An assessment of pirfenidone's tolerability profile in real-world settings is considered, the authors of a recent study concluded "*pirfenidone is well tolerated, and the most common adverse events are gastrointestinal, skin-related events and weight loss*" (Harari and Caminati, 2015).
- Pirfenidone has a different tolerability and adverse event profile compared to nintedanib. The most frequently reported adverse reactions associated with nintedanib are diarrhoea, nausea and vomiting, abdominal pain, decreased appetite, weight loss and elevation of hepatic enzymes (EMC 2015b). A recent meeting of the Pharmacovigilance Risk Assessment Committee (PRAC) also recommended the product label should be updated to include cases of haemorrhage and epistaxis (EMA 2015b). Pirfenidone is the only approved IPF treatment without a special warning or precaution for use in patients at risk of cardiovascular disease or bleeding (EMC 2015a).

Evidence from pirfenidone clinical trial programme

ASCEND study (King 2014)

The most commonly reported treatment-emergent AEs are summarised in Table 60. The most common AEs with higher incidence in the pirfenidone group were primarily gastrointestinal and skin-related events. These AEs were generally mild to moderate in severity, manageable, reversible, and without clinical significance. Elevations in alanine or aspartate aminotransferase levels (≥3 x ULN) occurred in 2.9% (n=8) patients receiving pirfnidone compared with 0.7% (n=2) patients receiving placebo.

Table 60. Adverse events in ≥15% of patients in either treatment group in ASCEND

Adverse event, n (%)	Pirfenidone 2403 mg/day (n=278)	Placebo (n=277)
Nausea	100 (36)	37 (13.4)
Rash	78 (28.1)	24 (8.7)
Headache	72 (25.9)	64 (23.1)
Cough	70 (25.2)	82 (29.6)
Diarrhoea	62 (22.3)	60 (21.7)
Upper respiratory tract infection	61 (21.9)	56 (20.2)
Fatigue	58 (20.9)	48 (17.3)
Dizziness	49 (17.6)	36 (13)
Dyspepsia	49 (17.6)	17 (6.1)
Anorexia	44 (15.8)	18 (6.5)
Dyspnoea	41 (14.7)	49 (17.7)
Worsening of IPF	26 (9.4)	50 (18.1)

There were 55 patients (19.8%) and 69 patients (24.9%) in pirfenidone and placebo groups, respectively, who experienced a serious adverse event. The most common serious AE was worsening of IPF which was reported in 7 patients (2.5%) in the pirfenidone group, and 27 patients (9.7%) in the placebo group.

The proportion of patients discontinuing treatment due to an AE was 14.4% (n=40) in the pirfenidone group and 10.8% (n=30) in the placebo group. The most common AE leading to treatment discontinuation was worsening IPF (1.1% [n=3] in the pirfenidone group vs. 5.4% [n=15] in the placebo group). The only other AEs leading to discontinuation of treatment in ≥1% of patients in the pirfenidone group were elevated hepatic enzymes levels, pneumonia, rash and decreased weight in 3 patients (1.1%) each.

There was no significant difference in the rates of death from any cause in the pirfenidone group vs placebo (4% [n=11] vs 7.2% [n=20]; HR: 0.55 [95% CI: 0.26-1.15] p=0.10), or from rates of death from IPF (1.1% [n=3] vs 2.5% [n=7]; HR: 0.44 [95% CI: 0.11-1.72] p=0.23).

CAPACITY 1 & 2 studies (Noble 2011)

The most common treatment-emergent AEs in the pooled CAPACITY 1 & 2 population are summarised in Table 61. The most commonly reported AEs in the pooled pirfenidone 2403 mg/day group, with at least 1.5 times increased incidence relative to placebo, were gastrointestinal events, skin-related events, and dizziness. A dose response in frequency was observed, and these AEs were generally mild or moderate in severity and did not result in clinically significant consequences. Stevens-Johnson syndrome or toxic epidermal necrosis were not reported.

Table 61. Treatment-emergent adverse events in ≥10% of patients from CAPACITY 1 & 2*

Adverse event, n (%)	Pirfenidone 2403 mg/day (n=345)	Placebo (n= 347)
Nausea	125 (36)	60 (17)
Rash	111 (32)	40 (12)
Dyspepsia	66 (19)	26 (7)
Dizziness	63 (18)	35 (10)
Vomiting	47 (14)	15 (4)
Photosensitivity reaction	42 (12)	6 (2)
Anorexia	37 (11)	13 (4)
Arthralgia	36 (10)	24 (7)
Insomnia	34 (10)	23 (7)
Abdominal distension	33 (10)	20 (6)
*Occurring in ≥10% of patients give pirfenidone 2403 mg/day and with an incidence of 1.5 x greater than that in patients receiving placebo		

Study treatment was discontinued due to AEs in 15% (n=51) of 345 patients in the pooled pirfenidone 2403 mg/day group vs. 9% (n=30) of 347 patients in the placebo group. The most common AE leading to discontinuation was worsening of IPF (3% in both groups). Substantial laboratory abnormalities (Grade 4 or a shift of 3 grades e.g. from 0 to 3) occurring more frequently in the pooled pirfenidone 2403 mg/day group vs placebo, were hyperglycemia (1% [n=4] vs <1% [n=3], respectively), hyponatraemia (1% [n=5] vs 0%), hypophosphatemia (2% [n=6] vs <1% [n=3]), and

lymphopenia (1% [n=5] vs 0); however, none were associated with clinically significant consequences. More patients in the pooled pirfenidone-treated group than in the pooled placebo group had elevations in alanine aminotransferase and aspartate aminotransferase of more than 3x the upper limit of normal (4% [n=14] vs. <1% [n=2]). However, all reports were reversible and without clinical sequelae.

SP3 study (Taniguchi 2010)

Photosensitivity, anorexia, dizziness, and elevated gamma-GTP were reported more frequently in patients treated with pirfenidone 1,800 mg/day compared with placebo, while respiratory infections were more common in patients treated with placebo.

Table 62. Adverse events with an incidence of ≥5% during the SP3 study

Adverse event, n (%)	High dose	Low dose	Placebo	p-value[#]		
				High dose vs placebo	Low dose vs placebo	High dose vs low dose
Subjects	109	55	107			
Any adverse event	109 (100.0)	54 (98.2)	106 (99.1)	0.50	1.00	0.34
Photosensitivity	56 (51.4)	29 (52.7)	24 (22.4)	<0.01	<0.01	1.00
Eszema asteatotic	0 (0.0)	3 (5.5)	0 (0.0)		0.04	0.04
Anorexia	18 (16.5)	6 (10.9)	3 (2.8)	<0.01	0.06	0.48
Abdomial discomfort	3 (2.8)	4 (3.7)	0 (0.0)	0.25	0.01	0.23
Dizziness	8 (7.3)	0 (0.0)	1 (0.9)	0.04	1.00	0.05
Nasopharyngitis	54 (49.5)	30 (54.5)	70 (65.4)	0.02	0.23	0.62
Upper respiratory tract infection	1 (0.9)	3 (5.5)	9 (8.4)	<0.01	0.75	0.11
γ-GTP elevation	25 (22.9)	12 (21.8)	10 (9.3)	<0.01	0.05	1.00
WBC decrease	4 (3.7)	3 (5.5)	0 (0.0)	0.12	0.04	0.69

Events which were observed with an incidence of ≥ 5% during the study period and for which a significant difference was detected between the placebo group and each of the pirfenidone treatment groups, high dose or low dose (p<0.05).
[#]Using Fisher's exact test.
γ-GTP, gamma glutamyl-transpeptidase; WBC, white blood cell

SP2 study (Azuma 2005)

Skin photosensitivity was the major adverse event for discontinuing or reducing pirfenidone dose. Adverse events occurring in >20% of patients treated with pirfenidone 1800 mg/day include photosensitivity (43.8%), stomach discomfort (30.1%), anorexia (31.5%), elevation of gamma-guanosine triphosphate (27.4%), drowsiness (23.3%), nausea (21.9%), fatigue (21.9%), and elevation of C-reactive

protein (20.5%). Most of the adverse events disappeared with decrease of the dose or temporarily holding the medication.

Table 63. Adverse events with an incidence of ≥10% at six months during the SP2 study

Adverse Events Observed with the Frequency of ⩾ 10% at 6 Months			
Adverse Events	Pirfenidone Number of Patients (%)	Placebo	p Value*
Any adverse events	72 (98. 6)	32 (88.9)	0.0400
Photosensitivity	32 (43.8)	0 (0.0)	0.0000
Stomach discomfort	22 (30.1)	3 (8.3)	0.0143
Anorexia	23 (31.5)	2 (5.6)	0.0030
Nausea	16 (21.9)	2 (5.6)	0.0314
Heartburn	12 (16.4)	1 (2.8)	0.0566
Drowsiness	17 (23.3)	6 (16.7)	0.4672
Fatigue	16 (21.9)	1 (2.8)	0.0102
Upper respiratory tract infections	12 (16.4)	3 (8.3)	0.3767
Fever	6 (8.2)	4 (11.1)	0.7271
Elevation of GOT	4 (5.5)	6 (16.7)	0.0785
Elevation of γ-GTP	20 (27.4)	3 (8.3)	0.0249
Urinary occult blood positive	6 (8.2)	4 (11.1)	0.7271
Elevation of CRP	15 (20.5)	10 (27.8)	0.4694

Long-term safety of pirfenidone in IPF (Lancaster 2016)

A comprehensive analysis of the safety of pirfenidone in IPF was conducted using the three Phase III ASCEND, CAPACITY 1 & 2 trials, and two open-label trials (PIPF-002, PIPF-012 [RECAP]). Safety outcomes were assessed from baseline until 28 days after study drug discontinuation.

PIPF-002 is an ongoing open-label compassionate-use study in US patients with either IPF or secondary pulmonary fibrosis. RECAP is an ongoing open-label extension study in patients who completed either ASCEND, CAPACITY 1 or 2 studies (See Section 4.11).

The latest interim analyses of the same integrated population (ASCEND, CAPACITY 1 & 2, PIPF-002, RECAP) was conducted using a data cut-off date of 17 January 2014 and was presented at ATS 2015. A total of 1299 patients were included in the integrated population. The cumulative total exposure to pirfenidone was 3160 person exposure years (PEY). The median duration of exposure was 1.7 years (range, 1

week–9.9 years); 545 (42%) patients received pirfenidone for ≥2 years and 325 (25%) patients received pirfenidone for ≥4 years. The majority of patients (964 [74.2%]) received a mean daily dose between 1800 mg and 2600 mg. Cumulative safety outcomes in the pooled pirfenidone 2403 mg/day and placebo treatment groups in the Phase 3 studies are presented in Table 64 for comparison.

Table 64. Treatment emergent AEs in the integrated population compared with the pooled pirfenidone 2403 mg/day and placebo groups in the Phase III trials*

	Integrated population† (N=1299)	Pooled ASCEND, CAPACITY 1 & 2 population	
		Pirfenidone 2403 mg/day (N=623)	Placebo (N=624)
Median duration of exposure, years (range)	1.7 (>0, 9.9)	1.0 (>0, 2.3)	1.0 (>0, 2.3)
Treatment-emergent adverse event, %			
Nausea	37.6	36.1	15.5
Cough	35.1	27.8	29.2
Dyspnoea	30.9	16.9	20.2
Upper respiratory tract infection	30.6	26.8	25.3
Idiopathic pulmonary fibrosis	29.3	13.0	19.9
Fatigue	28.2	26.0	19.1
Diarrhoea	28.1	25.8	20.4
Rash	25.0	30.3	10.3
Bronchitis	23.8	14.1	15.4
Headache	21.6	22.0	19.2
Nasopharyngitis	21.3	16.7	17.9
Dizziness	21.2	18.0	11.4
Dyspepsia	18.4	18.5	6.9
Vomiting	15.9	13.3	6.3
Weight decreased	15.6	10.1	5.4
Back pain	15.4	10.4	10.4
Anorexia	15.2	13.0	5.0

*Occuring in 15% of patients in the cumulative clinical database
†Includes two patients from PIPF-002 with a diagnosis of “pulmonary fibrosis”

Consistent with prior observations, gastrointestinal and skin-related events were among the most common treatment emergent adverse events. However, these were mainly mild to moderate in severity, reversible, and rarely led to treatment discontinuation. Elevations in liver enzymes (ALT or AST >3 x ULN) occurred in 40/1299 (3.1%) patients in the integrated population, compared with 23/623 (3.7%) and 5/624 (0.8%) in the pooled pirfenidone and placebo groups in the phase 3 trials. All elevations were reversible without clinical sequelae. Respiratory adverse events

were more common in the integrated population than the placebo and pirfenidone-treated patients from the pooled Phase 3 trials. This is finding is expected from a chronic progressive respiratory disease followed over a long period of observation.

Overall, this comprehensive safety analysis demonstrates that long-term treatment with pirfenidone is generally well tolerated in IPF patients, with no new or unexpected adverse events.

Safety of the technology in relation to the Decision Problem

The safety and efficacy of pirfenidone in patients with IPF was evaluated in 3 multinational, randomised, double-blind, placebo-controlled Phase 3 studies (ASCEND, CAPACITY 1 & 2) and two smaller Japanese studies (SP2 and SP3).

- In ASCEND, CAPACITY 1 & 2, GI and skin-related adverse events were more frequently reported in the pirfenidone 2403 mg/day group vs placebo. The incidence of GI-related AEs was highest during the initial 3 months of treatment, and decreased over time.
- In ASCEND, Grade 3 GI-related AEs were reported in 5.4% of patients in the pirfenidone 2403 mg/day group and 1.4% of patients in the placebo group. No Grade 4 GI-related AEs were reported. Rash (any grade) was reported in 28.1% of pirfenidone-treated patients vs 8.7% of patients in the placebo group.
- In the pooled CAPACITY 1 & 2 studies, commonly reported AEs in the pirfenidone 2,403 mg/day group with at least 1.5 times the incidence of placebo included GI events. The incidence of rash was 32% vs 12%, and the incidence of photosensitivity was 12% vs 2% in the pirfenidone 2,403 mg/day group (n=345) vs placebo (n=347), respectively.
- In a long-term safety analysis of an integrated population (n=789) from the CAPACITY studies and 2 ongoing open-label studies (PIPF-002 and RECAP), GI AEs were among the most commonly reported AEs. These AEs were mostly mild to moderate in severity and rarely led to treatment discontinuation.

The incidence of treatment-emergent nausea, dyspepsia, and vomiting was 41%, 21%, and 19%, respectively. The majority of new-onset treatment-emergent GI AEs (i.e. nausea, vomiting, and diarrhea) decreased substantially over time. The incidence of skin-related AEs in the integrated population was consistent with the incidence from the pooled CAPACITY studies. The majority of new-onset skin-related AEs occurred within the first 6 months of treatment initiation and decreased substantially over time.

- Liver enzyme elevations have been reported in pirfenidone-treated patients in ASCEND, CAPACITY 1 & 2 trials. Increases in alanine and aspartate aminotransferase ≥3× ULN were reversible with dose modification or treatment discontinuation. No cases of liver transplant or death due to liver failure related to pirfenidone were reported in these studies.
- Changes to liver enzymes, GI and skin-related adverse events were also reported in the Japanese SP2 and SP3 studies.

Overall, the results from the 3 large Phase III studies show that treatment with pirfenidone 2403 mg/day is generally well-tolerated. The most common treatment-emergent adverse events observed in clinical trials were gastrointestinal and skin-related adverse events. These adverse events were generally mild to moderate in severity and rarely resulted in treatment discontinuation. These safety profile of pirfenidone are supported by the two Japanese SP2 and SP3 studies.Long-term clinical safety data of pirfenidone in IPF patients reflect the expected safety profile of pirfenidone, with no new or unexpected adverse events.

The safety and adverse event profile of pirfenidone is different to that of nintedanib. The most frequently reported adverse reactions from the nintedanib clinical trial programme were diarrhoea, nausea and vomiting, abdominal pain, decreased appetite, weight loss and elevation of hepatic enzymes (EMC 2015b).

It is also noted that, at the latest Pharmacovigilance Risk Assessment Committee (PRAC) meeting of the EMA in November, the following recommendation was made: "*the product information should be updated to include cases of haemorrhage that have been reported in the post-marketing period, including in patients with or without anticoagulant therapy or other drugs that could cause bleeding, in the special*

warnings and precautions for use section. In addition the product information should be updated to add epistaxis as new undesirable effect with a common frequency" (EMA 2015b). Pirfenidone is the only approved IPF treatment without a special warning or precaution for use in patients at risk of cardiovascular disease or bleeding (EMC 2015a).

A recent publication assessed the use of treatment for IPF in a real-world setting through review of 7 observationsal studies conducted across 4 countries, including the UK (Hanari 2015). The authors concluded that the findings of these studies support the results of clinical trials for pirfenidone: "*pirfenidone is well tolerated, and the most common adverse events are gastrointestinal, skin-related events and weight loss*".

The authors go on to state that: "*No real-life studies on nintedanib are available yet. However, the real-life scenario seems to be quite different from the one of randomised trials: patients often have comorbidities, have more severe disease, take concomitant medications and have a higher mortality. For these reasons, it seems important that in the future clinical trials better reflect the general population of patients on whom the drugs will be used and prescribed long term*" (Hanari 2015).

4.13 Interpretation of clinical effectiveness and safety evidence

Interpretation of clinical effectiveness and safety of pirfenidone in IPF

Results from the three large double-blind, randomised, placebo-controlled Phase 3 clinical studies (ASCEND, CAPACITY 1 & 2) provide strong collective evidence that pirfenidone has a favourable benefit in patients with IPF. Further supportive evidence of a therapeutic effect for pirfenidone comes from the Japanese studies (SP2 and SP3).

Patients enrolled in the ASCEND, CAPACITY 1 & 2 studies had mild-to-moderate IPF which is relevant to the decision problem. The enrolment criteria were based on international IPF guidelines applicable at the time and recruited a sample representative of UK practice. Whilst the pivitol RCTs were multi-national studies with UK study sites, comparisons of IPF patients from multiple real-world studies provides reassurance that the results are applicable in the UK; "*The profile of these patients* [included in the observational studies] *seems to be quite similar all over the world, as does their clinical management*"(Harari 2015). Baseline characteristics of the patients across all three studies were similar, despite slight differences in study design.

In ASCEND, treatment with pirfenidone for 52 weeks significantly reduced disease progression (as measured by changes in FVC, 6MWD, and PFS) compared with placebo in patients with IPF. At Week 52, there was a relative reduction of 47.9% in the proportion of patients with a ≥10% decline in % predicted FVC or death ($p<0.001$) in the pirfenidone group as compared with placebo. There was also a relative increase of 132.5% in the proportion of patients with no decline in FVC in the pirfenidone group vs. placebo (see Table 18: 63 patients [22.7%] vs. 27 patients [9.7%]) ($p<0.001$). The treatment effect on FVC emerged early and increased during the course of the trial, resulting in an approximate halving in the rate of decline at 1 year.

The finding with respect to the primary end point was supported by the favorable effect on rates of death from any cause and from IPF. At Week 52, there was also a relative reduction of 27.5% of the proportion of patients who had a decrease of ≥50m in 6MWD or who died in the pirfenidone group as compared with placebo ($p=0.04$).

Pirfenidone also reduced the relative risk of death or disease progression by 43% compared with placebo (HR: 0.57 [95% CI: 0.43-0.77] p<0.001). (King 2014)

CAPACITY 2 reached its primary endpoint at Week 72; pirfenidone 2403 mg/day was associated with significantly reduced decline FVC compared with placebo (-8.0% vs. -12.4%; p=0.001). Mean change in % FVC in the pirfenidone 1197 mg/day group were intermediate to the pirfenidone 2403 mg/day and placebo groups. CAPACITY 1 did not reach its primary endpoint: there was no significant difference between pirfenidone 2403 mg/day and placebo in the mean change in % predicted FVC was evident at Week 72 (-9.0% vs. -9.6%; p=0.501). Nonetheless, the study provided supportive data on treatment effect of pirfenidone in patients with IPF. A significant treatment effect was evident at every timepoint from Week 12 to Week 48, and in repeated-measures analysis of % predicted FVC change over all assessment timepoints (p=0.007). At Week 72, the pooled analysis of pirfenidone 2403 mg/day across the CAPACITY studies showed that it prolonged PFS by 26% compared with placebo (HR: 0.57 [95% CI: 0.57-0.96] p=0.025). In the pooled analysis, a 31.2% relative difference was noted between pirfenidone 2403 mg/day and placebo in reducing the declining in 6MWD at Week 72 (p=0.0009). (Noble 2011)

There were two sources of variability identified in CAPACITY 1 which may explain why the primary endpoint was not met. Firstly, a significant imbalance was observed resulting in a higher proportion of patients trending towards borderline obstructive disease (defined as FEV1/FVC <0.8) in the CAPACITY 1 placebo arm. Secondly, adherence in the pirfenidone arm diminished after Week 48. Once both sources of variability are accounted for, the heterogeneity of pirfenidone's treatment effect in the CAPACITY studies attenuated and closely resembled the treatment effect in CAPACITY 2, making the results of pirfenidone on functional parameters in patients with IPF highly consistent, statistically significant and clinically relevant.

Due to the inconsistency in results of the primary outcome in the CAPACITY trials, a third phase III trial was requested by the FDA to confirm the efficacy of pirfenidone in patients with IPF. In its consideration of the totality of evidence from ASCEND, CAPACITY 2 and CAPACITY 1 (including the pooled analysis), the FDA concluded: "*Efficacy data shows consistent positive benefit of pirfenidone in the treatment of IPF.*

Statistically significant improvement of FVC was seen in Studies 004 (CAPACITY 2) and 016 (ASCEND). Benefit in FVC was supported by a numerical trend in favor of mortality for pirfenidone compared to placebo. There was also benefit noted in other secondary measures such as PFS, and 6MWT distance"(FDA, 2014a)

The efficacy results from the 3 large studies (ASCEND, CAPACITY 1 & 2) show that pirfenidone 2403 mg/day in IPF patients is more effective in reducing disease progression compared with placebo.

The mortality analyses were prespecified to be conducted in both the ASCEND population and in the pooled population from the ASCEND and CAPACITY trials. The magnitude of the treatment effect on mortality was large and internally consistent across analyses and populations — an important clinical finding. In the pre-specified pooled analysis of ASCEND, CAPACITY 1 & 2 at 52 weeks, patients treated with pirfenidone 2403 mg/day had an approximate (and statistically significant) 50% reduction risk of death compared to the placebo arm. (HR=0.52; 95% CI: 0.31-0.87; p=0.011)

A post-hoc analysis of the pooled CAPACITY and ASCEND datasets has demonstrated that patients on pirfenidone with a ≥10% reduction in FVC have significantly less mortality than those on placebo with a similar change. These findings are suggestive of a salutary treatment effect with pirfenidone, beyond just the slowing in the rate of loss of lung function. The effects of pirfenidone on physical function (as measured by 6MWD) could also play a role in longer-term survival (Nathan 2015a, Nathan 2016).

The two Japanese studies (SP2 and SP3) provide supportive efficacy and safety data to the three large trials summarised above, however, the doses used in these studies are lower than the licensed dose in the UK, due to the lower body weight of the Japanese population has compared to their European counterparts.

In a subgroup analysis of the pooled ASCEND, CAPACITY 1 & 2, data looking at the treatment effect of pirfenidone stratified by baseline disease severity, results showed that pirfenidone reduces disease progression with no significant differences between earlier (FVC ≥80%) and later/more advanced disease groups (FVC <80%) (Albera

2015). These findings support the prompt initiation of treatment with pirfenidone after IPF diagnosis, irrespective of disease severity (see Section 4.8). The use of pirfenidone in patients with mild IPF has been recommended by national agencies in a number of countries, including Canada, Sweden and Switzerland.

Analysis of the RECAP population – those with a % predicted FVC <50% and those with ≥ 50% – also supports early intervention with pirfenidone, and its continued use in patients, with a similar decline over 180 weeks experienced in both groups (Section 4.8).

Clinical trial results are supported by UK observational data reporting clinical experience with pirfenidone 6 and 9 months before and after treatment initiation (Chaudrhuri 2014). The investigators reported a reduction in the decline of mean percentage change of FVC. After 9 months there was a difference in the gradient of slope of decline in FVC before and after pirfenidone commencement from −1.043±1.605 to −0.197±0.231. Although retrospective, observational data such as this supports the previous efficacy data from clinical trials.

Overall, the results from the 3 large Phase III studies (ASCEND, CAPACITY 1 & 2) show that treatment with pirfenidone 2403 mg/day is generally well-tolerated with an acceptable side effect profile. EMA has recently requested BI to update the nintedanib SmPC with a warning on the risk of haemorrhage and epistaxis and new data on mild/moderate hepatic impairment (EMA, 2015b). Pirfenidone is the only approved IPF treatment without a special warning or precaution for use in patients at risk of cardiovascular disease or bleeding. The most common treatment-emergent adverse events observed in clinical trials were gastrointestinal and skin-related adverse events. These adverse events were generally mild to moderate in severity, more common during the first months of treatment, and rarely resulted in treatment discontinuation. Clinical elevations in liver enzyme levels were also more common in the pirfenidone-treated group compared with placebo, but were reversible with no clinical consequences. These safety findings are supported by the two Japanese SP2 and SP3 studies. Long-term clinical safety data of pirfenidone in IPF patients (n=789) with data collated from 2059 patient exposure years, reflect the expected safety profile of pirfenidone, with no new or unexpected adverse events.

Strengths of the clinical evidence base

Marketing authorisation status

Both the FDA and EMA approve the use of pirfenidone for the treatment of IPF. Pirfenidone was granted EU marketing authorisation on the 28th February 2011 by the European Commission (EU/1/11/667/001; EU/1/11/667/002; EU/1/11/667/003) (EMA public summary/SMPC). Prior to this, pirfenidone was designated as an 'orphan medicine' (a medicine used in rare diseases) on 16th November 2004 by the European Commission (EMA EU/3/04/241) (EMA orphan). The FDA granted pirfenidone fast track, priority review, orphan product, and breakthrough designations in October 2014.

Study design

ASCEND, CAPACITY 1 & 2 were large, rigorous, randomised double-blind, placebo controlled studies and were very similar in design. The planned schedules for CAPACITY 1 & 2 were identical. Entry criteria CAPACITY 1 & 2 were in general based on the diagnostic criteria in the ATS/ERS guidelines (ATS/ERS 2000) thus ensuring consistency and robustness in the diagnosis of patients for inclusion into the studies. ASCEND had minor differences in study design to CAPACITY 1 & 2, but despite this, the baseline characteristics were similar across all three studies.

Collection of study data

Patients who participated in the ASCEND, CAPACITY 1 & 2 studies were recruited from specialist centres around the world, including the UK. It is anticipated that in clinical practice IPF patients will be managed at specialist centres and therefore the results should be applicable to UK practice.

Baseline characteristics

In the CAPACITY 1 & 2 studies, patients were eligible to enter if their FVC was ≥50% predicted and their DLco ≥35% of predicted value and the licensed indication states that pirfenidone is indicated for mild to moderate IPF. The indication itself does not define mild and moderate disease in terms of FVC thresholds (Nathan 2011). In ASCEND, patients with major airflow limitation were excluded, in order to enrol

patients at higher risk for disease progression. Despite minor differences in eligibility criteria, baseline characteristics were similar between intervention and control groups in all five trials (ASCEND, CAPACITY 1 & 2, SP2 and SP3).

Relevant clinical endpoints

Whilst overall survival is a critical end point for IPF studies, there were relatively few deaths (<10%) in the three large Phase III studies (ASCEND and CAPACITY 1 & 2), in either treatment group. None of the studies were powered to assess the effect of pirfenidone on mortality, so a pre-specified pooled analysis of the three trials was perfomed to increase statistical power. In the prespecified analysis of all-cause mortality in the pooled population of 1247 patients (555 from the ASCEND study and 692 from the CAPACITY studies), pirfenidone reduced the risk of death at 1 year by 48%, as compared with placebo (hazard ratio, 0.52; 95% CI, 0.31 to 0.87; P=0.01)

Change in FVC was selected as the primary endpoint because of its widespread clinical use, and the clinical relevance of irreversible loss of lung function. FVC was selected as the primary endpoint because it is a reliable, valid, and responsive measurement of irreversible morbidity in IPF, and is highly predictive of survival. An assessment of the proportion of patients with a decline of ≥10% is a threshold widely accepted as clinically meaningful and prognostic of death is more directly clinically meaningful than is the assessment of differences in treatment group means. FVC is also easily measured in routine clinical practice. It is a reliable and reproducible test which is used by many IPF specialist centres.

Weakness of the clinical evidence

Generalisability to patients with severe IPF

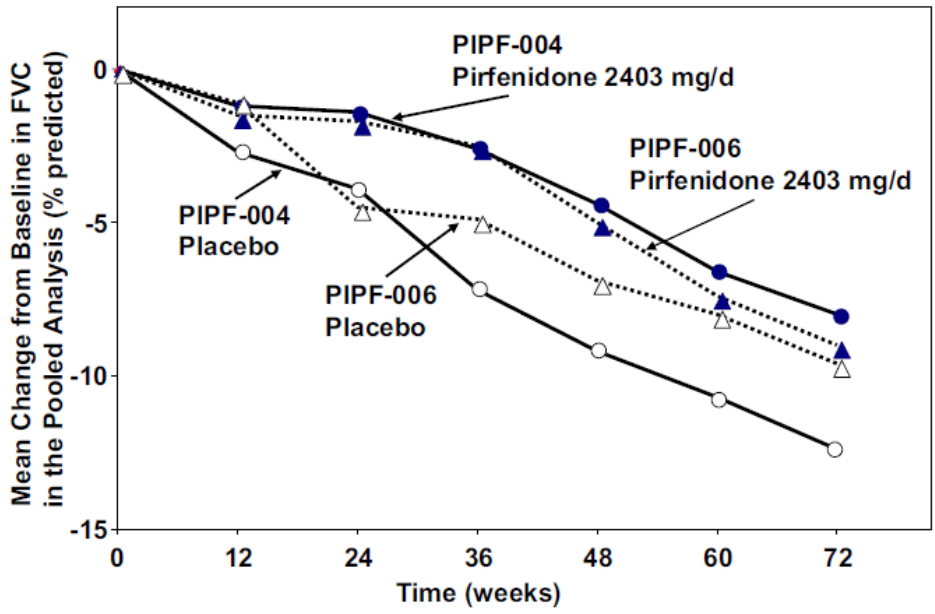
As only patients with mild to moderate IPF, and relatively few comorbidities, were enrolled in ASCEND, CAPACITY 1 & 2 the results need to be carefully interpreted for the broader population of patients. Since concomitant administration of other treatments for IPF was generally prohibited, the effect of these therapies in patients given pirfenidone is not known but other commonly used supportive therapies were administered as well as medicines for other conditions.

A review article on real-world evidence for pirfenidone in IPF support the efficacy and safety findings from the clinical trials. Results from a Named Patient Programme conducted in the UK (n=40), showed a reduction in FVC and DLco decline after starting pirfenidone treatment. Pirfenidone was well-tolerated, and the most common adverse events were GI, skin-related, and weight loss. Real-world evidence also suggests pirfenidone could be effective in patients who are more severe those in the RCTs (Harari and Caminati, 2015).

Difference outcomes in CAPACITY 1 & 2

Different FVC outcomes across studies may be a product of the natural variability in rates of FVC decline in this heterogeneous disease. This is supported by an analysis of the placebo groups in the two studies, each of which behaved quite differently. As shown in Figure 38, the CAPACITY 1 placebo group relative to the CAPACITY 2 placebo group appears to have an attenuated rate of decline in percent predicted FVC, beginning around Week 36 and persisting through Week 72 (FDA, 2010). In addition, the slope of the percent predicted FVC curve for the CAPACITY 1 placebo group is different from that observed for the CAPACITY 1 pirfenidone group and both groups in CAPACITY 2.

Figure 38. Mean Change from Baseline in Percent Predicted FVC in CAPACITY 2 (PIPF-004) and CAPACITY 1 (PIPF-006) (ITT population)



Disease severity based on % predicted FVC

Whilst using FVC threshold is potentially attractive to define the severity of IPF, it is recognised that this is an arbitrary cut-off to some extent and that other factors are taken into account in clinical practice as to when to treat a patient. For instance, if the predicted FVC was 81% at diagnosis (potentially mild disease), it would not be known whether that patient started with a predicted FVC at 81% and had lost no lung function or started at 120% and has lost one third of their FVC. Given that most patients die with an FVC of about 40%, in the second scenario, the patient has progressed halfway from onset to death. Clinicians would also expect to take into consideration other factors to reach a truly accurate measure of disease severity including HRCT and DLco.

Japanese studies

SP2 is the smallest and weakest of the pirfenidone studies. It used a non-validated primary endpoint, oxygen desaturation during a treadmill exercise test and failed to distinguish active from placebo treatment. However, there was a statistically significant difference in the rate of decline of FVC and by nine months 14% of placebo patients and none of the pirfenidone group had experienced an acute exacerbation.

SP3 is a larger, randomised, placebo controlled trial of pirfenidone in 275 patients who were diagnosed with IPF according to ATS criteria. Significant differences were found in the decrease in vital capacity over 52 weeks which was the primary endpoint and the progression free survival (Taniguchi 2010). The main challenge with SP3 is that the initial primary endpoint was oxygen desaturation during exercise. However, during the course of this multi-year trial, the academic community's views on appropriate primary endpoints in IPF evolved, and the data safety and monitoring board recommended a change of the primary endpoint to VC after a discussion of blinded interim comparative data.

The SP3 (and SP2) study used the last observation carried forward to compensate for missing data in about one third of the subjects and this approach might have contributed to the apparent difference between the groups. Using the last observation carried forward to replace missing data may make the group with the

greater dropout rate appear to do better, if lung function decreases progressively with time. Importantly, this technique may underestimate the true variability of missing data and inflate the type 1 error rate (i.e. the rate of finding a statistically significant difference when a difference does not truly exist). In studies such as this one, with a small treatment effect and marginal p-value, significance may hinge on the method of statistical adjustment used.

Relevance to clinical practice

ASCEND and CAPACITY 2 showed statistically significant evidence of benefit in the primary outcome variable of change in lung function, with a relative difference from placebo of 35% in FVC change. The primary endpoint was not met in CAPACITY 1; however, supportive analyses (including a repeated measures analysis of overall treatment effect during the study) provide evidence of a pirfenidone treatment effect on FVC.

Scientific literature and expert opinion support the clinical meaningfulness of changes in FVC and recognise a decline of 10% in percent predicted FVC as an independent predictor of mortality in patients with IPF (Collard 2003; Latsi 2003; Flaherty 2003; Zappala 2010). In the recent study by du Bois the one-year risk of death in patients with IPF was more than 2-fold higher ($p<0.001$) in patients with a 24-week decline in FVC between 5-10%. The estimated minimal clinically important difference in this study was estimated at 2-6% (du Bois 2011b). In ASCEND, at Week 52, 16.5% of patients in the pirfenidone group experienced an FVC decline of ≥10% or death, compared with 31.8% in the placebo group, representing a clinically significant 47.9% reduction in the proportion of patients who experienced a meaningful change in FVC or death. In CAPACITY 2, there was a reduction of 42% in the proportion of patients with a decrement of ≥10% in percent predicted FVC at Week 72; in CAPACITY 1, there was a 14% reduction and in the pooled analysis (representing the most stable estimate of the magnitude of effect), a 30% difference between pirfenidone- and placebo-treated patients. This magnitude of reduction in the proportion of patients with this clinically meaningful decrement is inherently relevant in a disease manifested by ongoing and irreversible loss of lung function.

In ASCEND, there was a relative reduction of 27.5% of the proportion of patients who had a decrease of ≥50m in 6MWD or who died in the pirfenidone 2403 mg/day group as compared with placebo at Week 52 (p=0.04). In the pooled analysis of CAPACITY 1 & 2 studies, the mean decline from baseline in 6MWD was markedly reduced in the pirfenidone 2403 mg/day group compared to the placebo group at Week 72 (absolute difference of 24 m; p<0.001). The selection of the 50 m decrement was based on recent data supporting this threshold as clinically meaningful and prognostic for survival where a 24-week decrement of 50 m was associated with a four-fold increase in the risk of death over the subsequent 12 months (p<0.001) (du Bois 2011a).

Other findings also provide evidence of a clinically meaningful benefit of pirfenidone in patients with IPF. In ASCEND, pirfenidone also reduced the relative risk of death or disease progression by 43% compared with placebo (HR: 0.57 [95% CI: 0.43-0.77] p<0.001). In CAPACITY 2, pirfenidone 2403 mg/day was associated with a relative reduction of 36% in the risk of disease progression or death compared with placebo. While the studies excluded patients with advanced disease and, therefore, had low power to assess survival, in both CAPACITY 1 & 2, there were fewer deaths overall and fewer IPF-related deaths in patients treated with pirfenidone 2403 mg/day than in those receiving placebo.

The evidence from the phase 3 studies investigating pirfenidone is relevant as it provides a comparison with placebo which is representative of 'best supportive care' as outlined in national guidelines by the BTS (Wells 2008a) as outlined in the final scope. In the ASCEND, CAPACITY 1 & 2 studies patients could receive supplemental oxygen, pulmonary rehabilitation, opiates and anti-reflux therapy. In CAPACITY 1 & 2 patients were allowed short courses of azathioprine, cyclophosphamide, corticosteroids, or acetylcysteine for protocol defined acute exacerbation of idiopathic pulmonary fibrosis, acute respiratory decompensation, or progression of disease (Noble 2011). In SP3 concomitant use of corticosteroid was permitted during the study period, concomitant use of immunosuppressants and other experimental agents under investigation was not allowed.

The majority of the evidence base comes from patients that were treated with the UK licensed dose of pirfenidone 2403 mg/day. The 2403 mg/day dose used in the ASCEND and CAPACITY 1 & 2 studies was derived by normalisation of the 1800 mg/day dose used in the Japanese studies. This was based on a calculation that adjusted for the differences in mean weights in the North American and European patient population compared with the Japanese patient population.

4.14 Ongoing studies

PASSPORT is an ongoing, multicentre, prospective, 2-year safety registry enrolling European patients with IPF in a real-world setting to evaluate the long-term safety profile of newly prescribed pirfenidone (≤30 days of treatment at registry entry). A full manuscript with primary results will be submitted for peer-review in Q1 2016.

5 Cost effectiveness

- A Markov health state model was constructed to be reflective of all available clinical evidence while allowing results to be easily reproduced.
- Pirfenidone significantly reduces IPF-related and all-cause mortality (Section 4). This survival benefit is demonstrated in the de novo economic model, with pirfenidone facilitating a 3.29 life year gain for the ITT population, vs. BSC.
- Pirfenidone was associated with a substantial QALY gain of 1.87 for the ITT population, vs. BSC. This benefit was observed in both the pre-progression and post-progression health states, demonstrating the lack of clinical evidence to support the use of a progression status-based stopping rule in practice.
- In the model base case, the ICER for pirfenidone vs. BSC in the ITT population was [redacted] (list price) per QALY gained.
- As expected, similar cost-effectiveness results applied to both the mild and moderate populations when modelled in isolation. For the mild population, the ICER for pirfenidone vs. BSC was [redacted] (list price). For the moderate population, the ICER for pirfenidone vs. BSC was [redacted] (list price).
- Pirfenidone was also cost-effective vs. nintedanib in patients with moderate disease, with ICERs of [redacted] (list price).
- Validation against alternative sources demonstrated the plausibility of long-term survival assumptions vs. real-world evidence.
- The model was shown to be particularly sensitive to options regarding the estimation of long-term survival and duration of treatment effect for all treatments. This is likely due to the lack of equivalently robust long-term follow-up data for comparator therapies.

5.1 Published cost-effectiveness studies

Identification of studies

A systematic review was conducted to identify cost effectiveness studies of pirfenidone for adult patients with mild to moderate IPF in England. Full economic evaluations were included as well as relevant economic data reported in technology assessments, including those produced for NICE. Given the extent of the previous searches in the 2011 pirfenidone STA submission, only records published from 2010 onwards were screened [InterMune 2011].

A single search strategy was used to identify cost effectiveness studies, health related quality of life (Section 5.4.3) and resource use data (Section 5.5.2) in November 2015. The full strategy is included in Appendix 17. The search was structured using the following concepts:

Idiopathic Pulmonary Fibrosis AND (Resource use OR Cost-effectiveness OR Utilities).

The strategy used a combination of subject indexing terms and free text search terms in the title, abstract or keyword fields to capture the concepts.

The following databases were searched in November 2015: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present; Embase 1974 to 2015 November 16; Econlit 1886 to October 2015; Health Technology Assessment Database : Issue 4 of 4, October 2015 (Cochrane Library); NHS Economic Evaluation Database: Issue 2 of 4, April 2015 (Cochrane Library); Cost-Effectiveness Analysis registry; PROQOLID; ScHARRHUD and the EuroQol database. The following websites were searched: NICE; Pharmaceutical Benefits Scheme (PBS); Canadian Agency for Drugs and Technologies in Health (CADTH); the SMC. NICE submissions were hand searched, along with a recent systematic review, Google Scholar and seven conferences. Full strategies are provided in Appendix 17 to this submission.

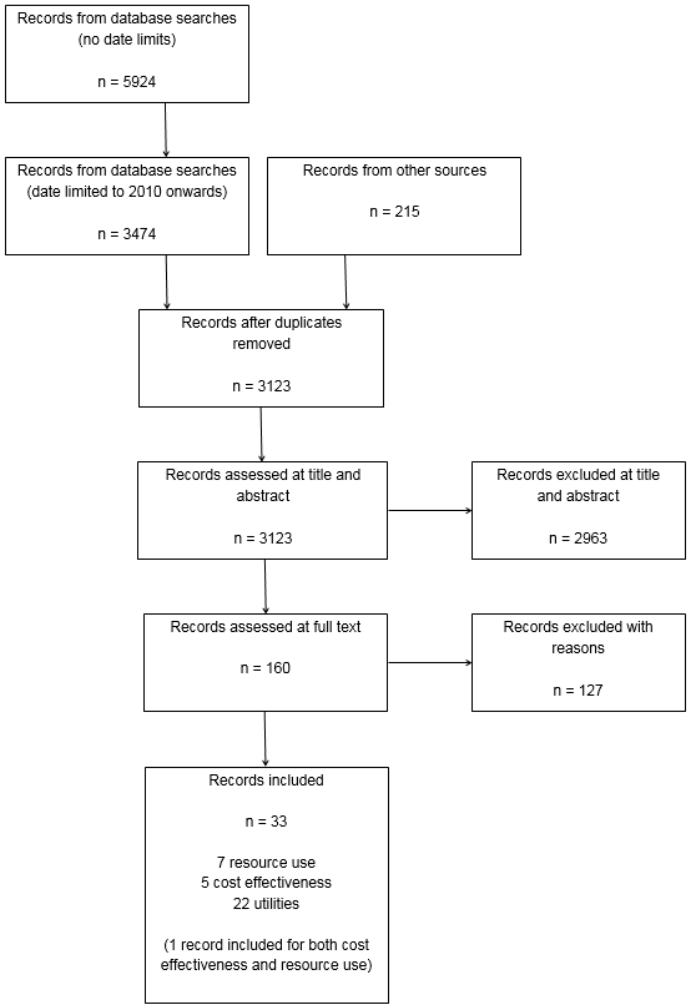
Following the searches, obviously irrelevant records (such as animal studies and studies about ineligible populations) were removed by a single reviewer. The titles and abstracts of the remaining records were assessed for relevance by one reviewer

and checked independently by a second reviewer. Once the full text of studies was obtained, two independent researchers assessed studies in detail for relevance to the systematic review's eligibility criteria and made the final selection of studies to inform the systematic review. The record selection process is shown in Figure 39.

Data on the study design and cost effectiveness outcomes were extracted and quality assessment was conducted by one reviewer and checked by a second reviewer. Quality assessment was conducted using the Drummond and Jefferson criteria [Drummond 1996].

The combined searches for economic evaluations of pirfenidone, resource use and costs in IPF and utilities in IPF retrieved 5924 records from the database searching. Of these, 3474 records were published in 2010 or later. 215 records were retrieved from other sources. 3123 records were assessed after removal of duplicates. 4 studies (in 5 reports) were eligible (Figure 39). A list of excluded studies appears in Appendix 18.

Figure 39: PRISMA diagram showing the record selection process for the systematic literature reviews of economic evaluations, resource use and costs and utilities



Description of identified studies

4 studies (in 5 reports) were identified that met the inclusion criteria, summarised in Table 65 [CADTH 2014, CADTH 2015, InterMune 2011, Loveman 2014, Loveman 2015].

The manufacturer's 2011 pirfenidone NICE submission and the associated Evidence Review Group (ERG) report met the inclusion criteria [Cooper 2012, InterMune 2011]. A patient simulation model was submitted by the manufacturer. The ERG stated that whilst this was a satisfactory approach, they did not believe that there was any reason why a Markov cohort model could not have been used.

The manufacturer's 2011 submission reported an ICER of £25,969/QALY for patients with FVC of ≤80%. The ERG reported concerns about the quality of the evidence to support the ICER for patients with FVC of ≤80% which was due to data from the ASCEND study not being ready to be included in the submission. The ERG suggested that triple therapy should have been considered as a comparator, although this is not included in the scope to this appraisal. The ERG was also not convinced by the data that pirfenidone was clinically superior to BSC.

Loveman et al. produced a UK health technology assessment of all available treatments for IPF for the National Institute for Health Research [Loveman 2014, Loveman 2015]. The model used efficacy data from a systematic review of 14 studies, 10 of which were pharmaceutical treatments. The studies were of all patients with a diagnosis of IPF regardless of severity.

The findings of the 10 pharmaceutical treatment studies were used to inform a network meta-analysis (NMA) and to populate a Markov state transition economic model with four health states: unprogressed IPF, progressed IPF, lung transplantation and death. In all live states patients could have an exacerbation. Movement into the progressed state occurred after a decline in FVC of at least 10%.

Six interventions in total were assessed: pirfenidone; azathioprine and prednisolone; BIBF 1120 (nintedanib); NAC triple therapy; inhaled NAC; and sildenafil. Utility data were taken from a separate systematic review with costs from the BNF and NHS reference costs.

An incremental analysis was performed producing an efficiency frontier comprising BSC, inhaled NAC compared to BSC at £5,037/QALY and nintedanib compared to inhaled NAC at £209,246/QALY. Pirfenidone was dominated by inhaled NAC.

Although the study was of high quality when judged against the Drummond checklist, with no areas of concern, its relevance to the UK and to this submission is limited. The systematic review of the effectiveness evidence did not include the ASCEND and INPULSIS trials and so did not take into account three large, recent and highly relevant evidence sources. In addition, the NMA includes a trial of severe IPF patients where the placebo arm could influence overall results in the network.

The utility values chosen for the un-progressed and progressed states were not from the UK and that may also limit the generalisability of findings. Similarly, efficacy data were taken from studies predominantly outside of the UK and for pirfenidone the data were taken from two Japanese studies and two multi-national studies (of which the UK was one country). The authors acknowledged that this may limit the generalisability of the findings to the UK.

Two of the eligible studies were Common Drug Reviews (CDR) for the CADTH Canadian Drug Expert Committee (CDEC). One of the CDRs was for nintedanib and the other for pirfenidone (an update of a previous CDR following publication of data from the ASCEND) [CADTH 2014, CADTH 2015].

The pirfenidone CDR used pooled data from the ASCEND, CAPACITY and RECAP studies for effectiveness [CADTH 2014]. Pirfenidone was compared to BSC (symptom relief, pulmonary rehabilitation, management of comorbidities, and end of life care, including oxygen therapy) for which survival data were drawn from an unreferenced observational study. The CDEC questioned the way the model extrapolated short term onto long-term survival without evidence that short-term gains with pirfenidone persisted over a patient's lifetime.

The nintedanib CDR used data from the INPULSIS and ASCEND trials [CADTH 2015]. The manufacturer's model used data from two indirect comparisons that suggested similar efficacy between nintedanib and pirfenidone. Although the CDEC concluded that there remains uncertainty, if efficacy is equal then the lower daily drug

acquisition cost of nintedanib (CAN$109) compared to pirfenidone (CAN$117) would mean that nintedanib dominates pirfenidone.

The CDRs provide only a brief summary of the cost effectiveness results and therefore score poorly against most areas of the Drummond quality assessment check list. In addition, the costs provided relate to drug acquisition prices in Canada. In the absence of the full economic models and reports provided to and produced by CADTH as part of the CDRs and the use of Canadian drug prices, the generalisability of the results of the CDRs to the UK is limited.

It is noted that the model for the 2015 nintedanib NICE submission did not meet the inclusion criteria for this review as it was for all patients with IPF and not just those patients with mild to moderate disease [Boehringer Ingelheim 2015]. Although the ICERs reported are not relevant to this submission, the ERG's comments on the model structure and parameter values have been taken into account when constructing the *de novo* model described in the following sections.

Table 65: Summary of eligible cost-effectiveness studies

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
InterMune 2011	2011	Microsimulation model. Efficacy data and QoL taken from CAPACITY trials. Cost data taken from NHS reference costs and published literature. Discontinuation data taken from RECAP extension. Time horizon was lifetime.	NR	NR	NR	£25,969/QALY compared to BSC in patients with predicted FVC <80% or less
CADTH 2014	2014	Cost utility model. Efficacy data drawn from ASCEND and CAPACITY trials and RECAP extension. Utility taken from mapping of SGRQ onto EQ5D. Assumed 50% discontinuation on pirfenidone at four years and 85% at 10 years. Assumed ongoing efficacy of Pirfenidone after discontinuation	NR	NR	Daily cost of Pirfenidone CAN$115 (CAN$12.77 per 267mg capsule. 3 capsules taken three times daily)	CAN$78,024/QALY vs BSC
CADTH 2015	2015	Cost utility model. Efficacy data for nintedanib drawn from INPULSIS trials and for pirfenidone from the ASCEND trial via	NR	NR	Daily cost of Pirfenidone CAN$117 (annual cost CAN$41,983 year one and CAN$42,804 subsequent years). Daily cost of	As the model assumed pirfenidone and nintedanib were equally efficacious and nintedanib was CAN$8 per day less costly, nintedanib

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
		an NMA. QoL was determined for changes in FVC from INPULSIS trials. Time horizon was lifetime.			nintedanib:CAN$109 (annual cost CAN$39,683).	dominated pirfenidone although this was not explicitly stated in the CDR
Loveman 2014 & Loveman 2015	2015	Markov state transition cost utility model. Four health states: unprogressed IPF, progressed IPF, lung transplantation and death. In all live states patients can have an exacerbation. Time horizon of 30 years. Movement into the progressed state occurs after a decline in FVC of at least 10%. Effectiveness data from systematic review and NMA. Utilities from systematic review. Costs from BNF and NHS reference costs.	Unprogressed IPF mean age 66	Pirfenidone: 3.34; BSC: 2.98; Azathioprine & prednisolone: 2.66; NAC triple therapy: 3.03; Inhaled NAC: 3.37; Sildenafil: 3.11; Nintedanib: 4.01	Pirfenidone: £70,118; BSC: £3,084; Azathioprine & prednisolone: £4,313; NAC triple therapy: £5,021; Inhaled NAC: £5,029; Sildenafil: £12,008; Nintedanib: £139,613	Pirfenidone: Dominated by inhaled NAC. Inhaled NAC: £5,037/QALY. All other comparators dominated of extendedly dominated by inhaled NAC except Nintedanib with an ICER compared to inhaled NAC of £209,246/QALY

Key: BSC, Best Supportive Care; CADTH, Canadian Agency for Drugs and Technologies in Health; FVC, Forced Vital Capacity; ICER, Incremental Cost Effectiveness Ratio; IPF, Idiopathic Pulmonary Fibrosis; NICE, National Institute for Health and Care Excellence; NMA, Network Meta Analysis; NR, Not Reported; QALY, Quality Adjusted Life-Year; QoL, Quality of Life; SGRQ, St George's Respiratory Questionnaire.

The detailed quality assessment of the eligible studies against the 36 point Drummond and Jefferson checklist is provided in Appendix 19.

Loveman was assessed as being of high quality with no concerns relating to study design, data collection or the analysis and interpretation of results [Loveman 2015]. The CDRs provided only a brief summary of the cost effectiveness results and therefore scored poorly against most areas of the Drummond quality assessment check list [CADTH 2014 CADTH 2015]. The 2011 pirfenidone submission was a detailed economic model designed to meet NICE's requirements and presented as required by NICE [InterMune 2011].

5.2 De novo analysis

Patient population

Pirfenidone is indicated in adults for the treatment of mild to moderate IPF [EMA 2015a]. In accordance with the information most easily available across clinical trials, the severity of IPF is derived from patient FVC as shown in **Table 66**.

Table 66: IPF severity by FVC

IPF severity	% predicted FVC category
Mild	> 80%
Moderate	50 – 80%
Severe	< 50%

Key: FVC, forced vital capacity; IPF, idiopathic pulmonary fibrosis.

As described in Section 3 there is no agreed staging system for IPF. The definition of severity of IPF is multifactorial with key clinical parameters such as FVC, DLco, the presence of co-existing emphysema and the existence of progressive symptoms taken into account when determining the best course of treatment for a patient. This multifactorial definition has led to some differences in the clinical trial definitions used making comparison between different sources of evidence complex [NICE 2015e].

The EMA licence for pirfenidone was granted on the basis of the finding of the pivotal phase III studies, therefore the modelled population reflects the trial and licence population. Baseline patient characteristics for this population are presented in **Table 67**, taken from pooled data from the Phase III clinical trials: CAPACITY (CAPACITY I

& II) and ASCEND. Full details regarding patient characteristics are presented in Section 4.5.

Table 67: Baseline patient characteristics

Parameter		Pooled data from CAPACITY and ASCEND
Age		67
% male		74%
% predicted FVC		72%
Distribution of % predicted FVC category	**> 80%**	25.3%
	50 – 80%	73.6%
	< 50%	1.0%
SGRQ		418
6MWD		37
Key: 6MWD, 6 minute walking distance; FVC, forced vital capacity; SGRQ, St. George's respiratory questionnaire.		

Due to the irreversible nature of IPF, the treatment goal should be to stabilise the disease, if possible, or at the least reduce the rate of progression [du Bois 2011c]. Therefore, it is anticipated that pirfenidone will be given from first line, in line with several European guidelines which state that pirfenidone should be offered to patients as a first line therapy within its licensed indication for the treatment of mild to moderate IPF [Raghu 2015].

It should be noted that since pirfenidone was previously recommended in patients with moderate IPF, treatment with pirfenidone also has the potential to keep people well rather than to treat once disease is established and progressive [NICE 2015e]. As described in Section 4.8, the efficacy of pirfenidone has been demonstrated to be similar whether initiated in patients with FVC %-predicted >80%, or ≤80% [Albera 2015].

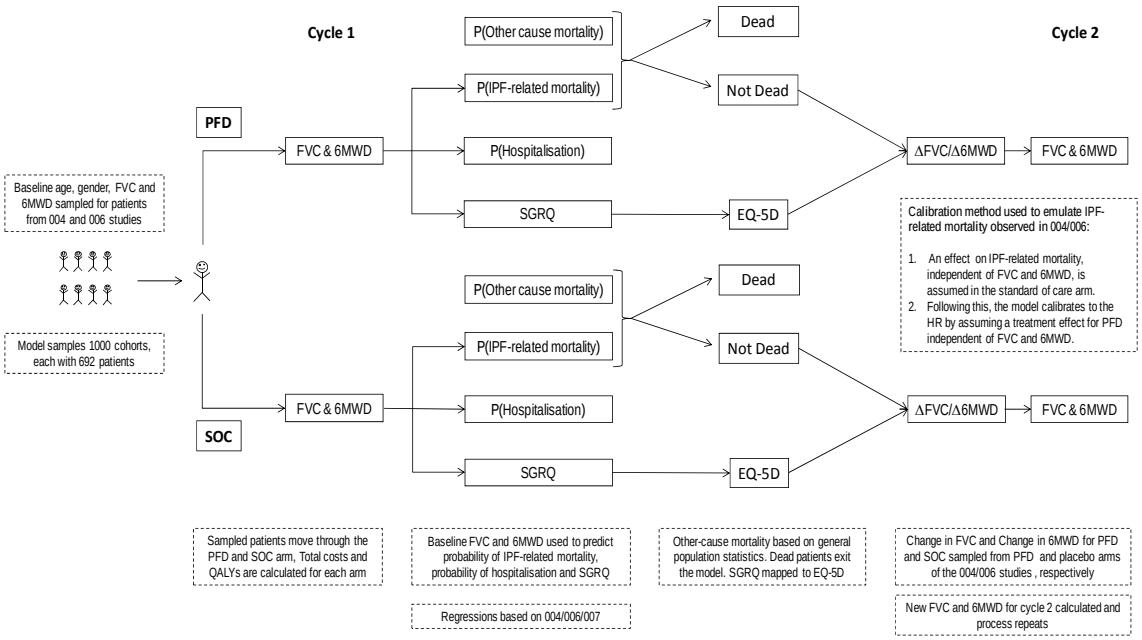
Model structure: Markov vs simulation

We consulted clinical and health economic experts in developing the model structure, as well as reviewing all available literature (see Section 5.1).

The structure of the economic model has been designed with the objective of reflecting all available clinical evidence while allowing results to be easily

reproduced. In the original pirfenidone NICE manufacturer submission (TA282), a micro-simulation was chosen (see Figure 40) [InterMune 2011].

Figure 40: Previous economic model structure from TA282 submission



Key: EQ-5D, EuroQol 5-Dimension; FVC, forced vital capacity; IPF, idiopathic pulmonary fibrosis; PFD, pirfenidone; SGRQ, St George's Respiratory Questionnaire; SOC, standard of care; 6MWD, 6 minute walking distance.

This structure was considered necessary at the time because of the immaturity of the survival data in the CAPACITY studies. However, with long term follow up data from RECAP, along with new survival data from the ASCEND trial, we now firmly believe the best structure of the model is a cohort-based partitioned survival model.

The micro-simulation model that was previously submitted, and accepted by NICE, has additional complexity than the preferred cohort simulation model. The ERG report from the TA282 appraisal of pirfenidone stated "*Although individual patient simulation is a valid modelling approach, the ERG highlights that these concerns could have been addressed with a cohort approach and the reason for not using this more common approach is not made clear*." [NICE 2016f].

In developing our evidence dossier, we considered whether the added complexity of the micro-simulation model provided any more precise estimates on the cost-effectiveness of pirfenidone. In reviewing both models, we do not believe the micro-

simulation model better estimates the cost-effectiveness, but does contribute considerably more complexity.

Selection of outcomes to include within the model

When selecting which outcomes to base the model structure upon the following factors were taken into consideration:

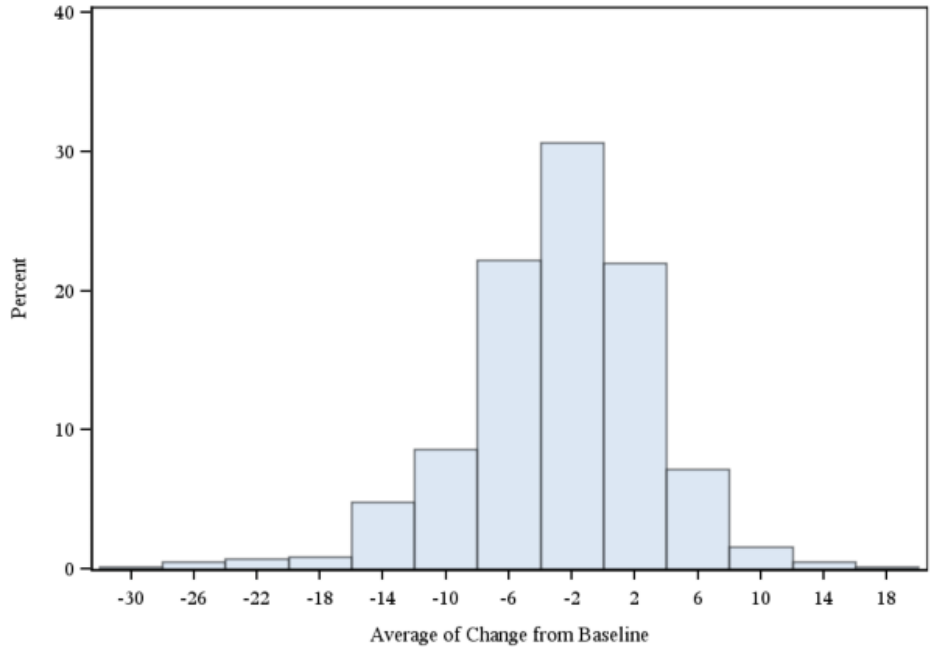
- Importance of the measure to the prognosis, costs and quality of life of IPF patients
- Correlation between measures
- Availability of data within the pirfenidone clinical trials
- Availability of data to allow comparison with both nintedanib and BSC
- Precedent in terms of previously accepted models

The previously accepted pirfenidone model utilised a model structure based upon both FVC and 6MWD. The nintedanib NICE manufacturer submission to NICE utilised a model structure comprising of FVC categories as model health states [Boehringer Ingelheim 2015]. Each percentage predicted FVC health state within the nintedanib NICE manufacturer submission considered a separate utility value, based on nintedanib Phase III trial post-hoc analysis. The Loveman model used health states based upon progression free survival [Loveman 2015].

Whilst FVC is an important clinical outcome, the quantity and quality of the evidence required to support the relatively complex approach within the nintedanib NICE manufacturer submission is insufficient. This is similar to the issues surrounding whether or not the additional complexity was justified for the microsimulation model within TA282. Specifically, the large number of health states requires a number of assumptions in order to incorporate transition probabilities.

In addition to the number of assumptions required for a model with this structure, patients would be expected to progress multiple times throughout the model. Pooled data from the pirfenidone Phase III trials suggests that the number of patients experiencing multiple progressions within a 12 month period is negligible, as shown in **Figure 41**. This further confirms the inappropriateness of considering a model with FVC-specific health states.

Figure 41: Histogram of change in FVC% predicted at 12 months



In IPF, FVC trends are the most accurate serial surrogate for mortality [Wells 2008b]. A decline in percentage predicted FVC of ≥10% is a decrement that has long been recognised as both clinically significant and highly predictive of mortality [Collard 2003, Flaherty 2003, Latsi 2003, Zappala 2010]. Smaller changes in FVC of 5-10% have also been shown to predict mortality [du Bois 2011a; du Bois 2011b; Zappala 2010; Taniguchi 2011], and the minimal clinically important difference established for FVC is between 2-6% [du Bois 2011b].

The study by du Bois et al. (2011) assessing the usage of percentage predicted FVC in IPF patients found that a decline in percent-predicted FVC was consistently greater for patients with larger declines in levels of physiologic function, functional status, and HRQoL [du Bois 2011a]. Equivalently, HRQoL is assumed to deteriorate over time, in accordance with a decline in percentage predicted FVC being associated with large declines in levels of HRQoL.

FVC has the additional benefit that it reflects the absolute state of a patient's condition, adjusted for body capacity; age, gender, and height removing the potential for some of the heterogeneity of the health-state members.

Comments received from stakeholders during the scoping consultation to this appraisal suggested that FVC alone may not provide sufficient representation of

disease course [NICE 2015e]. Based on these comments, two other potential surrogates for disease severity and progression in IPF were considered for use within the model:

- DLco
- 6MWD

DLco is a frequently reported outcome measure in IPF trials, but was not used in the model structure for the following reasons:

- Lack of comparable data
    - Four of the pirfenidone trials (CAPACITY I & II, SP3, SP2) reported data in relation to the change from baseline in DLco. The CAPACITY trials reported the change in % predicted DLco, while SP2 and SP3 reported the mean decline (mL/min/mmHG). Change from baseline in DLco was not reported in the ASCEND trial [King 2014a].
    - Absolute change from baseline in DLco was reported for the duration of the trial in the TOMORROW study, but for 52 weeks only in the INPULSIS trials [Boehringer Ingelheim 2015].
- Collineraity of DLco and FVC
    - When the DLco has been used in clinical trials, it has mostly been the absolute DLco measure without correction for the alveolar volume. Therefore, the DLco will tend to track with the FVC if it is used uncorrected for lung volumes, which raises the issue of collinearity between these two pulmonary function measurements [Nathan 2014].
- Serial gas transfer trends have a lower prognostic value in IPF and may be confounded by pulmonary vascular disease in systemic sclerosis [Wells 2008b].
    - Compared to FVC, DLco is more difficult to measure, requires a breath hold that can be difficult for more symptomatic patients and has greater intrinsic variability. The variability has commonly been recognised as being as high as 15%, which is the threshold that has typically been utilised to signify a significant change [Nathan 2014].
    - DLco is also a less consistent predictor of mortality. Coupled with the complexities of the procedure and variability in measurements between

laboratories, it is generally considered a less robust predictor of prognosis than FVC [du Bois 2011b; Flaherty 2003].

Based upon the available evidence, 6MWD is a more appropriate secondary measure of disease severity to include within the economic model. Within the original micro-simulation model, inclusion of 6MWD within predictive equations considerably improved model fit indicating its prognostic value. Additionally, evidence from the CAPACITY trials indicates that 6MWD measures different functional domains of the IPF disease process to FVC [Valeyre 2010a]. Decline in 6MWD has been shown to be predictive of mortality; a 24-week decline >50m is associated with a 3-fold increase in risk of death at 1 year [du Bois 2014, Raghu 2011]. Importantly, 6MWD has also been shown to be a predictor of mortality, independent of FVC [du Bois 2014].

Since there is increasing evidence indicating that the fundamental hallmarks of cancer biology are comparable to those of IPF, PFS (which is usually employed for lung cancer studies) could be an appropriate endpoint for IPF studies using predictive endpoints such as categorical changes in FVC and distance walked in 6MWD [Albera 2011].

The model structure presented within this submission was therefore chosen to be structured around PFS, using the definition decided by the EMA within the most recent clinical trial for pirfenidone (ASCEND). The definition of PFS in the ASCEND trial was patients who had not experienced any of the following events:

- Confirmed ≥10% absolute decline in percent predicted FVC
- Confirmed ≥50m decline in 6MWD
- Death

Whilst it is recognised that disease progression is variable from patient to patient, available literature indicates that declines in FVC≥10% and 6MWD≥50m are both prognostic markers for progressing disease, lower quality of life and increased mortality in the ensuing years. Therefore, it was decided to adopt a heath state structure similar to that adopted by Loveman et al. which included health states defined according to a patient's progression status (non-progressed vs. progressed) [Loveman 2014].

In addition the model has the option to include 'Lung transplant' as an additional health state. NICE guidance currently recommends lung transplantation as a treatment option for patients with IPF who do not have absolute contraindications [NICE 2013b]. Therefore, it is expected that some patients undergo lung transplantation if they are eligible. In practice, however, very few patients with IPF ever receive a transplant; between 2011 and 2012 there were 190 lung transplantations carried out in the UK, meaning that only a very small number of IPF patients eligible for a transplant could have received one [NHS 2013]. Lung transplantation is therefore included as an additional health state as sensitivity analysis.

The model structure also considers acute exacerbations as an important feature. Acute exacerbations of IPF are dramatic, singular events that are often fatal and a major cause of mortality and morbidity in IPF [Boehringer Ingelheim 2015]. As a result of this, acute exacerbations were included in the NMA, and consequently the model, to accurately establish the impact of treatment on the incidence of these events.

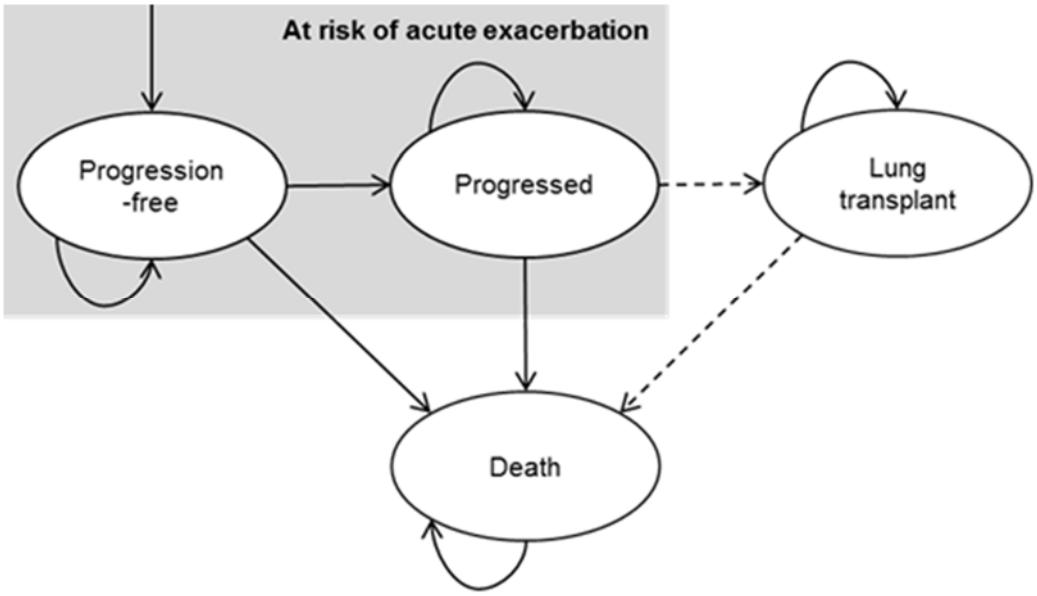
Model structure selected

The chosen model structure uses a cohort based partitioned survival model and was constructed in Microsoft Excel®. The model consists of the following health states:

- Progression-free
- Progressed
- Lung transplant
- Death

The possible routes patients may flow through the model are presented in Figure 42.

Figure 42: Model structure



Patients enter the model in the ‘Progression-free’ health state. Based on the risk of disease progression, patients can remain in the ‘Progression-free’ health state or move into the ‘Progressed’ or ‘Death’ health state. The rates of progression and the mortality have been derived from patient level data from the pivotal Phase III studies. Patients are able to transition to the ‘Death’ health state from any other health state. ‘Death’ is an absorbing state.

Within the sensitivity analysis, patients could only enter the lung transplant state from the progressed disease state and if they were under the age of 65. This assumption was in line with criteria patients must meet to be eligible for lung transplant in clinical practice.

Acute exacerbation was applied in the model as a risk per cycle based on current treatment received regardless of prior experience. The application of the risk of acute exacerbation per cycle allows the model to include initial and subsequent acute exacerbations, in line with the previously accepted nintedanib NICE manufacturer submission.

The key features of the de novo analysis are presented in **Table 68**.

Table 68: Features of the de novo analysis

Factor	Chosen values	Justification
Time horizon	34 years	Lifetime horizon – after this time <1% of patients are alive.
Cycle length	3 months	Assumption
Were health effects measured in QALYs; if not, what was used?	Yes	NICE reference case
Discount of 3.5% for utilities and costs	Yes	NICE reference case
Perspective (NHS/PSS)	NHS	NICE reference case

Key: NICE, National Institute of Health and Care Excellence; NHS, National Health Service; PSS, personal social services; QALYs, quality-adjusted life years.

Intervention technology and comparators

Pirfenidone is indicated in adults for the treatment of mild to moderate IPF and is currently recommended by NICE for the treatment of patients with FVC ≤80%. This review of NICE guidance was precipitated by the publication of data from the ASCEND trial, which confirmed the efficacy of pirfenidone in patients with FVC>80%. In line with the scope, comparison is therefore presented according to FVC subgroup.

NICE clinical guideline 163 on the diagnosis and management of suspected IPF recommends that BSC (including symptom relief, management of co-morbidities, withdrawal of therapies suspected to be ineffective or causing harm and end of life care) should be offered to people from diagnosis and be tailored according to disease severity, rate of progression and the person's preference [NICE 2013b].

Nintedanib has recently gained NICE approval for the treatment of adults with moderate IPF, hence it is included as the primary comparator for patients in this subgroup of the population [NICE 2016]. Nintedanib is available in 100mg and 150mg capsules. The recommended dose of nintedanib is 150 mg twice daily reduced to 100mg twice daily if not tolerated [EMC 2015b].

The model therefore compares pirfenidone across its marketing authorisation (adults with mild to moderate IPF) with the relevant comparator based on FVC subgroup, as shown in **Table 69**.

Table 69: Comparators included in the de novo analysis

Subgroup	Comparator	Justification
ITT	BSC	NTB is not recommended by NICE for the treatment of mild IPF
Mild (FVC >80%)	BSC	The only treatment currently recommended in this subgroup is BSC
Moderate (FVC 50 – 80%)	NTB, BSC	NTB is currently recommended by NICE for the treatment of moderate IPF, and therefore would be displaced by PFN

Key: BSC, best supportive care; FVC, forced vital capacity; IPF, idiopathic pulmonary fibrosis; NICE, National Institute of Health and Care Excellence; NHS, National Health Service; NTB, nintedanib; PFN, pirfenidone.

Pirfenidone is available in 267mg capsules, and is administered orally. Upon treatment initiation, the dose should be titrated to the recommended daily dose of nine capsules per day over a 14-day period as follows:

- Days 1 to 7: one capsule, three times a day (801 mg/day)
- Days 8 to 14: two capsules, three times a day (1,602 mg/day)
- Day 15 onward: three capsules, three times a day (2,403 mg/day)

The recommended daily dose of pirfenidone for patients with IPF is three 267 mg capsules three times a day with food for a total of 2,403 mg/day [EMC 2015a].

Based on prevailing NICE guidance, treatment with both pirfenidone and nintedanib may be terminated via the application of a stopping rule: treatment should be stopped if a patient's predicted FVC declines by 10% or more in any 12 month period [NICE 2013a, NICE 2016].

Patients who are not eligible for, cannot tolerate treatment with, or experience the 10% drop in 12 months whilst on pirfenidone or nintedanib, currently receive BSC suggesting a high unmet need for IPF patients. Imposing the stopping rule may deny treatment to patients who could derive a morbidity/mortality benefit, as there is no information to indicate that these benefits are limited to patients whose lung function declines at a slower rate (see Section 4.8) [NICE 2015e, Nathan 2015, Nathan 2016]. Therefore, the stopping rule is not applied in the model base-case to patients treated with pirfenidone for both mild and moderate IPF.

The stopping rule is applied for patients with moderate IPF treated with nintedanib in the model base-case in line with NICE guidance on the usage of nintedanib within the NHS.

5.3 Clinical parameters and variables

Patient level data were taken from the trials given in Table 70 to inform the clinical parameters and variables used in the de novo analysis. Further information regarding these trials are presented in depth in Section 4. Registry data were also used for patients receiving BSC, with details presented in Table 71.

Table 70: Sources of clinical parameters and variables: clinical trials

Clinical evidence	Brief description	Use in the model
CAPACITY [Noble 2011]	CAPACITY I and II (CAPACITY) were Phase III, double-blind, randomised, placebo-controlled, multicentre studies evaluating the safety and efficacy of pirfenidone in IPF.	Patient level data were used to fit OS and TTD parametric curves using CAPACITY, ASCEND and RECAP data based upon the 7 year data cut from RECAP PFS data were taken from the CAPACITY and ASCEND trials as PFS was not measured in RECAP Used in NMA for model outcomes: ≥10% categorical decline in FVC % predicted All-cause discontinuation of treatment All-cause mortality Acute exacerbations Patient characteristics were used for the average age of patients entering the model. Patient level data from CAPACITY I and II were used to model quality of life based upon published mapping algorithms for SGRQ.
ASCEND [King 2014]	ASCEND was a Phase III, double-blind, randomised, placebo-controlled, 52-week study designed to provide additional evidence of the effect of pirfenidone on disease progression in IPF.	
RECAP [Costabel 2012]	RECAP is an open-label extension study designed to assess the long-term safety of pirfenidone 2403 mg/day in patients with IPF who received ≥80% of scheduled doses and completed the Week 72 final study visit in either of the two CAPACITY studies. Patients in the ASCEND study were also eligible to roll-over into RECAP.	
SP3 [Taniguchi 2010]	SP3 was a Phase III, double-blind, randomised, placebo-controlled, multicentre study evaluating the efficacy and safety of pirfenidone in IPF patients in 73 centres in Japan.	Used in NMA for model outcomes: ≥10% categorical decline in FVC % predicted 6MWD (TOMORROW) All-cause discontinuation of treatment All-cause mortality Acute exacerbations
INPULSIS [Richeldi 2014]	INPULSIS I and II (INPULSIS) were multicentre, randomised, double-blind, placebo-controlled replicate Phase III studies to evaluate the efficacy and safety of treatment with 150 mg of nintedanib, twice daily in patients with IPF.	
TOMORROW [Richeldi 2011]	TOMORROW was a multicentre, randomised, double-blind, placebo-controlled, Phase II study to evaluate the efficacy and safety of four different dose strategies of nintedanib treatment for 52 weeks in IPF.	

Clinical evidence	Brief description	Use in the model
Key: FVC, forced vital capacity; IPF, idiopathic pulmonary fibrosis; NMA, network meta-analysis; OS, overall survival; PFS, progression-free survival; TTD, time to treatment discontinuation; 6MWD, 6-minute walking distance.		

Table 71: Sources of clinical parameters and variables: registry data

Registry	Brief description (see Section 4.11 for full details)	Use in the model
INOVA	US IPF registry n=286 following matching process Good match in terms of patient characteristics to the CAPACITY and ASCEND trials	Validation of long-term survival outcomes Sensitivity analysis regarding the shape of long-term survival Data from INOVA is presented as primary comparison as the larger dataset with the better overlap in terms of patient characteristics
Edinburgh	UK IPF registry n=182 following matching process Reasonable match in terms of patient characteristics to the CAPACITY and ASCEND trials	

In the model, efficacy data for the pirfenidone arm were taken from the pivotal Phase III clinical trials (CAPACITY I, CAPACITY II and ASCEND), along with SP3 and RECAP. Based on the committee's consideration / conclusions in the second nintedanib NICE appraisal committee meeting held in October 2015, SP2 was excluded as it was considered to be an outlier [NICE 2016].

Standard parametric survival analysis was used to fit curves for the pirfenidone arm using the clinical trial data. To select the most appropriate curve fit, goodness of fit statistics were used as well as visual inspection and assessment of long-term plausibility. The process of fitting standard parametric survival curves was conducted in accordance to the NICE DSU TSD 14 [NICE DSU 2011].

For the BSC and nintedanib arms, efficacy inputs were derived from an NMA since no head-to-head RCT data were available for the direct comparison of nintedanib and pirfenidone. The NMA was incorporated in the model to inform the following key features:

- Overall survival
- Progression-free survival
- Time to treatment discontinuation
- Acute exacerbations

The NMA base case ("base case network, random effects model") is presented in Section 4.10. In the base-case, the NMA considers a random-effects model including all Phase II and Phase III trials. Alternative scenarios considered for the NMA are presented in Appendix 14. Additional clinical parameters and variables (such as IPF-related deaths) were also incorporated using outcomes separate to the NMA. These are discussed in the relevant sections below.

Overall survival

Log cumulative hazard and residual plots were inspected and the results of a log-rank test conducted determining whether there was a significant interaction between treatment effect and time considered. The results of these analyses are presented in

Appendix 20. Based upon the results of these analyses, the assumption of proportional hazards between pirfenidone and BSC was supported.

Parametric survival curves were therefore fit to the pirfenidone data and OS estimated for both comparators within the base-case using the HR from the NMA.

The curve fits applied to the OS data and their associated goodness of fit statistics are presented in **Table 72**.

Table 72: Overall survival – goodness of fit statistics

Model	AIC	BIC
EXPONENTIAL	865.47	878.77
WEIBULL	**844.15**	**861.89**
LNORMAL	853.23	870.97
GAMMA	845.78	867.95
LLOGISTIC	844.54	862.28
GOMPERTZ	851.70	869.44

Key: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion.

The Weibull distribution was selected for the model base-case based upon both visual and statistical goodness of fit to the observed portion of the pirfenidone survival curve.

Median survival for patients with IPF is stated as 3 years from diagnosis [Navaratnam 2011]. However, the clinical course of IPF is variable and consequently this median survival does not reflect the variability seen in subgroups where both slowly progressive and rapidly progressive patient types are recognised. For patients staged using GAP severity assessment survival at 3 years is >80% for Stage I, >50% for Stage II and <25% for Stage III. Patients classified as mild-moderate IPF will generally have a median survival in excess of 3 years. Compared to the longer term data available for BSC, the Weibull curve follows closely the observed data from all three registries until approximately 10% of patients are still alive, at which time the Kaplan-Meier curves become less reliable (**Figure 52**). Registry data appears to indicate a flattening of the survival curves at this time, however, due to

low patient numbers to support longer tails the more conservative Weibull curve was selected for the model base-case.

Alternative curve fits were explored as sensitivity analyses, and are presented in Appendix 21.

As discussed in Section 4.4, reliable data are only available for BSC for 52 weeks from the 3 clinical trials (ASCEND, CAPACITY I and II).

For the BSC arm, three options are therefore available for OS:

- Application of a hazard ratio to the OS for patients receiving pirfenidone based upon the results of the NMA.
- Incorporation of trial data up until 52 weeks after which a hazard ratio is applied based upon the results of the NMA.
- Use of real-world data from the INOVA or Edinburgh registries.

In the base-case, a hazard ratio is applied to the pirfenidone arm to inform overall survival. Alternative options for the OS of BSC patients are explored as sensitivity analyses.

For the nintedanib arm, a hazard ratio is applied to the OS of pirfenidone patients based upon the results of the NMA. The base-case settings used in the model for OS are summarised in **Table 73**.

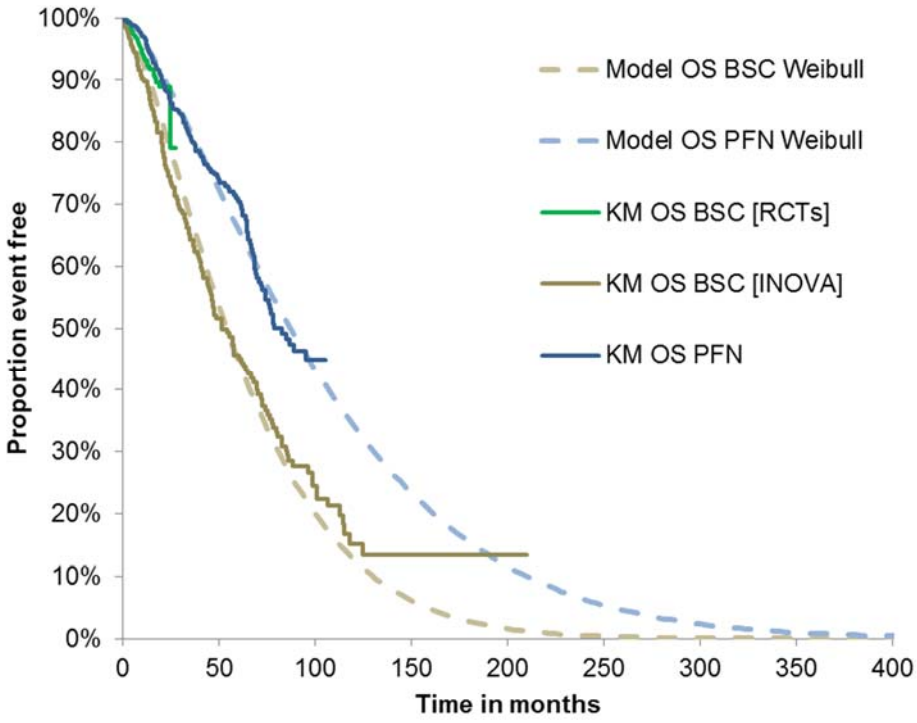
Table 73: Overall survival – base-case settings

Treatment	Model setting	Base-case HR (PFN vs comp)
PFN	Standard parametric survival analysis applied to trial data	
NTB	HR applied to PFN efficacy	0.72
BSC	HR applied to PFN efficacy	0.52

Key: BSC, best supportive care; comp, comparator; HR, hazard ratio; PFN, pirfenidone; NTB, nintedanib.

Figure 43 shows OS for pirfenidone and BSC in the ITT population as it is applied in the model base case. Also shown are the Kaplan-Meier plots for the trial data available, as well as the registry data used to validate the long-term outcomes for patients on BSC.

Figure 43: Overall survival – curve fits applied in model



Key: BSC, best supportive care; KM, Kaplan-Meier; OS, overall survival; PFN, pirfenidone; RCT, randomised controlled trial.

IPF-related mortality

To incorporate the cost for end of life care into the model, the proportion of deaths related to IPF for patients receiving pirfenidone were taken from pooled clinical data from CAPACITY and ASCEND trials. IPF-related deaths were derived for patients receiving BSC and nintedanib using the NMA (random-effects model including all Phase II and III trials; see Section 4.10).

Within the model deaths were apportioned into those related to IPF or those unrelated to IPF, since costs associated with IPF are greatly increased in the last year of life due to increased resource use, home care and length of stay in hospital.

The proportion of deaths related to IPF for each treatment are presented in **Table 74**.

Table 74: IPF-related deaths

Treatment	Proportion of deaths IPF-related	Reference
PFN	57.89%	Pooled trial data
BSC	72.22%	NMA
NTB	68.57%	
Key: BSC, best supportive care; IPF, idiopathic pulmonary fibrosis; NMA, network meta-analysis; NTB, nintedanib; PFN, pirfenidone.		

Progression-free survival

As outlined in Section 4, PFS was defined differently between the CAPACITY and ASCEND studies:

- In the CAPACITY studies, disease progression was defined as time to confirmed ≥10% decline in percentage predicted FVC or ≥15% decline in percentage predicted DLCO or death.
- In the ASCEND study, disease progression was defined as time to confirmed ≥10% decline in percentage predicted FVC or ≥50m decline in 6MWD or death.

As 6MWD was considered the more robust outcome for inclusion within the model (Section 5.2.3), a common definition of PFS was applied to the pooled ASCEND/CAPACITY data, which aligned with the ASCEND criteria, classifying disease progression as having experienced a ≥10% decline in FVC or a ≥50m decline in 6MWD, or death.

Log cumulative hazard plots were inspected and the results of a log-rank test conducted determining whether there was a significant interaction between treatment effect and time considered. Based upon the results of these analyses the assumption of proportional hazards between pirfenidone and BSC was supported.

The curve fits applied to the PFS data and their associated goodness of fit statistics are presented in **Table 72**.

Table 75: Progression-free survival – goodness of fit statistics

Model	AIC	BIC
EXPONENTIAL	2422.29	2432.53
WEIBULL	**2342.93**	**2358.28**
LNORMAL	2312.07	2327.43
GAMMA	2312.73	2333.21
LLOGISTIC	2329.21	2344.57
GOMPERTZ	2382.81	2398.16

Key: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion.

Consistent with the fit selected for OS, the Weibull distribution was selected for the model base-case based upon both visual and statistical goodness of fit to the observed portion of the pirfenidone PFS curve. Alternative curve fits were explored as sensitivity analyses, and are presented in Appendix 21.

For the BSC and nintedanib arms, a hazard ratio is applied to the PFS of pirfenidone patients based upon the results of the NMA. The base-case settings used in the model for PFS are summarised in **Table 76**.

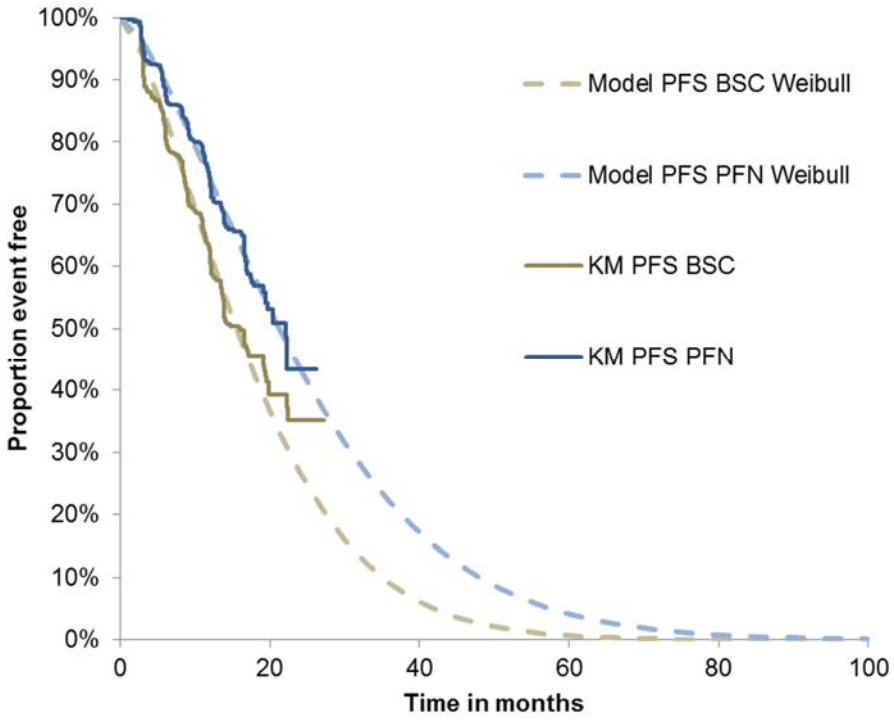
Table 76: Progression-free survival – base-case settings

Treatment	Model setting	Base-case HR (PFN vs comp)
PFN	Standard parametric survival analysis applied to trial data	
NTB	HR applied to PFN efficacy	0.85
BSC	HR applied to PFN efficacy	0.63

Key: BSC, best supportive care; comp, comparator; NTB, nintedanib; HR, hazard ratio; PFN, pirfenidone.

Figure 44 shows PFS for pirfenidone and BSC in the ITT population as it is applied in the model. Registry data are not shown as PFS and/or FVC decline were not available in the registries.

Figure 44: Progression-free survival – curve fits applied in model



Key: KM, Kaplan-Meier; PFS, Progression-free survival; PFN, Pirfenidone; BSC, Best supportive care.

Time to treatment discontinuation

In each cycle, the proportion of patients on and off treatment are calculated based upon curves fitted to patient level data from ASCEND/CAPACITY/RECAP. Those who had discontinued therapy with pirfenidone without dying (or receiving a lung transplant in the sensitivity analysis) were assumed to continue to receive BSC. Since deaths and lung transplant were captured independently in the model, discontinuations relating to death and lung transplant were censored in the Kaplan Meier estimator.

The curve fits applied to the TTD data and their associated goodness of fit statistics are presented in **Table 77**.

Table 77: Time to treatment discontinuation – goodness of fit statistics

Model	AIC	BIC
EXPONENTIAL	1642.15	1655.46
WEIBULL	**1632.51**	**1650.25**
LNORMAL	1666.42	1684.16
GAMMA	1633.82	1655.99
LLOGISTIC	1644.32	1662.06
GOMPERTZ	1637.31	1655.04

Key: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion.

Consistent with the fit selected for OS and PFS, the Weibull distribution was selected for the model base-case based upon both visual and statistical goodness of fit to the observed portion of the pirfenidone curve. Alternative curve fits were explored as sensitivity analyses, and are presented in Appendix 21.

TTD for nintedanib was derived using the relative risk for all cause treatment discontinuation applied to the TTD of pirfenidone patients based upon the results of the NMA.

Additionally, stopping rules may be applied to both, either or neither treatment arm. The stopping rule currently included within NICE guidance for both pirfenidone and nintedanib is that treatment is discontinued for patients who experience predicted FVC declines by 10% or more in any 12 month period.

The same discontinuation rule was highly criticised by clinicians at an advisory board held by Boehringer Ingelheim on 23rd April 2014, as it was considered difficult to impose [Boehringer Ingelheim 2015]. These opinions were reiterated by clinicians during the Appraisal Committee meeting for nintedanib in August 2015, the consultation responses to the ACD, and the scoping consultation for this re-appraisal of pirfenidone [NICE 2015d, NICE 2015e]. Consequently, the stopping rule is not applied for pirfenidone in the model base-case, but is applied for nintedanib in line with current NICE guidance [NICE 2016].

Due to the absence of comprehensive stopping rule data, the stopping rule was applied in the model using tunnel states to estimate the proportion of patients who

progress within 12 months. Patients who progress within 12 months will no longer be on treatment, and will therefore not incur treatment costs for nintedanib.

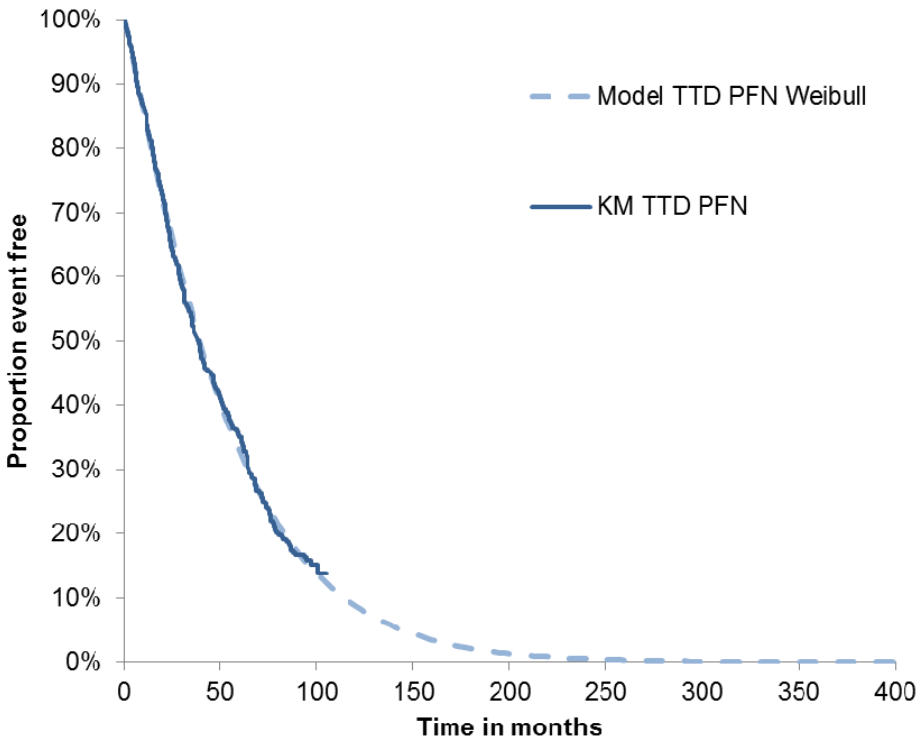
The base-case settings used in the model for TTD are summarised in **Table 78**.

Table 78: Time to treatment discontinuation – base-case settings

Treatment	Model setting	Base-case RR (PFN vs NTB)
PFN	Standard parametric survival analysis applied to trial data	
NTB	RR applied to PFN efficacy – stopping rule applied separately	1.08
Key: NTB, nintedanib; PFN, pirfenidone; RR, relative risk.		

Figure 45 shows TTD for pirfenidone in the ITT population as it is applied in the model.

Figure 45: Time to treatment discontinuation – curve fits applied in model



Key: KM, Kaplan-Meier; PFN, pirfenidone; TTD, time to treatment discontinuation.

Acute exacerbations

Because of its clinical significance, acute exacerbation of IPF has become a key endpoint in clinical trials of novel drug agents for IPF [Collard 2013]. Acute exacerbations are termed periods of acute deterioration in respiratory function when a cause cannot be identified [Kim 2013]. The natural history of IPF is highly variable amongst individuals. Some patients will experience relatively stable periods with bouts of acute exacerbations, which require hospitalisation for respiratory failure [Ley 2011]. As a result of this, acute exacerbations were incorporated into the model.

The pathological findings in patients with acute exacerbation of IPF represent a variety of pathological conditions, making a diagnosis of acute exacerbation of IPF difficult [Oda 2014].

To ensure consistent capture of an acute exacerbation, head-to-head acute exacerbation data would be required, where a common definition could be applied. Since these data are not currently available, but recognising the potential significance of this outcome, the NMA included acute exacerbations in order to estimate the approximate frequency of events across treatments. However, as highlighted in Section 4.7, the results of these comparisons should be interpreted with caution, as trials providing data to inform the NMA used different definitions for acute exacerbation.

The baseline risk of acute exacerbations was applied for patients on BSC. These patients had a 1.46% chance of experiencing acute exacerbation per model cycle (3 months). This was calculated using the following equation (taken from the nintedanib NICE manufacturer submission):

$$1 - EXP(-EXP(-constant) * cycle\ length\ in\ days)$$

The constant figure was taken from the adjudication committee reported exacerbation table from the STATA analysis conducted in the nintedanib NICE manufacturer submission. The full table from the submission is shown in **Table 79**.

Table 79: STATA survival analysis output – time to acute exacerbation (Table 125 in the nintedanib NICE manufacturer submission)

Investigator reported exacerbation						
Variable	**Coef.**	**Std. Err.**	**z**	**P>z**	**95% Conf. Interval**	
cons	8.430	0.177	47.69	<0.001	8.083	8.776
Adjudication committee reported exacerbation						
Variable	**Coef.**	**Std. Err.**	**z**	**P>z**	**95% Conf. Interval**	
cons	8.724	0.204	42.74	<0.001	8.324	9.124

An odds ratio based upon the results of the NMA was then applied to patients on pirfenidone and nintedanib. The risk of acute exacerbation by treatment is presented in **Table 80**.The incidence of acute exacerbation is hard to establish due to variation in the methodology used for assessment in different studies, but it is believed to occur in between 5 and 10% of patients with IPF every year (in line with our model estimate for baseline risk of 1.46% per model cycle) [Kim 2013].

Table 80: Acute exacerbation incidence

Treatment	Model setting	Base-case OR (BSC vs comp)
BSC	Baseline risk of 1.46% per model cycle	-
PFN	OR applied for patients on PFN	0.62
NTB	OR applied for patients on NTB	0.55

Key: BSC, best supportive care; comp, comparator; NTB, nintedanib; OR, odds ratio; PFN, pirfenidone.

Lung transplantation

Interstitial lung disease, and in particular IPF, carries the worst prognosis among the common disease indications for lung transplantation [Weill 2015]. However, it is expected that some IPF patients undergo lung transplantation if they are eligible. Evidence from the International Society for Heart and Lung Transplantation suggests that an age of >65 years would be considered to be a relative contraindication for lung transplantation, in association with low physiologic reserve and/or other relative contraindications [Weill 2015].

The numbers presented in **Table 81** were used in the model to estimate the uptake and outcomes of patients undergoing lung transplantation in this population. These cycle-specific probabilities were applied to the proportion of patients who were aged <65 for each model cycle (taking into account the aging of patients over time). In the model base-case, the 'Lung transplant' health state is not included, and is instead explored as sensitivity analysis.

Table 81: Parameters for assessing the probability of lung transplant

Parameter	Value	Source
Number of patients currently diagnosed with IPF	15,000	Navaratnam 2011
Number of transplants between 2012-2013		NHS Blood and Transplant 2012/13 [NHS 2013]
Single lung transplants	28	
Double lung transplants	159	
Heart/lung transplants	3	
Total	190	
Proportion of lung transplants in IPF patients	28.6%	Mendonca 2013
Proportion of patients aged <65 upon entering the model	35.28%	ASCEND/CAPACITY
Annual probability of a lung transplant in IPF patients <65 years of age	1.03%	= (190*28.6%) / (15000*35.47%)
Quarterly probability of a lung transplant in IPF patients <65 years of age	0.26%	= 1 - EXP(LN(1-1.02%) / 4)
5-year probability of death following lung transplantation	61%	Thabut 2003 (5-year survival = 39%)
Quarterly probability of death following lung transplantation	4.60%	= 1 - EXP(LN(1-61%) / 20)
Key: IPF, idiopathic pulmonary fibrosis.		

Adverse event rates

Adverse events (AEs) were chosen as per the nintedanib NICE manufacturer submission previously accepted by NICE and the ERG, who considered the company's approach to the inclusion of AEs in the economic model as reasonable and justified [NICE 2015f]. In summary, the inclusion criteria for AEs used in this submission was as follows:

AEs were selected to be incorporated into the model based on an AE satisfying all of the following criteria in at least one of the clinical studies considered:

- AEs with a significant impact on costs and QALYs: assumed to be those that were severe or serious
- AEs with an incidence of greater than or equal to 5%
- AEs with an incidence of 1.5 times greater between the two arms

In addition to these criteria, the following AEs were of particular interest to clinicians, and hence were implemented into the model regardless of the above criteria:

- For pirfenidone: photosensitivity and rash
- For nintedanib: gastrointestinal perforation

The observed rates of incidence for these adverse events are presented in **Table 82**.

Table 82: Adverse event incidence rates per model cycle

Adverse event	Incidence (per cycle)	
Serious cardiac event	PFN	0.0176
	BSC	0.0139
	NTB	0.0129
Serious gastrointestinal event	PFN	0.0025
	BSC	0.0042
	NTB	0.0098
Gastrointestinal perforation	NTB	0.0008
Photosensitivity reaction	PFN	0.0232
Rash	PFN	0.0679
Key: BSC, best supportive care; NTB, nintedanib; PFN, pirfenidone.		

The cost of treating AEs is applied per model cycle according to the incidence. These costs are discussed further in Section 5.5.

5.4 Measurement and valuation of health effects

Health-related quality-of-life data from clinical trials

Patients with IPF suffer progressive breathlessness, and may eventually become dependent on supplementary oxygen. As the disease progresses, a patients' ability

to carry out day to day activities deteriorates and patients become more dependent on family and carers. It is therefore expected that patients HRQoL will worsen throughout the course of their illness.

Tomioka et al. (2007) found a significant decline in HRQoL in 2 sub-scales of the SF-36 physical function ($p<0.001$) and bodily pain ($p<0.05$), during a 12 to 31 month follow-up of IPF patients [Tomioka 2007]. A decline in HRQoL has been correlated with disease progression; one study which analysed data from the BUILD (the Bosentan Use in ILD-1 or BUILD-1) clinical trial found subjects whose clinical status changed most had the greatest changes (in the appropriate direction) in SF-36 and SGRQ scores; subjects whose clinical status did not change had essentially no change in HRQoL scores; and subjects whose clinical status changed minimally had minimal changes in HRQoL scores [King 2008].

Accordingly, HRQoL is assumed to deteriorate over time in the model, and is accounted for by the change in utility from the 'Progression-free' to the 'Progressed' health state and by the decrease in utility for patients experiencing acute exacerbations.

Section 4.7.3 details the HRQoL data collected in the clinical trials. SGRQ and WHO-QOL instruments were used and measured every 12 weeks in the CAPACITY trials. Both instruments were consistent with the reference case in that the source of data for measurement of HRQoL was reported directly from the patients of the CAPACITY trials.

No EQ-5D data were available from the CAPACITY or ASCEND trials, which is the preferred measure for NICE of HRQoL. When EQ-5D data are not available, NICE guidance recommends that these data may be estimated by mapping other health-related quality of life measures or health-related benefits observed in the relevant clinical trial(s) to EQ-5D, thus a mapping study was employed.

Prior to mapping a generalised estimating equation (GEE) a regression model was developed in STATA to obtain the mean SGRQ score for patients in each model health state (pre- and post-progression). GEE regression was used as this method

accounts for potential autocorrelation of patient quality of life scores. The results of the analysis are presented in **Table 83**.

Table 83: GEE model for SGRQ by model health state

Health state	Coefficient	SE	p-value	95% CI
Progression-free	37.30935	0.6360919	0.000	[36.06264, 38.55607]
Progressed	5.095259	0.3099897	0.000	[4.487691, 5.702828]

Key: SE, standard error; CI, confidence interval.

The analysis gives SGRQ values of 37.31 and 42.40 for the pre- and post-progression model health states, respectively.

Mapping

For the model base-case, a mapping study by Freemantle et al. was chosen [Freemantle 2015].

The study utilised data from a double-blind multicentre study conducted in England and Wales of 181 IPF patients. In total, 202 pairs of data were collected recording both SGRQ and the EQ-5D-3L. The model considered explanatory variables such as age, BMI, sex and FVC.

Generalised mixed models, accounting for repeated measurements within subjects as residual or generalised random effects, were fit to the data. The following explanatory variables were evaluated to aid model fit: age; gender; BMI; FEV1; percentage predicted FVC, and; FEV1/FVC ratio. Additionally, the potential non-linearity of the relationship between explanatory and response variables was explored by investigating transformations and fitting restricted cubic splines with four knot points. Root mean square error (RMSE), AIC and BIC were used to determine the best fitting model.

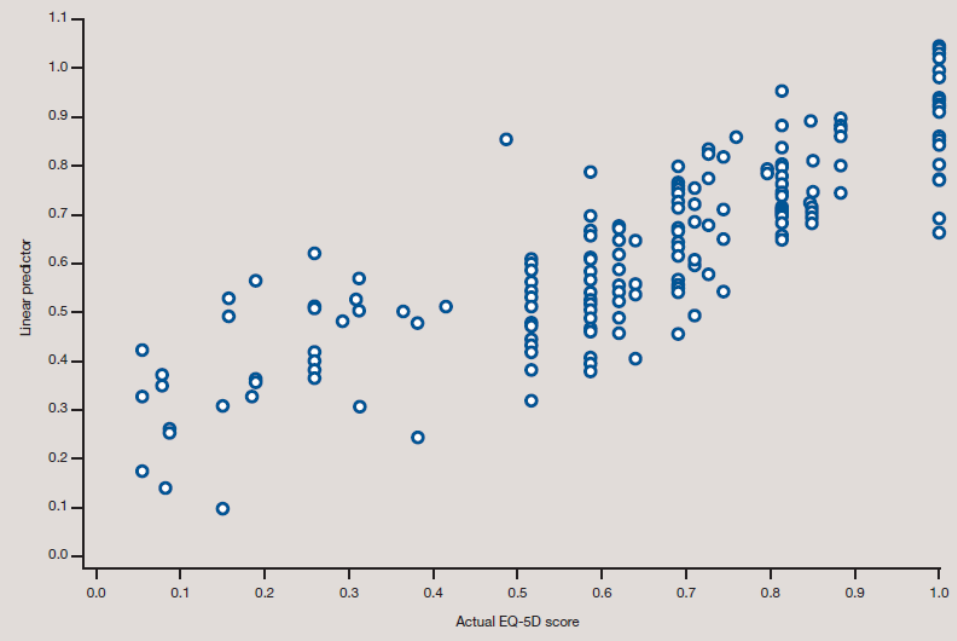
The generalised random effects (GRE) model was determined to be the best fitting model, which had a RMSE of 0.1391. The GRE model produced the following final mapping algorithm:

$$\boldsymbol{EQ\text{-}5D} = 1.3246 - 0.01276 * \boldsymbol{SGRQ}$$

Key: EQ-5D, EuroQoL 5-Dimension; SGRQ, St George's respiratory questionnaire

Inspection of the scatter plot showed that EQ-5D utilities predicted by the mapping algorithm closely approximated observed values from the study, shown in **Figure 46**.

Figure 46: Scatter plot of observed versus predicted EQ-5D values



Key: EQ-5D, EuroQoL 5-Dimension

The previous pirfenidone NICE manufacturer submission utilised a different mapping study by Starkie et al. [Starkie 2011]. This mapping study was performed in Chronic Obstructive Pulmonary Disease (COPD) patients from the TORCH (Towards a Revolution in COPD Health) trial. The trial collected EQ-5D and SGRQ between weeks 24 to 3 years. The best mapping equation as measured by goodness of fit was a simple OLS model, which had a RMSE of 0.1723. The OLS model produced the following final mapping algorithm:

$$\boldsymbol{EQ\text{-}5D} = 0.9617 - 0.0013 * \boldsymbol{SGRQ} - 0.0001 * \boldsymbol{SGRQ}^2 + 0.0231 * \boldsymbol{Male}$$

Validation of the mapping equation was conducted within the mapping study; QALYs derived using the entire EQ-5D dataset were compared to QALYs derived using the predicted EQ-5D from the mapping equation. QALYs estimated using the mapping equation were slightly higher than the QALYs using the EQ-5D data directly, but the results were largely comparable [Starkie 2011].

The study by Starkie et al. (2011) was conducted in patients with COPD (versus IPF patients), and was associated with larger RMSE (0.1723 versus 0.1391). Therefore, the model utilises the Freemantle et al. (2015) study in the model base-case, with the Starkie et al. (2011) study explored as sensitivity analysis [Freemantle 2015]. Using this mapping study, utility values were derived for the 'Progression-free' and 'Progressed' health states.

Within sensitivity analysis, the utility value assigned to the lung transplant state was assumed equivalent to patients in the progression-free disease state, since there was sparse literature to inform lung transplant utilities in IPF. However, it would appear that a successful lung transplant may result in a similar if not improved quality of life to progression-free disease based on a study by Groen et al. evaluating the cost-effectiveness of lung transplantation in relation to type of end-stage pulmonary disease. In this study, average utility values ranged from 0.69 (1 month after transplantation) to between 0.83 and 0.85 (from 3 to 12 months after transplantation). During the next year, utility rose to 0.91 [Groen 2004].

Health-related quality-of-life studies

A systematic review was conducted to identify relevant health related quality of life data for adult patients with mild to moderate IPF in England. Data presented in economic evaluations, utility elicitation studies, RCTs and technology assessments were eligible for inclusion in the review and reviews were mined for references. Given the extent of the previous searches in the 2011 pirfenidone submission, only studies published from 2010 onwards were screened [InterMune 2011].

A single search strategy was used to identify cost effectiveness studies (Section 5.1.1), health related quality of life and resource use data (Section 5.5.2). Full strategies are provided in Appendix 17.

The study selection process is described in Section 5.1.1. Detailed data on the study design and utility outcomes were extracted. Quality assessment was conducted by one researcher and checked by a second. The quality and relevance of each utility report and utility elicitation study was assessed by considering a range of issues:

- Selection of participants – was there clear inclusion and exclusion criteria?
- How representative of the target population were the participants in the study, considering age, disease severity, co-morbidities?
- Was a reasonable pilot testing approach used?
- Were the utility elicitation methods reasonable/valid?
- Have the researchers tried to reduce potentials for bias, for example interviewer bias?
- Have the researchers offered assessments of the limitations of the study approach?

In the combined searches for cost-effectiveness studies, resource use and costs in IPF and utilities in IPF, 5924 records were retrieved from database searching and 3474 of these records were published from 2010 onwards. 215 records were found from other sources. After removing duplicates, 3123 records were assessed. 22 relevant studies were found in the health related quality of life review. Appendix 17 presents the study selection process.

Description of Identified Studies

A detailed analysis of the 22 eligible studies is presented in Appendix 22.

4 studies reported EQ-5D scores for patients with IPF and were reported in 5 papers [Boehringer Ingelheim 2015 , King 2011b, Pittrow 2010, Richeldi 2014, Zisman 2010]. EQ-5D data were reported in the 2015 STA for nintedanib for the treatment of adults with IPF from Richeldi et al. but which are not directly available from that article [Boehringer Ingelheim 2015 , Richeldi 2014].

Richeldi et al. was based on 2 multinational RCTs that compared nintedanib with placebo [Richeldi 2014]. The study had a large sample of 1,066 patients from 24 countries. EQ-5D data and other outcomes (including SGRQ) were reported. Utility decrements associated with AEs were also considered.

Pittrow et al. analysed a prospective registry to assess HRQoL using EQ-5D in 421 patients through time trade off (TTO) [Pittrow 2010]. SGRQ outcomes were also reported. These data were reported in an abstract with limited reporting on data and

methods. King et al. reported a RCT comparing bosentan and placebo in 616 patients where SF-36 was used to assess HRQoL [King 2011b]. Finally, Zisman et al. used the EQ-5D questionnaire in 180 patients randomised to receive either sildenafil or placebo [Zisman 2010].

Of the remaining 17 studies, 12 administered a generic instrument and 13 used a disease-specific questionnaire. None of the studies mapped these outcomes to EQ-5D (see Appendix 18).

3 of the 4 studies that used EQ-5D scores for patients with IPF were based on data from RCTs. Richeldi et al. plus the 2015 NICE nintedanib STA [12] were based on the combination of two RCTs (INPULSIS-1 and INPULSIS-2) [Boehringer Ingelheim 2015 , Richeldi 2014]. EQ-5D results (with standard deviation) from this reference study are shown in Table 84.

Table 84: EQ-5D utility values by FVC percentage predicted

FVC percentage predicted	Mean EQ-5D utility	SD
90 and above	0.8380	0.1782
80 to 89.9	0.8105	0.2051
70 to 79.9	0.7800	0.2244
60 to 69.9	0.7657	0.2380
50 to 59.9	0.7387	0.2317
40 to 49.9	0.6634	0.2552

King et al. was a RCT that provided reference data in a sample of patients with mild to moderate disease (Table 85) [King 2011b].

Table 85: Changes from EQ-5D at baseline to 1 year

	Bosentan		Placebo	
	Mean	**SD**	**Mean**	**SD**
Baseline	0.758	0.185	0.718	0.242
1 year	0.660	0.386	0.656	0.366

Finally, Pittrow et al used a prospective registry with a population of patients with idiopathic lung fibrosis and a mean % of predicted FVC of 72 ± 20 [Pittrow 2014].

The mean EQ-5D TTO score was 0.8 ± 0.2, consistent with reference data from RCTs.

Adverse reactions

Overall treatment compliance in the CAPACITY and ASCEND trials was high: the CAPACITY studies reported 380 (88%) of 432 patients in the pirfenidone groups and 323 (93%) of 347 in the placebo groups adhered to treatment (i.e. received ≥80% of scheduled doses). It is therefore expected that treatment-related AEs such as photosensitivity reactions and gastrointestinal side effects such as nausea are manageable and would not impact significantly on HRQoL, particularly in the context of a disease where median survival is 2-5 years [Meltzer 2008.

However, consistent with the nintedanib NICE manufacturer submission previously accepted by NICE, AE-related disutilities were incorporated into the model.

The ERG for the nintedanib appraisal stated that the duration used for AE disutility was for one year, and considered that the duration of the AE would be significantly less than one year for GI events and skin disorders [NICE 2015f]. However, the duration of disutilities was assumed to be one model cycle (3 months) in the nintedanib NICE manufacturer submission [Boehringer Ingelheim 2015].

The ERG noted that based upon pirfenidone clinical trial data, rash and photosensitivity reaction in most cases are resolved within 15 days through pirfenidone dose reduction, therefore the duration for which skin disorders were applied for has been assumed to be 15 days [NICE 2015f]. The ERG report also noted that photosensitivity can be managed by avoiding exposure to the sun.

The duration of all other AE-related disutilities has been assumed to be three months (or one model cycle) in accordance with the nintedanib NICE manufacturer submission. The duration of disutility was explored as sensitivity analysis (see Section 5.8). The disutilities applied in the model are presented in Table 86.

Table 86: Adverse event-related disutilities per event

Adverse event	Disutility	Duration of disutility	Reference
Serious cardiac event	-0.20	3 months (1 model cycle)	Disutilities taken from nintedanib NICE manufacturer submission
Serious gastrointestinal event	-0.07		
Gastrointestinal perforation	-0.12		
Photosensitivity reaction	-0.032	15 days	[Handorf 2012]
Rash	-0.030		Centre for Clinical Practice at NICE [NICE 2010]
Key: NICE, National Institute for Health and Care Excellence.			

These AE disutilities were combined with the incidence rates of AEs to produce per model cycle disutilities, presented in **Table 87**.

Table 87: Adverse event-related disutilities per model cycle

Pirfenidone	BSC	Nintedanib
-0.0041	-0.0030	-0.0033

Health-related quality-of-life data used in cost-effectiveness analysis

The utility values applied per model health state based upon the mapping study by Freemantle are shown in Table 88 [Freemantle 2015].

Table 88: Summary of health state utility values applied in the model

Parameter	Value	Source
Progression-free	0.847	Trial data, and mapping algorithm [Costabel 2012, Freemantle 2015, King 2014, Noble 2011]
Progressed	0.782	
Lung transplant*	0.847	Assumed equivalent to progression-free disease
Death	0	NICE reference case
Key: NICE, National Institute for Health and Care Excellence.* lung transplant health state included only as a sensitivity analysis.		

In addition to the utilities used in the model base-case, alternative utility values are also available for use in the model:

- Derived in the same manner using the Starkie et al. (2011) mapping algorithm
- Nintedanib NICE STA
    - Utility values by patient FVC group taken from the nintedanib NICE manufacturer submission [Boehringer Ingelheim 2015]
    - Spread of patient FVC group from pooled CAPACITY and ASCEND trial data
    - Progression defined as moving FVC subgroup (i.e. having a decline in predicted FVC of approximately 10%)
- Taken from the PANTHER and ACE studies identified in the literature review [Martinez 2014, Noth 2012]

The utility values elicited using these methods may be used as sensitivity analyses within the model, and are presented for each health state in **Table 89**. For all sources, the utility for the 'Lung transplant' health state is assumed to be equal to the utility for the 'Progression-free' health state.

Utility values are reasonably consistent across all the data sources identified.

Table 89: Alternative utility values available in the model

Parameter	Starkie et al. (2011)	Nintedanib NICE STA	PANTHER and ACE
Progression-free	0.79	0.78	0.82
Progressed	0.74	0.75	0.74
Key: NICE, National Institute for Health and Care Excellence; STA, single technology appraisal.			

Incidence of acute exacerbation is associated with a decrement in HRQoL, given that it is associated with substantial morbidity and mortality, requiring hospitalisation for respiratory failure [Collard 2007, Ley 2011]. Consequently, the decrement in HRQoL associated with acute exacerbation was included in the model, as shown in Table 90. The disutility applied for acute exacerbation in the model was taken from the nintedanib NICE manufacturer submission. The per cycle disutility was calculated based on monthly disutilities of -0.274 [SD: 0.059] for the first month of acute exacerbation, and -0.033 [SD: 0.053] for subsequent months of acute exacerbation. The total duration over which this disutility was applied was 3 months,

and therefore the weighted average of the disutility associated with the first and subsequent months was applied in the model.

Table 90: Disutility for acute exacerbation

Time frame	Utility [SE]	Reference
First month	-0.274 [0.059]	Nintedanib NICE manufacturer submission
Subsequent months	-0.033 [0.053]	
Per model cycle	-0.113	First month + 2 * Subsequent months

Key: Dist, distribution; N/A, not applicable; NICE, National Institute for Health and Care Excellence; SE, standard error.

As no additional information was found to inform this parameter within the literature conducted for this submission this disutility was applied using the same rate as in the nintedanib NICE manufacturer submission previously accepted by NICE.

In the nintedanib NICE manufacturer submission, the risk of recurrent exacerbations was assumed to be the same as the risk for patients who had not previously experienced an acute exacerbation [Boehringer Ingelheim 2015]. In our model, patients are also assumed to be at the same risk of acute exacerbation irrespective of prior incidence.

Table 91 provides a summary of the utility values used within the base-case of the model.

Table 91: Summary of utility values for cost-effectiveness analysis

State (adverse event)	Utility value	Varied by	Reference in submission	Justification
Progression-free	0.8470	Normal	Table 88	Trial data and mapping algorithm
Progressed	0.7818			
Lung transplant	0.8470			
Serious cardiac event	-0.1980	Gamma	Table 86	Nintedanib NICE manufacturer submission
Serious gastrointestinal event	-0.0680			
Gastrointestinal perforation	-0.1180			
Photosensitivity reaction	-0.0320			[Handorf 2012]
Rash	-0.0300			Centre for Clinical Practice at NICE [NICE 2010]
Acute exacerbation	-0.1133	Normal	Table 90	Nintedanib NICE manufacturer submission
Dead	0	-	Table 88	NICE reference case

Key: N/A, Not applicable; NICE, National Institute of Health and Care Excellence.
Note: The Gamma distribution was used to vary adverse event-related disutilities as information regarding the uncertainty around these parameters were unavailable.

5.5 Cost and healthcare resource use identification, measurement and valuation

Resource identification, measurement and valuation studies

A systematic review was conducted to identify relevant resource use and cost data for adult patients with mild to moderate IPF in England. Eligible studies were any type of study design reporting on resource use and costs for the treatment of IPF (generally and for specific treatments) which were relevant to the UK. Given the extent of the previous searches in the 2011 pirfenidone STA submission only records published from 2010 onwards were screened [InterMune 2011].

A single search strategy was used to identify cost effectiveness studies (Section 5.1.1), health related quality of life (Section 5.4.3) and resource use data. Full strategies are provided in Appendix 17. The study selection process is described in Section 5.1.1. Detailed data on the study design and resource use and cost outcomes were extracted. Quality assessment was conducted by one researcher and checked by a second.

In the combined searches for cost-effectiveness studies, resource use and costs in IPF and utilities in IPF, 5924 records were retrieved from database searching and 3474 of these records were published from 2010 onwards. 215 records were found from other sources. After removing duplicates, 3123 records were assessed. 7 relevant studies were found in the resource use and costs review. Full details of the identified publications are reported in Appendix 23 presents the study selection process.

7 studies were included reporting resource use data InterMune 2011, Navaratnam 2013, NICE 2013c, Parfrey 2013, Whiting 2014, Wilson 2014]. 2 studies were cohort studies, one was an economic evaluation alongside an RCT in patients with IPF, one was a cost-effectiveness analysis in patients with cystic fibrosis (CF; providing lung transplant data) and three were technology appraisals by national bodies (NICE and the SMC).

In a real world assessment of pirfenidone via the UK Named Patient Programme, retrospective cohort data were collected to determine the burden on healthcare resources as a consequence of IPF (UK NHS perspective) 100 patients were included (mean age 69.3 [standard deviation (SD) 7.5] years; 76% male) starting pirfenidone for IPF [Parfrey 2013].

In an economic evaluation (UK NHS and society perspective) alongside a 12-month RCT comparing co-trimoxazole (n=64) versus placebo (n=72) in IPF, the resources used were reported for GP visits, outpatient appointments, day case appointments, nurse surgery visits, nurse home visits, A&E attendances and inpatient admissions [Wilson 2014].

A study of a national, secondary care dataset was included, which aimed to determine the trend in hospital admissions for IPF in England between 1998 and 2010 [Navaratnam 2013].

A health technology assessment report (UK NHS perspective; applicable to UK) of ivacaftor for the treatment of CF in patients aged ≥ 6 years included the procedure cost of bilateral lung transplantation [Whiting 2014]. However, patients in these studies had CF, not IPF (the two diseases have a different age profile and natural history). In particular, the cost of lung transplants may vary for paediatric versus adult patients, and follow up data are based on patients still alive in subsequent (follow-up) years - this may differ between patients with cystic fibrosis and those with IPF due to the age and disease profiles.

The NICE costing report for IPF looks at the resource impact of implementing the NICE guideline ‘Idiopathic pulmonary fibrosis’ in England (perspective of direct costs to UK NHS) [NICE 2013c].

The NICE technology appraisal guidance for pirfenidone for treating IPF included the manufacturer's submission (NHS and personal and social services perspective) [InterMune 2011]. It included data from a number of trials: PIPF-004, PIPF-006 (CAPACITY-2 and CAPACITY-1); SP3 and SP2; PIPF-012 (RECAP, an open-label extension study of PIPF-004 and PIPF-006); and PIPF-002 (a safety study).

The SMC assessment of nintedanib 100mg and 150mg involved a lifetime Markov model [SMC 2015]. Data from two double-blind phase III studies (INPULSIS-1 and -2) and a double-blind phase II study (TOMORROW) were used along with studies of pirfenidone and NAC in the NMA.

As the nintedanib NICE submission is the most recent source of information reviewed in an English context, this source has been used in preference to the sources identified in the literature for unit costs consistent between the two models.

Intervention and comparators' costs and resource use

The model uses a price year of 2014/2015. Treatment costs were taken from the Monthly Index of Medical Specialities (MIMS) website. Where available, additional costs were taken from the 2014-2015 NHS national schedule of reference costs and the Personal Social Service Research Unit (PSSRU) Unit Costs of Health & Social Care 2015 [Curtis 2015]. Where costs were not available from these sources, other available evidence were used and inflated where appropriate using the PSSRU healthcare inflation indices [Curtis 2015].

Technology costs

Pirfenidone

Pirfenidone is available in 267mg capsules, and is administered orally. Upon treatment initiation, the dose is titrated to the recommended daily dose as discussed in Section 5.2.5. The recommended daily dose of pirfenidone for patients with IPF is three 267 mg capsules three times a day with food for a total of 2403 mg/day [EMC 2015a].

Pirfenidone is linearly priced, with pack size costs for 267mg capsules of £501.92 for 63, £2,007.70 for 252 and £2,151.10 for 270. The average cost of pirfenidone used in the model is based upon the calculation of the average number of pills received by a patient per day, shown in **Table 92**.

Table 92: Calculation of the cost of treatment – pirfenidone

Trial	RDD (mg)	N	ADR [SE] (mg)	NPD	Source
CAPACITY	2,403.00	343	2,086.60 [451.37]	7.81	[Noble 2011]
ASCEND		276	2,127.10 [395.58]	7.97	[King 2014]
Weighted		519	2,104.66 [17.0209]	7.88	
Key: ADR, actual dose received; NPD, number of pills per day; RDD, recommended daily dose					

The average number of pills received by a patient per day was calculated to be 7.88 based upon data from both CAPACITY and ASCEND (varied in sensitivity analysis). Applying this to the cost of one pill gives a cost per day for pirfenidone of £62.80, and therefore a cost per model cycle of £5,730.62.

Nintedanib

As per the nintedanib NICE manufacturer submission, the cost per day of nintedanib is assumed to be at parity with the cost per day of pirfenidone in the UK [Boehringer Ingelheim 2015].

Best supportive care

Best supportive care was not assumed to incorporate any technology costs, as it is representative of the placebo arms of the clinical trials from which its efficacy data were derived.

In summary, the treatment costs applied per model cycle are shown in **Table 93**. It should be noted that treatment costs only apply to those patients still on treatment, which is affected by the TTD.

Table 93: Cost of treatment per model cycle

Pirfenidone	Best supportive care	Nintedanib
£5,730.62	£0	£5,730.62

Administration costs

Administration for both pirfenidone and nintedanib is oral and no additional visits are needed for administration on top of monitoring visits.

Resource use costs

Resource use costs were applied in the model for all patients according to progression status, and consisted of the costs related to the continuous treatment and monitoring of disease progression (excluding treatment costs discussed above).

Resource use items were identified through consultation with a panel of clinical experts (see Section 1) on management of patients with IPF in UK clinical practice, and in line with the pirfenidone SmPC [EMC 2015a]. Resource use was assumed to be the same for pirfenidone, BSC and nintedanib, except for liver function tests. Liver function tests are required for pirfenidone patients in line with the SmPC: monthly liver function tests during the first 6 months of treatment and then 3 monthly thereafter [EMC 2015a].

The nintedanib SmPC does not pre-define a required regimen of liver function tests, but since administration of nintedanib was associated with elevations of liver enzymes (ALT, AST, ALKP, GGT), the resource use has been assumed to be the same as for pirfenidone patients (also in line with the nintedanib NICE manufacturer submission) [Boehringer Ingelheim 2015, EMC 2015a, EMC 2015b].

A summary of the resource use of patients is presented in **Table 94**.

Table 94: Resource use per model cycle

Resource use item	At treatment initiation	Subsequent MRU
Liver function test	TRUE	*
Gas transfer	TRUE	every 4 months
Lung volume study	TRUE	none
Full pulmonary (covers spirometry)	TRUE	every 4 months
Field exercise test	TRUE	every 6 months
Oxygen	FALSE	for all patients with <80% FVC
Healthcare professional visit	TRUE	every 4 months if FVC >60%, every 3 months if FVC<60%
GP visit	FALSE	based upon FVC

Key: FVC, forced vital capacity; GP, general practitioner; MRU, medical resource use; SmPC, summary of product characteristics.

* Liver function tests were administered as per the pirfenidone SmPC for pirfenidone and nintedanib patients (every month for the first 6 months of treatment, then every 3 months), and for BSC are administered according to clinician opinion (every 1.5 months for the first 6 months of treatment, then 0.3 times per model cycle).

The costs of function tests and healthcare professional visits were sourced from the NHS reference costs (2014-15). Upon reviewing the nintedanib NICE manufacturer submission, the ERG considered an appropriate approach for the cost of oxygen per model cycle is to use a home oxygen costing tool from the Department of Health to obtain a cost per year of £824.30 per patient [Boehringer Ingelheim 2015, NICE 2015f]. This gives a cost per model cycle of £206.08. Unit costs for resource use items are presented in **Table 95**.

Table 95: Unit costs of resource use items

Resource use item	Unit cost	Reference
Liver function test	£1.87	NHS reference costs 2014-15: DAPS05 Haematology
Gas transfer	£202.08	NHS reference costs 2014-15: DZ56Z Carbon monoxide transfer factor test
Lung volume study	£170.54	NHS reference costs 2014-15: DZ45Z Lung volume studies
Full pulmonary (covers spirometry)	£165.85	NHS reference costs 2014-15: DZ52Z Full pulmonary function testing
Field exercise test	£177.13	NHS reference costs 2014-15: DZ32Z Field exercise testing
Oxygen	£206.08	ERG report from the nintedanib NICE manufacturer submission
First healthcare professional visit	£248.17	NHS reference costs 2014-15: WF02B Multiprofessional Non-Admitted Face to Face Attendance, First: Respiratory medicine 340
Follow up healthcare professional visit	£177.53	NHS reference costs 2014-15: WF02A Multiprofessional Non-Admitted Face to Face Attendance, Follow-up: Respiratory medicine 340
GP visit	£37.00	PSSRU 2015 GP visit, per patient contact lasting 11.7 minutes, including direct care staff costs, excluding qualifications

Key: ERG, Evidence Review Group; GP, general practitioner; NHS, National Health Service; NICE, National Institute of Health and Care Excellence; PSSRU, Personal Social Services Research Unit.

In summary, the monitoring costs applied per model cycle are shown in **Table 96**. As expected, the costs are the same across treatments other than the slight differences for the per-cycle based costs due to the frequency of liver function tests, of which the cost impact is minimal.

Table 96: Cost of resource use per model cycle

Cost type		PFN	BSC	NTB*
Cycle specific costs	**Cycle 1**	£969.38	£964.71	£969.38
	Cycle 2	£5.61	£0.94	£5.61
	Cycle ≥ 3	£1.87	£0.56	£1.87
Progression status specific costs	**Pre-progression**	£513.22	£513.22	£513.22
	Post-progression	£525.44	£525.44	£525.44

Key: BSC, best supportive care; NTB, nintedanib; PFN, pirfenidone.
* only for patients with moderate IPF

Lung transplant costs

The cost of lung transplantation is included in sensitivity analysis. When the health state is considered, lung transplant costs are applied as a one-off cost to all new patients entering the 'Lung transplant' health state.

The average cost of lung transplant was calculated from the cost of lung/heart transplants reported in the UK, weighted by the frequency of each type of transplant in the UK as shown in **Table 97**.

Table 97: Calculation of lung transplant costs

	Number of transplants	Unit cost
Single lung transplants	28	£34,530
Double lung transplants	159	£34,530
Heart/lung transplants	3	£64,650
Weighted average	£35,006	
Reference	NHS Blood and Transplant 2013	DoH 2013
Key: DoH, Department of Health; NHS, National Health Service.		

This cost was uplifted using the PSSRU hospital and community health services pay and prices index, to give a cost of £35,712 for a lung transplant [Curtis 2015].

Acute exacerbation costs

The cost of acute exacerbations was applied in the model as the cost of hospitalisation (that is, based upon the incidence of acute exacerbation, the associated cost of hospitalisation is applied). This cost was taken from the NHS reference costs 2014-15, as the weighted cost of respiratory failure without interventions by complication and comorbidity (CC) score.

Table 98: Calculation of bed day cost

Cost of a bed day (Non-Elective Excess Bed Days)		N	Cost
DZ27S	**Respiratory Failure without Interventions, with CC Score 11+**	1617	£240.47
DZ27T	**Respiratory Failure without Interventions, with CC Score 6-10**	3057	£273.62
DZ27U	**Respiratory Failure without Interventions, with CC Score 0-5**	1338	£282.63
Average cost of a bed day			£266.71
Key: CC, complication and comorbidity; N, number of occurrences (activity).			

Using the average cost for a bed day of £266.71, the total cost of hospitalisation due to acute exacerbation was calculated using the probability of hospitalisation derived from the rate of hospitalisation in the pooled trial data. This calculation is shown in **Table 99**.

Table 99: Calculation of hospitalisation cost

	PFN	BSC
Number of cycle-length intervals observed [a]	3768	3771
Number of subjects with hospitalisation [b]	195	202
Rate of hospitalisation per cycle [c = a/b]	0.052	0.054
Probability of hospitalisation per cycle [d = 1-exp[c]]	0.050	0.052
Average length of stay in hospital [e]	8.48	16.27
Total cost of hospitalisation [f = e * cost of bed day]	£2,261.70	£4,339.37
Hospitalisation cost applied per cycle [g = d * f]	£114.08	£226.34

Due to lack of comparable data, the hospitalisation cost applied per cycle for nintedanib (for patients with moderate IPF) was assumed to be the same as the hospitalisation cost applied per cycle for pirfenidone.

Health-state unit costs and resource use

The costs associated with each health state in the model are presented in **Table 100**.

Table 100: Health states and associated costs per model cycle

Health state	Items		Value			Reference in submission
			PFN	BSC	NTB*	
Progression-free	Technology		£5,730.62	£0	£5,730.62	Table 93
	RU	Cycle 1	£969.38	£964.71	£969.38	Table 96
		Cycle 2	£5.61	£0.94	£5.61	
		Cycle ≥ 3	£1.87	£0.56	£1.87	
		Health state	£513.22			
Progressed	Technology		£5,730.62	£0	£5,730.62	Table 93
	RU	Cycle 2	£5.61	£0.94	£5.61	Table 96
		Cycle ≥ 3	£1.87	£0.56	£1.87	
		Health state	£525.44			
Lung transplant			£35,712†			Table 97
Exacerbation cost (hospitalisation)			£114.08	£226.34	£114.08	Table 99

Key: BSC, best supportive care; NTB, nintedanib; PFN, pirfenidone; RU, resource use.
* only for patients with moderate IPF; † applied as a lump sum at the time of transplantation

Adverse reaction unit costs and resource use

AEs were identified based on the criteria used in the nintedanib NICE manufacturer submission (as discussed in Section 5.3). The cost of treating these AEs are presented in **Table 101**. Costs were taken from the NHS reference costs (2014-15), with cost codes that were consistent with those used in the nintedanib NICE manufacturer submission [Boehringer Ingelheim 2015]. Costs were calculated as a weighted average of the cost items given in the reference column.

Table 101: Adverse event unit costs

Adverse event	Cost	Reference	
Serious cardiac event	£3,729.23	NHS reference costs 2014-15: Actual or Suspected Myocardial Infarction	EB10A (CC Score 13+)
	£2,838.79		EB10B (CC Score 10-12)
	£2,236.11		EB10C (CC Score 7-9)
	£1,808.45		EB10D (CC Score 4-6)
	£1,505.11		EB10E (CC Score 0-3)
	£2,200.15		**Weighted average**
Serious GI event	£2,722.34	NHS reference costs 2014-15: Inflammatory Bowel Disease without Interventions	FZ37P (CC Score 5+)
	£2,093.95		FZ37Q (CC Score 3-4)
	£1,804.38		FZ37R (CC Score 1-2)
	£1,629.11		FZ37S (CC Score 0)
	£1,910.91		**Weighted average**
GI perforation	£2,695.63	NHS reference costs 2014-15: Gastrointestinal Bleed without Interventions	FZ38M (CC Score 9+)
	£1,895.28		FZ38N (CC Score 5-8)
	£1,370.09		FZ38P (CC Score 0-4)
	£1,583.03		**Weighted average**
Photosensitivity reaction (Rash)	£467.62	NHS reference costs 2014-15: Skin Disorders without Interventions	JD07E (CC Score 19+)
	£575.87		JD07F (CC Score 14-18)
	£507.00		JD07G (CC Score 10-13)
	£470.71		JD07H (CC Score 6-9)
	£429.65		JD07J (CC Score 2-5)
	£389.56		JD07K (CC Score 0-1)
	£428.63		**Weighted average**

Key: CC, complication and comorbidity; GI, gastrointestinal; NHS, National Health Service.

In the model base-case, all costs used the whole range of CC scores. In summary, the AE costs applied per model cycle are shown in **Table 96**. These costs take into account the incidence of AEs (presented in Section 5.3) and the costs presented in **Table 101**.

Table 102: Adverse event costs per model cycle by treatment

Pirfenidone	Best supportive care	Nintedanib
£82.48	£38.60	£48.11

Miscellaneous unit costs and resource use

End of life costs

As discussed in Section 5.3, deaths were apportioned in the model into those related to IPF or those unrelated to IPF. This is because costs associated with IPF are greatly increased in the last year of life due to increased resource use, home care and length of stay in hospital. Therefore, a one-off end of life cost was applied to patients at death that died due to IPF-related causes.

The cost of end of life care was taken from a report by Hatziandreu *et al.*, in conjunction with the National Audit Office (as per the nintedanib NICE company submission). The report gives an annual end of life cost of £9,098 for home care and hospice care [Hatziandreu 2008].

This cost was inflated using the PSSRU hospital and community health services pay and prices index, to give a total end of life cost of £9,996 [Curtis 2015].

5.6 Summary of base-case de novo analysis inputs and assumptions

Summary of base-case de novo analysis inputs

A summary of the base-case de novo analysis inputs is presented in **Table 103**. Further details on the inputs used in the model base-case may be found in the relevant sections given in the reference column.

Table 103: Summary of base-case de novo analysis inputs

Parameter	Value	Varied by	Reference
Model settings			
Time horizon	34 years	Fixed	Lifetime
Cycle length	3 months		-
Discount rate: Costs	3.50%		NICE reference case
Discount rate: Lys	0%		
Discount rate: QALYs	3.50%		
Starting age	67 years		Section 5.2.1
Utilities			
Progression-free	0.8470	Beta	Section 0
Progressed	0.7818		
Lung transplant	0.8470		
Disutility due to acute exacerbation	-0.1133	Gamma	
Disutility due to adverse events: PFN	-0.0041		
Disutility due to adverse events: BSC	-0.0030		
Disutility due to adverse events: NTB	-0.0033		
OS			
Distribution used	Weibull	Scenario analysis	Section 0
Lambda (λ)	[redacted]	Covariance matrix	
Gamma (γ)	[redacted]		
NMA scenario	Base case network, random effects 52 weeks	Scenario analysis	Section 4.10
PFS			
Distribution used	Weibull	Scenario analysis	Section 0
Lambda (λ)	0.0078	Covariance matrix	
Gamma (γ)	1.4675		

NMA scenario	Base case network, random effects 52 weeks	Scenario analysis	Section 4.10
IPF-related mortality			
PFN	57.89%	Beta	Section 0
BSC	80.38%		
NTB	75.49%		
TTD			
Distribution used	Weibull	Scenario analysis	Section 0
Lambda (λ)	[redacted]	Covariance matrix	
Gamma (γ)	[redacted]		
NMA scenario	Base case network, random effects	Scenario analysis	Section 4.10
Stopping rule	Nintedanib only		Section 0
Acute exacerbations			
Risk per model cycle	1.46%	Fixed	Section 0
NMA scenario	Base case network, random effects	Scenario analysis	Section 4.10
Treatment costs			
PFN - per cycle (list price)	£5,730.62	Fixed	Section 0
PFN - capsules per day	7.88	Normal	
BSC - per cycle (list price)	£0.00	Fixed	
NTB - per cycle (list price)	£5,730.62		
Resource use costs			
PFN - Cycle 1	£969.38	Gamma	Section 0
PFN - Cycle 2	£5.61		
PFN - Cycle >= 3	£1.87		
PFN - Pre-progression	£513.22		
PFN - Post-progression	£525.44		
BSC - Cycle 1	£964.71		
BSC - Cycle 2	£0.94		

BSC - Cycle >= 3	£0.56		
BSC - Pre-progression	£513.22		
BSC - Post-progression	£525.44		
NTB - Cycle 1	£969.38		
NTB - Cycle 2	£5.61		
NTB - Cycle >= 3	£1.87		
NTB - Pre-progression	£513.22		
NTB - Post-progression	£525.44		
Adverse event costs			
Adverse event costs PFN	£82.48	Gamma	Section 0
Adverse event costs BSC	£38.60		
Adverse event costs NTB	£48.11		
Other costs			
Lung transplant	£35,712.46	Gamma	Section 0
Hospitalisation (exacerbation): PFN	£114.08		
Hospitalisation (exacerbation): BSC	£226.34		
Hospitalisation (exacerbation): NTB	£114.08		
End of life care	£9,996.14		Section 0

Key: BSC, best supportive care; IPF, idiopathic pulmonary fibrosis; LY, life year; NICE, National Institute for Health and Care Excellence; NMA, network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; QALY, quality-adjusted life year; TTD, time to treatment discontinuation.

A summary of the scenario analyses performed on the base case is provided in **Table 104**.

Table 104: Scenario analysis summary

Category	Purpose	Base case	Scenarios
Time horizon	To assess the impact of varying the model time horizon on the results of the model.	Lifetime (34 years)	10 years
			15 years
			20 years
			25 years
			30 years
Utilities	To assess the impact of varying the source of utility values used for the model health state utilities within the model.	Freemantle *et al.* (2015) mapping algorithm.	Utilities from the PANTHER and ACE studies identified in the literature review.
			Utilities derived from FVC model health states in the nintedanib NICE company submission.
			Alternative mapping algorithm by Starkie *et al.* (2012)
Treatment effect	To assess the impact of the duration of treatment effect on the model results.	Treatment effect applied for the lifetime horizon.	Treatment effect applied for up to 7 years: the duration of survival data available for pirfenidone from RECAP. After this time, a HR of 1 is applied to survival projections using pirfenidone data (the most reliable source of survival information) for the efficacy of both BSC (all populations) and nintedanib (moderate only) patients.
			Treatment effect applied for up to 10 years
			Treatment effect applied for up to 14 years
			Treatment effect applied for lifetime
OS	To assess the impact of varying the statistical model applied to the OS data for	Weibull	Exponential
			Log-normal

	pirfenidone patients.		Gamma
			Log-Logistic
			Gompertz
	To assess the impact of directly using the limited 52-week RCT data available for patients on best supportive care using Kaplan Meier data from CAPACITY and ASCEND rather than the network meta-analysis hazard ratio	Not applied in the model base case	Applied for the first 52 weeks of the model, after which OS is derived using a HR applied to the OS of pirfenidone patients.
	To assess the impact of using registry data to inform the OS for BSC patients.	Not applied in the model base case	INOVA registry used for BSC OS
			Edinburgh registry used for BSC OS
PFS	To assess the impact of varying the statistical model applied to the PFS data for pirfenidone patients.	Weibull	Exponential
			Log-normal
			Gamma
			Log-Logistic
			Gompertz
	To assess the impact of directly using the limited 52-week RCT data available for patients on best supportive care using Kaplan Meier data from CAPACITY and ASCEND rather than the network meta-analysis hazard ratio	Not applied in the model base case	Applied for the first 52 weeks of the model, after which PFS is derived using a HR applied to the PFS of pirfenidone patients.
TTD	To assess the impact of varying the statistical model applied to the TTD data for pirfenidone patients.	Weibull	Exponential
			Log-normal
			Gamma
			Log-Logistic
			Gompertz
Stopping rule	To assess the impact of applying the	Applies for	Applied for pirfenidone and nintedanib patients (moderate

	stopping rule for pirfenidone (all populations) and nintedanib (moderate only) patients.	nintedanib patients only (moderate population only) or is not applied at all (mild and ITT populations)	population) or for pirfenidone patients only (mild and ITT population)
			Not applied for pirfenidone or nintedanib patients (moderate population only)
OS - NMA	To assess the impact of varying the NMA scenario for the HRs (OS and PFS) and RRs (TTD; moderate population only) applied to estimate outcomes of BSC (all populations) and nintedanib (moderate only) patients	Base case network, random effects, 52 weeks cut off	Restricted network, random effects, 52 weeks cut off
			Restricted network, fixed effects, 52 weeks cut off
			Base case network, fixed effects, 52 weeks cut off
			Restricted network, random effects, 72 weeks cut off
			Base case network, random effects, 72 weeks cut off
			Restricted network, fixed effects, 72 weeks cut off
			Base case network, fixed effects, 72 weeks cut off
PFS - NMA*		Base case network, random effects, 52 weeks cut off	Restricted network, random effects, 52 weeks cut off
			Restricted network, fixed effects, 52 weeks cut off
			Base case network, fixed effects, 52 weeks cut off
			Restricted network, random effects, 72 weeks cut off
			Base case network, random effects, 72 weeks cut off
			Restricted network, fixed effects, 72 weeks cut off
			Base case network, fixed effects, 72 weeks cut off
TTD - NMA		Base case network, random effects	Restricted network, random effects
			Restricted network, fixed effects
			Base case network, fixed effects
AcExs- NMA	To assess the impact of varying the NMA scenario for the ORs applied to estimate the incidence of acute exacerbations per model cycle for BSC (all populations) and nintedanib (moderate only) patients	Base case network, random effects	Restricted network, random effects
			Restricted network, fixed effects
			Base case network, fixed effects
			Base case network, random effects with adjustments in data for

			differences in end point (discussed in detail in NMA appendix)
			Restricted network, fixed effects with adjustments in data for differences in end point (discussed in detail in NMA appendix)
Lung transplant	To assess the impact of including the lung transplant health state in the model structure.	Not applied in the model base case	Applied as scenario analysis
Gas Transfer	To assess the impact of varying the frequency of MRU in subsequent model cycles after treatment initiation.	Test administered every 4 months	Test administered every 6 months
Full pulmonary		Test administered every 4 months	Test administered every 6 months
Lung Volume Transfer		No subsequent tests administered	Test administered every 4 months
Field exercise test		Test administered every 6 months	Test administered every 12 months
			Test administered every 3 months if on oxygen
Healthcare professional visit		Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Test administered every 6 months if FVC >60%, and every 3 months if FVC<60%

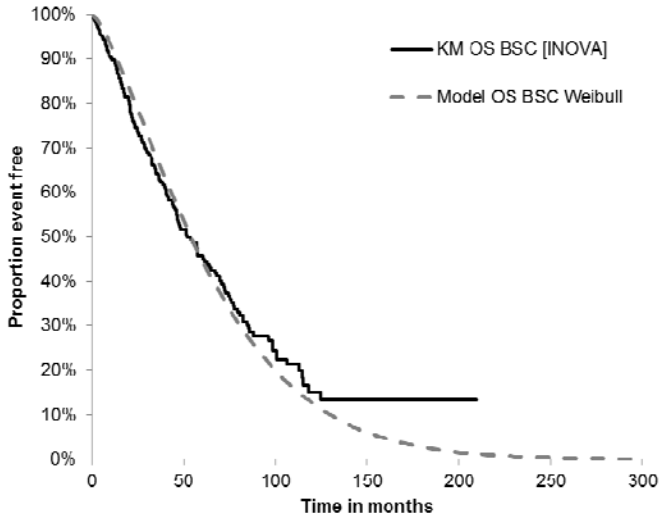
Key: AcEx, acute exacerbation; BSC, best supportive care; FVC, forced vital capacity; HR, hazard ratio; ITT, intention to treat; MRU, medical resource use; NICE, National Institute for Health and Care Excellence; NMA, network meta-analysis; OR, odds ratio; OS, overall survival; PFS, progression-free survival; RCT, randomised controlled trial; RR, relative risk; TTD, time to treatment discontinuation.

* The mean values of each HR for the PFS scenarios were all the same, therefore this set of scenarios yields the same results regardless of choice of scenario. However, these scenarios have been included in this table for completeness.

Assumptions

Table 105 contains a list of all assumptions made in the de novo economic model along with justification for each assumption and references where appropriate.

Table 105: Key model assumptions

Assumption	Rationale/ reference
Cycle length = 3 months.	Consistency with clinical trials and previous economic modelling [Boehringer Ingelheim 2015, InterMune 2011].
Outcomes for BSC arm taken from placebo arms of relevant trials.	The NMA considered a number of trials in order to accurately ascertain the expected outcomes for patients on BSC. Placebo and BSC are defined similarly in the trials considered. The end result for overall survival was reflective of registry data observed in a similar population:  No technology costs were associated with BSC given that its efficacy was derived from the placebo arms of clinical trials, and that no active treatment is administered. Any concomitant medications used in the placebo arm would be associated with a small cost burden, and would be similar across patients on all lines of treatment, and therefore would not be displaced in the NHS by implementation of pirfenidone.

Assumption	Rationale/ reference
FVC cannot improve over time i.e. patients cannot move from 'Progressed' to 'Progression-free'.	Similar assumptions were made in the previously accepted nintedanib NICE manufacturer submission, as well as in the previous pirfenidone NICE manufacturer submission [Boehringer Ingelheim 2015, InterMune 2011]. Alongside this, the assumption was recently validated by clinical experts in the nintedanib NICE manufacturer submission [Boehringer Ingelheim 2015].
FVC and 6MWD are significant predictors of IPF-related mortality and quality of life, and therefore have been used throughout the model.	A decline in percentage predicted FVC of ≥10% is a decrement that has long been recognised as both clinically significant and highly predictive of mortality [Collard et al., 2003; Flaherty et al., 2003; Latsi et al., 2003; King et al., 2005; Zappala et al., 2010]. The previously accepted nintedanib NICE manufacturer submission presented a Markov model with transition probabilities and quality of life based upon FVC and incidence of first acute exacerbations over the cohort's lifetime. Decline in 6MWD has been shown to be predictive of mortality; a 24-week decline >50m is associated with a 3-fold increase in risk of death at 1 year [du Bois 2014, Raghu 2011]. Importantly, 6MWD has also been shown to be a predictor of mortality independent of FVC.[du Bois 2014] Equivalently, HRQoL has been demonstrated to deteriorate over time, in accordance with a decline in percentage predicted FVC and 6MWD being associated with large declines in levels of HRQoL (Section 5.4).
Proportional hazards between BSC, nintedanib and pirfenidone for OS, PFS and TTD	This assumption implies that the survival curves fitted to each treatment group have a similar shape. This assumption is supported by log-cumulative hazard plots, residual plots, log-rank tests for the significance of the interaction between treatment and time and validation against long-term survival data for BSC.
'Lung transplant' health state utility assumed to be the same as the 'Progression-free' health state utility.	There is a lack of available literature to implement the expected utility of an IPF patient post-lung transplant. However, based on a study by Groen et al. (2004) evaluating the cost-effectiveness of lung transplantation in relation to type of end-stage pulmonary disease, it would appear that a successful lung transplant may result in a similar if not improved quality of life to progression-free disease [Groen 2004].
SGRQ mapped onto EQ-5D	NICE guidelines suggest mapping is the second best option when EQ-5D is not available within the clinical trial data. The mapping algorithm selected is expected to be representative of patients in this setting.
Adverse events included in the model	Adverse events were identified as per the previously accepted nintedanib NICE manufacturer submission. These adverse events were deemed to have a significant impact on costs and QALYs [Boehringer Ingelheim 2015]. The duration of adverse event-related disutilities were assumed to be the length of one model cycle (3 months). This assumption was made based on the absence of data to inform the true duration each disutility should be applied, and that the ERG from the nintedanib NICE manufacturer submission suggested that the duration of adverse event-related disutilities was too large (at one year). Finally, a disutility applied for the length of a model cycle is a structurally simple assumption which may be easily and transparently implemented.
Liver function test requirement for nintedanib patients assumed	The nintedanib SmPC does not pre-define a required regimen of liver function tests, but since administration of nintedanib is associated with elevations of liver enzymes, the resource use has been assumed to be the same as for pirfenidone

Assumption	Rationale/ reference
the same as for pirfenidone patients	patients which requires frequent liver function tests [EMC 2015a, EMC 2015b].

Key: BSC, best supportive care; ERG, Evidence Review Group; FVC, forced vital capacity; HRQoL, health-related quality of life; IPF, idiopathic pulmonary fibrosis; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; OS, overall survival; PFS, progression-free survival; QALY, quality-adjusted life year; SmPC, summary of medicinal product characteristics; SGRQ, St George's Respiratory Questionnaire; TTD, time to treatment discontinuation; 6MWD, six minute walking distance.

5.7 Base-case results

Base-case incremental cost effectiveness analysis results

Base-case incremental cost effectiveness analysis results are presented below. These results consider the list price of pirfenidone vs. BSC for ITT (mild and moderate IPF) patients (Table 3). Subgroup analyses are presented in Section 5.9

Table 106: Base-case results – ITT population, list price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	5.38	3.81				
PFN	[redacted]	8.67	5.68	[redacted]	3.29	1.87	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years.

The discounted base-case results for the list price of pirfenidone versus BSC are shown in Table 3. Pirfenidone is associated with a 3.29 LYG, 1.87 incremental QALYs, and incremental costs of [redacted] per patient, compared with BSC. The ICER is [redacted] per additional QALY gained. Undiscounted base case results are presented in Table 107.

Table 107: Base-case results – ITT population, list price (undiscounted)

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	5.38	4.28				
PFN	[redacted]	8.67	6.88	[redacted]	3.29	2.60	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years.

Clinical outcomes from the model

The proportion of patients alive at given time points are presented in Table 108. The clinical trial data used to compare to model results were pooled data from CAPACITY, ASCEND and RECAP, as used in the economic model. The model results were also compared to registry data from the INOVA and Edinburgh registries

(see Section 4.11). Clinical trial data, registry data and modelled results are consistent across the time points measured.

Table 108: Percentage of patients alive over time

Time	Clinical trial result		Registry data		Model result (includes half cycle correction)	
			INOVA	Edinburgh		
	PFN	BSC	BSC	BSC	PFN	BSC
1 year	96.4%	93.2%	89.8%	88.0%	95.5%	91.5%
2 years	87.7%	-	74.4%	75.9%	88.8%	79.5%
5 years	70.4%	-	45.6%	43.6%	65.9%	44.9%
7 years	56.8%	-	38.6%	32.2%	51.7%	28.1%

Key: BSC, best supportive care; PFN, pirfenidone

Markov traces are presented for pirfenidone and BSC in Figure 47 and Figure 48, respectively.

Figure 47: Markov trace for pirfenidone

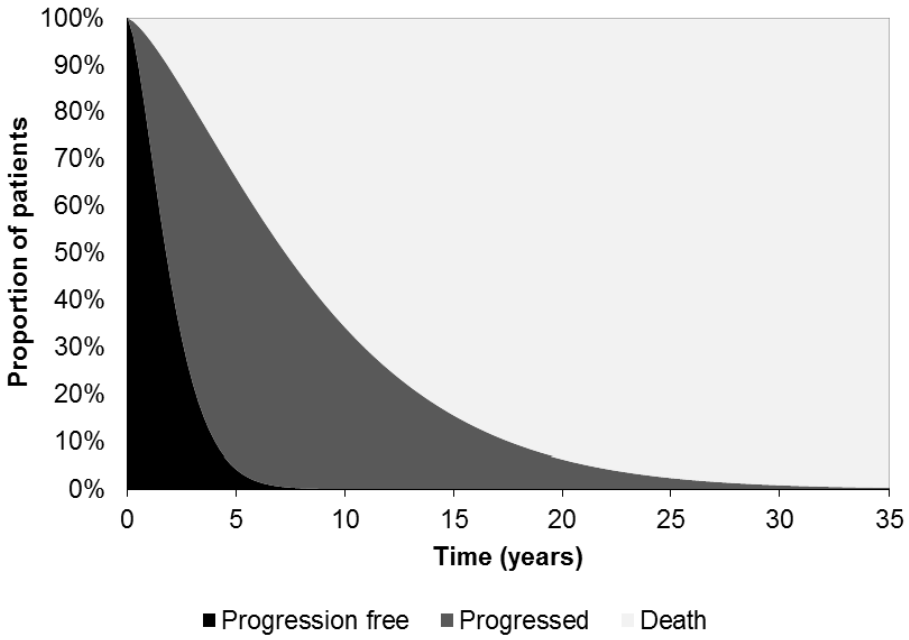
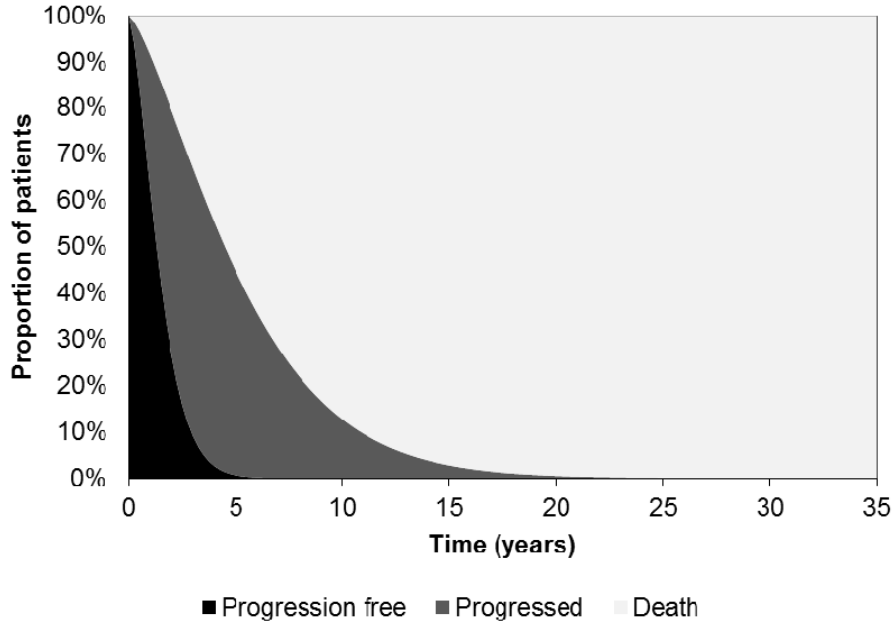


Figure 48: Markov trace for BSC



Disaggregated results of the base-case incremental cost effectiveness analysis

A summary of the LY gain by health state for the entire model time horizon is presented in **Table 109**.

Table 109: Summary of LY gain by health state

Health state	LY PFN	LY BSC	Increment	Absolute increment	% absolute increment
Progression-free	2.05	1.50	0.55	0.55	17%
Progressed	6.62	3.88	2.74	2.74	83%
Total LYs	8.67	5.38	3.29	3.29	100%

Key: LY, life year.

A summary of discounted and undiscounted QALY gain by health state for the entire model time horizon is presented in Table 110 and Table 111, respectively. Pirfenidone provides benefits to patients both pre and post progression; supporting the assertion that the economic stopping rule defined in previous IPF submissions does not have a sound clinical basis.

Table 110: Summary of QALY gain by health state – discounted

Health state	QALY PFN	QALY BSC	Increment	Absolute increment	% absolute increment
Progression-free	1.68	1.24	0.44	0.44	23%
Progressed	4.03	2.59	1.44	1.44	76%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.02	-0.01	-0.01	0.01	1%
Total QALYs	5.68	3.81	1.87	1.89	100%

Key: BSC, best supportive care; PFN, pirfenidone; QALY, quality-adjusted life year.

Table 111: Summary of QALY gain by health state – undiscounted

Health state	QALY PFN	QALY BSC	Increment	Absolute increment	% absolute increment
Progression-free	1.74	1.27	0.47	0.47	18%
Progressed	5.18	3.04	2.14	2.14	82%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.03	-0.02	-0.01	0.01	1%
Total QALYs	6.88	4.28	2.60	2.62	100%

Key: BSC, best supportive care; PFN, pirfenidone; QALY, quality-adjusted life year.

A summary of discounted and undiscounted costs by health state for the entire model time horizon is presented in Table 112 and Table 113, respectively.

Table 112: Summary of costs by health state – discounted

Health state	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£3,256	£4,324	-£1,068	£1,068	1%
End of life	£5,717	£7,899	-£2,182	£2,182	3%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	100%

Key: BSC, best supportive care; PFN, pirfenidone.

Table 113: Summary of costs by health state – undiscounted

Health state	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£3,958	£4,873	-£915	£915	1%
End of life	£5,717	£7,899	-£2,182	£2,182	3%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	100%

Key: BSC, best supportive care; PFN, pirfenidone.

A summary of discounted and undiscounted predicted resource use by category of cost for the entire model time horizon is presented in Table 114 and Table 115, respectively.

Table 114: Summary of predicted resource use by cost category - discounted

Item	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£3,256	£4,324	-£1,068	£1,068	1%
Disease management costs	£15,900	£10,935	£4,965	£4,965	6%
Terminal care	£5,717	£7,899	-£2,182	£2,182	3%
Total	[redacted]	[redacted]	[redacted]	[redacted]	100%

Key: BSC, best supportive care; PFN, pirfenidone.

Table 115: Summary of predicted resource use by cost category - undiscounted

Item	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£3,958	£4,873	-£915	£915	1%
Disease management costs	£19,133	£12,208	£6,925	£6,925	8%
Terminal care	£5,717	£7,899	-£2,182	£2,182	3%
Total	[redacted]	[redacted]	[redacted]	[redacted]	100%

Key: BSC, best supportive care; PFN, pirfenidone.

5.8 Sensitivity analyses

Sensitivity analysis was conducted to address areas of uncertainty in the model.

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed to explore the uncertainty around key model inputs. PSA was conducted by varying these inputs simultaneously and recording the mean model results. 5,000 PSA iterations were ran in order to obtain a stable estimate of the mean model results.

PSA results are presented below alongside the base-case incremental cost effectiveness analysis results. These results consider the list price of pirfenidone vs. BSC for ITT (mild and moderate IPF) patients (Table 116).

Table 116: PSA results – ITT population, list price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
Deterministic model results							
BSC	[redacted]	5.38	3.81				
PFN	[redacted]	8.67	5.68	[redacted]	3.29	1.87	[redacted]
Mean probabilistic model results							
BSC	[redacted]	5.31	3.739				
95% CI	[redacted]	(3.19; 7.90)	(2.42; 5.23)				
PFN	[redacted]	8.70	5.68	[redacted]	3.39	1.94	
95% CI	[redacted]	(7.41; 10.14)	(5.03; 6.38)				[redacted]
% chance of being cost effective		**£20,000 per QALY**		**£25,000 per QALY**			**£30,000 per QALY**
BSC		100%		95%			78%
PFN		0%		5%			22%

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years.

The mean PSA results lie close to the deterministic model results. In these results, pirfenidone is associated with a 3.39 LYG, 1.94 incremental QALYs, and incremental costs of [redacted] per patient, compared with BSC. The ICER is [redacted] per additional QALY gained.

A PSA scatterplot of the results is presented in Figure 49. The cost-effectiveness acceptability curve is presented in

Figure **50**.

Figure 49: PSA scatterplot – ITT population, list price

The PSA scatterplot demonstrates greater certainty around the estimate of QALYs for pirfenidone patients, likely due to the availability of long-term survival data from which more accurate estimates of the benefit associated with pirfenidone may be estimated.

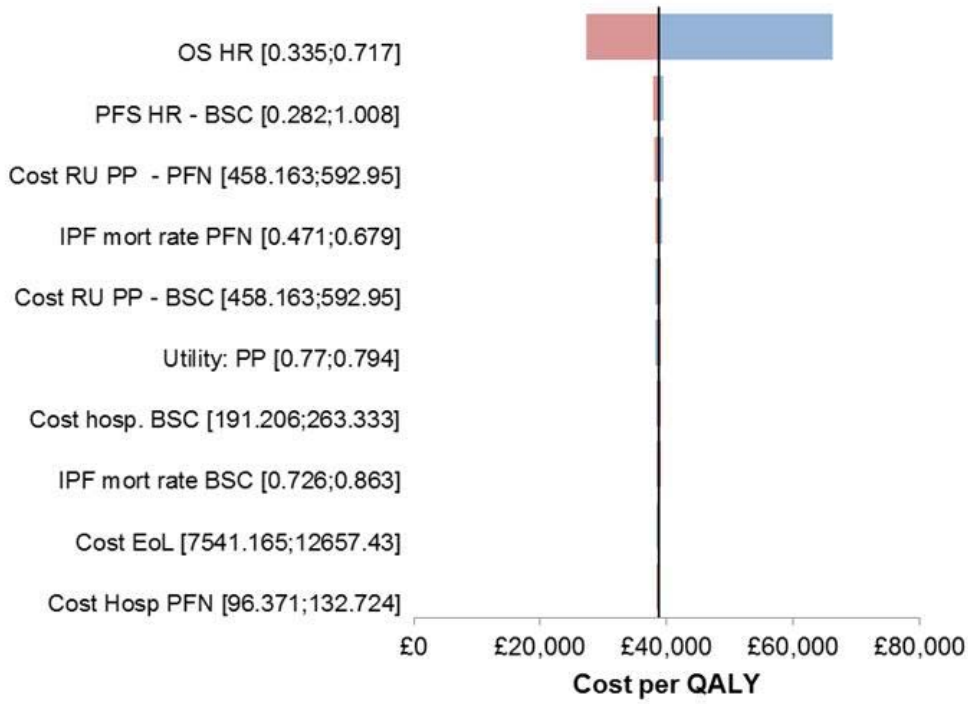
Figure 50: Cost-effectiveness acceptability curve – ITT population, list price

[redacted]

Deterministic sensitivity analysis

One-way sensitivity analysis (OWSA) was performed to assess the impact of individual parameters on the model results.

Figure 51: Tornado diagram – ITT population, list price



Key: BSC, best supportive care; EoL, end of life; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; OS, overall survival; PF, progression-free; PFN, pirfenidone; PFS, progression-free survival; PP, post-progression; QALY, quality-adjusted life year; RU, resource use.

Scenario analysis

Scenario analysis was performed to investigate the uncertainty around alternate model settings and structural assumptions. Results of the scenario analysis are presented in Table 117.

Table 117: Scenario analysis

	Category	Base case setting	Model change	PFN			BSC			ICER vs. BSC (£)
				Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Base Case				[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
1	Time horizon	Lifetime (34 years)	10 years	[redacted]	6.67	4.73	[redacted]	4.97	3.60	[redacted]
2			15 years	[redacted]	7.86	5.35	[redacted]	5.30	3.77	[redacted]
3			20 years	[redacted]	8.38	5.57	[redacted]	5.37	3.80	[redacted]
4			25 years	[redacted]	8.58	5.65	[redacted]	5.38	3.81	[redacted]
5			30 years	[redacted]	8.65	5.67	[redacted]	5.38	3.81	[redacted]
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	[redacted]	8.67	5.41	[redacted]	5.38	3.63	[redacted]
7			Nintedanib NICE company submission	[redacted]	8.67	5.34	[redacted]	5.38	3.58	[redacted]
8			Starkie et al. (2012) mapping algorithm	[redacted]	8.67	5.37	[redacted]	5.38	3.60	[redacted]
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	[redacted]	8.67	5.68	[redacted]	6.22	4.21	[redacted]
10			Treatment effect applied for up to 10 years	[redacted]	8.67	5.68	[redacted]	5.74	3.96	[redacted]
11			Treatment effect applied for up to 14 years	[redacted]	8.67	5.68	[redacted]	5.48	3.85	[redacted]
12	OS	Weibull	Exponential	[redacted]	11.91	7.01	[redacted]	6.62	4.37	[redacted]
13			Log-normal	[redacted]	12.71	7.25	[redacted]	6.89	4.47	[redacted]
14			Gamma	[redacted]	9.49	6.01	[redacted]	5.59	3.91	[redacted]
15			Log-Logistic	[redacted]	11.19	6.64	[redacted]	6.12	4.12	[redacted]
16			Gompertz	[redacted]	7.60	5.21	[redacted]	5.16	3.70	[redacted]
17		BSC trial data	BSC trial data applied for the first 52 weeks	[redacted]	8.67	5.68	[redacted]	5.45	3.85	[redacted]

		not applied for the first 52 weeks of the model.	of the model.							
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	[redacted]	8.67	5.68	[redacted]	5.75	3.96	[redacted]
19			Edinburgh registry used for BSC OS	[redacted]	8.67	5.68	[redacted]	5.03	3.59	[redacted]
20	PFS	Weibull	Exponential	[redacted]	8.67	5.73	[redacted]	5.38	3.83	[redacted]
21			Log-normal	[redacted]	8.67	5.73	[redacted]	5.38	3.83	[redacted]
22			Gamma	[redacted]	8.67	5.75	[redacted]	5.38	3.84	[redacted]
23			Log-Logistic	[redacted]	8.67	5.73	[redacted]	5.38	3.82	[redacted]
24			Gompertz	[redacted]	8.67	5.66	[redacted]	5.38	3.80	[redacted]
25		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
26	TTD	Weibull	Exponential	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
27			Log-normal	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
28			Gamma	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
29			Log-Logistic	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
30			Gompertz	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
31	Stopping rule	Not applied	Applied for pirfenidone patients	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]

32										
33	OS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.31	3.76	[redacted]
34			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
35			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
36			Phase III trials only, random effects, 72 weeks cut off	[redacted]	8.67	5.68	[redacted]	6.12	4.24	[redacted]
37			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	8.67	5.68	[redacted]	6.12	4.24	[redacted]
38			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	8.67	5.68	[redacted]	6.12	4.24	[redacted]
39			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	8.67	5.68	[redacted]	6.12	4.24	[redacted]
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
41			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
42			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
43			Phase III trials only, random effects, 72 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]

44			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
45			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
46			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
47	AcExs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
48			Phase III trials only, fixed effects	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
49			Phase II and III trials, fixed effects	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
50			Phase II and III trials, random effects with adjustments in data for differences in end point	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
51			Phase III trials only, fixed effects with adjustments in data for differences in end point	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
52	Lung transplant	Not applied	Applied	[redacted]	8.63	5.68	[redacted]	5.40	3.82	[redacted]
53	Gas Transfer	Test administered every 4 months	Test administered every 6 months	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
54	Full pulmonary	Test administered every 4 months	Test administered every 6 months	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
55	Lung Volume	No subsequent	Test administered every 4 months	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]

	Transfer	tests administered								
56	Field exercise test	Test administered every 6 months	Test administered every 12 months	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
57			Test administered every 3 months if on oxygen	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]
58	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	[redacted]	8.67	5.68	[redacted]	5.38	3.81	[redacted]

Key: AcEx, acute exacerbation; BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use; NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

As shown in the table, the largest driver of the ICER in the model is the time horizon over which the model considers the costs and impact on patients. This is caused by the majority of costs being accrued over a relatively short time period, with the long-term benefits associated with treatment accrued over the longer-term (which are not captured by the choice of a shorter model time horizon). The model is not, however, overly sensitive to this parameter with the vast majority of both benefits and costs accrued within a 15 year time horizon.

In addition to this, other key drivers of model results relate to the choices surrounding the application of OS in the model. As previously discussed, pirfenidone is the only treatment with proven benefits to patient survival, therefore variation around the estimate of this survival benefit has an expectedly large impact on the model results.

The duration of treatment effect assumed for pirfenidone is also influential on model results as would be expected. Sensitivity to this parameter was explored to investigate the impact of a potential lack of treatment effect following the end of available trial data for pirfenidone at 7 years, continuation of treatment effect through to 10 years and 14 years. As expected, assuming equivalent efficacy beyond 7 years impacts model results, as the benefit of pirfenidone lies in both pre- and post-progression; however as with the time horizon the model does not demonstrate extreme sensitivity to this parameter indicating that the ICERs presented are not contingent upon the assumption that the treatment effect of pirfenidone continues long beyond observed data.

All other model parameters show marginal change based on the choice of parameter. The overall range of ICERs at list price is [redacted] to [redacted].

Summary of sensitivity analyses results

Pirfenidone is shown to demonstrate similar levels of cost effectiveness against BSC for all sensitivity analyses completed in the sections above. PSA results demonstrated much less variation in the expected QALYs for patients on pirfenidone versus patients on BSC, driven by the improved estimate of long-term efficacy of these patients due to the availability of new long-term follow up trial data.

OWSA showed that the key driver of cost effectiveness results in the model was the HR applied to the OS of pirfenidone patients to estimate the OS of BSC patients. This is due to the uncertainty surrounding long-term estimates of survival in BSC patients, where long-term estimates are more limited (compared to pirfenidone). This uncertainty is, however, somewhat mitigated by the availability of a large number of different registry sources which provide confirmation of the expected long-term survival profile for BSC.

Scenario analysis further explored specific model settings and their associated impact on the model results. The most influential model parameters were the time horizon over which the model considers the costs and impact on patients (since the majority of costs are accrued over a relatively short time period, but the long-term benefits associated with treatment accrue over the longer-term), and options regarding OS as previously discussed. Compared to many models presented for similarly impactful diseases in oncology, however, the sensitivity of the model to time horizon is relatively limited indicating that model outcomes are not dependent upon considerable extrapolation beyond outcomes observed in RECAP.

5.9 Subgroup analysis

The model base-case results using the list prices of pirfenidone and nintedanib are presented for the mild (pirfenidone vs. BSC) and moderate (pirfenidone vs. nintedanib vs. BSC) populations in Table 116 and Table 117 to Table 119, respectively.

Given the importance of these subgroups full sensitivity analysis was undertaken around both subgroups, the results of which are presented in Appendix 24.

As expected, the results of the subgroup analysis are reflective of those seen in the ITT population. This is also in line with clinical results which demonstrate no significant differences in the efficacy of pirfenidone across the FVC subgroups.

The ability to compare pirfenidone to nintedanib in the moderate IPF subgroup is limited due to the lack of equivalently robust data to enable comprehensive comparison.

In the mild IPF population, pirfenidone is associated with a 4.15 LYG, 2.17 incremental QALYs, and incremental costs of [redacted] per patient, compared with BSC. The ICER is [redacted] per additional QALY gained.

In the moderate IPF population, pirfenidone is associated with a 2.87 LYG, 1.70 incremental QALYs, and incremental costs of [redacted] per patient, compared with BSC. The ICER is [redacted] per additional QALY gained. Compared with nintedanib, pirfenidone is associated with a 1.61 LYG, 0.92 incremental QALYs, and incremental costs of [redacted] per patient. The ICER is [redacted] per additional QALY gained.

Table 118: Base-case results – Mild IPF patients, list price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	7.11	4.84				
PFN	[redacted]	11.26	7.00	[redacted]	4.15	2.17	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 119: Base-case results – Moderate IPF patients, list price - incremental

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus BSC (QALYs)	Fully Incremental (QALYs)
BSC	22,475	4.80	3.45					
NTB	62,639	6.06	4.23	40,164	1.26	0.78	51,611	51,611
PFN	[redacted]	7.67	5.15	[redacted]	1.61	0.92	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 120: Base-case results – Moderate IPF patients, list price – pairwise (vs, best supportive care)

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	22,475	4.80	3.45				
PFN	[redacted]	7.67	5.15	[redacted]	2.87	1.70	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 121: Base-case results – Moderate IPF patients, list price – pairwise (vs. nintedanib)

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
NTB	62,639	6.06	4.23				
PFN	[redacted]	7.67	5.15	[redacted]	1.61	0.92	[redacted]

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

5.10 Validation

Validation of de novo cost-effectiveness analysis

Table 122 shows a comparison of the model outcomes from the nintedanib NICE manufacturer submission, the submission presented here, and the previous pirfenidone NICE manufacturer submission TA282 [Boehringer Ingelheim 2015, InterMune 2011].

Table 122: Comparison of LYs and QALYs – moderate population

Outcome	NTB submission			This submission			TA282	
	BSC	NTB	PFN	BSC	NTB	PFN	BSC	PFN
Total QALYs	3.27	3.67	3.62	3.15	3.77	4.46	3.18	4.30
Total LYs	4.36	4.86	4.86	4.33	5.30	6.47	4.40	5.96

Key: IPF, idiopathic pulmonary fibrosis; LY, life year; NTB, nintedanib; PFN, pirfenidone; QALY, quality-adjusted life year

LYs in the nintedanib NICE manufacturer submission were discounted, hence the LYs presented for the moderate IPF population considered in this submission have also been discounted. The moderate IPF population has been used to compare results as this is the population for which nintedanib is currently recommended for use in the NHS.

The LYs and QALYs gained in our modelling of the moderate IPF population are similar for BSC and larger for both nintedanib and pirfenidone than those presented in the nintedanib NICE manufacturer submission, with the estimates presented for pirfenidone being more consistent with those seen in TA282. It is reassuring that BSC estimates are similar across submissions. The difference in the estimates of long-term outcomes for nintedanib and pirfenidone are likely driven by availability of the long-term RECAP data (now available for 7 years) with which considerably more certainty can be gained regarding the size of the long-term survival benefit expected with pirfenidone, compared to the estimates presented in the nintedanib submissions and TA282.

The LYs and QALYs gained in our modelling of the moderate IPF population are similar to those presented in TA282 with the difference in LYs again likely driven by

the availability of additional data for RECAP. Outcomes for BSC patients are near identical.

Regarding the difference in the QALY to LY ratio for patients treated with pirfenidone between TA282 and the current modelling, the impact on patient quality of life has been conservatively included for one progression alone in the updated model. As patients on pirfenidone experience fewer progressions, and hence are expected to perform better than patients on BSC, the impact of no longer modelling additional progressions is more pronounced on the pirfenidone arm. Combining this with the more appropriate mapping algorithm causes the differences shown in **Table 122**.

For the ITT population, results from the model were compared with available published literature, as shown in Table 123. Fisher *et al.* used the Strand registry to estimate the long-term outcomes for BSC patients, whereas Roskell *et al.* used the CPRD registry [Fisher 2015, Roskell 2014].

Table 123: Comparison of OS and PFS – ITT population

Outcome	This submission		Fisher *et al.* [Fisher 2015]		Roskell *et al.* [Roskell 2014]	
	BSC	PFN	BSC	PFN	BSC	PFN
Mean OS	5.38	8.67	6.10	9.29	5.25	9.26
Mean PFS	1.50	2.05	1.28	3.28	NR	NR

Key: IPF, idiopathic pulmonary fibrosis; LY, life year; NR, not reported; NTB, nintedanib; PFN, pirfenidone; QALY, quality-adjusted life year.

The published sources demonstrate comparable estimates for the OS and PFS of both pirfenidone and BSC patients. Within Fisher *et al.* for BSC patients outcomes were derived using the Strand registry: as previously discussed, this is associated with some uncertainty due to the likelihood of misdiagnosis. Fisher *et al.* estimated PFS using the log-normal distribution. This distribution typically consists of a heavy tail and is therefore deemed inappropriate in some cases when estimating the long-term survival in diseases associated with rapid decline. Both Roskell *et al.* and Fisher *et al.* estimated either similar or greater incremental benefits for pirfenidone versus BSC.

Long-term prediction of survival

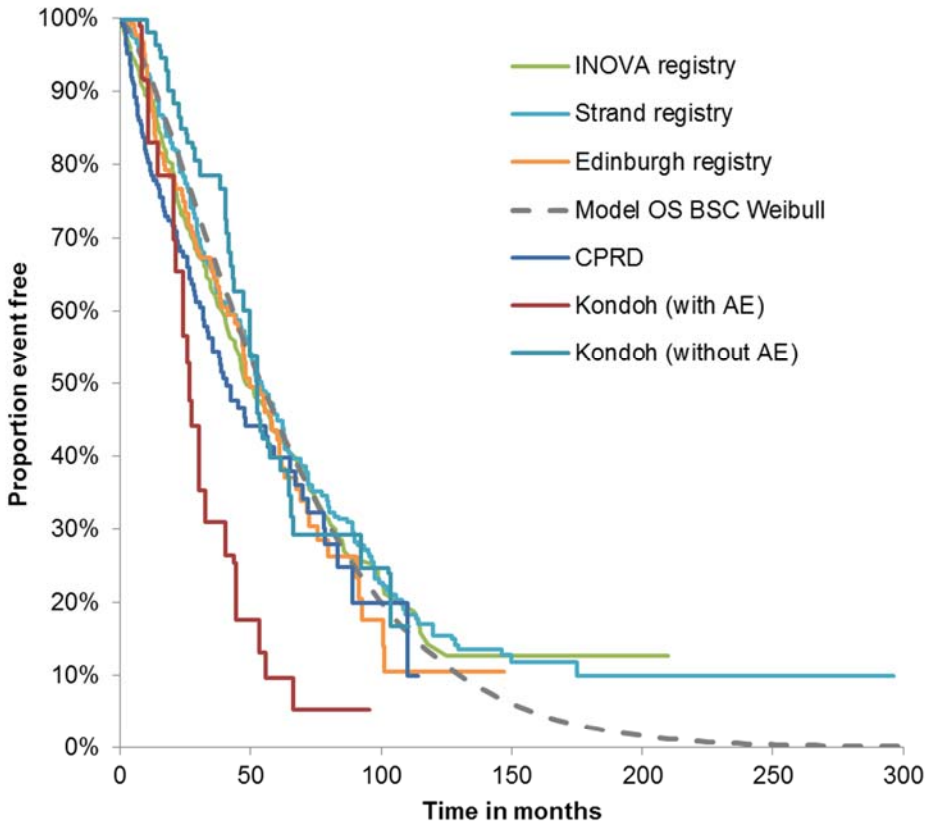
Figure 52 shows the long-term overall survival prediction for IPF patients treated with BSC compared with available long-term registry data.

The figure shows the extrapolated data closely follows the registries considered most relevant until approximately 10% of patients are still alive, at which time the Kaplan-Meier curves become less reliable.

A noteworthy caveat linked to the Strand registry data is that clinical guidance indicates that some patients may have been misdiagnosed, causing over-prediction of long-term survival. Additionally, it should be noted that the Kondoh registry (for patients with acute exacerbation) considered 74 Japanese patients, and therefore are potentially not reflective of patients in the UK.

Also shown in Figure 52 are data extracted from the CPRD (see Section 4.11). CPRD is a national database containing outcomes and diagnostic information for patients in the UK with a variety of diagnoses. Using this database, patients were identified using IPF-related READ codes [Roskell 2014]. The CPRD database has a large number of patients, but is not IPF-specific due to the derivation of likely IPF-diagnosis. However, outcomes for these patients are similar to those observed in other registries, as shown in Figure 52.

Figure 52: Long-term overall survival for BSC IPF patients – ITT population



Number at risk	0	50	100	150	200	250	300
Strand	321	174	72	41	31	31	0
Edinburgh	125	44	6	0	0	0	0
CPRD	193	31	2	0	0	0	0
Kondoh with AE	23	4	0	0	0	0	0
Kondoh without AE	51	27	12	0	0	0	0
INOVA	234	82	22	8	8	0	0

Key: AE, acute exacerbation; BSC, best supportive care; OS, overall survival.

5.11 Interpretation and conclusions of economic evidence

The economic evidence presented in this document demonstrates the cost effectiveness of pirfenidone versus the relevant standard of care in each subpopulation of patients with IPF.

In the base case analysis at list price, pirfenidone was shown to be associated with an ICER of [redacted] versus BSC, with a 3.29 LYG, 1.87 incremental QALYs, and incremental costs of [redacted] per patient. Sensitivity analysis around this result shows that this estimate of the cost effectiveness of pirfenidone is robust around all key model parameters.

Similar results are also seen in the mild and moderate subpopulations, with uncertainties observed for patients treated with nintedanib due to the lack of comparably robust estimates of long-term survival.

The benefits of pirfenidone have been demonstrated in both the pre-progression and post-progression model health states, demonstrating the lack of clinical rationale regarding the use of a stopping rule for patients on pirfenidone.

Pirfenidone has been shown to be associated with a QALY gain of 1.87, demonstrating the long-term survival benefit in these patients, as shown in the newly-available long-term follow up data from the RECAP study [Costabel 2014, Kreuter 2014, Roche 2016a].

The main model sensitivity was the choice of efficacy for BSC patients. This is due to the lack of similarly robust long-term efficacy data available for these patients, hence the reliance on outcomes from the NMA to consider as wide a range of efficacy sources as possible.

The HRQoL data used to inform the model for all patients was derived using a mapping study based on a patient population of IPF patients [Freemantle 2015]. It was noted in the previous submission that a mapping study in an orphan disease such as IPF was previously unavailable, and therefore estimates of HRQoL in this submission are expected to be truly reflective of the IPF patient population [InterMune 2011].

In addition to the inclusion of improved HRQoL data, the costs of exacerbations and treatment-related adverse events have now been explicitly incorporated into the modelling to further reflect the true course of the disease.

6 Assessment of factors relevant to the NHS and other parties

6.1 Eligible patient population

On 1 January, 2016, the total population of England was 56,324,887 individuals [Quality and Outcomes Framework 2013-14, HSCIC]. The prevalence of IPF in England is estimated to be 23.6 per 100,000 individuals [BTS 2014]. Hence, on 1 January, 2016, approximately 13,293 individuals in England had prevalent IPF. An estimated 30.5% of patients with prevalent IPF have mild IPF, 54.0% have moderate IPF, and 15.5% have severe IPF [Roche 2016a]. This corresponds to 4,054 patients with mild IPF, 7,178 patients with moderate IPF, and 2,060 patients with severe IPF on 1 January, 2016. Patients with severe IPF are not eligible for treatment with pirfenidone or nintedanib.

Patients with moderate IPF are currently recommended by NICE for treatment with pirfenidone or nintedanib before 2016. Therefore, we assumed that no new treatment-eligible patients would arise from this cohort aside from increasing market share over time. Depending on the recommendations made for this re-review, patients with mild IPF may become eligible for treatment with pirfenidone. Since these patients were not recommended for treatment before 2016, we assumed that all patients with mild IPF on 1 January 2016 would be eligible to start treatment in 2016 if treatment were recommended. We assumed that the uptake of any form of treatment in mild IPF patients (pirfenidone or BSC) will be 40% in 2016. This is the same uptake as was observed in patients with moderate IPF when NICE first recommended pirfenidone treatment [NICE 2013c]. Hence, 1,622 patients (0.4*4,054) with mild IPF on 1 January, 2016 would potentially take up treatment with pirfenidone in 2016 if it were recommended.

The incidence of IPF in England is estimated to be 8.0 per 100,000 individuals. Hence, during 2016 another 4,506 individuals will be diagnosed with IPF. We assumed that the distribution of these cases across disease categories will be identical to that observed among prevalent cases (i.e. 30.5%: mild disease; 54.0%: moderate disease; 15.5%: severe disease). This results in 1,374 patients diagnosed

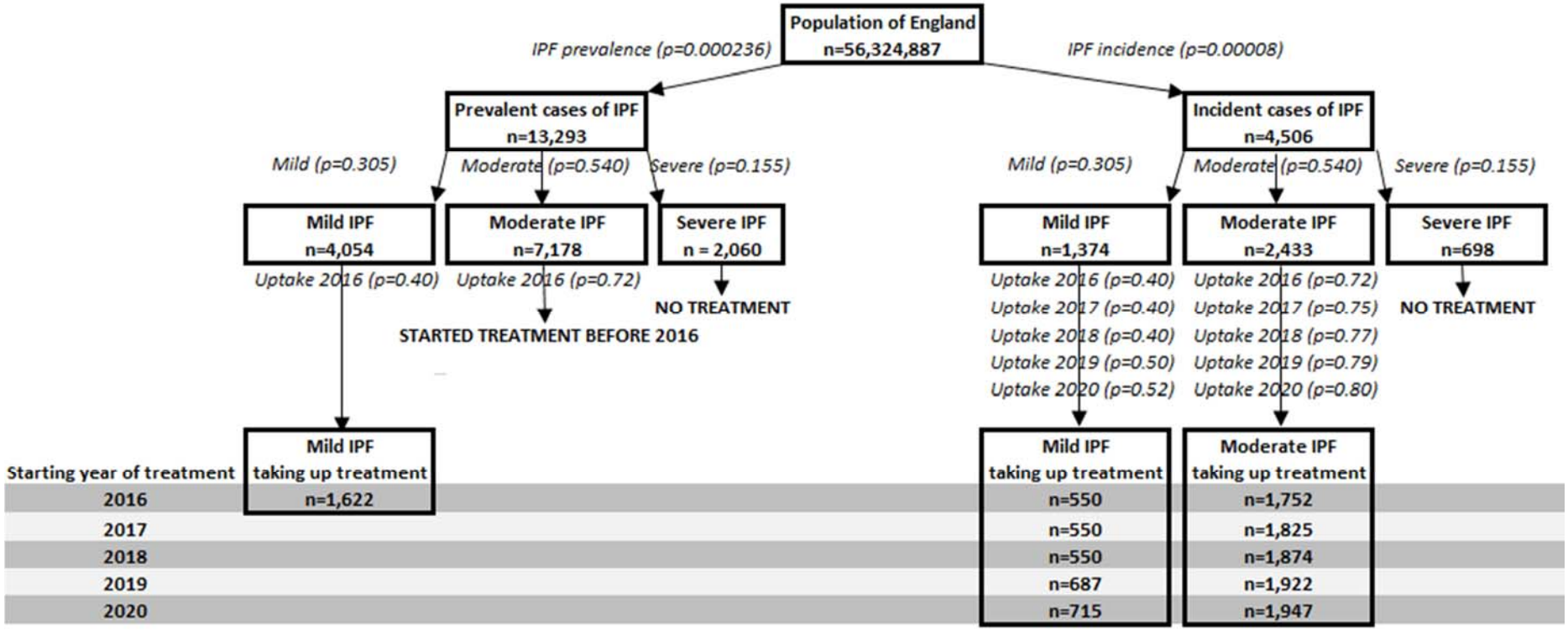
with mild disease, 2,433 patients diagnosed with moderate disease and 698 patients diagnosed with severe disease.

We assumed that population size, IPF incidence, and distribution of IPF across disease categories will remain constant over time.

We have also assumed that the uptake of any form of treatment in these patients (i.e. treatment with pirfenidone, nintedanib or BSC) will increase over time: from an observed 68% in 2015 to an estimated 72%, 75%, 77%, 79%, and 80% in 2016, 2017, 2018, 2019, and 2020, respectively [Roche 2016a]. This results in an estimated uptake of treatment in 1,752, 1,825, 1,874, 1,922 and 1,947 patients with incident, moderate IPF in 2016, 2017, 2018, 2019 and 2020, respectively.

In the scenario in which treatment with pirfenidone is recommended for individuals with mild IPF, we assumed that the uptake of any form of treatment in patients with mild IPF (pirfenidone or BSC) will increase from 40% in 2016-2018 to 50% in 2019 and 52% in 2020: a similar increase as observed in patients with moderate IPF when they were first recommended treatment. This results in an uptake of treatment in 550, 550, 550, 687 and 715 patients with incident, mild IPF in 2016, 2017, 2018, 2019 and 2020, respectively.

Figure 53: Numbers of patients that could potentially take up treatment with pirfenidone in 2016-2020.



6.2 Uptake of technologies

All patients with mild or severe IPF currently receive BSC. For patients with moderate IPF, 3 treatment options are currently available: treatment with pirfenidone, treatment with nintedanib, and BSC. Based upon current recommendations for treatment with pirfenidone and nintedanib, a stopping rule applies: treatment should be discontinued if there is evidence of disease progression (i.e. a decline in percent predicted FVC of 10% or more within any 12-month period).

We determined the budget impact of 3 scenarios, as well as the net budget impact of scenarios 2 and 3 compared with scenario 1:

Scenario 1: A scenario in which the recommendations for treatment with pirfenidone remain unchanged (the 'existing recommendations' scenario);

Scenario 2: A scenario in which treatment with pirfenidone is also recommended for patients with mild IPF (the 'expanded treatment' scenario);

Scenario 3: A scenario in which treatment with pirfenidone is not recommended for any population (the 'no pirfenidone scenario').

6.3 Market share

The expected market shares for the treatment of mild and moderate IPF under the three scenarios described above are based on uptake seen previously for pirfenidone in the moderate IPF population and displayed in Table 124.

Table 124: Expected market shares for the treatment of mild and moderate IPF under the three scenarios

Scenario 1: Existing recommendations					
Patients with mild IPF					
	Calendar year				
Treatment option	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone	0%	0%	0%	0%	0%
Nintedanib	0%	0%	0%	0%	0%
BSC	100%	100%	100%	100%	100%
Patients with moderate IPF					
	Calendar year				
Treatment option	**2016**	**2017**	**2018**	**2019**	**2020**

Pirfenidone	37.8%	34.8%	33.0%	31.0%	32.0%
Nintedanib	16.2%	23.2%	27.0%	31.0%	32.0%
BSC	46.0%	42.0%	40.0%	38.0%	36.0%
Scenario 2: Expanded treatment					
Patients with mild IPF					
	Calendar year				
Treatment option	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone	16%	35%	38%	38%	39%
Nintedanib	0%	0%	0%	0%	0%
BSC	84%	65%	62%	62%	61%
Patients with moderate IPF					
	Calendar year				
Treatment option	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone	38%	36%	34%	33%	33%
Nintedanib	16%	24%	28%	29%	31%
BSC	46%	40%	38%	38%	36%
Scenario 3: No pirfenidone					
Patients with mild IPF					
	Calendar year				
Treatment option	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone	0%	0%	0%	0%	0%
Nintedanib	0%	0%	0%	0%	0%
BSC	100%	100%	100%	100%	100%
Patients with moderate IPF					
	Calendar year				
Treatment option	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone	9.5%	0.0%	0.0%	0.0%	0.0%
Nintedanib	44.6%	58.0%	60.0%	60.0%	60.0%
BSC	46.0%	42.0%	40.0%	40.0%	40.0%

6.4 Other costs included in budget impact

Treatment with pirfenidone, treatment with nintedanib, and BSC are associated with different risks for hospitalisation, and hence, different hospitalisation costs. Therefore, hospitalisation costs are also included in this analysis.

6.5 Unit costs used within the budget impact model

We used the cost-effectiveness model described in Section 5 to determine the average drug and hospitalisation costs per patient by year following diagnosis for the three treatment options available: treatment with pirfenidone, treatment with nintedanib, and BSC (Table 125). Drug costs are assessed at list price.

Table 125: Average drug and hospitalisation costs per patient by year following diagnosis for the treatment options available.[1]

Year following diagnosis	Treatment with pirfenidone		Treatment with nintedanib		BSC	
	Drug costs	Hospitalisation costs	Drug costs	Hospitalisation costs	Drug costs	Hospitalisation costs
1	18,799	447	18,325	444	0	872
2	10,152	421	8,264	409	0	775
3	4,482	388	2,194	367	0	663
4	1,717	354	408	323	0	554
5	585	318	58	280	0	453

Notes: [1]Costs were determined using the cost-effectiveness model described in Chapter 5.

All costs decrease with increasing time after diagnosis. This is a result of IPF-related mortality and other cause mortality, which cause the number of IPF patients to decline over time. The decline in treatment costs is further driven by the decline in patients compliant with treatment and the discontinuation of treatment in individuals meeting the stopping rule for treatment. Hospitalisation costs are substantially lower if pirfenidone/nintedanib treatment is provided as a result of the associated decrease in the risk for acute exacerbations requiring hospitalisation.

6.6 Resource use savings

Aside from hospitalisations no additional resource savings are expected to result from the interventions considered.

6.7 Estimated budget impact

Detailed information on the budget impact of IPF treatment in England under the three scenarios presented in Table 126.

Table 126: Budget impact of the treatment of patients with IPF in England under the three scenarios– results by patient category, calendar year and cost category

Scenario 1: Existing recommendations					
Patients with mild IPF					
	Calendar year				
Cost category	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone – drugs	0	0	0	0	0
Pirfenidone – hospitalisation	0	0	0	0	0
Nintedanib – drugs	0	0	0	0	0
Nintedanib – hospitalisation	0	0	0	0	0
BSC – drugs	0	0	0	0	0
BSC – hospitalisation	1,893,493	2,162,229	2,345,070	2,592,696	2,808,410
Subtotal	*1,893,493*	*2,162,229*	*2,345,070*	*2,592,696*	*2,808,410*
Patients with moderate IPF					
	Calendar year				
Cost category	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone – drugs	12,449,248	18,661,715	21,038,527	21,462,591	22,008,601
Pirfenidone – hospitalisation	296,016	562,677	800,685	1,007,503	1,204,613
Nintedanib – drugs	5,200,865	10,103,919	13,391,581	16,145,058	17,638,383
Nintedanib – hospitalisation	126,013	304,061	501,929	718,533	922,168
BSC – drugs	0	0	0	0	0
BSC – hospitalisation	702,734	1,292,924	1,781,825	2,172,403	2,463,744
Subtotal	*18,774,876*	*30,925,295*	*37,514,547*	*41,506,088*	*44,237,511*
TOTAL	**20,668,369**	**33,087,524**	**39,859,617**	**44,098,784**	**47,045,921**
Scenario 2: Expanded treatment					
Patients with mild IPF					
	Calendar year				
Cost category	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone – drugs	6,510,925	7,143,458	7,417,599	8,489,530	9,367,369
Pirfenidone – hospitalisation	154,816	232,062	308,502	401,954	493,975
Nintedanib – drugs	0	0	0	0	0
Nintedanib – hospitalisation	0	0	0	0	0

BSC – drugs	0	0	0	0	0
BSC – hospitalisation	1,591,481	1,725,555	1,784,703	1,883,610	1,960,712
Subtotal	*8,257,222*	*9,101,075*	*9,510,804*	*10,775,094*	*11,822,056*
Patients with moderate IPF					
	Calendar year				
Cost category	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone – drugs	12,449,248	19,073,396	21,648,285	22,492,701	22,894,807
Pirfenidone – hospitalisation	296,016	572,466	819,117	1,041,861	1,245,249
Nintedanib – drugs	5,200,865	10,371,453	13,821,233	15,611,933	17,006,913
Nintedanib – hospitalisation	126,013	310,543	515,387	713,718	908,706
BSC – drugs	0	0	0	0	0
BSC – hospitalisation	702,734	1,261,097	1,720,864	2,119,163	2,418,680
Subtotal	*18,774,876*	*31,588,955*	*38,524,885*	*41,979,376*	*44,474,355*
TOTAL	**27,032,098**	**40,690,030**	**48,035,689**	**52,754,470**	**56,296,412**
Scenario 3: No pirfenidone					
Patients with mild IPF					
	Calendar year				
Cost category	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone – drugs	0	0	0	0	0
Pirfenidone – hospitalisation	0	0	0	0	0
Nintedanib – drugs	0	0	0	0	0
Nintedanib – hospitalisation	0	0	0	0	0
BSC – drugs	0	0	0	0	0
BSC – hospitalisation	1,893,493	2,162,229	2,345,070	2,592,696	2,808,410
Subtotal	*1,893,493*	*2,162,229*	*2,345,070*	*2,592,696*	*2,808,410*
Patients with moderate IPF					
	Calendar year				
Cost category	**2016**	**2017**	**2018**	**2019**	**2020**
Pirfenidone – drugs	3,112,312	1,680,738	742,028	284,262	96,851
Pirfenidone – hospitalisation	74,004	69,700	64,236	58,607	52,647
Nintedanib – drugs	14,302,378	25,846,153	31,059,607	33,065,909	33,877,563
Nintedanib – hospitalisation	346,535	789,173	1,218,471	1,612,418	1,963,275
BSC – drugs	0	0	0	0	0

BSC – hospitalisation	702,734	1,292,924	1,781,825	2,205,927	2,561,436
Subtotal	*18,537,964*	*29,678,688*	*34,866,169*	*37,227,123*	*38,551,772*
TOTAL	**20,431,457**	**31,840,917**	**37,211,238**	**39,819,819**	**41,360,183**

The overall budget impact of IPF treatment in England under the three scenarios is summarised in Table 127. Expanding treatment with pirfenidone to patients with mild IPF would increase the budget impact of treatment of patients with IPF from £20.7 million to £27.0 million in 2016: an increase of £6.4 million. The net budget impact of expanding treatment would show a slight increase over time: up to +£9.3 million in 2020.

Not recommending treatment with pirfenidone in any population would also increase the budget impact of treatment of patients with IPF. The net budget impact of not recommending pirfenidone in any population would result in little change to net budget impact at list price reduction of £237k in 2016; £569k in 2020.

Table 127: Effects of changing recommendations for pirfenidone on the budget impact of the treatment of patients with IPF in England

Budget impact					
	Calendar year				
Scenario	**2016**	**2017**	**2018**	**2019**	**2020**
1. Existing recommendations	20,668,369	33,087,524	39,859,617	44,098,784	47,045,921
2. Expanded treatment	27,032,098	40,690,030	48,035,689	52,754,470	56,296,412
3. No pirfenidone	20,431,457	31,840,917	37,211,238	39,819,819	41,360,183
Budget impact compared with Scenario 1: Existing recommendations (% change)					
	Calendar year				
Scenario	**2016**	**2017**	**2018**	**2019**	**2020**
2. Expanded treatment	6,363,729 (+31)%	7,602,506 (+23)%	8,176,072 (+21)%	8,655,686 (+20)%	9,250,491 (+20)%
3. No pirfenidone	-236,912 (-1)%	-1,246,607 (-4)%	-2,648,379 (-7)%	-4,278,965 (-10)%	-5,685,738 (-12)%

References

Albera 2010. Albera C, Bradford W, Costabel U, King T, Noble P, Sahn S, et al. The magnitude of pirfenidone treatment effect in patients with idiopathic pulmonary fibrosis (IPF): A pooled analysis of outcomes in the CAPACITY (CAP) studies. In: European Respiratory Society International Congress. Barcelona; 2010. p. [A389].

Albera 2011. Albera, C. Challenges in idiopathic pulmonary fibrosis trials: the point on end-points. Eur.Respir.Rev. 2011; 20(121): 195-200.

Albera 2015. Albera C. Pirfenidone Is Efficacious in Patients With Idiopathic Pulmonary Fibrosis (IPF) and Mild Restrictive Disease: Benefit of Early Intervention. Oral presentation presented at ATS 2015 annual congress; 15-20 May, Denver, USA

ATS/ERS, 2000. American Thoracic Society and European Respiratory Society. Idiopathic pulmonary fibrosis: diagnosis and treatment: international consensus statement. Am J Respir Crit Care Med 2000; 161: 646–664.

Azuma 2002. Azuma, A., et al. A placebo control and double blind phase II clinical study of pirfenidone in patients with idiopathic pulmonary fibrosis in Japan [abstract]. American Journal of Respiratory and Critical Care Medicine, 2002. A729.

Azuma 2005. Azuma A, Nukiwa T, Tsuboi E et al. Double-blind, Placebo-controlled Trial of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis. Am J Respir Crit Care 2005; 171: 1040–1047

Azuma 2011. Azuma A, Taguchi Y, Ogura T, Ebina M, Taniguchi H, Kondoh Y, et al. Exploratory analysis of a phase III trial of pirfenidone identifies a subpopulation of patients with idiopathic pulmonary fibrosis as benefiting from treatment. Respiratory Research. 2011;12:[143].

Baba 2009. Baba T. Safety analysis of pirfenidone in clinical use. In: Asian Pacific Society of Respirology 14th Congress. Seoul; 2009. p. [A375].

BLF 2014. British Lung Foundation. Pulmonary fibrosis services in the UK. September 2014. Available on: https://www.blf.org.uk/Page/IPF-Services-around-the-UK [date accessed: Dec 2015]

BNF 2016. British New Formulary. Pirfenidone. January 2016. Accessed via: www.medicinescomplete.com [date accessed: Jan 2016]

BTS 2014. The British Thoracic Society Interstitial Lung Disease Registry Programme Annual Report 2014/15. BRITISH THORACIC SOCIETY REPORTS VOL 7 ISSUE 1 2015

Boehringer Ingelheim 2015. Boehringer Ingelheim Ltd. Single technology appraisal: Nintedanib for the treatment of adults with idiopathic pulmonary fibrosis: Company evidence submission. Bracknell: Boehringer Ingelheim Ltd; 2015.

Burge 2015. Burge G, Petkova D, Ghani S et al. Pirfenidone treatment is only available in the UK for a minority of patients with usual interstitial pneumonitis. Thorax 2015;70(Suppl 3):A79

Brusasco 1997. Brusasco, V., A. Pellegrino, and J. Rodarte, Vital capacities in acute and chronic airway obstruction: dependence on flow and volume histories. European Respiratory Journal, 1997. 10: p. 1316–1320

CADTH 2014. Canadian Agency for Drugs and Technologies in Health. CDEC Final Recommendation - Pirfenidone Resubmission. Ottwa: Canadian Agency for Drugs and Technologies in Health; 2014. Available from: https://www.cadth.ca/pirfenidone-12.

CADTH 2015. Canadian Agency for Drugs and Technologies in Health. CDEC Final Recommendation - Nintedanib. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2015. Available from: https://www.cadth.ca/nintedanib.

Chaudhuri 2013. Chaudhuri N, Duck A, Frank R, et al. Real world experiences: pirfenidone is well tolerated in patients with idiopathic pulmonary fibrosis. Respir Med 2014; 108: 224–226.

Chhabra 1998. Chhabra SK. Forced Vital Capacity, Slow Vital Capacity, or Inspiratory Vital Capacity: Which Is the Best Measure of Vital Capacity? Journal ot Asthma 1998, 35(4), 361-365 (1998)

Collard 2003. Collard, H. R., King, T. E., Jr., Bartelson, B. B. et al. Changes in clinical and physiologic variables predict survival in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2003; 168(5): 538-542

Collard 2007. Collard HR, Moore BB, Flaherty KR, Brown KK, Kaner RJ, King TE Jr, Lasky JA, Loyd JE, Noth I, Olman MA, et al.: Acute exacerbations of idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2007, 176:636-643

Collard 2012. Collard HR, Ward AJ, Lanes S et al. Burden of illness in idiopathic pulmonary fibrosis. Journal of Medical Economics 2012; 15 (5): 829–835

Collard 2013. Collard HR, Yow E, Richeldi L et al. Suspected acute exacerbation of idiopathic pulmonary fibrosis as an outcome measure in clinical trials. Respiratory Research 2013 14:73

Cooper 2012. Cooper K, Mendes D, Picot J, Loveman E. Evidence Review Group Report: Pirfenidone for the treatment of idiopathic pulmonary fibrosis. Southampton: Southampton Health Technology Assessments Centre,; 2012. Available from: https://www.nice.org.uk/guidance/ta282/documents/idiopathic-pulmonary-fibrosis-pirfenidone-evidence-review-group-report2.

Costabel 2010a. Costabel U, Albera C, Du Bois RM, Bradford WZ, King Jr TE, Noble PW, et al. Global assessment of pirfenidone-treatment outcomes over time in the capacity studies of patients with IPF. American Journal of Respiratory and Critical Care Medicine. 2010;181 (Meeting Abstracts 1)

Costabel 2010b. Costabel U, Albera C, Bradford W, King T, Noble P, Sahn S. Pirfenidone dose-response in patients with idiopathic pulmonary fibrosis (IPF): A comprehensive analysis of outcomes in CAPACITY 2 In: European Respiratory Society Annual Congress, Barcelona, Spain, September 18-22; 2010. p. [388].

Costabel 2014. Costabel U, Albera C, Bradfprd WZ et al. Analysis of lung function and survival in RECAP: An open-label extension study of pirfenidone in patients with idiopathic pulmonary fibrosis. Sarcoidosis Vasc Diffuse Lung Dis 2014; 31: 198-205

Curtis 2015. Curtis L. Unit Costs of Health and Social Care. 2015

De Vries 2000. De Vries J, Seebregts A, Drent M. Assessing health status and quality of life in idiopathic pulmonary fibrosis: Which measure should be used? Respiratory Medicine. 2000;94(3):273-78.

De Vries 2001. De Vries J, Kessels BL, Drent M. Quality of life of idiopathic pulmonary fibrosis patients. European Respiratory Journal. 2001;17(5):954-61.

Drummond 1996. Drummond M, Jefferson T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996;313(7052):275-83.

Du Bois 2009. Du Bois R, Albera C, Bradford W, Costabel U, King T, Sahn S, et al. Pirfenidone treatment for idiopathic pulmonary fibrosis (IPF): A comprehensive analysis of safety. In: European Respiratory Society International Congress. Vienna; 2009. p. [A2823].

du Bois 2011a . du Bois, R. M., Weycker, D., Albera, C. et al. Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference (a). Am.J.Respir.Crit Care Med. 2011c; 183(9): 1231-1237.

du Bois 2011b. du Bois, R. M., Weycker, D., Albera, C. et al. Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis (b). Am.J.Respir.Crit Care Med. 2011b; 184(4): 459-466.

du Bois 2011c. du Bois RM. Idiopathic pulmonary fibrosis: present understanding and future options (d). Eur Respir Rev 2011; 20(121):132-133.

du Bois 2014. du Bois RM, Albera C, Bradford WZ, et al. Idiopathic pulmonary fibrosis: present understanding and future options. Eur Respir J. 2014 May;43(5):1421-9.

Ebina 2009. Ebina M, Kimura Y, Ohta H, Ohkouchi S, Hisata S, Hirota N, et al. Enhanced effects of pirfenidone on the early phase of idiopathic pulmonary fibrosis (IPF). In: European Respiratory Society International Congress. Vienna; 2009. p. [P666].

Ebina 2010. Ebina M, Kimura Y, Ohta H, Hisata S, Tamada T, Nukiwa T. Pirfenidone, a novel anti-fibrotic agent, provides higher benefit for idiopathic pulmonary fibrosis (IPF) patients with better pulmonary function. American Journal of Respiratory and Critical Care Medicine. 2010;181 (Meeting Abstracts 1):[A3988].

EMA 2011. European Medicines Agency. EMA/CHMP/115147/2011. CHMP assessment report. Available on: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002154/human_med_001417.jsp&mid=WC0b01ac058001d125 [date accessed: Dec 2015]

EMA 2015a. European Medicines Agency. EU/3/04/241: Public summary of positive opinion for orphan designation: Pirfenidone for the treatment of idiopathic pulmonary fibrosis. Available on: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/orphans/2011/03/human_orphan_000905.jsp&mid=WC0b01ac058001d12b [date accessed: Dec 2015]

EMA 2015b. European Medicines Agency. EMA/PRAC/799709/2015. Pharmacovigilance risk assessment committee minutes. 3 Dec 2015.

EMC 2015a. Esbriet Summary of Product Characteristics. Available on: www.medicines.org.uk [date of revision of text: 08 Sept 2015]

EMC 2015b. Ofev Summary of Product Characteristics. Available on: www.medicines.org.uk [date of revision of text: 30 Mar 2015]

FDA 2010. U.S. Food and Drug Administration. Pirfenidone Capsules [NDA 22-535]. Sponsor's Briefing Document for the FDA Pulmonary-Allergy Drugs Advisory Committee Meeting Scheduled for

09 March 2010. Available on: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/Pulmonary-AllergyDrugsAdvisoryCommittee/UCM203083.pdf

FDA 2014a. U.S. Food and Drug Administration. Medical Review for pirfenidone. Application No. 022535. Available on: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/022535Orig1s000toc.cfm [date accessed: Dec 2015]

FDA 2014b. U.S. Food and Drug Administration. FDA approves Esbriet to treat idiopathic pulmonary fibrosis. 15 Oct 2014. Press release available on: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm418991.htm [date accessed: Dec 2015]

Fisher 2015. Fisher M. Disease-progression modeling in idiopathic pulmonary fibrosis: Prediction of time to disease progression and life expectancy with pirfenidone. Oral presentation presented at BTS 2015 annual congress; 2-4 Dec, London, UK

Flaherty 2003. Flaherty, K. R., Mumford, J. A., Murray, S. et al. Prognostic implications of physiologic and radiographic changes in idiopathic interstitial pneumonia. Am.J.Respir.Crit Care Med. 2003; 168(5): 543-548.

Freemantle 2015. Freemantle N, Wilson A, Fisher M. Mapping The St George's Respiratory Questionnaire To The Euroqol 5 Dimensions: A Study In Patients With Idiopathic Pulmonary Fibrosis. Value in Health. 2015;18(7):A503.

Gries 2013. Gries, K. S., Esser, D., & Wiklund, I.. Content validity of CASA-Q cough domains and UCSD-SOBQ for use in patients with IPF. Glob.J Health Sci. 2013; 5(6): 131-141.

Groen 2004. Groen H, van der Bij W, Koëter GH, et al. Cost-effectiveness of lung transplantation in relation to type of end-stage pulmonary disease. Am J Transplant. 2004 Jul;4(7):1155-62.

Handorf 2012. Handorf, E. A., McElligott, S., Vachani, A et al. Cost Effectiveness of Personalized Therapy for First-Line Treatment of Stage IV and Recurrent Incurable Adenocarcinoma of the Lung. Journal of Oncology Practice 2012, 8(5), 267–274. http://doi.org/10.1200/JOP.2011.000502

Hatziandreu 2008. Hatziandreu. The potential cost savings of greater use of home- and hospice based end of life care in England (2008).

Harari 2015. Harari S and Caminati A. Idiopathic pulmonary fibrosis: from clinical trials to real-life experiences. Eur Respir Rev 2015; 24: 420–427

Health Canada 2012. Health Canada. Summary basis of decision for pirfenidone. Available on: http://www.hc-sc.gc.ca/dhp-mps/prodpharma/sbd-smd/drug-med/sbd_smd_2012_esbriet_153849-eng.php#summary [date accessed: Dec 2015]

InterMune 2009a. InterMune. CAPACITY 2 PIPF-004 Final Clinical Study Report. Date of final report: 26 June 2009

InterMune 2009b . InterMune. CAPACITY 1 PIPF-006 Final Clinical Study Report. Date of final report: 30 June 2009

InterMune 2011. InterMune. Single technology appraisal: Pirfenidone for the treatment of idiopathic pulmonary fibrosis: Manufacturer's submission. Brisbane, CA: InterMune; 2011

InterMune 2014 . InterMune. ASCEND PIPF-016 Final Clinical Study Report. Date of final report: 22 May 2014.

Karimi-Shah 2015. Karimi-Shah BA, Chowdhury BA. Forced vital capacity in idiopathic pulmonary fibrosis-FDA review of pirfenidone and nintedanib. New England Journal of Medicine. 2015;372(13):1189-91.

Kim 2006. Kim S, Collard HR, and King TE. Classification and natural history of the idiopathic interstitial pneumonias. Proc Am Thorac Soc 2006; 3: 285-292

Kim 2013. Kim DS. Acute Exacerbations in Patients With Idiopathic Pulmonary Fibrosis. Respiratory Research. 2013;14(86)

King 2008. King TE, Jr., Behr J, Brown KK, du Bois RM, Lancaster L, de Andrade JA, et al. BUILD-1: a randomized placebo-controlled trial of bosentan in idiopathic pulmonary fibrosis. American Journal of Respiratory & Critical Care Medicine. 2008;177(1):75-81.

King 2011a. King TE, Albera C, Du Bois RM, Bradford W, Costabel U, Noble PW, et al. The effect of treatment with pirfenidone on longitudinal change in lung volume in patients with idiopathic pulmonary fibrosis (IPF): A meta-analysis of outcomes in four randomised controlled clinical trials. American Journal of Respiratory and Critical Care Medicine. 2011;183 (Meeting Abstracts 1):[A5302].

King 2011b. King TE, Jr., Brown KK, Raghu G, du Bois RM, Lynch DA, Martinez F, et al. BUILD-3: a randomized, controlled trial of bosentan in idiopathic pulmonary fibrosis. American Journal of Respiratory and Critical Care Medicine. 2011;184(1):92-9.

King 2014a. King TE, Bradford WZ, Castro-Bernardini S. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. N Engl J Med 2014; 370: 2083-2092

King 2014a Suppl. King TE, Bradford WZ, Castro-Bernardini S. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. N Engl J Med 2014; 370: 2083-2092 [Supplementary Appendix]

King 2014b. King TE, Bradford WZ, Castro-Bernadini S, Fagan EA, Glaspole I, Glassberg MK. The ASCEND study: a randomised, double-blind, placebo controlled trial of pirfenidone in patients with idiopathic pulmonary fibrosis (IPF). American Journal of Respiratory and Critical Care Medicine. 2014: A6602.

King 2014c. King TE, Jr., Albera C, Bradford WZ, Costabel U, du Bois RM, Leff JA, et al. All-cause mortality rate in patients with idiopathic pulmonary fibrosis. Implications for the design and execution of clinical trials. American Journal of Respiratory & Critical Care Medicine. 2014;189(7):825-31.

Kondoh 2010. Kondoh Y, Taniguchi H, Katsuta T, Kataoka K, Kimura T, Nishiyama O, et al. Risk factors of acute exacerbation of idiopathic pulmonary fibrosis. Sarcoidosis, vasculitis, and diffuse lung diseases : official journal of WASOG / World Association of Sarcoidosis and Other Granulomatous Disorders. 2010;27(2):103-10.

Kreuter 2014. Kreuter M. Pirfenidone: an update on clinical trial data and insights from everyday practice. European Respiratory Review Mar 2014, 23 (131) 111-117; DOI: 10.1183/09059180.00008513

Kreutzkamp 2014. Kreutzkamp B. Pirfenidone inhibits loss of vital capacity and disease progression [German] Pirfenidon hemmt vitalkapazitats-verlust und krankheitsprogression. Krankenhauspharmazie. 2014;35(9):2142-43

Krishnan 2008. Krishnan V, McCormack MC, Mathai SC, Agarwal S, Richardson B, Horton MR, et al. Sleep quality and health-related quality of life in idiopathic pulmonary fibrosis. Chest. 2008;134(4):693-98.

Lancaster 2016. Lancaster L, Albera C, Bradford WZ et al. Safety of pirfenidone in patients with idiopathic pulmonary fibrosis: integrated analysis of cumulative data from 5 clinical trials. BMJ Open Resp Res 2016; 3: e000105 doi:10.1136/bmjresp-2015-000105

Latsi 2003. Latsi, P. I., du Bois, R. M., Nicholson, A. G. et al. Fibrotic idiopathic interstitial pneumonia: the prognostic value of longitudinal functional trends. Am.J.Respir.Crit Care Med. 2003; 168(5): 531-537.

Lederer 2014. Lederer DJ, Albera C, Bradford WZ et al. Effect of Treatment With Pirfenidone in Patients With Idiopathic Pulmonary Fibrosis (IPF): a Meta-analysis of 5 Randomized, Double-blind, Placebo Controlled Clinical Trial [A17]. Poster presented at ATS 2014 annual congress; 16-21 May, San Diego, USA.

Lederer 2015. Lederer DJ, Bradford WZ, Fagan EA, Glaspole I, Glassberg MK, Glasscock KF, et al. Sensitivity analyses of the change in forced vital capacity in a phase 3 trial of pirfenidone for idiopathic pulmonary fibrosis. Chest 2015; 148(1): 196‐201

Lee 2014 . Lee, A. S., Mira-Avendano, I., Ryu, J. H. et al. The burden of IPF: An unmet public health need. Respir Med. 2014.

Ley 2011. Ley B, Collard HR, King TE. Clinical Course and Prediction of Survival in Idiopathic

Pulmonary Fibrosis. Am J Respir Crit Care Med 2011; 183: 431–440

Ley 2012. Ley B, Ryerson CJ, E Vittinghoff. A Multidimensional Index and Staging System for Idiopathic Pulmonary Fibrosis. Ann Intern Med. 2012;156:684-691.

Lindell 2010. Lindell KO, Olshansky E, Song M-K, Zullo TG, Gibson KF, Kaminski N, et al. Impact of a disease-management program on symptom burden and health-related quality of life in patients with idiopathic pulmonary fibrosis and their care partners. Heart & Lung. 2010;39(4):304-13.

Loveman 2014. Loveman E, Copley VR, Colquitt JL, Scott DA, Clegg AJ, Jones J, et al. The effectiveness and cost-effectiveness of treatments for idiopathic pulmonary fibrosis: systematic review, network meta-analysis and health economic evaluation. BMC Pharmacology & Toxicology. 2014;15:63.

Loveman 2015. Loveman E, Copley VR, Colquitt J, Scott DA, Clegg A, Jones J, et al. The clinical effectiveness and cost-effectiveness of treatments for idiopathic pulmonary fibrosis: a systematic review and economic evaluation. Health Technology Assessment. 2015;19(20):1-336.

Martindale 2015. Martindale: The Complete Drug Reference. Pirfenidone. Accessed via: www.medicinescomplete.com [date accessed: Dec 2015]

Martinez 2000. Martinez TY, Pereira CA, dos Santos ML, Ciconelli RM, Guimaraes SM, Martinez JA. Evaluation of the short-form 36-item questionnaire to measure health-related quality of life in patients with idiopathic pulmonary fibrosis. Chest. 2000;117(6):1627-32.

Martinez 2002. Martinez JAB, Martinez TY, Lovetro Galhardo FP, de Castro Pereira CA. Dyspnea scales as a measure of health-related quality of life in patients with idiopathic pulmonary fibrosis. Medical Science Monitor. 2002;8(6):CR405-10.

Martinez 2014. The Idiopathic Pulmonary Fibrosis Clinical Research Network. Randomised trial of acetylcysteine in idiopathic pulmonary fibrosis. N Engl J Med 2014; 370: 2093-2101

Martinez 2014 Suppl. The Idiopathic Pulmonary Fibrosis Clinical Research Network. Randomised trial of acetylcysteine in idiopathic pulmonary fibrosis. N Engl J Med 2014; 370: 2093-2101 [Supplementary Appendix]

Meltzer 2008. Meltzer, E.B. and P.W. Noble, Idiopathic pulmonary fibrosis. Orphanet J Rare Dis, 2008. 3: p. 8.

Mendonca 2013. Mendonca, L., Perelman, J., Rodrigues, V. et al; Cost-effectiveness of lung transplantation and its evolution: the Portuguese case. Eur J Health Econ . 8-7-2013.

Nathan 2011. Nathan, S. D., Shlobin, O. A., Weir, N. et al. Long-term course and prognosis of idiopathic pulmonary fibrosis in the new millennium. Chest. 2011; 140(1): 221-229.

Nathan 2014. Nathan SD Meyer KC. IPF clinical trial design and endpoints. Curr Opin Pulm Med. 2014 Sep; 20(5): 463–471.

Nathan 2015a. Nathan SD. Effect of Continued Treatment With Pirfenidone Following a Clinically Meaningful Decline in Percent Predicted Forced Vital Capacity in Patients With Idiopathic Pulmonary Fibrosis. Oral presentation presented at ATS 2015 annual congress; 15-20 May, Denver, USA

Nathan 2016 [AiC]. Nathan SD, Albera C, Bradford WZ. Effect of continued treatment with pirfenidone following clinically meaningful declines in forced vital capacity: Analysis of data from three phase 3 trials in patients with idiopathic pulmonary fibrosis. In review with Thorax [AiC: unpublished]

Navaratnam 2011. Navaratnam V, Fleming KM, West J et al. The rising incidence of idiopathic pulmonary fibrosis in the UK. Thorax 2011; 66: 462-467

Navaratnam 2013. Navaratnam V, Fogarty AW, Glendening R et al. The Increasing Secondary Care Burden of Idiopathic Pulmonary Fibrosis: Hospital Admission Trends in England From 1998 to 2010. Chest 2013; 143 (4): 1078-1084

NHS 2013. NHS Blood and Transplant; Transplant activity in the UK, Activity Report 2012/2013. 2013. (link: https://nhsbtmediaservices.blob.core.windows.net/organ-donation-assets/pdfs/activity_report_2012_13.pdf

NICE 2010. Centre for Clinical Practice at NICE (UK). Barrett's Oesophagus: Ablative Therapy for the Treatment of Barrett's Oesophagus. London: National Institute for Health and Clinical Excellence (UK); 2010 Aug. (NICE Clinical Guidelines, No. 106.) Appendix 6, Cost effectiveness analysis for Barrett's oesophagus. Available from: http://www.ncbi.nlm.nih.gov/books/NBK109765/

NICE 2011. National Institute for Health and Care Excellence. Pirfenidone for the treatment of idiopathic pulmonary fibrosis. Appendix C - Institute's response to consultee and commentator comments on the draft scope. October 2011. Available on: http://www.nice.org.uk/guidance/TA282/documents/idiopathic-pulmonary-fibrosis-pirfenidone-appendix-c-institutes-response-to-consultee-and-commentator-comments-on-the-draft-scope2 [date accessed: Jan 2016]

NICE DSU 2011. NICE DSU TECHNICAL SUPPORT DOCUMENT 14: SURVIVAL ANALYSIS FOR ECONOMIC EVALUATIONS ALONGSIDE CLINICAL TRIALS - EXTRAPOLATION WITH PATIENT-LEVEL DATA. NICHOLAS LATIMER 2011

NICE DSU 2011a. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011, NICE Decision Support Unit (DSU): Sheffield, UK.

NICE 2013a. National Institute for Health and Care Excellence. Pirfenidone for treating idiopathic pulmonary fibrosis [TA282]. Published: 24 April 2013

NICE 2013b. National Institute for Health and Care Excellence. Idiopathic pulmonary fibrosis in adults: diagnosis and management. June 2013 [CG163].

NICE 2013c. National Institute for Health and Care Excellence. Costing report: Idiopathic pulmonary fibrosis. London: NICE; 2013. Available from: https://www.nice.org.uk/guidance/cg163/resources/costing-report-190080397.

NICE 2015a. National Institute for Health and Care Excellence. NICE Pathways: Idiopathic pulmonary fibrosis. Last updated: 15 Dec 2015. Available on: http://pathways.nice.org.uk/pathways/idiopathic-pulmonary-fibrosis#content=view-info-category%3Aview-about-menu [date accessed: Dec 2015]

NICE 2015b. National Institute for Health and Care Excellence. Idiopathic pulmonary fibrosis. NICE quality standard January 2015 [QS79]

NICE 2015c. National Institute for Health and Care Excellence. Idiopathic pulmonary fibrosis - nintedanib. NICEs response to comments on the draft scope and provisional matrix. March 2015. Available on: http://www.nice.org.uk/guidance/indevelopment/gid-tag491/documents [date accessed: Jan 2016]

NICE 2015d. National Institute for Health and Care Excellence. Idiopathic pulmonary fibrosis - nintedanib. Final appraisal determination. December 2015 Available on: http://www.nice.org.uk/guidance/indevelopment/gid-tag491/documents [date accessed: Jan 2016]

NICE 2015e. National Institute for Health and Care Excellence. Pirfenidone for treating idiopathic pulmonary fibrosis. Available on: https://www.nice.org.uk/guidance/GID-TAG504/documents/draft-scope-and-provisional-matrix-comments-table-post-referral [date accessed: Jan 2016]

NICE 2015f. National Institute for Health and Care Excellence. Idiopathic pulmonary fibrosis - nintedanib [ID752]: Evidence Review Group report (SHTAC). Available on: http://www.nice.org.uk/guidance/GID-TAG491/documents/idiopathic-pulmonary-fibrosis-nintedanib-id752-committee-papers2 [date accessed: Jan 2016]

NICE 2016. National Institute for Health and Care Excellence. Nintedanib for treating idiopathic pulmonary fibrosis [TA379]

Nishiyama 2005. Nishiyama O, Taniguchi H, Kondoh Y. Health-related quality of life in patients with idiopathic pulmonary fibrosis. What is the main contributing factor? Respiratory Medicine 2005; 99: 408–414

Noble 2009a. Noble P, Albera C, Bradford W, Costabel U, Kardatzke D, King T. The CAPACITY (CAP) trials: randomised, double-blind, placebo-controlled, phase III trials of pirfenidone (PFD) in patients with idiopathic pulmonary fibrosis (IPF) In: American Thoracic Society International Conference, May 15-20, San Diego; 2009. p. A1129 [Poster #216].

Noble 2009b. Noble PW. Pirfenidone in the treatment of idiopathic pulmonary fibrosis : results of two randomised, double-blind, placebo-controlled, phase III trials (The CAPACITY trials) In: American Thoracic Society International Conference, May 15-20, 2009 San Diego; 2009. p. [C98].

Noble 2010. Noble PW, Albera C, Bradford WZ, Costabel U, Kardatzke D, King TE, et al. Effect of pirfenidone on lung function and progression-free survival (PFS) in patients with idiopathic pulmonary fibrosis (IPF): A meta-analysis of three phase 3 studies. American Journal of Respiratory and Critical Care Medicine. 2010;181 (Meeting Abstracts 1):[A1257].

Noble 2011. Noble PW, Albera C, Bradford WZ et al. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. Lancet 2011; 377: 1760–69

Noble 2014a. Noble PW, Albera C, Bradford WZ et al. Analysis of Pooled Data from Three Phase 3, Multinational, Randomized, Double-blind, Placebo Controlled Trials Evaluating Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis [A11]. Poster presented at ATS 2014 annual congress; 16-21 May, San Diego, USA.

Noble 2014b [AiC]. Noble, P. W., Albera, C., Bradford, Z. W. et al. Pirfenidone for IPF: Analysis of pooled data from three multinational Phase III trials. [AiC: unpublished]

Noth 2011. Noth, I., Anstrom, K. J., calvert, S. B. et al; A Placebo-Controlled Randomized Trial of Warfarin in Idiopathic Pulmonary Fibrosis. Am J Respir Crit Care Med 186[1], 88-95. 2012.

O'Brien 2000. O'Brien, C., P. Guest, and S. Hill, Physiological and radiological characterization of patients diagnosed with chronic obstructive pulmonary disease in primary care. Thorax, 2000. 55: p. 635–642

Oda 2014. Oda K, Ishimoto H, Yamada S, Kushima H. Autopsy analyses in acute exacerbation of idiopathic pulmonary fibrosis. Respiratory Research 2014 15:109

Ogura 2008. Ogura T, Ebina M, Taniguchi H, Azuma A, Suga M, Taguchi Y, et al. A phase III double blind placebo controlled clinical trial of pirfenidone in patients with idiopathic pulmonary fibrosis in Japan In: American Thoracic Society International Conference, May 16-21, 2008, Toronto; 2008. p. [A768].

Oldham and Noth, 2014. Oldam JM and Noth I. Idiopathic pulmonary fibrosis: early detection and referral. Respiratory Medicine 2014; 108: 819-829

Parfrey 2013. Parfrey H, Leonard C, Gibbons MA, Armstrong E, Harris E, Frank R, et al. Healthcare utilisation by patients with idiopathic pulmonary fibrosis; observations from the UK pirfenidone named patient programme. Thorax. 2013;68(Suppl 3):A164-A65.

Peng 2008. Peng S, Li Z, Kang J, Hou X. Cross-sectional and longitudinal construct validity of the Saint George's Respiratory Questionnaire in patients with IPF. Respirology. 2008;13(6):871-9.

Pittrow 2014. Pittrow D, Klotsche J, Kreuter M, Hoeper MM, Wirtz H, Koschel D, et al. Symptom burden and health related quality of life in patients with idiopathic pulmonary fibrosis in clinical practice: Insights-IPF registry. Value in Health. 2014;17(7):A600.

Raghu 2006. Raghu G, Weycker D, Edelsberg J et al. Incidence and Prevalence of Idiopathic Pulmonary Fibrosis. Am J Respir Crit Care Med 2006, 174. pp 810–816

Raghu 2010. Raghu G, King TE, Jr., Behr J, Brown KK, du Bois RM, Leconte I, et al. Quality of life and dyspnoea in patients treated with bosentan for idiopathic pulmonary fibrosis (BUILD-1). European Respiratory Journal. 2010;35(1):118-23.

Raghu 2011. Raghu G, Collard HR, Egan JJ, et al. An Official ATS/ERS/JRS/ALAT Statement: Idiopathic Pulmonary Fibrosis: Evidence-based Guidelines for Diagnosis and Management. Issued 2011 (link: http://www.ncbi.nlm.nih.gov/pubmed/21471066)

Raghu 2012. The Idiopathic Pulmonary Fibrosis Clinical Research Network. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. N Engl J Med 2012; 366: 1968-1977

Raghu 2012 Suppl. The Idiopathic Pulmonary Fibrosis Clinical Research Network. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. N Engl J Med 2012; 366: 1968-1977 [Supplementary Appendix]

Raghu 2015. Raghu G, Rochwerg B, Zhang Y et al. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. Am J Respir Crit Care Med 2015; 192 (2): e3–e19

Richeldi 2011. Richeldi L, Costabel U, Selman M et al. Efficacy of a Tyrosine Kinase Inhibitor in Idiopathic Pulmonary Fibrosis. N Engl J Med 2011; 365: 1079-1087

Richeldi 2011 Suppl. Richeldi L, Costabel U, Selman M et al. Efficacy of a Tyrosine Kinase Inhibitor in Idiopathic Pulmonary Fibrosis. N Engl J Med 2011; 365: 1079-1087 [Supplementary Appendix]

Richeldi 2014. Richeldi L, du Bois R, Raghu G et al. Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis. N Engl J Med 2014; 370: 2071-82

Richeldi 2014 Suppl. Richeldi L, du Bois R, Raghu G et al. Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis. N Engl J Med 2014; 370: 2071-82 [Supplementary Appendix]

Roche, 2014. Hoffman-La Roche. Media release: Roche purchases shares in tender offer for InterMune, Inc. Available on: http://www.roche.com/media/store/releases/med-cor-2014-09-29b.htm

Roche 2016 [AiC]. Systematic Review and Network Meta-Analysis of Treatments for Idiopathic Pulmonary Fibrosis. [AiC: Manuscript in development; unpublished]

Roche 2016a. Data on file

Roskell 2014. Roskell R, Saunders O, Lee L, Fisher MJ. Long-term survival analysis: pirfenidone compared to standard care for the treatment of patients with idiopathic pulmonary fibrosis. In: International Congress 2014. Munich: European Respiratory Society; 2014. p. [A1905].

Sahn 2010. Sahn SA, Albera C, Du Bois RM, Bradford WZ, Costabel U, King TE, et al. Clinical outcomes with pirfenidone therapy in treatment-adherent (TA) patients with idiopathic pulmonary fibrosis (IPF). American Journal of Respiratory and Critical Care Medicine. 2010;181 (Meeting Abstracts 1):[A6025].

Sahn 2011. Sahn SA, Albera C, Du Bois RM, Bradford W, Costabel U, King TE, et al. The effect of treatment with pirfenidone on progression-free survival in patients with idiopathic pulmonary fibrosis (IPF): Exploratory analysis of outcomes using novel criteria for disease progression. American Journal of Respiratory and Critical Care Medicine. 2011;183 (Meeting Abstracts 1):[A3810].

SMC 2013. Scottish Medicines Consortium. Detailed advice for pirfenidone (Esbriet®) [SMC no. 835/13]. Available on: https://www.scottishmedicines.org.uk/SMC_Advice/Advice/835_13_pirfenidone_Esbriet/pirfenidone_Esbriet [date accessed: Dec 2015]

SMC 2015. Scottish Medicines Consortium. Detailed advice for nintedanib (Ofev®) [SMC no. 1076/15]. Available on: http://www.scottishmedicines.org.uk/SMC_Advice/Advice/1076_15_nintedanib_Ofev/nintedanib_Ofev [date accessed: Dec 2015]

Starkie 2011. Starkie HJ, Briggs AH, Chambers MG, et al. Predicting EQ-5D Values Using the SGRQ. Value in Health Volume 14, Issue 2, March–April 2011, Pages 354–360 doi:10.1016/j.jval.2010.09.011

Swigris 2005. Swigris JJ, Steward AL, Gould MK et al. Patients' perspectives on how idiopathic pulmonary fibrosis affects the quality of their lives. Health and Quality of Life Outcomes 2005; 3:61

Swigris 2005a. Swigris JJ, Kuschner WG, Jacobs SS, Wilson SR, Gould MK. Health-related quality of life in patients with idiopathic pulmonary fibrosis: a systematic review. Thorax. 2005;60(7):588-94.

Swigris 2010. Swigris JJ, Brown KK, Behr J, du Bois RM, King TE, Raghu G, et al. The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. Respiratory Medicine. 2010;104(2):296-304.

Swigris 2012. Swigris JJ, Han M, Vij R et al. The UCSD Shortness of Breath Questionnaire has Longitudinal Construct Validity in Idiopathic Pulmonary Fibrosis. Respir Med 2012; 106(10): 1447–1455

Swigris 2014. Swigris, J. J., Streiner DL, Brown, K. K. et al. Assessing exertional dyspnea in patients with IPF. 2014; 108: 181-188.

Swigris 2015. Swigris JJ, Stewart AL, Gould MK. Patients' perspectives on how idiopathic pulmonary fibrosis affects the quality of their lives. Health and Quality of Life Outcomes 2005, 3:61

Taguchi 2010. Taguchi Y, Ebina M, Hashimoto S, Ogura T, Azuma A, Taniguchi H, et al. Extended analysis of phase III trial of pirfenidone: More pronounced effect in mild impairment of patients with idiopathic pulmonary fibrosis (IPF). In: European Respiratory Society International Congress. Barcelona; 2010. p. [A390].

Taniguchi 2010. Taniguchi H, Ebina M, Kondoh Y et al. Pirfenidone in idiopathic pulmonary fibrosis. Eur Respir J 2010; 35: 821–829

Taniguchi 2011a. Taniguchi H, Kondoh Y, Ebina M, Azuma A, Ogura T, Taguchi Y, et al. The clinical significance of 5% change in vital capacity in patients with idiopathic pulmonary fibrosis: extended analysis of the pirfenidone trial. Respiratory Research. 2011;12:[93].

Taniguchi 2011b. Taniguchi H, Kondoh Y, Ebina M, Azuma A, Ogura T, Suga M, et al. Risk factors of acute exacerbation of idiopathic pulmonary fibrosis-extended analysis of the pirfenidone trial in Japan. In: European Respiratory Society International Congress. Amsterdam; 2011. p. [A649].

Thabut 2003. Thabut G, Mal H, Castier Y, Groussard O et al. Survival benefit of lung transplantation for patients with idiopathic pulmonary fibrosis. J Thorac Cardiovasc Surg. 2003 Aug;126(2):469-75.

Tomioka 2007. Tomioka H, Imanaka K, Hashimoto K, Iwasaki H. Health-related quality of life in patients with idiopathic pulmonary fibrosis--cross-sectional and longitudinal study. Internal Medicine. 2007;46(18):1533-42.

Tomioka 2015. Tomioka H, Imanaka K, Hashimoto K, Iwasaki H. Health-related quality of life in patients with idiopathic pulmonary fibrosis--cross-sectional and longitudinal study. Internal Medicine. 2007;46(18):1533-42.

Tzanakis 2005. Tzanakis N, Samiou M, Lambiri I, Antoniou K, Siafakas N, Bouros D. Evaluation of health-related quality-of-life and dyspnea scales in patients with idiopathic pulmonary fibrosis. Correlation with pulmonary function tests. European Journal of Internal Medicine. 2005;16(2):105-12.

Valeyre 2010a. Valeyre D, Albera C, Du Bois RM, Bradford WZ, Costabel U, King Jr TE, et al. 6 minute walk distance (6MWD) and forced vital capacity (FVC) in patients with idiopathic pulmonary fibrosis (IPF): similar pattern of pirfenidone response. American Journal of Respiratory and Critical Care Medicine. 2010;181 (Meeting Abstracts 1):[A6026].

Valeyre 2010b. Valeyre D, Albera C, Bradford W, Costabel U, King T, Noble P, et al. The effect of treatment with pirfenidone on death or lung transplantation in patients with idiopathic pulmonary fibrosis (IPF): Analysis of outcomes in the CAPACITY (CAP) trials. In: European Respiratory Society International Congress. Barcelona; 2010. p. [A391].

Valeyre 2014. Valeyre D, Albera C, Bradford WZ, Costabel U, King TE, Jr., Leff JA, et al. Comprehensive assessment of the long-term safety of pirfenidone in patients with idiopathic pulmonary fibrosis. Respirology. 2014;19(5):740-47.

Vancheri 2010. Vancheri C, Failla M, Crimi N et al. Idiopathic pulmonary fibrosis: a disease with similarities and links to cancer biology. Eur Respir J 2010; 35: 496–504

Weill 2015. Weill D, Benden C, Corris PA. A consensus document for the selection of lung transplant candidates: 2014--an update from the Pulmonary Transplantation Council of the International Society for Heart and Lung Transplantation. J Heart Lung Transplant. 2015 Jan;34(1):1-15. doi: 10.1016/j.healun.2014.06.014. Epub 2014 Jun 26.

Wells 2003. Wells AU, Desai SR, MB Rubens et al. Idiopathic Pulmonary Fibrosis: A Composite Physiologic Index Derived from Disease Extent Observed by Computed Tomography. Am J Respir Crit Care Med 2003; 167: 962–969

Wells 2008a. Wells AU, Hirani N et al. Interstitial lung disease guideline: the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. Thorax 2008; 63: v1-v58

Wells 2008b. Wells AU, Behr J and Silver R. Outcome measures in the lung. Rheumatology 2008;47:v48–v50

Whiting 2014. Whiting P, Al M, Burgers L, Westwood M, Ryder S, Hoogendoorn M, et al. Ivacaftor for the treatment of patients with cystic fibrosis and the G551D mutation: a systematic review and cost-effectiveness analysis. Health Technology Assessment. 2014;18(18):1-106.

Wilson 2014. Wilson ECF, Shulgina L, Cahn AP, Chilvers ER, Parfrey H, Clark AB, et al. Treating idiopathic pulmonary fibrosis with the addition of co-trimoxazole: an economic evaluation alongside a randomised controlled trial. Pharmacoeconomics. 2014;32(1):87-99.

Yorke 2010. Yorke J, Jones PW, Swigris JJ. Development and validity testing of an IPF-specific version of the St George's Respiratory Questionnaire. Thorax. 2010;65(10):921-6.

Zappala 2010. Zappala, C. J., Latsi, P. I., Nicholson, A. G. et al. Marginal decline in FVC is associated with a poor outcome in IPF. Eur.Respir.J. 2010; 35(4): 830-83

Zibrak JD and Price D, 2014. Zibrak JD and Price D. Interstitial lung disease: raising the index of suspicion in primary care. NPJ Primary Care Respiratory Medicine 2014; 24: 14054

Zisman 2010. Zisman DA, Schwarz M, Anstrom KJ, Collard HR, Flaherty KR, Hunninghake GW. A controlled trial of sildenafil in advanced idiopathic pulmonary fibrosis. New England Journal of Medicine. 2010;363(7):620-8.

Appendices

Appendices to this evidence submission have been supplied as a separate file.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Technology appraisals

Patient access scheme submission template

October 2009

1 Introduction

The 2009 Pharmaceutical Price Regulation Scheme (PPRS) (www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpriceregulationscheme/2009PPRS) is a non-contractual scheme between the Department of Health and the Association of the British Pharmaceutical Industry. The purpose of the 2009 PPRS is to ensure that safe and cost-effective medicines are available on reasonable terms to the NHS in England and Wales. One of the features of the 2009 PPRS is to improve patients' access to medicines at prices that better reflect their value through patient access schemes.

Patient access schemes are arrangements which may be used on an exceptional basis for the acquisition of medicines for the NHS in England and Wales. Patient access schemes propose either a discount or rebate that may be linked to the number, type or response of patients, or a change in the list price of a medicine linked to the collection of new evidence (outcomes). These schemes help to improve the cost effectiveness of a medicine and therefore allow the National Institute for Health and Care Excellence (NICE) to recommend treatments which it would otherwise not have found to be cost effective. More information on the framework for patient access schemes is provided in the 2009 PPRS (www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpriceregulationscheme/2009PPRS.

Patient access schemes are proposed by a pharmaceutical company and agreed with the Department of Health, with input from the Patient Access Schemes Liaison Unit (PASLU) within the Centre for Health Technology Evaluation at NICE.

2 Instructions for manufacturers and sponsors

This document is the patient access scheme submission template for technology appraisals. If manufacturers and sponsors want the National Institute for Health and Care Excellence (NICE) to consider a patient access scheme as part of a technology appraisal, they should use this template. NICE can only consider a patient access scheme after formal referral from the Department of Health.

The template contains the information NICE requires to assess the impact of a patient access scheme on the clinical and cost effectiveness of a technology, in the context of a technology appraisal, and explains the way in which background information (evidence) should be presented. If you are unable to follow this format, you must state your reasons clearly. You should insert 'N/A' against sections that you do not consider relevant, and give a reason for this response.

Please refer to the following documents when completing the template:

- 'Guide to the methods of technology appraisal' (http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9)
- 'Specification for manufacturer/sponsor submission of evidence' (http://www.nice.org.uk/aboutnice/howwework/devnicetech/singletechnologyappraisalsubmissiontemplates.jsp) and
- Pharmaceutical Price Regulation Scheme 2009 (www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpriceregulationscheme/2009PPRS).

For further details on the technology appraisal process, please see NICE's 'Guide to the single technology appraisal (STA) process' and 'Guide to the multiple technology appraisal (MTA) process' (http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/technology_appraisal_process_guides.jsp). The

‘Specification for manufacturer/sponsor submission of evidence’ provides details on disclosure of information and equality issues.

Make the submission as brief and informative as possible. Only mark information as confidential when absolutely necessary. Sufficient information must be publicly available for stakeholders to comment on the full content of the technology appraisal, including details of the proposed patient access scheme. Send submissions electronically to NICE in Word or a compatible format, not as a PDF file.

Appendices may be used to include additional information that is considered relevant to the submission. Do not include information in the appendices that has been requested in the template. Appendices should be clearly referenced in the main submission.

When making a patient access scheme submission, include:

- an updated version of the checklist of confidential information, if necessary
- an economic model with the patient access scheme incorporated, in accordance with the ‘Guide to the methods of technology appraisal’ (http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9).

If you are submitting the patient access scheme at the end of the appraisal process, you should update the economic model to reflect the assumptions that the Appraisal Committee considered to be most plausible. No other changes should be made to the model.

3 Details of the patient access scheme

3.1 Please give the name of the technology and the disease area to which the patient access scheme applies.

Pirfenidone (Esbriet▼) is indicated for the treatment of mild-to-moderate Idiopathic Pulmonary Fibrosis (IPF).

3.2 Please outline the rationale for developing the patient access scheme.

The scheme rationale is to provide a cost effective treatment for patients with mild and moderate IPF, in a manner which creates no additional administrative burden to the NHS.

3.3 Please describe the type of patient access scheme, as defined by the PPRS.

The PAS has been applied by reducing the price of pirfenidone to [redacted] below the list price stated in the primary evidence submission. This equates to the net prices outlined in Table 2.

Table 1: List and net prices of pirfenidone presentations

Presentation	List price	Discount	Net price
267 mg capsules (x 63 caps)	£501.92	[redacted]	[redacted]
267 mg capsules (x 252 caps)	£2,007.70	[redacted]	[redacted]
267 mg capsules (x 270 caps)	£2,151.10	[redacted]	[redacted]

3.4 Please provide specific details of the patient population to which the patient access scheme applies. Does the scheme apply to the whole licensed population or only to a specific subgroup (for example, type of tumour, location of tumour)? If so:

- *How is the subgroup defined?*

- *If certain criteria have been used to select patients, why have these have been chosen?*
- *How are the criteria measured and why have the measures been chosen?*

The Scheme applies to both patients with mild and moderate IPF, in line with the licensed indication for pirfenidone.

3.5 *Please provide details of when the scheme will apply to the population specified in 3.4. Is the scheme dependent on certain criteria, for example, degree of response, response by a certain time point, number of injections? If so:*

- *Why have the criteria been chosen?*
- *How are the criteria measured and why have the measures been chosen.*

See above. The scheme is not dependent upon any criteria and is simply applied at the point of purchase.

3.6 *What proportion of the patient population (specified in 3.4) is expected to meet the scheme criteria (specified in 3.5)?*

The scheme will apply to all patients for whom pirfenidone is indicated.

3.7 *Please explain in detail the financial aspects of the scheme. How will any rebates be calculated and paid?*

The simple discount will be applied at the point of invoice and will not require any rebate

3.8 *Please provide details of how the scheme will be administered. Please specify whether any additional information will need to be collected, explaining when this will be done and by whom.*

No additional data will be required. The PAS discount will be applied at point of invoice and pass through the normal financial administration.

3.9 *Please provide a flow diagram that clearly shows how the scheme will operate. Any funding flows must be clearly demonstrated.*

Not applicable

3.10 *Please provide details of the duration of the scheme.*

Following positive guidance, Roche will only be entitled to terminate the PAS if pirfenidone is subject to a further NICE TA review. In the case of Roche terminating this Agreement, Roche will consult with the Department of Health and NICE prior to giving any notice of termination to the Customer.

3.11 *Are there any equity or equalities issues relating to the scheme, taking into account current legislation and, if applicable, any concerns identified during the course of the appraisal? If so, how have these been addressed.*

The proposed Scheme is consistent with and does not infringe applicable competition law, including but not limited to Articles 101 and 102 of the Treaty on the Functioning of the EU and the equivalent provisions of the Competition Act 1998

3.12 *If available, please list any scheme agreement forms, patient registration forms, pharmacy claim forms/rebate forms, guides for pharmacists and physicians and patient information documents. Please include copies in the appendices.*

Not applicable. No registration or claim forms are required.

3.13 *In the exceptional case that you are submitting an outcome-based scheme, as defined by the PPRS, please also refer to appendix B.*

Not applicable

4 Cost effectiveness

4.1 *If the population to whom the scheme applies (as described in sections 3.4 and 3.5) has not been presented in the main manufacturer/sponsor submission of evidence for the technology appraisal (for example, the population is different as there has been a change in clinical outcomes or a new continuation rule), please (re-)submit the relevant sections from the 'Specification for manufacturer/sponsor submission of evidence' (particularly sections 5.5, 6.7 and 6.9). You should complete those sections both with and without the patient access scheme. You must also complete the rest of this template.*

Not applicable: the PAS applies to the population considered in our primary evidence submission.

4.2 *If you are submitting the patient access scheme at the end of the technology appraisal process, you should update the economic model to reflect the assumptions that the Appraisal Committee considered to be most plausible. No other changes should be made to the model.*

Cost-effectiveness results, inclusive of the proposed PAS for pirfenidone (along with the assumed PAS for nintedanib in the relevant subgroup of patients) are presented below.

4.3 *Please provide details of how the patient access scheme has been incorporated into the economic model. If applicable, please also provide details of any changes made to the model to reflect the assumptions that the Appraisal Committee considered most plausible.*

The PAS has been applied by reducing the price of pirfenidone to [redacted] below the list price stated in the primary evidence submission. This equates to the net prices outlined in Table 2.

Table 2: List and net prices of pirfenidone presentations

Presentation	List price	Discount	Net price
267 mg capsules (x 63 caps)	£501.92	[redacted]	[redacted]
267 mg capsules (x 252 caps)	£2,007.70	[redacted]	[redacted]
267 mg capsules (x 270 caps)	£2,151.10	[redacted]	[redacted]

4.4 *Please provide the clinical effectiveness data resulting from the evidence synthesis and used in the economic model which includes the patient access scheme.*

As the PAS is a simple discount the clinical effectiveness data provided in the primary evidence submission is unaffected by the proposal.

4.5 *Please list any costs associated with the implementation and operation of the patient access scheme (for example, additional pharmacy time for stock management or rebate calculations). A suggested format is presented in table 1. Please give the reference source of these costs. Please refer to section 6.5 of the 'Specification for manufacturer/sponsor submission of evidence'.*

The PAS is a simple discount introduced at the point of invoicing. It is therefore not subject to operational or implementation costs.

4.6 *Please provide details of any additional treatment-related costs incurred by implementing the patient access scheme. A suggested format is presented in table 2. The costs should be provided for the intervention both with and without the patient access scheme. Please give the reference source of these costs.*

The PAS is a simple discount introduced at the point of invoicing. It is therefore not subject to operational or implementation costs.

Summary results

Base-case analysis

4.7 *Please present in separate tables the cost-effectiveness results as follows.*[1]

- *the results for the intervention without the patient access scheme*
- *the results for the intervention with the patient access scheme.*

All results provided below (4.7 to 4.13) are those associated with revised base case analysis, as per our response to the ERG clarification questions which were submitted on 3 March. The revised base case includes a number of updates to the analyses which were supplied as part of our initial evidence submission, and are outlined in the introduction to Section B of our response to the clarification questions (page 38 of the document submitted on 3 March).

As previously described, the PAS is a simple discount (a [redacted] discount on the current list price of pirfenidone).

A simple PAS is available for nintedanib, which is a relevant comparator In the subgroup of patients with moderate IPF. In these analyses, a PAS of [redacted] has been assumed for nintedanib.

Table 3: Discounted base case model results – ITT population – without the PAS

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	5.38	3.80				
PFN	[redacted]	8.67	5.67	[redacted]	3.29	1.87	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

[1] For outcome-based schemes, please see section 5.2.8 in appendix B.

Table 4: Discounted base case model results – ITT population – with the PAS

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	26,627	5.38	3.80				
PFN	66,638	8.67	5.67	40,010	3.29	1.87	21,387

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 5: Discounted base case model results – Mild population – without the PAS

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	7.11	4.82				
PFN	[redacted]	11.26	6.99	[redacted]	4.15	2.17	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 6: Discounted base case model results – Mild population – with the PAS

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	31,729	7.11	4.82				
PFN	84,209	11.26	6.99	52,480	4.15	2.17	24,187

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 7: Discounted base case model results – Moderate population (vs. BSC) – without the PAS

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	4.80	3.44				
PFN	[redacted]	7.67	5.14	[redacted]	2.87	1.70	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 8: Discounted base case model results – Moderate population (vs. BSC) – with the PAS

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	24,868	4.80	3.44				
PFN	61,012	7.67	5.14	36,145	2.87	1.70	21,318

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 9: Discounted base case model results – Moderate population (vs. NTB) – without the PAS

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
NTB	65,065	6.06	4.23				
PFN	[redacted]	7.67	5.14	[redacted]	1.61	0.91	[redacted]

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 10: Discounted base case model results – Moderate population (vs. NTB) – with the PAS (both treatments)

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
NTB	[redacted]	6.06	4.23				
PFN	[redacted]	7.67	5.14	[redacted]	1.61	0.91	[redacted]

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

4.8 Please present in separate tables the incremental results as follows. [2]

- *the results for the intervention without the patient access scheme*
- *the results for the intervention with the patient access scheme.*

List the interventions and comparator(s) from least to most expensive. Present the incremental cost-effectiveness ratios (ICERs) in comparison

[2] For outcome-based schemes, please see section 5.2.9 in appendix B.

with baseline (usually standard care), and the incremental analysis ranking technologies in terms of dominance and extended dominance.

Results for the moderate subgroup are presented in this section. Results for the ITT and mild populations are presented in response to 4.7.

Table 11: Discounted base case model results – Moderate population – without the PAS

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)	Incremental (QALYs)
BSC	24,868	4.80	3.44					
NTB	65,065	6.06	4.23	40,197	1.26	0.78	51,331	51,331
PFN	[redacted]	7.67	5.14	[redacted]	1.61	0.91	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 12: Discounted base case model results – Moderate population – with the PAS

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)	Incremental (QALYs)
BSC	24,868	4.80	3.44					
NTB	[redacted]	6.06	4.23	[redacted]	1.26	0.78	[redacted]	[redacted]
PFN	[redacted]	7.67	5.14	[redacted]	1.61	0.91	[redacted]	[redacted]

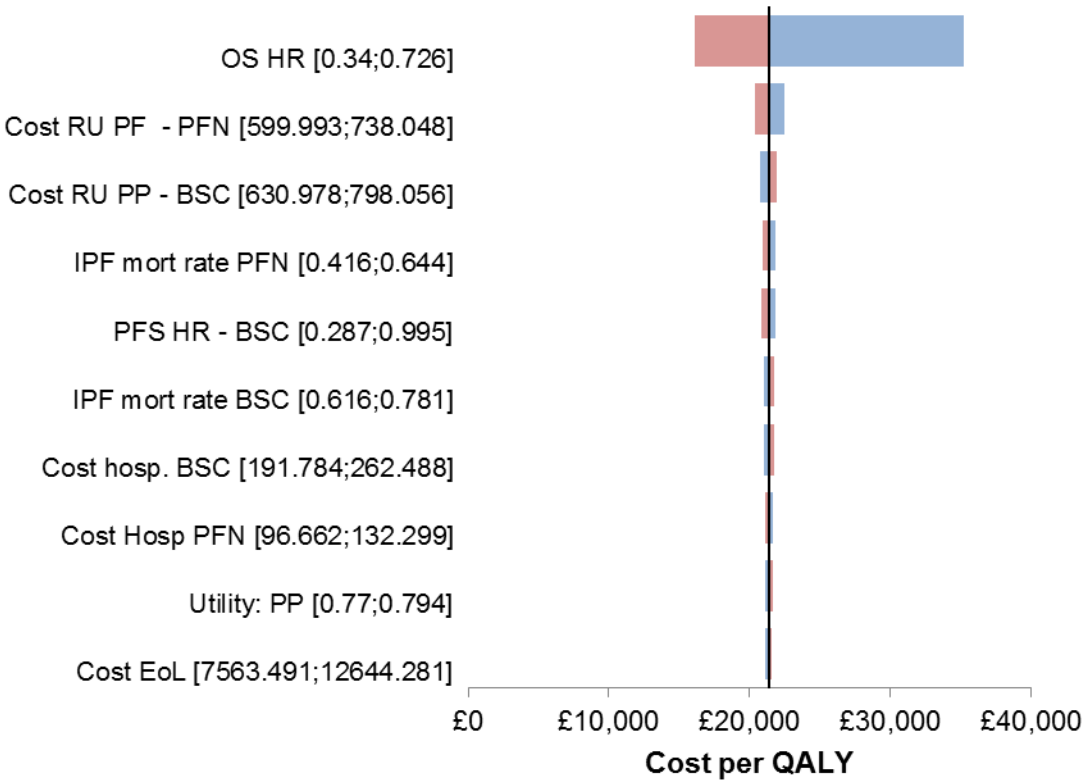
Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

; QALY: quality-adjusted life-year; ICER: incremental cost-effectiveness ratio.

Sensitivity analyses

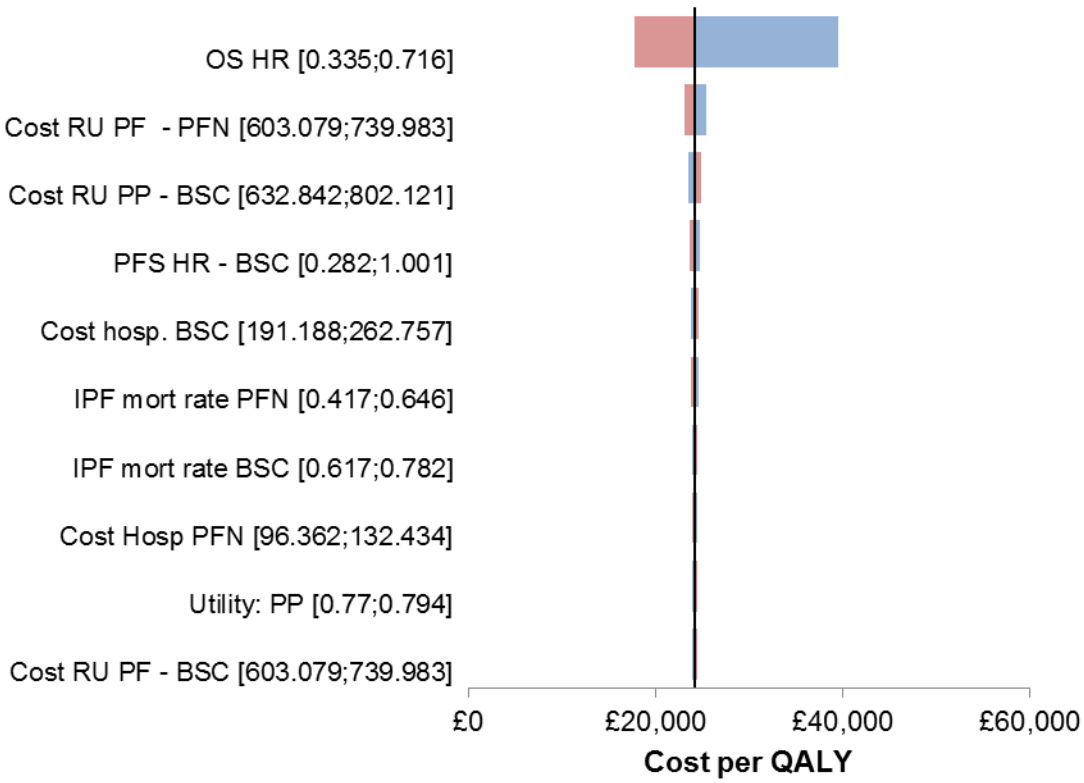
4.9 *Please present deterministic sensitivity analysis results as described for the main manufacturer/sponsor submission of evidence for the technology appraisal. Consider using tornado diagrams.*

Figure 1: Tornado diagram – ITT population – with the PAS



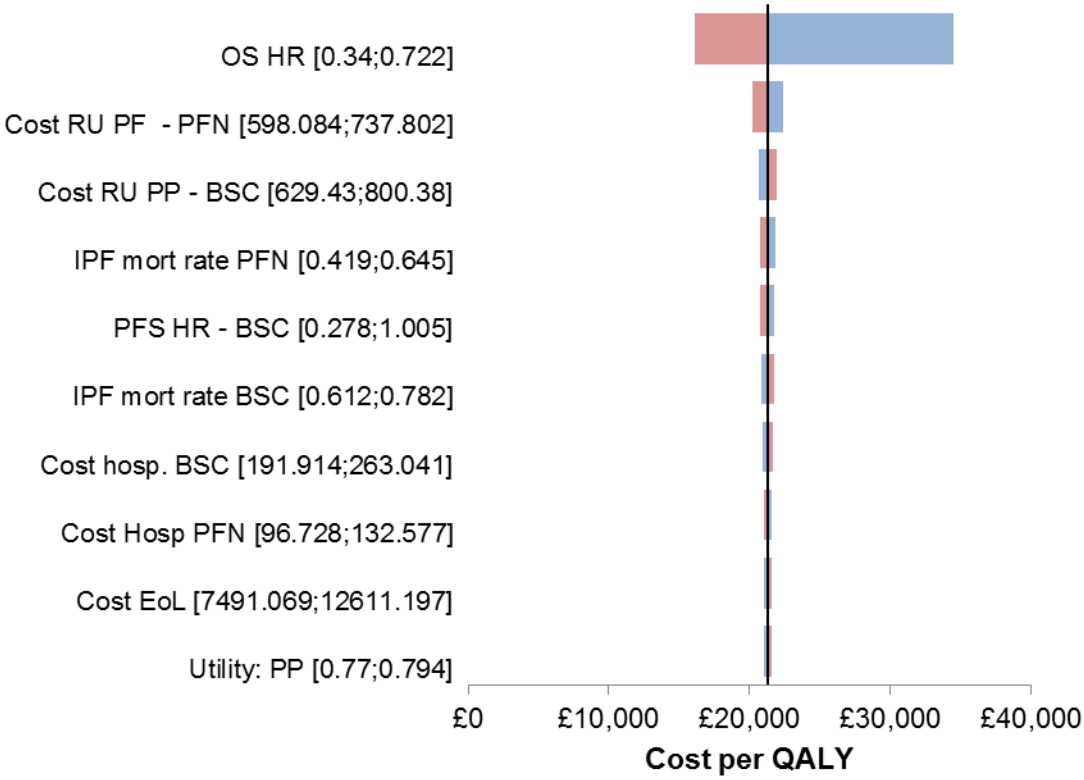
Key: BSC, best supportive care; EoL, end of life; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; mort, mortality; PF, progression-free; PFN, pirfenidone; PP, post-progression; QALY, quality-adjusted life year; RU, resource use.

Figure 2: Tornado diagram – Mild population – with the PAS



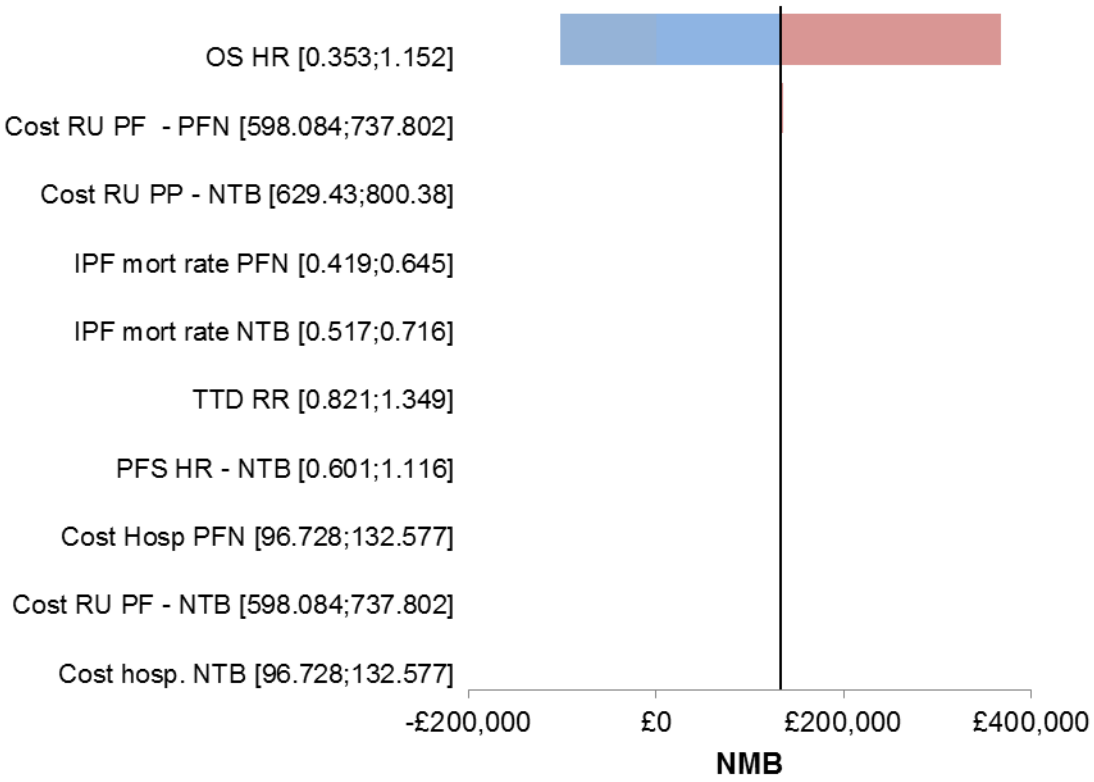
Key: BSC, best supportive care; EoL, end of life; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; mort, mortality; PF, progression-free; PFN, pirfenidone; PP, post-progression; QALY, quality-adjusted life year; RU, resource use.

Figure 3: Tornado diagram – Moderate population (vs. BSC; Cost per QALY) – with the PAS



Key: BSC, best supportive care; EoL, end of life; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; mort, mortality; PF, progression-free; PFN, pirfenidone; PP, post-progression; QALY, quality-adjusted life year; RU, resource use.

Figure 4: Tornado diagram – Moderate population (vs. NTB; NMB) – with the PAS



Key: BSC, best supportive care; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; mort, mortality; NMB, net monetary benefit; PF, progression-free; PFN, pirfenidone; PP, post-progression; QALY, quality-adjusted life year; RU, resource use; TTD, time to treatment discontinuation.

4.10 *Please present any probabilistic sensitivity analysis results, and include scatter plots and cost-effectiveness acceptability curves.*

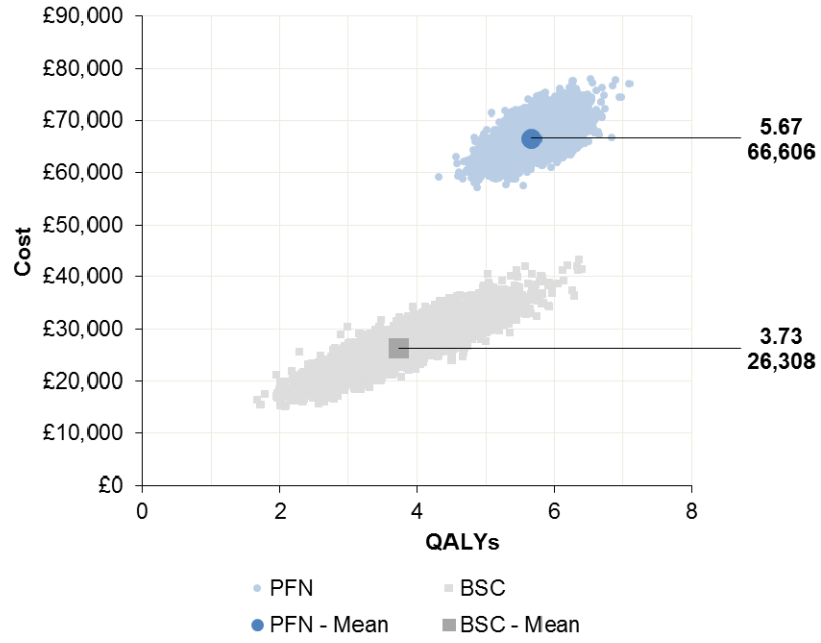
Table 13: PSA results – ITT population – with the PAS

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Vs.baseline (QALYs)
Deterministic model results							
BSC	26,627	5.38	3.80				
PFN	66,638	8.67	5.67	40,010	3.29	1.87	21,387
Mean probabilistic model results							
BSC	26,308	5.31	3.729				
95% CI	(18,899; 34,868)	(3.28; 7.88)	(2.47; 5.20)				
PFN	66,606	8.69	5.67	40,298	3.38	1.94	20,794
95% CI	(60,830; 72,826)	(7.37; 10.16)	(5.01; 6.37)				

% chance of being cost effective	£20,000 per QALY	£25,000 per QALY	£30,000 per QALY
BSC	55%	28%	15%
PFN	45%	72%	85%

Key: BSC, best supportive care; CI, confidence interval; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

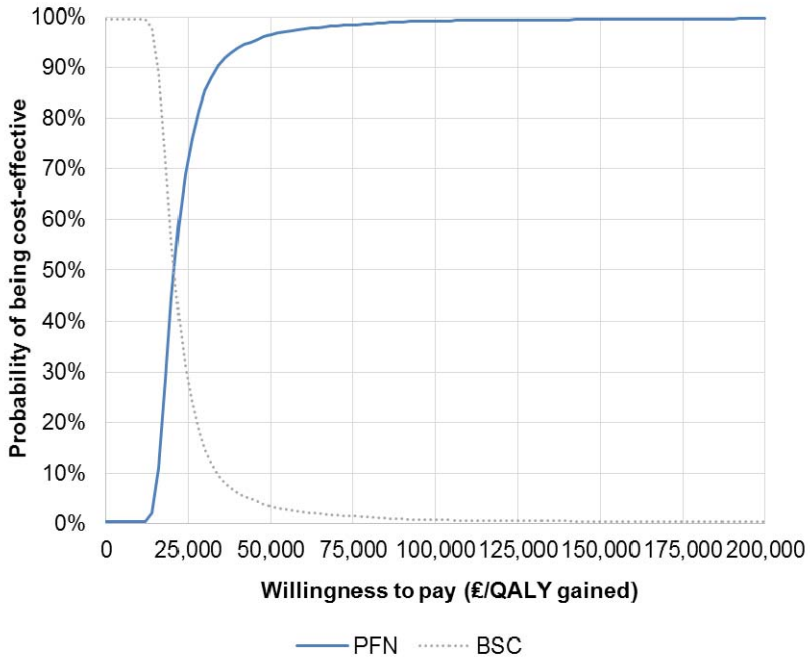
Figure 5: PSA scatterplot – ITT population – PAS price



Key: BSC, best supportive care; PFN, pirfenidone; QALY, quality-adjusted life year.

The Figure above is Commercial in Confidence

Figure 6: CEAC – ITT population – PAS price



Key: BSC, best supportive care; PFN, pirfenidone; QALY, quality-adjusted life year.

The Figure above is Commercial in Confidence

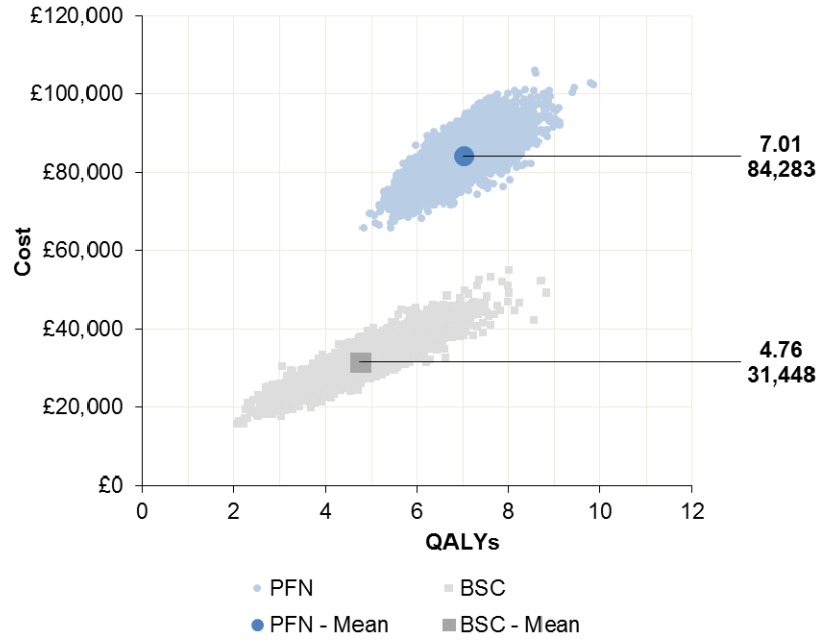
Table 14: PSA results – Mild population – with the PAS

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Vs.baseline (QALYs)
Deterministic model results							
BSC	31,729	7.11	4.82				
PFN	84,209	11.26	6.99	52,480	4.15	2.17	24,187
Mean probabilistic model results							
BSC	31,448	7.06	4.758				
95% CI	(22,026; 42,393)	(4.16; 10.75)	(3.07; 6.68)				
PFN	84,283	11.32	7.01	52,835	4.26	2.25	23,476
95% CI	(73,500; 95,448)	(8.68; 14.32)	(5.75; 8.35)				

% chance of being cost effective	£20,000 per QALY	£25,000 per QALY	£30,000 per QALY
BSC	73%	42%	24%
PFN	27%	58%	76%

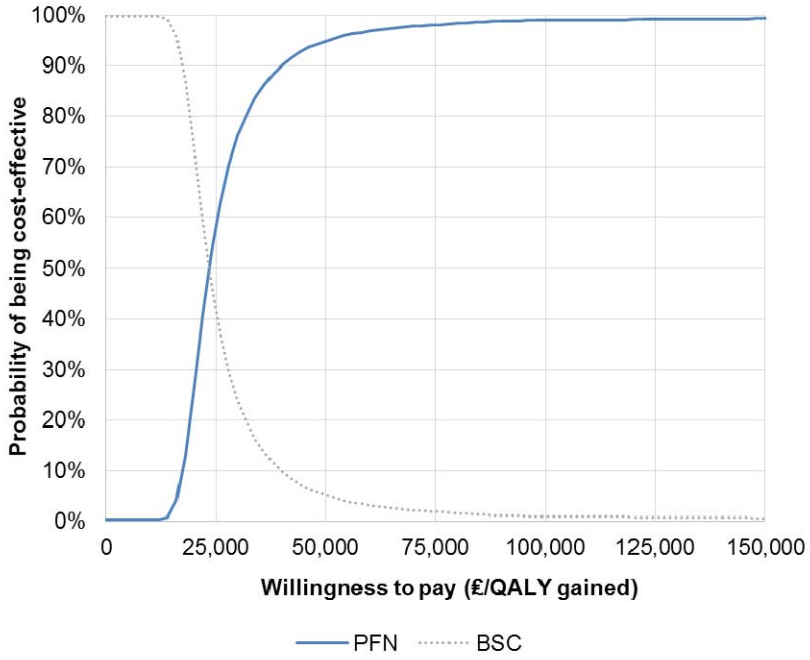
Key: BSC, best supportive care; CI, confidence interval; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Figure 7: PSA scatterplot – Mild population – PAS price



Key: BSC, best supportive care; PFN, pirfenidone; QALY, quality-adjusted life year.

Figure 8: CEAC – Mild population – PAS price



Key: BSC, best supportive care; PFN, pirfenidone; QALY, quality-adjusted life year.

Table 15: PSA results – Moderate population – with the PAS

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)	Incremental (QALYs)
Deterministic model results								
BSC	24,868	4.80	3.44					
NTB	[redacted]	6.06	4.23	[redacted]	1.26	0.78	[redacted]	[redacted]
PFN	[redacted]	7.67	5.14	[redacted]	1.61	0.91	21,318	[redacted]
Mean probabilistic model results								
BSC	24,651	4.77	3.40					
95% CI	(17,757; 32,580)	(2.94; 7.02)	(2.22; 4.72)					
NTB	[redacted]	6.00	4.13	[redacted]	1.23	0.73	[redacted]	[redacted]
95% CI	[redacted]	(2.72; 10.36)	(2.11; 6.55)					
PFN	[redacted]	7.69	5.14	[redacted]	1.69	1.01	20,863	[redacted]
95% CI	[redacted]	(6.51; 9.01)	(4.50; 5.80)					

% chance of being cost effective	£20,000 per QALY	£25,000 per QALY	£30,000 per QALY
BSC	29%	17%	10%
NTB	56%	49%	43%
PFN	16%	34%	47%

Key: BSC, best supportive care; CI, confidence interval; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Figure 9: PSA scatterplot – Moderate population – PAS price

[redacted]

Figure 10: CEAC – Moderate population – PAS price

[redacted]

4.11 *Please present scenario analysis results as described for the main manufacturer/sponsor submission of evidence for the technology appraisal.*

Table 16: Scenario analysis – ITT population – PAS price

	Category	Base case setting	Model change	PFN			BSC			ICER vs. BSC (£)
				Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Base Case				66,638	8.67	5.67	26,627	5.38	3.80	21,387
1	Time horizon	Lifetime (34 years)	10 years	60,683	6.67	4.72	24,946	4.97	3.59	31,540
2			15 years	64,678	7.86	5.34	26,340	5.30	3.76	24,300
3			20 years	66,025	8.38	5.56	26,587	5.37	3.79	22,244
4			25 years	66,467	8.58	5.64	26,622	5.38	3.80	21,619
5			30 years	66,603	8.65	5.66	26,627	5.38	3.80	21,433
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	66,638	8.67	5.40	26,627	5.38	3.62	22,481
7			Nintedanib NICE company submission	66,638	8.67	5.33	26,627	5.38	3.57	22,676
8			Starkie et al. (2012) mapping algorithm	66,638	8.67	5.36	26,627	5.38	3.59	22,580
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	66,638	8.67	5.67	28,653	6.22	4.19	25,776
10			Treatment effect applied for up to 10 years	66,638	8.67	5.67	27,420	5.74	3.95	22,865
11			Treatment effect applied for up to 14 years	66,638	8.67	5.67	26,823	5.48	3.84	21,731
12	OS	Weibull	Exponential	72,388	11.91	6.99	29,338	6.62	4.36	16,340
13			Log-normal	73,714	12.71	7.24	29,979	6.89	4.45	15,690
14			Gamma	68,127	9.49	6.00	27,141	5.59	3.89	19,508
15			Log-Logistic	70,895	11.19	6.63	28,242	6.12	4.11	16,965

16			Gompertz	64,362	7.60	5.20	25,996	5.16	3.69	25,360
17		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	66,638	8.67	5.67	26,899	5.45	3.84	21,797
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	66,638	8.67	5.67	27,287	5.75	3.94	22,839
19			Edinburgh registry used for BSC OS	66,638	8.67	5.67	25,495	5.03	3.58	19,698
20	PFS	Weibull	Exponential	66,638	8.67	5.72	26,558	5.38	3.82	21,085
21			Log-normal	66,638	8.67	5.72	26,572	5.38	3.82	21,031
22			Gamma	66,638	8.67	5.75	26,547	5.38	3.82	20,869
23			Log-Logistic	66,638	8.67	5.72	26,581	5.38	3.81	21,020
24			Gompertz	66,638	8.67	5.65	26,645	5.38	3.79	21,473
25		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	66,638	8.67	5.67	26,631	5.38	3.80	21,369
26	TTD	Weibull	Exponential	66,672	8.67	5.67	26,627	5.38	3.80	21,405
27			Log-normal	68,096	8.67	5.67	26,627	5.38	3.80	22,162
28			Gamma	66,728	8.67	5.67	26,627	5.38	3.80	21,435
29			Log-Logistic	68,145	8.67	5.67	26,627	5.38	3.80	22,188

30			Gompertz	66,461	8.67	5.67	26,627	5.38	3.80	21,293
31	Stopping rule	Not applied	Applied for pirfenidone patients	54,360	8.67	5.66	26,627	5.38	3.80	14,847
32										
33	OS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	66,638	8.67	5.67	26,392	5.31	3.75	21,001
34			Phase III trials only, fixed effects, 52 weeks cut off	66,638	8.67	5.67	26,627	5.38	3.80	21,387
35			Phase II and III trials, fixed effects, 52 weeks cut off	66,638	8.67	5.67	26,627	5.38	3.80	21,387
36			Phase III trials only, random effects, 72 weeks cut off	66,638	8.67	5.67	28,872	6.12	4.23	26,309
37			Phase II and III trials, random effects, 72 weeks cut off	66,638	8.67	5.67	28,872	6.12	4.23	26,309
38			Phase III trials only, fixed effects, 72 weeks cut off	66,638	8.67	5.67	28,872	6.12	4.23	26,309
39			Phase II and III trials, fixed effects, 72 weeks cut off	66,638	8.67	5.67	28,872	6.12	4.23	26,309
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	66,638	8.67	5.67	26,627	5.38	3.80	21,387
41			Phase III trials only, fixed effects, 52 weeks cut off	66,638	8.67	5.67	26,627	5.38	3.80	21,387
42			Phase II and III trials, fixed effects, 52 weeks cut off	66,638	8.67	5.67	26,627	5.38	3.80	21,387

43			Phase III trials only, random effects, 72 weeks cut off	66,638	8.67	5.67	26,627	5.38	3.80	21,387
44			Phase II and III trials, random effects, 72 weeks cut off	66,638	8.67	5.67	26,627	5.38	3.80	21,387
45			Phase III trials only, fixed effects, 72 weeks cut off	66,638	8.67	5.67	26,627	5.38	3.80	21,387
46			Phase II and III trials, fixed effects, 72 weeks cut off	66,638	8.67	5.67	26,627	5.38	3.80	21,387
47	AEs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	66,638	8.67	5.67	26,627	5.38	3.80	21,387
48			Phase III trials only, fixed effects	66,638	8.67	5.67	26,627	5.38	3.80	21,387
49			Phase II and III trials, fixed effects	66,638	8.67	5.67	26,627	5.38	3.80	21,387
50			Phase II and III trials, random effects with adjustments in data for differences in end point	66,638	8.67	5.67	26,627	5.38	3.80	21,387
51			Phase III trials only, fixed effects with adjustments in data for differences in end point	66,638	8.67	5.67	26,627	5.38	3.80	21,387
52	Lung transplant	Not applied	Applied	66,741	8.63	5.67	26,812	5.40	3.81	21,430
53	Gas Transfer	Test administered every 4 months	Test administered every 6 months	65,196	8.67	5.67	25,662	5.38	3.80	21,132
54	Full pulmonary	Test administered	Test administered every 6 months	65,454	8.67	5.67	25,835	5.38	3.80	21,177

		every 4 months								
55	Lung Volume Transfer	No subsequent tests administered	Test administered every 4 months	70,288	8.67	5.67	29,071	5.38	3.80	22,032
56	Field exercise test	Test administered every 6 months	Test administered every 12 months	65,374	8.67	5.67	25,781	5.38	3.80	21,163
57			Test administered every 3 months if on oxygen	67,898	8.67	5.67	27,861	5.38	3.80	21,401
58	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	65,626	8.67	5.67	26,114	5.38	3.80	21,120

Key: BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use; NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

* Result driven by low number of patients with mild disease in the INOVA patient registry: numbers at risk: 53 at baseline; 24 at 50 months; 16 at 100 months; 2 at 150 months.

Table 17: Scenario analysis – Mild population – PAS price

	Category	Base case setting	Model change	PFN			BSC			ICER vs. BSC (£)
				Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Base Case				84,209	11.26	6.99	31,729	7.11	4.82	24,187
1	Time horizon	Lifetime (34 years)	10 years	72,892	7.53	5.30	27,827	6.03	4.30	44,812
2			15 years	79,802	9.38	6.26	30,691	6.78	4.69	31,295
3			20 years	82,519	10.39	6.70	31,490	7.02	4.79	26,762
4			25 years	83,622	10.91	6.89	31,681	7.09	4.82	25,057
5			30 years	84,062	11.16	6.97	31,722	7.10	4.82	24,401
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	84,209	11.26	6.66	31,729	7.11	4.60	25,412
7			Nintedanib NICE company submission	84,209	11.26	6.58	31,729	7.11	4.53	25,662
8			Starkie et al. (2012) mapping algorithm	84,209	11.26	6.61	31,729	7.11	4.56	25,544
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	84,209	11.26	6.99	35,288	8.70	5.53	33,316
10			Treatment effect applied for up to 10 years	84,209	11.26	6.99	33,499	7.96	5.17	27,864
11			Treatment effect applied for up to 14 years	84,209	11.26	6.99	32,345	7.44	4.95	25,338
12	OS	Weibull	Exponential	90,940	15.16	8.47	35,562	9.02	5.61	19,323
13			Log-normal	90,717	15.11	8.40	35,245	8.87	5.52	19,267
14			Gamma	85,402	11.93	7.24	32,219	7.33	4.92	22,879
15			Log-Logistic	88,187	13.64	7.85	33,551	8.01	5.19	20,525

16			Gompertz	79,543	9.09	6.09	30,124	6.46	4.52	31,379
17		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	84,209	11.26	6.99	31,531	7.05	4.79	23,892
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	84,209	11.26	6.99	36,672	9.27	5.81	40,161
19			Edinburgh registry used for BSC OS	84,209	11.26	6.99	29,025	6.18	4.30	20,470
20	PFS	Weibull	Exponential	84,209	11.26	7.07	31,623	7.11	4.86	23,812
21			Log-normal	84,209	11.26	7.07	31,641	7.11	4.85	23,720
22			Gamma	84,209	11.26	7.09	31,637	7.11	4.86	23,579
23			Log-Logistic	84,209	11.26	7.06	31,663	7.11	4.85	23,752
24			Gompertz	84,209	11.26	6.95	31,816	7.11	4.79	24,337
25		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	84,209	11.26	6.99	31,756	7.11	4.81	24,069
26	TTD	Weibull	Exponential	84,624	11.26	6.99	31,729	7.11	4.82	24,376
27			Log-normal	87,753	11.26	6.99	31,729	7.11	4.82	25,810
28			Gamma	84,583	11.26	6.99	31,729	7.11	4.82	24,358
29			Log-Logistic	87,210	11.26	6.99	31,729	7.11	4.82	25,561

30			Gompertz	83,604	11.26	6.99	31,729	7.11	4.82	23,909
31	Stopping rule	Not applied	Applied for pirfenidone patients	65,740	11.26	6.99	31,729	7.11	4.82	15,707
32										
33	OS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	84,209	11.26	6.99	31,449	7.01	4.77	23,720
34			Phase III trials only, fixed effects, 52 weeks cut off	84,209	11.26	6.99	31,729	7.11	4.82	24,187
35			Phase II and III trials, fixed effects, 52 weeks cut off	84,209	11.26	6.99	31,729	7.11	4.82	24,187
36			Phase III trials only, random effects, 72 weeks cut off	84,209	11.26	6.99	34,377	8.06	5.34	30,133
37			Phase II and III trials, random effects, 72 weeks cut off	84,209	11.26	6.99	34,377	8.06	5.34	30,133
38			Phase III trials only, fixed effects, 72 weeks cut off	84,209	11.26	6.99	34,377	8.06	5.34	30,133
39			Phase II and III trials, fixed effects, 72 weeks cut off	84,209	11.26	6.99	34,377	8.06	5.34	30,133
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	84,209	11.26	6.99	31,729	7.11	4.82	24,187
41			Phase III trials only, fixed effects, 52 weeks cut off	84,209	11.26	6.99	31,729	7.11	4.82	24,187
42			Phase II and III trials, fixed effects, 52 weeks cut off	84,209	11.26	6.99	31,729	7.11	4.82	24,187

43			Phase III trials only, random effects, 72 weeks cut off	84,209	11.26	6.99	31,729	7.11	4.82	24,187
44			Phase II and III trials, random effects, 72 weeks cut off	84,209	11.26	6.99	31,729	7.11	4.82	24,187
45			Phase III trials only, fixed effects, 72 weeks cut off	84,209	11.26	6.99	31,729	7.11	4.82	24,187
46			Phase II and III trials, fixed effects, 72 weeks cut off	84,209	11.26	6.99	31,729	7.11	4.82	24,187
47	AEs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	84,209	11.26	6.99	31,729	7.11	4.82	24,187
48			Phase III trials only, fixed effects	84,209	11.26	6.99	31,729	7.11	4.82	24,187
49			Phase II and III trials, fixed effects	84,209	11.26	6.99	31,729	7.11	4.82	24,187
50			Phase II and III trials, random effects with adjustments in data for differences in end point	84,209	11.26	6.99	31,729	7.11	4.82	24,187
51			Phase III trials only, fixed effects with adjustments in data for differences in end point	84,209	11.26	6.99	31,729	7.11	4.82	24,187
52	Lung transplant	Not applied	Applied	84,233	11.20	6.99	31,904	7.12	4.83	24,250
53	Gas Transfer	Test administered every 4 months	Test administered every 6 months	82,431	11.26	6.99	30,502	7.11	4.82	23,932
54	Full pulmonary	Test administered	Test administered every 6 months	82,750	11.26	6.99	30,722	7.11	4.82	23,978

		every 4 months								
55	Lung Volume Transfer	No subsequent tests administered	Test administered every 4 months	88,712	11.26	6.99	34,835	7.11	4.82	24,830
56	Field exercise test	Test administered every 6 months	Test administered every 12 months	82,650	11.26	6.99	30,654	7.11	4.82	23,964
57			Test administered every 3 months if on oxygen	85,763	11.26	6.99	33,299	7.11	4.82	24,179
58	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	82,961	11.26	6.99	31,078	7.11	4.82	23,912

Key: BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use; NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

* Result driven by low number of patients with mild disease in the INOVA patient registry: numbers at risk: 53 at baseline; 24 at 50 months; 16 at 100 months; 2 at 150 months.

Table 18: Scenario analysis – Moderate population – PAS price

	Category	Base case setting	Model change	PFN			BSC			ICER vs. BSC (£)	NTB			ICER vs. NTB (£)
				Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs		Costs (£)	LYs	QALYs	
Base Case				[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
1	Time horizon	Lifetime (34 years)	10 years	[redacted]	6.30	4.47	23,836	4.57	3.32	[redacted]	31,702	5.44	3.92	[redacted]
2			15 years	[redacted]	7.21	4.95	24,742	4.77	3.43	[redacted]	33,423	5.92	4.17	[redacted]
3			20 years	[redacted]	7.53	5.09	24,856	4.80	3.44	[redacted]	33,788	6.04	4.22	[redacted]
4			25 years	[redacted]	7.63	5.13	24,867	4.80	3.44	[redacted]	33,852	6.06	4.22	[redacted]
5			30 years	[redacted]	7.66	5.14	24,868	4.80	3.44	[redacted]	33,862	6.06	4.23	[redacted]
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	[redacted]	7.67	4.89	24,868	4.80	3.28	[redacted]	33,863	6.06	4.03	[redacted]
7			Nintedanib NICE company submission	[redacted]	7.67	4.83	24,868	4.80	3.23	[redacted]	33,863	6.06	3.97	[redacted]
8			Starkie et al. (2012) mapping algorithm	[redacted]	7.67	4.86	24,868	4.80	3.25	[redacted]	33,863	6.06	4.00	[redacted]
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	[redacted]	7.67	5.14	26,324	5.38	3.73	[redacted]	34,931	6.55	4.46	[redacted]
10			Treatment effect applied for up to 10 years	[redacted]	7.67	5.14	25,349	5.01	3.54	[redacted]	34,324	6.29	4.33	[redacted]
11			Treatment effect applied for up to 14 years	[redacted]	7.67	5.14	24,959	4.85	3.46	[redacted]	33,994	6.14	4.25	[redacted]

12	OS	Weibull	Exponential	[redacted]	10.83	6.48	27,374	5.92	3.97	[redacted]	37,720	8.09	5.14	[redacted]
13			Log-normal	[redacted]	11.68	6.76	28,052	6.18	4.06	[redacted]	38,719	8.62	5.31	[redacted]
14			Gamma	[redacted]	11.93	7.20	32,309	7.33	4.89	[redacted]	42,104	9.38	5.97	[redacted]
15			Log-Logistic	[redacted]	10.14	6.11	26,337	5.44	3.73	[redacted]	36,473	7.45	4.81	[redacted]
16			Gompertz	[redacted]	6.87	4.78	24,430	4.67	3.37	[redacted]	32,968	5.70	4.05	[redacted]
17		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	7.67	5.14	25,215	4.89	3.50	[redacted]	33,863	6.06	4.23	[redacted]
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	[redacted]	7.67	5.14	26,064	5.35	3.70	[redacted]	33,863	6.06	4.23	[redacted]
19			Edinburgh registry used for BSC OS	[redacted]	7.67	5.14	20,493	3.48	2.60	[redacted]	33,863	6.06	4.23	[redacted]
20	PFS	Weibull	Exponential	[redacted]	7.67	5.19	24,806	4.80	3.46	[redacted]	34,232	6.06	4.26	[redacted]
21			Log-normal	[redacted]	7.67	5.19	24,820	4.80	3.46	[redacted]	34,141	6.06	4.26	[redacted]
22			Gamma	[redacted]	7.67	5.13	24,865	4.80	3.44	[redacted]	33,889	6.06	4.22	[redacted]
23			Log-Logistic	[redacted]	7.67	5.18	24,827	4.80	3.46	[redacted]	34,087	6.06	4.26	[redacted]
24			Gompertz	[redacted]	7.67	5.13	24,865	4.80	3.44	[redacted]	33,889	6.06	4.22	[redacted]
25		BSC trial data not applied for the first 52 weeks of the	BSC trial data applied for the first 52 weeks of the model.	[redacted]	7.67	5.14	24,865	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]

		model.												
26	TTD	Weibull	Exponential	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,570	6.06	4.23	[redacted]
27			Log-normal	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,552	6.06	4.23	[redacted]
28			Gamma	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,835	6.06	4.23	[redacted]
29			Log-Logistic	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,776	6.06	4.23	[redacted]
30			Gompertz	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,773	6.06	4.23	[redacted]
31	Stopping rule	Applied for nintedanib patients only	Not applied	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	37,272	6.06	4.23	[redacted]
32			Applied for pirfenidone and nintedanib patients	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
33	OS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	7.67	5.14	24,655	4.74	3.40	[redacted]	34,199	6.18	4.30	[redacted]
34			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	34,199	6.18	4.30	[redacted]
35			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	34,031	6.12	4.26	[redacted]
36			Phase III trials only, random effects, 72 weeks cut off	[redacted]	7.67	5.14	26,896	5.45	3.84	[redacted]	36,416	7.00	4.76	[redacted]
37			Phase II and III trials, random effects, 72	[redacted]	7.67	5.14	26,896	5.45	3.84	[redacted]	36,112	6.89	4.70	[redacted]

			weeks cut off											
38			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	7.67	5.14	26,896	5.45	3.84	[redacted]	36,416	7.00	4.76	[redacted]
39			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	7.67	5.14	26,896	5.45	3.84	[redacted]	36,265	6.94	4.73	[redacted]
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
41			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
42			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
43			Phase III trials only, random effects, 72 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
44			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
45			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]

46			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
47	TTD - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,884	6.06	4.23	[redacted]
48			Phase III trials only, fixed effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,906	6.06	4.23	[redacted]
49			Phase II and III trials, fixed effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,884	6.06	4.23	[redacted]
50	AEs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
51			Phase III trials only, fixed effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
52			Phase II and III trials, fixed effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
53			Phase II and III trials, random effects with adjustments in data for differences in end point	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
54			Phase III trials only, fixed effects with adjustments in data for differences in end point	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	33,863	6.06	4.23	[redacted]
55	Lung	Not applied	Applied	[redacted]	7.64	5.15	25,053	4.82	3.46	[redacted]	34,030	6.08	4.24	[redacted]

	transplant													
56	Gas Transfer	Test administered every 4 months	Test administered every 6 months	[redacted]	7.67	5.14	23,993	4.80	3.44	[redacted]	32,791	6.06	4.23	[redacted]
57	Full pulmonary	Test administered every 4 months	Test administered every 6 months	[redacted]	7.67	5.14	24,150	4.80	3.44	[redacted]	32,983	6.06	4.23	[redacted]
58	Lung Volume Transfer	No subsequent tests administered	Test administered every 4 months	[redacted]	7.67	5.14	27,082	4.80	3.44	[redacted]	36,577	6.06	4.23	[redacted]
59	Field exercise test	Test administered every 6 months	Test administered every 12 months	[redacted]	7.67	5.14	24,101	4.80	3.44	[redacted]	32,923	6.06	4.23	[redacted]
60			Test administered every 3 months if on oxygen	[redacted]	7.67	5.14	25,983	4.80	3.44	[redacted]	35,231	6.06	4.23	[redacted]
61	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	[redacted]	7.67	5.14	24,401	4.80	3.44	[redacted]	33,292	6.06	4.23	[redacted]

Key: BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use;

NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

4.12 *If any of the criteria on which the patient access scheme depends are clinical variable (for example, choice of response measure, level of response, duration of treatment), sensitivity analyses around the individual criteria should be provided, so that the Appraisal Committee can determine which criteria are the most appropriate to use.*

Not applicable

Impact of patient access scheme on ICERs

4.13 *For financially based schemes, please present the results showing the impact of the patient access scheme on the ICERs for the base-case and any scenario analyses. A suggested format is shown below (see table 5). If you are submitting the patient access scheme at the end of the appraisal process, you must include the scenario with the assumptions that the Appraisal Committee considered to be most plausible.*

Table 19: Scenario analysis – ITT population

	Category	Base case setting	Model change	ICER vs. BSC (without PAS)	ICER vs. BSC (with PAS)
Base Case				[redacted]	21,387
1	Time horizon	Lifetime (34 years)	10 years	[redacted]	31,540
2			15 years	[redacted]	24,300
3			20 years	[redacted]	22,244
4			25 years	[redacted]	21,619
5			30 years	[redacted]	21,433
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	[redacted]	22,481
7			Nintedanib NICE company submission	[redacted]	22,676
8			Starkie et al. (2012) mapping algorithm	[redacted]	22,580
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	[redacted]	25,776
10			Treatment effect applied for up to 10 years	[redacted]	22,865
11			Treatment effect applied for up to 14 years	[redacted]	21,731
12	OS	Weibull	Exponential	[redacted]	16,340
13			Log-normal	[redacted]	15,690
14			Gamma	[redacted]	19,508
15			Log-Logistic	[redacted]	16,965
16			Gompertz	[redacted]	25,360

17		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	21,797
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	[redacted]	22,839
19			Edinburgh registry used for BSC OS	[redacted]	19,698
20	PFS	Weibull	Exponential	[redacted]	21,085
21			Log-normal	[redacted]	21,031
22			Gamma	[redacted]	20,869
23			Log-Logistic	[redacted]	21,020
24			Gompertz	[redacted]	21,473
25		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	21,369
26	TTD	Weibull	Exponential	[redacted]	21,405
27			Log-normal	[redacted]	22,162
28			Gamma	[redacted]	21,435
29			Log-Logistic	[redacted]	22,188
30			Gompertz	[redacted]	21,293
31	Stopping rule	Not applied	Applied for pirfenidone patients	[redacted]	14,847
32					
33	OS - NMA	Phase II and III trials,	Phase III trials only, random effects, 52 weeks cut off	[redacted]	21,001

34		random effects, 52 weeks cut off	Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	21,387
35			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	21,387
36			Phase III trials only, random effects, 72 weeks cut off	[redacted]	26,309
37			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	26,309
38			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	26,309
39			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	26,309
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	21,387
41			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	21,387
42			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	21,387
43			Phase III trials only, random effects, 72 weeks cut off	[redacted]	21,387
44			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	21,387
45			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	21,387
46			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	21,387
47	AEs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	21,387
48			Phase III trials only, fixed effects	[redacted]	21,387
49			Phase II and III trials, fixed effects	[redacted]	21,387
50			Phase II and III trials, random effects with adjustments in data for differences in end point	[redacted]	21,387
51			Phase III trials only, fixed effects with adjustments in data for differences in end point	[redacted]	21,387
52	Lung	Not applied	Applied	[redacted]	21,430

	transplant				
53	Gas Transfer	Test administered every 4 months	Test administered every 6 months	[redacted]	21,132
54	Full pulmonary	Test administered every 4 months	Test administered every 6 months	[redacted]	21,177
55	Lung Volume Transfer	No subsequent tests administered	Test administered every 4 months	[redacted]	22,032
56	Field exercise test	Test administered every 6 months	Test administered every 12 months	[redacted]	21,163
57			Test administered every 3 months if on oxygen	[redacted]	21,401
58	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	[redacted]	21,120

Key: BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use; NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

* Result driven by low number of patients with mild disease in the INOVA patient registry: numbers at risk: 53 at baseline; 24 at 50 months; 16 at 100 months; 2 at 150 months.

Table 20: Scenario analysis – Mild population

	Category	Base case setting	Model change	ICER vs. BSC (without PAS)	ICER vs. BSC (with PAS)
Base Case				[redacted]	24,187
1	Time horizon	Lifetime (34 years)	10 years	[redacted]	44,812
2			15 years	[redacted]	31,295
3			20 years	[redacted]	26,762
4			25 years	[redacted]	25,057
5			30 years	[redacted]	24,401
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	[redacted]	25,412
7			Nintedanib NICE company submission	[redacted]	25,662
8			Starkie et al. (2012) mapping algorithm	[redacted]	25,544
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	[redacted]	33,316
10			Treatment effect applied for up to 10 years	[redacted]	27,864
11			Treatment effect applied for up to 14 years	[redacted]	25,338
12	OS	Weibull	Exponential	[redacted]	19,323
13			Log-normal	[redacted]	19,267
14			Gamma	[redacted]	22,879
15			Log-Logistic	[redacted]	20,525
16			Gompertz	[redacted]	31,379

17		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	23,892
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	[redacted]	40,161
19			Edinburgh registry used for BSC OS	[redacted]	20,470
20	PFS	Weibull	Exponential	[redacted]	23,812
21			Log-normal	[redacted]	23,720
22			Gamma	[redacted]	23,579
23			Log-Logistic	[redacted]	23,752
24			Gompertz	[redacted]	24,337
25		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	24,069
26	TTD	Weibull	Exponential	[redacted]	24,376
27			Log-normal	[redacted]	25,810
28			Gamma	[redacted]	24,358
29			Log-Logistic	[redacted]	25,561
30			Gompertz	[redacted]	23,909
31	Stopping rule	Not applied	Applied for pirfenidone patients	[redacted]	15,707
32					
33	OS - NMA	Phase II and III trials,	Phase III trials only, random effects, 52 weeks cut off	[redacted]	23,720

34		random effects, 52 weeks cut off	Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	24,187
35			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	24,187
36			Phase III trials only, random effects, 72 weeks cut off	[redacted]	30,133
37			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	30,133
38			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	30,133
39			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	30,133
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	24,187
41			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	24,187
42			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	24,187
43			Phase III trials only, random effects, 72 weeks cut off	[redacted]	24,187
44			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	24,187
45			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	24,187
46			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	24,187
47	AEs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	24,187
48			Phase III trials only, fixed effects	[redacted]	24,187
49			Phase II and III trials, fixed effects	[redacted]	24,187
50			Phase II and III trials, random effects with adjustments in data for differences in end point	[redacted]	24,187
51			Phase III trials only, fixed effects with adjustments in data for differences in end point	[redacted]	24,187
52	Lung	Not applied	Applied	[redacted]	24,250

	transplant				
53	Gas Transfer	Test administered every 4 months	Test administered every 6 months	[redacted]	23,932
54	Full pulmonary	Test administered every 4 months	Test administered every 6 months	[redacted]	23,978
55	Lung Volume Transfer	No subsequent tests administered	Test administered every 4 months	[redacted]	24,830
56	Field exercise test	Test administered every 6 months	Test administered every 12 months	[redacted]	23,964
57			Test administered every 3 months if on oxygen	[redacted]	24,179
58	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	[redacted]	23,912

Key: BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use; NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

* Result driven by low number of patients with mild disease in the INOVA patient registry: numbers at risk: 53 at baseline; 24 at 50 months; 16 at 100 months; 2 at 150 months.

Table 21: Scenario analysis – Moderate population

	Category	Base case setting	Model change	ICER vs. BSC (without PAS)	ICER vs. BSC (with PAS)	ICER vs. NTB (without PAS)	ICER vs. NTB (with PAS)
Base Case				[redacted]	[redacted]	[redacted]	[redacted]
1	Time horizon	Lifetime (34 years)	10 years	[redacted]	[redacted]	[redacted]	[redacted]
2			15 years	[redacted]	[redacted]	[redacted]	[redacted]
3			20 years	[redacted]	[redacted]	[redacted]	[redacted]
4			25 years	[redacted]	[redacted]	[redacted]	[redacted]
5			30 years	[redacted]	[redacted]	[redacted]	[redacted]
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	[redacted]	[redacted]	[redacted]	[redacted]
7			Nintedanib NICE company submission	[redacted]	[redacted]	[redacted]	[redacted]
8			Starkie et al. (2012) mapping algorithm	[redacted]	[redacted]	[redacted]	[redacted]
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	[redacted]	[redacted]	[redacted]	[redacted]
10			Treatment effect applied for up to 10 years	[redacted]	[redacted]	[redacted]	[redacted]
11			Treatment effect applied for up to 14 years	[redacted]	[redacted]	[redacted]	[redacted]
12	OS	Weibull	Exponential	[redacted]	[redacted]	[redacted]	[redacted]
13			Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
14			Gamma	[redacted]	[redacted]	[redacted]	[redacted]
15			Log-Logistic	[redacted]	[redacted]	[redacted]	[redacted]

16			Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
17		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	[redacted]	[redacted]	[redacted]
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	[redacted]	[redacted]	[redacted]	[redacted]
19			Edinburgh registry used for BSC OS	[redacted]	[redacted]	[redacted]	[redacted]
20	PFS	Weibull	Exponential	[redacted]	[redacted]	[redacted]	[redacted]
21			Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
22			Gamma	[redacted]	[redacted]	[redacted]	[redacted]
23			Log-Logistic	[redacted]	[redacted]	[redacted]	[redacted]
24			Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
25		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	[redacted]	[redacted]	[redacted]
26	TTD	Weibull	Exponential	[redacted]	[redacted]	[redacted]	[redacted]
27			Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
28			Gamma	[redacted]	[redacted]	[redacted]	[redacted]
29			Log-Logistic	[redacted]	[redacted]	[redacted]	[redacted]
30			Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
31	Stopping rule	Applied for nintedanib patients	Not applied	[redacted]	[redacted]	[redacted]	[redacted]
32			Applied for pirfenidone and nintedanib patients	[redacted]	[redacted]	[redacted]	[redacted]

33	OS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
34			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
35			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
36			Phase III trials only, random effects, 72 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
37			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
38			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
39			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
41			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
42			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
43			Phase III trials only, random effects, 72 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
44			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
45			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
46			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	[redacted]	[redacted]	[redacted]
47	TTD - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	[redacted]	[redacted]	[redacted]
48			Phase III trials only, fixed effects	[redacted]	[redacted]	[redacted]	[redacted]
49			Phase II and III trials, fixed effects	[redacted]	[redacted]	[redacted]	[redacted]
50	AEs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	[redacted]	[redacted]	[redacted]
51			Phase III trials only, fixed effects	[redacted]	[redacted]	[redacted]	[redacted]
52			Phase II and III trials, fixed effects	[redacted]	[redacted]	[redacted]	[redacted]

53			Phase II and III trials, random effects with adjustments in data for differences in end point	[redacted]	[redacted]	[redacted]	[redacted]
54			Phase III trials only, fixed effects with adjustments in data for differences in end point	[redacted]	[redacted]	[redacted]	[redacted]
55	Lung transplant	Not applied	Applied	[redacted]	[redacted]	[redacted]	[redacted]
56	Gas Transfer	Test administered every 4 months	Test administered every 6 months	[redacted]	[redacted]	[redacted]	[redacted]
57	Full pulmonary	Test administered every 4 months	Test administered every 6 months	[redacted]	[redacted]	[redacted]	[redacted]
58	Lung Volume Transfer	No subsequent tests administered	Test administered every 4 months	[redacted]	[redacted]	[redacted]	[redacted]
59	Field exercise test	Test administered every 6 months	Test administered every 12 months	[redacted]	[redacted]	[redacted]	[redacted]
60			Test administered every 3 months if on oxygen	[redacted]	[redacted]	[redacted]	[redacted]
61	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	[redacted]	[redacted]	[redacted]	[redacted]

Key: BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use; NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

* Result driven by low number of patients with mild disease in the INOVA patient registry: numbers at risk: 53 at baseline; 24 at 50 months; 16 at 100 months; 2 at 150 months.

5 Appendices

5.1 Appendix A: Additional documents

5.1.1 If available, please include copies of patient access scheme agreement forms, patient registration forms, pharmacy claim forms/rebate forms, guides for pharmacists and physicians, patient information documents.

Not applicable

5.2 Appendix B: Details of outcome-based schemes

5.2.1 If you are submitting a proven value: price increase scheme, as defined in the PPRS, please provide the following information:

- *the current price of the intervention*
- *the proposed higher price of the intervention, which will be supported by the collection of new evidence*
- *a suggested date for when NICE should consider the additional evidence.*

Not applicable

5.2.2 If you are submitting an expected value: rebate scheme, as defined in the PPRS, please provide the following details:

- *the current price of the intervention (the price that will be supported by the collection of new evidence)*
- *the planned lower price of the intervention in the event that the additional evidence does not support the current price*
- *a suggested date for when NICE should consider the additional evidence.*

Not applicable

5.2.3 If you are submitting a risk-sharing scheme, as defined in the PPRS, please provide the following details:

- *the current price of the intervention (the price that will be supported by the collection of new evidence)*
- *the proposed relationship between future price changes and the evidence to be collected.*

Not applicable

5.2.4 *For outcome-based schemes, as defined in the PPRS, please provide the full details of the new information (evidence) planned to be collected, who will collect it and who will carry the cost associated with this planned data collection. Details of the new information (evidence) may include:*

- *design of the new study*
- *patient population of the new study*
- *outcomes of the new study*
- *expected duration of data collection*
- *planned statistical analysis, definition of study groups and reporting (including uncertainty)*
- *expected results of the new study*
- *planned evidence synthesis/pooling of data (if applicable)*
- *expected results of the evidence synthesis/pooling of data (if applicable).*

Not applicable

5.2.5 *If you are submitting a risk-sharing scheme, please specify the period between the time points when the additional evidence will be considered.*

Not applicable

5.2.6 *Please provide the clinical effectiveness data resulting from the evidence synthesis and used in the economic modelling of the patient access scheme at the different time points when the additional evidence is to be considered.*

Not applicable

5.2.7 *Please provide the other data used in the economic modelling of the patient access scheme at the different time points when the additional evidence is to be considered. These data could include cost/resource use, health-related quality of life and utilities.*

Not applicable

5.2.8 *Please present the cost-effectiveness results as follows.*

- *For proven value: price increase schemes, please summarise in separate tables:*
  - *the results based on current evidence and current price*
  - *the anticipated results based on the expected new evidence and the proposed higher price.*
- *For expected value: rebate schemes, please summarise in separate tables:*
  - *the results based on the expected new evidence and the current price (which will be supported by the additional evidence collection)*
  - *the results based on the current evidence and the lower price (if the new evidence is not forthcoming).*
- *For risk-sharing schemes, please summarise in separate tables:*
  - *the results based on current evidence and current price*
  - *the results based on the expected new evidence and the current price (which will be supported by the additional evidence collection)*
  - *the results based on the current evidence and the lower price (if the new evidence is not forthcoming)*
  - *the anticipated results based on the expected new evidence and the proposed higher price.*

A suggested format is shown in table 3, section 4.7.

Not applicable

5.2.9 *Please present in separate tables the incremental results for the different scenarios as described above in section 5.2.8 for the type of outcome-based scheme being submitted.*

List the interventions and comparator(s) from least to most expensive. Present the incremental cost-effectiveness ratios (ICERs) in comparison with baseline (usually standard care), and the incremental analysis ranking technologies in terms of dominance and extended dominance. A suggested format is presented in table 4, section 4.8.

Not applicable

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Dear [redacted],

The Evidence Review Group, the School of Health and Related Research Technology Assessment Group (ScHARR-TAG), and the technical team at NICE have looked at the submission received on 1 February 2016 from Roche. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter). The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **3 March 2016**. Your response and any supporting documents should be uploaded to NICE Docs.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Sophie Laurenson, Technical Lead (sophie.laurenson@nice.org.uk). Any procedural questions should be addressed to Jeremy Powell, Project Manager (jeremy.powell@nice.org.uk).

Yours sincerely

Melinda Goodall
Associate Director – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for confidential information

Section A: Clarification on effectiveness data

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Systematic literature review

A1. Pages 50-53: Were any searches undertaken for any ongoing trials in research registers? For example, metaRegister of Controlled Trials (http://www.isrctn.com/page/mrct), US National Institute of Health Ongoing Trials Register (clinicaltrials.gov) or the World Health Organization (http://www.who.int/ictrp/search/en/) or the EU Clinical Trials Register (https://www.clinicaltrialsregister.eu/).

A2. Page 56: The submission states that the company was aware of the following publications that were not captured by the literature search:

- Noble PW, Albera C, Bradford WZ. Pirfenidone for idiopathic pulmonary fibrosis: analysis of pooled data from three multinational Phase 3 trials. European Respiratory Journal 2015; doi: 10.1183/13993003.00026-2015
- Albera C. Pirfenidone is efficacious in Patients With Idiopathic Pulmonary Fibrosis (IPF) and Mild Restrictive Disease: Benefit of Early Intervention. Oral presentation presented at ATS annual congress 2015; 15-20 May; Denver, USA

Please clarify how these references were identified, given that Noble et al. was published after the company's searches were conducted.

A3. Appendix 3 Section A.1 (pp.15 and 17): A filter for randomised controlled trials was applied to the results of the Medline and Embase searches for clinical effectiveness. This filter has been slightly modified from the Cochrane highly sensitive search strategy published by Lefebvre et al. Please explain the reason for modifying the filter and explain the basis for the changes made. In order to increase the transparency of the reporting, please provide the source of the adapted search filter.

A4. Section 4.1, p.51: Please provide justification for using only a single systematic reviewer to assess and select relevant papers at the title and abstract stage. What are the limitations of this method? For the full-text screening process was there an assessment of inter-rater reliability between two independent reviewers?

A5. Please confirm if data extraction and quality assessment was undertaken independently by a minimum of 2 reviewers for the clinical effectiveness reviews (including the network meta-analysis [NMA] section). If this was not done, please explain why.

A6. Section 4.1, p.53 and Figure 2, p.54: Please provide the missing information from this figure. The reported numbers are: 4394 citations, of which 2407 were excluded

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

(leaving 1987 citations), leading to 116 citations. Please clarify how 1987 citations were reduced to 116.

A7. Section 4.6, p.86-88, Table 17 and Appendix 6: Please specify the critical appraisal tool used. Please explain why the critical appraisals reported in Appendix 6 were conducted using the Cochrane risk of bias tool, but a different tool was used and its findings reported in Section 4.6.

Clinical trial design

A8. **Priority question:** Please clarify whether trial patients receiving pirfenidone stopped treatment if their percent predicted forced vital capacity (FVC) decreased to below 50% (ASCEND, CAPACITY 1 & 2 and RECAP). Please comment on whether the discontinuation in trial is reflective of use of pirfenidone in clinical practice in England.

A9. **Priority question**: Please explain why carbon monoxide diffusing capacity of the lungs (DLco) was excluded from the outcomes measured in ASCEND when it had been previously collected in CAPACITY 1 & 2, SP2 and SP3.

A10. Section 4.5, p.82: Please clarify the statement: "the data from the three studies [ASCEND, CAPACITY 1 & 2] can be regarded as inclusive of the overall range of patients likely to present with mild to moderate IPF in routine clinical practice". This statement appears to contradict the following statement (p.182): "As only patients with mild to moderate IPF, and relatively few comorbidities, were enrolled in ASCEND, CAPACITY 1 & 2 the results need to be carefully interpreted for the broader population of patients".

A11. Section 4.3, pp.60-1: Please clarify the primary endpoint of the ASCEND and CAPACITY trials, in more specific terms. The company submission defines the primary endpoint as a change in FVC, but this could be defined more specifically in one of two ways: ≥10% decline in percent predicted FVC (or death), or absolute change in percent predicted FVC.

A12. Please provide original trial protocols (with any dated changes) for ASCEND, CAPACITY 1 & 2, SP2 and SP3.

Definition of clinical trial outcomes

A13. Section 4.7, p.90: Please define and justify the use of the composite outcome: ≥10% FVC or death. For example, does this refer to all-cause mortality?

A14. Section 4.7, p.99, Section 4.10, p.143: Please clarify if progression-free survival (PFS) has been assessed in any other IPF trials. If so, how did the definition of PFS differ from the definitions used in the pirfenidone trials? In particular, please describe the definitions of PFS used in INPULSIS, SP3 and PANTHER.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A15. Section 4.7, p.103: The company states, “For ASCEND, acute exacerbations were identified via a post-hoc analysis of adverse events.” This appears to differ from the approach taken in CAPACITY 1 & 2, SP2 and SP3 (specific definitions of acute IPF exacerbation have been provided for these trials on pp.103-4). Please clarify:

a. how acute IPF exacerbations were defined in ASCEND

b. whether data on acute IPF exacerbations were prospectively recorded in ASCEND

c. why the definition used in CAPACITY 1 & 2 was not used in ASCEND.

A16. Section 4.7, pp.110-112: Please provide further details of the minimal clinically important difference for the St. Georges Respiratory Questionnaire (SGRQ) and the Shortness of Breath Questionnaire (SOBQ).

A17. Section 4.10, p.137: Please confirm if the definition of overall survival (OS) is the same across all the trials in the NMA.

A18. Section 4.12, p.169: Please define the adverse event category “worsening of IPF”. Is this worsening due to the study drug or worsening due to lack of efficacy?

A19. Section 4.12: Please clarify how the adverse events grading of mild, moderate and severe were defined.

A20. Section 4.12, p.170: Please define “treatment emergent” as used in this context. How is this different from adverse events reported for the ASCEND trial in Table 60?

Analysis of clinical trial data

A21. **Priority question:** Section 4.4, p.69 and Section 4.7, p.90, Table 18: Please refer to the NICE guide to the methods of technology appraisal and provide additional detail on the method of pooling data from ASCEND, CAPACITY 1 and CAPACITY 2 (including data synthesis methods and the rationale, software and models used, and any methodological limitations).

A22. **Priority question:** Section 4.4, p.69 and Section 4.7, p.90 and 96: The published trial protocols for ASCEND and CAPACITY 1 & 2 do not mention pooled analyses of the trial data for the composite outcome (≥10% FVC or death) or for OS. Please clarify whether pooling data from these 3 trials for the endpoints listed above was “pre-specified”.

A23. Section 4.4, p.69: The company states, “The primary outcome of ASCEND compared change in percentage predicted FVC between the 2403 mg/day and placebo groups (King 2014).” By contrast, on p.92 it states, “The ASCEND

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

manuscript did not report the change in % predicted FVC, but this was analysed to inform the NMA." Please clarify and provide details of the sources and data in full.

Clinical trial results

A24. **Priority question:** Section 4.5, pp.76-80: Please provide full details of the nature and grade/severity of adverse events that led to discontinuation in each of the trial arms.

A25. **Priority question:** Section 4.12, p.171: Please provide rates of adverse events for pirfenidone relative to placebo, and related p values, for ASCEND and CAPACITY 1 & 2. Please provide data on all severe adverse events in these trials at all available time-points, including those that led to treatment discontinuation.

A26. **Priority question:** Section 1.3, p.22: An imbalance in baseline characteristics in CAPACITY 1 is given as a possible reason for a failure to demonstrate a significant treatment effect on the trial's primary outcome. The company suggests that an example of this imbalance is that "numerically more patients in CAPACITY 1 had been diagnosed with IPF ≥1 year". This example appears to be a comparison between CAPACTIY 1 and another trial rather than a description of a baseline imbalance within CAPACITY 1. Please provide further clarification on how this statement relates to baseline imbalances. If instead it relates to differences between the trial population between CAPACITY 1 and CAPACITY 2, please explain how this would lead to differences in the primary outcomes of these trials and which direction the difference would be expected to go based on current evidence.

A27. Section 4.7, p.104: Please clarify the consistency of the following statements: on p.104 the company states, "In SP2, the incidence of acute exacerbation of IPF was 14% (n=5) in the placebo group and was none in the pirfenidone group during the 9 months (p=0.0031) (Azuma 2005)." However, on p.105 the company states, "Whilst the incidence of hospitalisation was similar, the duration of these hospital stays was consistently numerically longer in the placebo arms. In SP2, five patients in the pirfenidone treatment arm were hospitalised due to exacerbations (Azuma 2005)."

A28. Section 4.12, p.171: Please provide data on serious adverse events for SP3.

Subgroup analysis

A29. **Priority question:** Section 4.8, p.113-117: For each subgroup analyses, please provide detailed results for each outcome (including event rates, hazard ratios, confidence intervals and p values). Please include enough information to support the statement that exploratory findings confirm the robustness and consistency of the findings across the study population. Please state whether these analyses were considered *a priori* or *post hoc*.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A30. **Priority question:** Section 4.8, p.113: The company present a subgroup analysis of the CAPACITY trials using 3 categories of baseline percent predicted FVC: <70%, 70-80% and ≥80%. Please provide a rationale for investigating these 3 subgroups.

A31. **Priority question**: Section 4.8, p.114-5: Please provide additional results from the subgroup analysis stratified by baseline percent predicted FVC (<80% and ≥80%):

a. Provide the results in figure 17 separately for each of the 3 individual trials (ASCEND, CAPACITY 1 and CAPACITY 2).

b. In addition, provide OS and PFS results for both subgroups at both 52 and 72 weeks. Please provide these for the 3 individual trials (ASCEND, CAPACITY 1 and CAPACITY 2) and for the pooled dataset.

Network meta-analysis (NMA)

A32. **Priority question:** Section 4.10 p.125. A network diagram is provided for all studies that contribute to the NMA, however different trials are included in the NMA for each outcome. Please present separate network diagrams for each outcome.

A33. **Priority question:** Section 4.7, p.99, Section 4.10, p.143: For CAPACITY 1 & 2, PFS was reanalysed using the definition used in the ASCEND trial. For other trials (INPULSIS, SP3, PANTHER), it is "assumed that they will lead to similar hazard ratios and odds ratios between a given pair of treatments, and thus that it is appropriate to combine them in an NMA" (p.143). Please validate this assumption by reanalysing the individual patient data from CAPACITY 1 & 2 and ASCEND, using the PFS definition(s) used in INPULSIS, SP3, and PANTHER trials.

A34. **Priority question:** Section 4.10 p.133-155. Please provide the following additional information for the NMA results:

a. Estimates of the between-study heterogeneity for all random effects models.

b. The 95% predictive intervals (PrI) in addition to the credible intervals (CrI) that are currently presented. The NICE Decision Support Unit (in Technical Support Document 2) recommends that the predictive distribution, rather than the distribution of the mean treatment effect, better represents uncertainty about comparative effectiveness in the presence of heterogeneity.

c. Model fit statistics (total residual deviance and deviance information criteria [DIC]) to allow a comparison of the random and fixed-effects analyses.

A35. **Priority question**: Section 4.10. Please re-run the NMA with the changes below.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

a. The evidence network includes one 3 arm trial (PANTHER), however no correction for multi-arm trials was implemented. The ERG considers that although the correlation between arms may be reduced, the assumption of zero correlation is not appropriate. Please implement correction for multi-arm trials for the NMA where required.

b. The NMAs of survival outcomes used a pooled hazard ratio for the INPULSIS trials. The ERG notes that results from the individual trials are available and should be used to inform the NMA. Please repeat the analyses with the trial level estimates of treatment effect.

A36. **Priority question**: Section 4.10, p.125: The individual trials report outcomes at different time points. These are synthesised under the assumption that the treatment effects are constant over time.

a. Please provide evidence to justify this assumption, by considering the effect of including a covariate for trial duration through meta-regression.

b. If time allows, for binary outcomes please consider the use of a complementary log-log (cloglog) link function (as described in the NICE Decision Support Unit Technical Support Document 2) or the piece-wise constant hazard model of Lu et al (2007). These approaches do not rely on the assumption of constant treatment effects.

A37. **Priority question:** Section 4.10, p.126: For the NMA of the survival outcomes, the principal analysis assumed that the "proportional hazards is an acceptable assumption up to 52 weeks, but not beyond". A sensitivity analysis explored the assumption that proportional hazards apply until week 72. Please justify choosing the weaker assumption for the base case analyses. Is there reason to believe that the hazards may not be proportional beyond 52 weeks?

A38. Section 4.10, p.124: Please provide a sensitivity analysis, excluding PANTHER and SP3 from the NMA. The ERG notes that there are questions over the generalisability of the population in the SP3 trial, and over the inclusion of the PANTHER trial because the active treatments are not listed as comparators in the scope.

A39. Section 4.10 p.146: The company states, "To mitigate the differences in definitions, we reanalysed our IPD to match BI's definition, adjusted for different base case by meta-regression, and corrected actual data based on the baseline prevalence of adverse events as an additional sensitivity analysis." Please provide further details on how this analysis was conducted.

Section B: Clarification on cost-effectiveness data

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Non-randomised trial evidence

B1. Section 4.11, p.158: Please provide the actual numbers of people who entered the RECAP study from each of the trials: CAPACITY 1, CAPACITY 2 and ASCEND.

B2. According to the study design summarised in Table 56, patients who did not take >80% of study drug in ASCEND and CAPACITY were excluded from RECAP. Please provide the rationale for excluding these patients from RECAP.

B3. Section 4.12, p.172-174: Please explain and justify the inclusion of the PIPF-002 trial, which includes a population and pirfenidone dosing regimen outside of the scope of this appraisal.

Systematic literature review

B4. Section 5.1, p.190: The combined searches were run without a date limit. However, the submission states that only data published after 2010 was screened. A large number of the sources (including NHS EED and the Cost-Effectiveness Analysis registry) were not searched in the original submission. Please explain how the company ensured key data were not missed by limiting the results by date.

B5. Appendix 17 (pp.195-201): The company does not provide a reference to any published filter that has been used; however, the utilities search filter appears to have been directly derived (with no variation) from Arber et al (2015) (http://www.yhec.co.uk/yhec-content/uploads/2015/06/Poster-374-Sensitivity-Of-A-Search-Filter.pdf). The cost-effectiveness filter appears to be a modified version of the NHS EED search filter (line 56 has been amended and lines 25-27 have been omitted). Please confirm the sources of published search filters and explain why they have been modified.

Model structure and key assumptions

B6. **Priority question:** Given the company's modelling approach (partitioned-survival) OS, PFS and discontinuations are modelled independently of each other. The effect of this is that increasing the discontinuation rate (for example, by applying the stopping rule) only affects costs – it does not impact the effectiveness in terms of PFS and OS. Please comment on whether this is appropriate and provide evidence to support the assumption that OS and PFS are independent of the time on treatment.

B7. **Priority question**: Following pirfenidone discontinuation, the company assumes that patients receive best supportive care even if the patient still has moderate IPF (percent predicted FVC 50–80%). Nintedanib has recently been recommended by NICE for use in people with moderate disease. Please comment on whether

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

nintedanib could be used following pirfenidone, and whether pirfenidone could be used following discontinuation of nintedanib (in patients who still have moderate IPF).

B8. Section 5.3, p.218: The company states, "the stopping rule was applied in the model using tunnel states to estimate the proportion of patients who progress within 12 months", but no further details are provided regarding the tunnel states. Please describe in more detail how the tunnel states operate within the model.

Mortality rates

B9. Section 5.3, p.215: In the base case, the company assumes that 57.89% of deaths are attributed to IPF for patients receiving pirfenidone. Please clarify the data source used to inform this estimate, including information on the trials and time points used.

B10. Section 5.3, p.215 (table 74): Please explain how the proportions of deaths that are IPF-related for best supportive care and nintedanib were calculated using the data from the NMA. The ERG cannot see how the hazard ratios from the NMA for OS and IPF-related survival can be used to estimate the proportion of deaths which are IPF-related.

Survival modelling

B11. **Priority question**: In Table 70 (p.209) it is stated that individual patient data (IPD) from CAPACITY, ASCEND and RECAP were used to fit the OS curve applied in the model. Please clarify which patients from RECAP were included. In particular:

a. Were patients enrolled in RECAP from the placebo arms of CAPACITY included in the IPD?

b. Were patients enrolled in RECAP from the non-licensed dose arm of CAPACITY 2 included in the IPD?

c. Were all patients enrolled in ASCEND censored at 52 weeks in this analysis?

B12. **Priority question**: Justification of the assumption of proportional hazards is provided in Appendix 20, however it is not clear which data were used for the test of interaction.

a. Please confirm which data were used for the test of interaction and for the test for the proportionality assumption for both OS and PFS.

b. Please clarify whether the analysis is based on data from 52 or 72 weeks.

c. If the analyses presented are based on data pooled across multiple trials, please repeat the analysis for each separate trial (using 72 week data, where available).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Calculation of costs

B13. **Priority question**: In Table 92 of the company submission, the mean actual pirfenidone dose received in the ASCEND and CAPACITY trials is reported. The company states that pirfenidone needs to be titrated in the first 2 weeks. Please provide the mean (and standard error) for the actual dose received:

a. estimated for the first 3 months of treatment only (to represent the first cycle of the model)

b. excluding the first 3 months of treatment (to represent subsequent cycles in the model).

B14. End of life costs are included for IPF-related deaths but not for deaths from other causes. It is stated on p.246, "This is because costs associated with IPF are greatly increased in the last year of life due to increased resource use, home care and length of stay in hospital". Please provide evidence which demonstrates that the costs in the last year of life for IPF-related deaths are higher than the costs in the last year of life for deaths from other causes, to support this assumption.

B15. The company uses data on hospitalisations from the trials to inform the costs associated with acute exacerbations (Table 99) but uses estimates from the published literature combined with data from the NMA to inform the disutility associated with acute exacerbation. Please explain and justify why the same data for incidence of exacerbation is not used to inform the costs and disutilities from exacerbations.

B16. Section 5.5, p.239: The company states that resource use was based upon advice on UK clinical practice in IPF from a panel of UK clinicians, stratified by treatment type and progression status. Please provide any materials used to elicit this information from the panel, and the analysis of these data.

Utility estimates

B17. Please explain how the Freemantle (2015) and Starkie (2011) mapping studies, which were used to map from the St George's Respiratory Questionnaire (SGRQ) to EQ-5D, were identified. In particular, was a systematic search conducted to identify all relevant mapping algorithms? How were these 2 studies selected from those identified in the search?

B18. In the base case, the company estimates the mean SGRQ for patients in the progression-free and progression health states using a generalised estimating equation (GEE) model; it estimates the mean EQ-5D score based on a linear mapping algorithm estimated in IPF patients (Freemantle, 2015). However, the

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

mapping algorithm used in the base case (EQ-5D = 1.3246 – 0.01276*SGRQ) predicts an EQ-5D value over 1 when the SGRQ becomes below 25.

a. Please provide the proportion of patients with a SGRQ score below 25 at baseline and last follow-up in the CAPACITY trials.

b. Please provide more information on how the GEE model was estimated, in particular please comment on whether each patient contributed data at multiple time points. If so, how those time points were selected?

c. Finally, please provide the following analyses:

   i. Apply the unconstrained mapping algorithm (EQ-5D = 1.3246 – 0.01276*SGRQ) to the individual patient level data from the CAPACITY trials, then estimate the mean EQ-5D score using a GEE model (similar to the approach used for SGRQ in the original submission).

   ii. Apply the mapping algorithm (EQ-5D = 1.3246 – 0.01276*SGRQ) to the individual patient level data from the CAPACITY trials and truncate the predicted EQ-5D to a maximum of 1, then estimate the mean EQ-5D score using a GEE model (similar approach used for SGRQ in the original submission).

Subgroup analyses

B19. **Priority question**: In addition to the intention to treat population, the company reports results for two subgroups; mild (percent predicted FVC ≥80%) and moderate (percent predicted FVC 50–80%). Whilst the ERG understands these analyses are post-hoc (not pre-specified), could the company provide the following:

a. Kaplan–Meier curves for OS for best supportive care from the pooled ASCEND and CAPACITY trials. Provide these separately for patients with mild and moderate IPF (as defined in the model) at baseline.

b. Evidence that the proportional hazard assumption holds for patients with moderate and mild IPF separately (at 72 weeks) for PFS and OS.

B20. For the subgroup analyses (mild and moderate IPF), the company fit a series of parametric curves to the OS and time to discontinuation for the pirfenidone arms using percent predicted FVC at baseline as covariates (FVC<50%, FVC 50–80% and FVC≥80%). However, for PFS, the company appears to use a different approach, fitting parametric curves to the 2 subgroups separately. Please justify the use of different approaches for OS and PFS?

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Executable model

B21. **Priority question**: In the economic model, in sheets "KM TTOT", "KM PFS" and "KM OS" the number of patients at risk in the Kaplan–Meier plots for pirfenidone appear to be different depending on the outcome (PFS n=618, OS n=623 and TTOT n=490). Could the company:

a. Clarify the source of data used for these Kaplan–Meier plots (ASCEND/CAPACITY/RECAP?)

b. Confirm whether only patients receiving 2,403 mg daily for pirfenidone are included for each of the outcomes.

c. Explain the reasons for the different numbers of patients at risk for these outcomes.

B22. In the economic model, in sheet "KM TTOT", please:

a. Clarify whether the time to discontinuation (column B-I) is for the pooled intention to treat population from the ASCEND/CAPACITY/RECAP trial in patients receiving 2,403 mg daily for pirfenidone.

b. Provide correct label for the subsequent Kaplan–Meier curves (columns K-BK)

B23. The submission states that costs for supplemental oxygen are applied to patients with a percent predicted FVC <80%. There appears to be some data in the model to facilitate this calculation but the costs applied beyond the first cycle always remain at zero due to a sumproduct calculation (in G63:H63 of Sheet named 'costs') that refers to a set of blank cells. Please clarify why costs for supplemental oxygen are not implemented in the model.

Section C: Textual clarifications and additional points

C1. Section 6.1, p.283: The company states, "An estimated 30.5% of patients with prevalent IPF have mild IPF, 54.0% have moderate IPF, and 15.5% have severe IPF [Roche 2016a]." Please provide more information regarding the source of data for these proportions and discuss whether they are representative of the distribution of severity within the population of England.

C2. The baseline characteristics for 6MWD and SGRQ in Table 67 do not look realistic. Please confirm if these figures have been mislabelled

C3. Different ICERs for pirfenidone compared with BSC (base case analysis, list prices) are reported in sections 5.7 (£38,779) and 5.11 (£38,644). Please confirm that £38,644 is a typographical error.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

C4. The penultimate bullet point in the summary box on page 48 of the submission appears to be incomplete. Please provide the complete sentence.

C5. The company submission refers to a “decline in FVC”, when it would be more appropriate to refer to a “decline in percent predicted FVC”, for example on pages 60 and 89 (which report the primary outcome of the ASCEND trial). Please confirm or clarify.

C6. Please clarify the statement in the summary box on page 89, because this analysis does not appear to be described elsewhere in the submission: “When considering patients with earlier (FVC ≥ 80% predicted) vs. later (FVC <80% predicted) disease …. There was also a numerically lower risk of FVC decline ≥10% or death in those with FVC ≥ 80% predicted, although this was not statistically significant.(p= 0.2403)”. Does the p value for the lower risk of FVC decline refer to the comparison between patients with earlier disease versus later disease, regardless of treatment received? That is, is the analysis unrelated to the effect of pirfenidone on FVC decline?

Section A: Clarification on effectiveness data

Systematic literature review

A1. *Pages 50-53: Were any searches undertaken for any ongoing trials in research registers? For example, metaRegister of Controlled Trials (http://www.isrctn.com/page/mrct), US National Institute of Health Ongoing Trials Register (clinicaltrials.gov) or the World Health Organization (http://www.who.int/ictrp/search/en/) or the EU Clinical Trials Register (https://www.clinicaltrialsregister.eu/).*

This was not performed as part of the submission, which we recognise as an oversight.

A separate search was performed in the clinical trials registers, and is reported as Appendix A to this response. The Appendix includes details of the searches performed. The search findings show that no published trials were inadvertently missed in the earlier searches. The Appendix also presents a list of ongoing studies.

A2. *Page 56: The submission states that the company was aware of the following publications that were not captured by the literature search:*

- *Noble PW, Albera C, Bradford WZ. Pirfenidone for idiopathic pulmonary fibrosis: analysis of pooled data from three multinational Phase 3 trials. European Respiratory Journal 2015; doi: 10.1183/13993003.00026-2015*
- *Albera C. Pirfenidone is efficacious in Patients With Idiopathic Pulmonary Fibrosis (IPF) and Mild Restrictive Disease: Benefit of Early Intervention. Oral presentation presented at ATS annual congress 2015; 15-20 May; Denver, USA*

Please clarify how these references were identified, given that Noble et al. was published after the company's searches were conducted.

Whilst both these additional publications only became available after the full literature searches were conducted, these were internal analyses, which we believe were important to the submission and in line with the scope of the appraisal. On this basis, they were incorporated as additional sources of information.

A3. *Appendix 3 Section A.1 (pp.15 and 17): A filter for randomised controlled trials was applied to the results of the Medline and Embase searches for clinical effectiveness. This filter has been slightly modified from the Cochrane highly sensitive search strategy published by Lefebvre et al. Please explain the reason for modifying the filter and explain the basis for the changes made. In order to increase the transparency of the reporting, please provide the source of the adapted search filter.*

The Cochrane highly sensitive search strategy published by Lefebvre et al. is now relatively old and was optimised to retrieve indexed studies. Given the volume of unindexed studies on Medline currently, we introduced the wild card into the search for 'randomi?ed' to enhance retrieval of unindexed studies where words with both spellings may be found. We do not

label adapted filters as 'adapted from XXXX' because an adapted filter does not necessarily work in the same way as the original filter and it would not be correct to imply the adapted filter has the same performance as the original filter: it is in effect a new filter. In this case, with the changes we made, the adapted filter is likely to be more sensitive and less precise than the original Cochrane RCT filter. The changes made maximised search sensitivity and hence, relevant studies were less likely to be missed.

A4. *Section 4.1, p.51: Please provide justification for using only a single systematic reviewer to assess and select relevant papers at the title and abstract stage. What are the limitations of this method? For the full-text screening process was there an assessment of inter-rater reliability between two independent reviewers?*

On review of the text on page 51, we agree that the detail of the methodology use in the selection of studies is not clear, or fully reflective of the employed approach: we apologise for this inaccuracy. To clarify, the titles and abstracts identified in the database searches were reviewed by two independent reviewers, both at abstract and full text stage. No assessment of inter-rater reliability between the two independent reviewers was done for the full-text screening process as this was not deemed necessary for the search at hand. Overall, the approach use was technically in 3 stages.

i. Obviously irrelevant studies (e.g. animal studies, case reports) were excluded by a single information specialist. These are studies that, for whatever reason, were picked up in the search but were clearly irrelevant to the review.

ii. Title and abstract screening were carried out by 2 independent reviewers

iii. Full text records were assessed by 2 independent reviewers

All potentially relevant records were therefore assessed by 2 independent reviewers.

A5. *Please confirm if data extraction and quality assessment was undertaken independently by a minimum of 2 reviewers for the clinical effectiveness reviews (including the network meta-analysis [NMA] section). If this was not done, please explain why.*

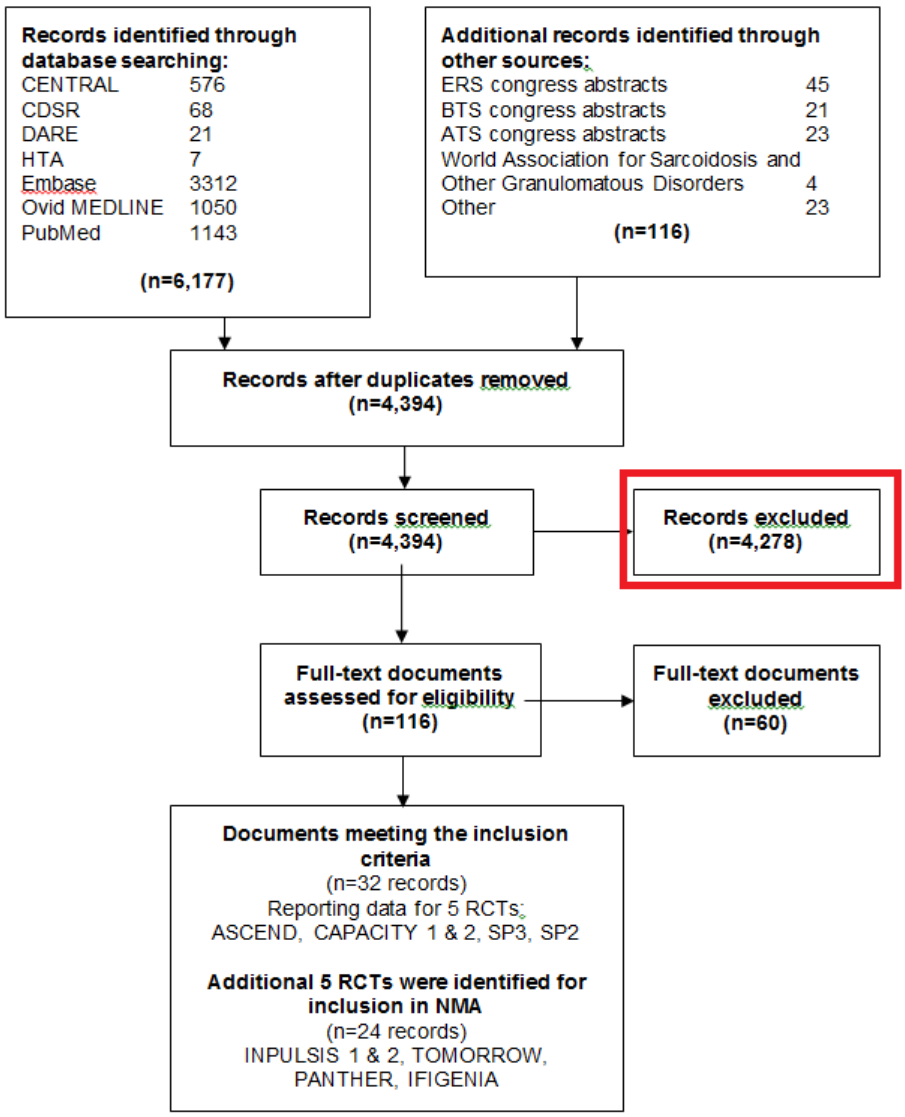
Data extraction and quality assessment was not undertaken independently by 2 reviewers. These tasks were, however, carried out by one reviewer with checking undertaken by a second reviewer. Any discrepancies were resolved by discussion with a third reviewer. The latter is the most thorough approach, short of independently extracting the data by 2 independent reviewers.

As described in the Search Objective on page 49 of the submission, the searches and data extraction for the NMA were combined with those to inform sections 4.1 and 4.2, with the same methodology employed. As described on page 123 of our submission, the searches informing the NMA were supplemented with additional data on nintedanib provided in the NICE manufacturer submission informing the appraisal of nintedanib [Boehringer Ingelheim 2015], with data extraction again carried out by one reviewer and checked by a second reviewer

A6. *Section 4.1, p.53 and Figure 2, p.54: Please provide the missing information from this figure. The reported numbers are: 4394 citations, of which 2407 were excluded (leaving 1987 citations), leading to 116 citations. Please clarify how 1987 citations were reduced to 116.*

There was an oversight in updating Figure 2 of the submission (p54), based on a previous version of this search. The number of records excluded at screening was actually 4,278 not 2,407. A corrected Figure 2 is represented below.

Figure 2. Flow diagram showing the study identification process



A7. *Section 4.6, p.86-88, Table 17 and Appendix 6: Please specify the critical appraisal tool used. Please explain why the critical appraisals reported in Appendix 6 were conducted using the Cochrane risk of bias tool, but a different tool was used and its findings reported in Section 4.6.*

The quality of randomised controlled studies reviewed was assessed using the quality elements suggested in the NICE STA guidance [NICE 2012]. A summary is presented in

Table 17 and the detailed assessment in Appendix 6 of the submission. For RCTs the following quality criteria were assessed:

- What was the method of randomisation?
- What the treatment allocation adequately concealed?
- Were the groups similar at the outset of the study in terms of prognostic factors (e.g. severity of disease)?
- Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?
- Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?
- Is there any evidence to suggest that the authors measured more outcomes than they reported?
- Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?

Additional guidance on the grading of the quality of each study against each criterion was taken from the Cochrane Handbook for Systematic Reviews of Interventions, but the risk of bias tool used is consistent with the NICE guidance.

Clinical trial design

A8. ***Priority question:*** *Please clarify whether trial patients receiving pirfenidone stopped treatment if their percent predicted forced vital capacity (FVC) decreased to below 50% (ASCEND, CAPACITY 1 & 2 and RECAP). Please comment on whether the discontinuation in trial is reflective of use of pirfenidone in clinical practice in England.*

In the ASCEND and CAPACITY trials, patients continued treatment even when FVC percent predicted decreased to below 50%. Patients had to have a FVC of >50% predicted at initiation of the study (inclusion criteria) but study drug was continued irrespective of FVC during the study period.

RECAP was an open-label, uncontrolled, Phase III extension study of ASCEND and CAPACITY trials: as such, some patients receiving pirfenidone with a predicted FVC <50% were enrolled into RECAP, as they had deteriorated to below this threshold during the blinded phase. As an open label extension treatment continuation was based on physician practice. Of those patients entering the RECAP study from the CAPACITY trials, less than [redacted] had a FVC <50% predicted, and they had [redacted] (Costabel 2016).

In clinical practice, there are no accepted thresholds of FVC percent predicted used to define the disease severity of a patient with IPF, and severity is not usually defined by FVC alone.

This is reflected within comments submitted during the scoping stage of this appraisal, along with evidence heard from clinicians during the NICE appraisal of nintedanib, which emphasise that FVC percent predicted can be hard to interpret in the presence of comorbidities, specifically it may be elevated in the presence of emphysema thus masking significant lung disease.

According to the expert panel consulted during the development of this submission, the decision to discontinue treatment with pirfenidone is usually based on a desire to transition to a palliative care approach, rather than being limited by FVC criteria alone.

A9. ***Priority question**: Please explain why carbon monoxide diffusing capacity of the lungs (DLco) was excluded from the outcomes measured in ASCEND when it had been previously collected in CAPACITY 1 & 2, SP2 and SP3.*

DLco has high intrinsic measurement variability, making demonstration of significant effects difficult. Moreover, DLco is driven not only by fibrosis in the lung, but is also influenced by concomitant emphysematous changes, or presence of pulmonary hypertension.

A decline in DLco has not been consistently associated with decreased survival. King et al. found that decreases in percent predicted DLco of ≥15% had no association with mortality [King 2005]. Similarly, in a multivariate analysis, 6- and 12-month decreases in percent predicted DLco of ≥10% were not found to be predictors of subsequent mortality in patients with IPF [Flaherty 2003]. In another study, percent predicted DLco and 24-week change in percent predicted DLco were found to be independent predictors of all-cause mortality, however, excluding baseline and longitudinal measures of percent predicted DLco from the model had no meaningful impact on model discrimination, suggesting that measures of DLco may not be incrementally informative in differentiating between IPF patients based on their mortality risk [du Bois 2011]. Therefore, changes in DLco are not good prognostic indicators for mortality, and appear to have no incremental prognostic value on top of change in FVC. Based on these observations, coupled with the complexities of the procedure and variability in DLco measurements between laboratories, it is generally considered a less robust predictor of prognosis than FVC [Maher, 2010].

Due to these difficulties in interpretation, it was decided to not use this endpoint in the ASCEND study.

A10. Section 4.5, p.82: Please clarify the statement: "the data from the three studies [ASCEND, CAPACITY 1 & 2] can be regarded as inclusive of the overall range of patients likely to present with mild to moderate IPF in routine clinical practice". This statement appears to contradict the following statement (p.182): "As only patients with mild to moderate IPF, and relatively few comorbidities, were enrolled in ASCEND, CAPACITY 1 & 2 the results need to be carefully interpreted for the broader population of patients".

These two statements are supportive of different sections of the submission, and we believe direct comparison of these statements, without the context of where they appeared, is misleading.

The first extract (p82 of the submission) was written to highlight that patients enrolled into the ASCEND and CAPACITY 1 & 2 studies are similar to mild to moderate patients seen in routine UK clinical practice, implying results are also generalisable to a UK population. The inclusion criteria represent a similar cohort of patients who present with IPF and be considered for treatment: to be enrolled, patients had to have definite usual interstitial pneumonia (UIP) pattern on radiology, or if there was a possible UIP pattern on radiology, patients had to have a lung biopsy which showed definite or probable UIP pattern on histology. This is the same classification as recommended in the ATS/ERS guidelines [Raghu 2011], and is routinely used in UK clinical practice.

The baseline demographics with regards to age, gender, FVC %-predicted, DLCO %-predicted, smoking status and time to diagnosis in the ASCEND and CAPACITY studies are similar as shown in IPF registries in the UK, as shown in Table 1 [BTS 2014]. This demonstrates that the demographics of patients in ASCEND/CAPACITY trials are similar to those treated in the UK.

The intention of the second extracted statement (p182) was to highlight the limitations of generalising the results from any trial population (whatever the trial) to patients seen in clinical practice. ASCEND and CAPACITY excluded patients with serious co-morbidities (which is common practice in IPF studies), and in line with the licensed indication, patients with a FVC<50% or DLco<30% were also excluded. As such, we suggest caution must be exercised in interpreting results from ASCEND and CAPACITY to severe patients or those with serious co-morbidities.

We would also like to make the general point that the generalisability of the CAPACITY studies was considered as part of the earlier NICE appraisal of pirfenidone, with the Committee willing to consider evidence provided by the studies when reaching their conclusions [NICE 2013].

Table 1. Comparison of baseline characteristics from ASCEND and CAPACITY trials to the BTS IPF registry data

	ASCEND [King 2014a]		CAPACITY 2 [Noble 2011; InterMune 2009b]			CAPACITY 1 [Noble 2011; InterMune 2009a]		BTS ILD registry [BTS 2014]
	PFN 2403 mg/day (n=278)	PBO (n=277)	PFN 1197 mg/day (n=87)	PFN 2403 mg/day (n=174)	PBO (n=174)	PFN (n=171)	PBO (n=173)	(n=660)
Age, mean years±SD	68.4±6.7	67.8±7.3	68.0±7.6	65.7±8.2	66.3±7.5	66.8±7.9	67.0±7.8	71±10
Male sex, n (%)	222 (79.9)	213 (76.9)	65 (75)	118 (68)	128 (74)	123 (72)	124 (72)	504 (76)
% predicted FVC, mean±SD*	67.8±11.2	68.6±10.9	76.4±14.4	74.5±14.5	76.2±15.5	74.9±13.2	73.1±14.2	54% (272/508) of patients had an FVC % predicted at presentation to the clinic of 50-80%
Former smoker, n (%)	184 (66.2)	169 (61.0)	57 (66)	110 (63)	114 (66)	112 (65)	101 (58)	443 (67)
Time since IPF diagnosis, mean±SD**	1.7±1.1	1.7±1.1	1.4±1.2	1.3±1.0	1.4±1.1	1.2±1.1	1.1±1.0	47% (308/660) of patients recorded **chest symptoms** for more than 24 months prior to first clinic visit
Surgical lung biopsy, n (%)***	86 (30.9)	79 (28.5)	32 (37)	86 (49)	85 (49)	94 (55)	94 (54)	66 (13/508)

* % predicted FVC of patients reported categorically in the BTS registry
**Time since diagnosis reported categorically in the BTS registry

A11. Section 4.3, pp.60-1: Please clarify the primary endpoint of the ASCEND and CAPACITY trials, in more specific terms. The company submission defines the primary endpoint as a change in FVC, but this could be defined more specifically in one of two ways: ≥10% decline in percent predicted FVC (or death), or absolute change in percent predicted FVC.

The protocol defined primary endpoint for ASCEND was the absolute change from baseline in percent predicted FVC at Week 52 [InterMune 2012]. The primary endpoint for both CAPACITY studies was the absolute change from baseline in percent predicted FVC at Week 72 [InterMune 2009a; InterMune 2009b].

For both studies, a hypothesis test and corresponding p-value were used to determine whether the absolute change from baseline in percent predicted FVC was significant. This was the primary analysis. On the other hand, absolute mean change (in percent predicted FVC or FVC litres) and categorical change ≥10% decline in percent predicted FVC (or death) both describe the magnitude of effect.

The ASCEND primary analysis was performed using a non-parametric rank analysis of covariance (ANCOVA) method, as missing FVC measurements due to death was imputed as 0 and the data are therefore skewed. For descriptive summaries, since the mean is not a robust measure for skewed data, the data were presented by number (%) of patients with change in % predicted FVC in the following 3 categories [InterMune 2014]:

- Decline of ≥10% or Death
- Decline of <10% to ≥0%
- No decline (>0%)

The p-values were calculated using the rank ANCOVA method.

A12. Please provide original trial protocols (with any dated changes) for ASCEND, CAPACITY 1 & 2, SP2 and SP3.

The trial protocols for ASCEND, CAPACITY 1, CAPACITY 2, SP2 and SP3 have been provided among a list of new references accompanying this response document [InterMune 2012; InterMune 2006a; InterMune 2006b; Shinogi 2001; Shinogi 2006]. Please note, SP2 and SP3 are English translations of the Japanese protocols.

Definition of clinical trial outcomes

A13. Section 4.7, p.90: Please define and justify the use of the composite outcome: ≥10% FVC or death. For example, does this refer to all-cause mortality?

The reason for using this composite outcome relates to the problem of imputation for missing FVC values in the case of death (as described in our response to A11). In order to calculate the proportion of patients who had a ≥10% FVC decline, it was assumed that death represented a ≥10% FVC decline. Death in this case refers to all-cause mortality. The reason death was assumed to constitute a ≥10% FVC decline is because once patients die

their FVC can be interpreted as 0%. Since no living patients have an FVC <10%, death can be viewed as an automatic FVC decline of 10% or more.

Therefore, we can confirm that the composite outcome of ≥10% FVC or death refers to all-cause mortality.

A14. Section 4.7, p.99, Section 4.10, p.143: Please clarify if progression-free survival (PFS) has been assessed in any other IPF trials. If so, how did the definition of PFS differ from the definitions used in the pirfenidone trials? In particular, please describe the definitions of PFS used in INPULSIS, SP3 and PANTHER.

Progression-free survival (PFS) has been assessed in other IPF trials. In these trials, PFS was assessed as a composite endpoint, and definitions varied between studies:

- ASCEND: confirmed ≥10% decline from baseline in %FVC, confirmed ≥50 m decline from baseline in 6MWT distance, or death;
- CAPACITY trials: confirmed ≥10% decline in % predicted FVC, ≥15% decline in % predicted DLco or death. In a post hoc analysis, the ASCEND definition of PFS was applied to the CAPACITY trials at 52 weeks and at 72 weeks, and used within the NMA;
- SP3: decline of 10% or more in VC or death;
- PANTHER: decline of 10% or more in FVC or death.

The differences in the definitions of PFS represent the evolving understanding of IPF and clinically meaningful endpoints over the last decade. As discussed in A9, DLco has been shown to be an inconsistent predictor of mortality whilst FVC and 6MWD have been shown to be independent predictors of mortality [du Bois 2011, du Bois 2014]. This may explain why the most recent studies (ASCEND and PANTHER) have favoured to use FVC and 6MWD as part of the definitions for PFS.

PFS for nintedanib from the INPULSIS trials was generated post-hoc and became publically available upon publication of the NICE manufacturer submission informing the appraisal of nintedanib [Boehringer Ingelheim 2015]. The nintedanib trial data was analysed to obtain a HR for nintedanib vs. placebo, by replicating the methods presented in the Noble et al. 2011 paper, i.e. CAPACITY trials. A composite endpoint of PFS was defined in the same way as the CAPACITY trials, as time to confirmed ≥10% decline in FVC percent predicted, or ≥15% decline in Dlco percent predicted, or death. A Cox proportional hazard model was run with geographic region (USA vs. non-USA) as a stratum (A12 of company response to clarification questions) [NICE 2015]. The choice of definition for PFS was likely driven by data availability; FVC and DLco were both collected whilst 6MWD was not collected.

Within the datasets available for pirfenidone, DLco is unique to the CAPACITY definition and was not collected in ASCEND. For this reason, the CAPACITY data was reanalysed to match the ASCEND definition: the inverse was not possible.

A further trial, TOMORROW, only reported the proportion of patients who progressed, rather than the proportion of patients who either progressed or died. Although the number of

deaths was reported, it was unclear how many patients progressed before they died and therefore PFS cannot be calculated.

To maintain similarity as far as possible, for CAPACITY 1 and 2, the PFS estimate based on the definition used in the ASCEND trial was included in the analysis. For the definitions of SP3, PANTHER and INPULSIS, it is assumed that they will lead to similar hazard ratios and odds ratios between a given pair of treatments, and thus that it is appropriate to combine them in an NMA. We believe this to be a reasonable assumption because in a comparison between the CAPACITY and ASCEND trials, the replacement of DLco by 6MWD led to an increase in qualifying events without changing the HR estimate.

In a previous Cochrane meta-analysis, the SP3 and CAPACITY definitions were analysed together and yielded similar results [Spagnolo 2010].

A15. *Section 4.7, p.103: The company states, "For ASCEND, acute exacerbations were identified via a post-hoc analysis of adverse events." This appears to differ from the approach taken in CAPACITY 1 & 2, SP2 and SP3 (specific definitions of acute IPF exacerbation have been provided for these trials on pp.103-4). Please clarify:*

   a. *how acute IPF exacerbations were defined in ASCEND*

   b. *whether data on acute IPF exacerbations were prospectively recorded in ASCEND*

   c. *why the definition used in CAPACITY 1 & 2 was not used in ASCEND.*

Part a: The effect of pirfenidone treatment on the incident number of episodes of acute exacerbations was not a pre-specified study endpoint in the ASCEND study: that is, 'acute exacerbation' was not specifically defined in the ASCEND study protocol. Acute exacerbations were identified via a post hoc analysis of adverse events, based on the MedDRA (version 11) lower level term 'acute exacerbation of IPF'.

Part b: Data on acute exacerbations were only collected through adverse event reporting. Adverse events were assessed at Screening, at Day 1, and at Weeks 2, 4, 8, 13, 26, 39, and 52, and were also evaluated during the Week 1 telephone call and at the restarting of treatment in patients who had an interruption. Adverse events were also collected and reported until 28 days after the last dose of study treatment in this study [InterMune 2014].

Part c: Acute exacerbations are notoriously difficult to diagnose, and there is no universally agreed definition of what constitutes an acute exacerbation of IPF. Moreover, the frequency of acute exacerbations is very low in Caucasian populations, as demonstrated by the low incidence reported in CAPACITY trials (CAPACITY 2: 3 events in placebo group, none in pirfenidone group; CAPACITY 1: 1 event in pirfenidone group, none in placebo group – Table 27 of submission). This is further supported by experience from the INPULSIS trials, which showed that adjudicated and confirmed acute exacerbations (i.e. not suspected cases) occur in less than 1% of patients [Richeldi 2014 Suppl]. Data on acute exacerbations were therefore only collected through adverse event reporting in the ASCEND study.

A16. Section 4.7, pp.110-112: Please provide further details of the minimal clinically important difference for the St. Georges Respiratory Questionnaire (SGRQ) and the Shortness of Breath Questionnaire (SOBQ).

The SGRQ has frequently been used to measure health-related quality of life in patients with IPF, but it was developed for patients with obstructive lung diseases. A triangulation approach has been used to determine a minimal important difference (MID) estimate for SGRQ scores in patients with IPF [Swigris 2010]. Using both distribution- and anchor-based approaches (using FVC, DLCO and the TDI as anchors), the MID for the SGRQ total scores was 7. Only CAPACITY trials reported data for this outcome: at Week 72, the mean change from baseline in respiratory status was similar across the pirfenidone 2403 mg/day and placebo groups in each trial (CAPACITY 1: mean score changes of 7.2 and 7.3, respectively; p=0.766; CAPACITY 2: mean score changes of 7.6 and 9.0, respectively; p=0.495) [InterMune 2009a; InterMune 2009b].

UCSD SOBQ has also been used extensively in IPF as a patient-reported outcome. A study analysing the validity of the UCSD SOBQ as a patient-reported outcome in patients with IPF examined associations between UCSD SOBQ scores and five external measures (anchors) at baseline and over time. Anchors included the Activity domain from SGRQ, Physical Function domain in SF-36, FVC, DLco, and 6MWD. The results showed the minimal important difference for the USCD SOBQ ranged from 5-11 [Swigris 2012]. Pooled data from ASCEND, CAPACITY 1 & 2 showed pirfenidone treatment reduced the proportion of patients who experienced a ≥20 point increase or death compared with placebo (p=0.0471) [Noble 2014a]

A17. Section 4.10, p.137: Please confirm if the definition of overall survival (OS) is the same across all the trials in the NMA.

The definition of OS was the same across all the trials in the NMA: patients who died due to any cause (all-cause mortality) in the intention-to-treat populations.

CAPACITY and INPULSIS trials reported two definitions of survival: all-cause mortality and treatment-emergent death. SP2 and SP3 reported deaths during the study, and the PANTHER and TOMORROW studies reported deaths from any cause.

The NMA for OS used the data described as 'all-cause mortality' from the CAPACITY and INPULSIS trials to maintain consistency.

A18. Section 4.12, p.169: Please define the adverse event category "worsening of IPF". Is this worsening due to the study drug or worsening due to lack of efficacy?

The listing "Worsening of IPF" in Table 60 of the submission was collated under the adverse event term "idiopathic pulmonary fibrosis". Since IPF was a criterion for enrolment, this category of adverse events refers to worsening of disease [King 2014]. If the investigator reported "worsening of IPF", then it would be coded to "idiopathic pulmonary fibrosis" using the MedDra dictionary.

The adverse event form was not specific in asking the investigator to assess whether the adverse event was due to lack of efficacy, although it did ask to assess whether or not it was due to the drug. In the ASCEND trial, only 5 patients in the placebo group experiencing "idiopathic pulmonary fibrosis" were considered treatment-related and none in the pirfenidone group, as described in Table 14.3.1-4 of the ASCEND CSR [InterMune 2014].

A19. Section 4.12: Please clarify how the adverse events grading of mild, moderate and severe were defined.

For the ASCEND and CAPACITY trials, the intensity (severity) of AEs was graded according to the Modified Common Terminology Criteria for Adverse Events. For events not listed in the respective protocol appendices, the following guidelines were used to evaluate the grade of intensity:

- Grade 1 – Mild: Transient or mild discomfort; no limitation in activity; no medical intervention/therapy required.
- Grade 2 – Moderate: Mild to moderate limitation in activity; some assistance may have been needed; no or minimal medical intervention/therapy required.
- Grade 3 – Severe: Marked limitation in activity; some assistance usually required; medical intervention/therapy required, hospitalisation possible

A20. Section 4.12, p.170: Please define "treatment emergent" as used in this context. How is this different from adverse events reported for the ASCEND trial in Table 60?

The CAPACITY trials used the same definition for "treatment-emergent adverse events" as ASCEND, which were defined as occurring after the first dose and within 28 days after the last dose of study treatment. For clarity, the adverse events reported from the ASCEND study in Table 60 were also treatment-emergent adverse events [InterMune 2014].

Analysis of clinical trial data

*A21. **Priority question:** Section 4.4, p.69 and Section 4.7, p.90, Table 18: Please refer to the NICE guide to the methods of technology appraisal and provide additional detail on the method of pooling data from ASCEND, CAPACITY 1 and CAPACITY 2 (including data synthesis methods and the rationale, software and models used, and any methodological limitations).*

Unfortunately, it has not been possible to collate the information required to provide a thorough and robust response to this question as part of this document. We will provide this response to NICE as soon as possible, and by no later than 18 March.

*A22. **Priority question:** Section 4.4, p.69 and Section 4.7, p.90 and 96: The published trial protocols for ASCEND and CAPACITY 1 & 2 do not mention pooled analyses of the trial data for the composite outcome (≥10% FVC or death) or for OS. Please clarify whether pooling data from these 3 trials for the endpoints listed above was "pre-specified".*

ASCEND was conducted as a confirmatory study for the FDA, and it was specifically requested by the FDA to pool all-cause mortality data of the new study (ASCEND) with CAPACITY-1 and CAPACITY-2 to provide supportive evidence of benefit. Therefore, ASCEND included all-cause mortality as a secondary endpoint and, because a single study would have too few events and power to show an unequivocal effect on mortality, pooling with the CAPACITY studies for mortality was pre-specified prospectively.

Consequently, the pooled analysis of ASCEND and CAPACITY 1 & 2 for OS was pre-specified in the Statistical Analysis Plan (SAP) of ASCEND, which was finalised on 01 January 2014. Additional pooled analyses were specified in the rest of world (ROW) integrated summary of efficacy (ISE) SAP, which was finalised on 16 February 2014. ASCEND was unblinded on 17 February 2014.

A23. *Section 4.4, p.69: The company states, "The primary outcome of ASCEND compared change in percentage predicted FVC between the 2403 mg/day and placebo groups (King 2014)." By contrast, on p.92 it states, "The ASCEND manuscript did not report the change in % predicted FVC, but this was analysed to inform the NMA." Please clarify and provide details of the sources and data in full.*

The primary outcome of ASCEND is described in A11.

The King 2014 publication did not report the mean absolute change from baseline in % predicted FVC at Week 52 from ASCEND. Therefore, mean absolute change from baseline in % predicted FVC at Week 52 was provided through data on file to inform the NMA analysis (Table 20 of the submission).

Clinical trial results

A24. ***Priority question:*** *Section 4.5, pp.76-80: Please provide full details of the nature and grade/severity of adverse events that led to discontinuation in each of the trial arms.*

REVISED QUESTION FROM NICE: *Please provide details of the adverse events that led to discontinuation, and whether any would be classified as severe, for ASCEND and CAPACITY 1 & 2. Data in a format similar to that presented in Taniguchi 2010 (table 2) for SP3 would be acceptable.*

As discussed with the NICE Project Manager, Associate Director and Technical Lead over email between 25 February to 1 March, these data are not readily available for presentation in a summary format by 3 March. Summary tables for ASCEND, CAPACITY 1 and CAPACITY 2 will be sent to NICE as soon as possible, and no later than 18 March.

A25. ***Priority question:*** *Section 4.12, p.171: Please provide rates of adverse events for pirfenidone relative to placebo, and related p values, for ASCEND and CAPACITY 1 & 2. Please provide data on all severe adverse events in these trials at all available time-points, including those that led to treatment discontinuation.*

REVISED QUESTION FROM NICE: *Please provide the numbers of serious adverse events, and p values to indicate whether there was any statistically significant difference between*

arms. Please see the data presented for CAPACITY 1 and 2 in the webappendix (pp.8-9) of the Noble 2011 publication and present the equivalent data for ASCEND, and SP2 and SP3.

As clarified with the NICE Technical Lead over email between 29 February and 1 March, the ERG would like information on serious adverse events (not severe adverse events). We have provided a tabulated summary of treatment-emergent serious adverse events from the ASCEND, CAPACITY 1 and CAPACITY 2 trials in Table 2 to Table 4, inclusive of p-values for differences between treatment arms. The number of patients with at least one serious adverse event was consistent across each arm of the CAPACITY trials (approximately 30%). Although in ASCEND, the number of patients with at least one serious adverse event was slightly higher in the placebo arm (24.9%) in comparison to the pirfenidone group (19.8%).

Summary data for SP3 are provided in response to A28.

Unfortunately, as confirmed with the NICE Technical Lead on 1 March, an equivalent summary for SP2 is not available, due to restrictions on access to data and analyses, as described on page 31 of the submission.

Table 2: Summary of All Treatment-Emergent Serious Adverse Events Reported by ≥2 Patients in Either Treatment Group (All Treated Patients) from ASCEND

Preferred Term	Number of patients, n (%)		Rate ratio (95% CI)	Pr>chi2
	Pirfenidone 2403 mg/d (N=278)	Placebo (N=277)		
Patients With at Least One TE SAE	**55 (19.8)**	**69 (24.9)**	**0.79 (0.58, 1.09)**	**0.147**
Idiopathic Pulmonary Fibrosis	7 (2.5)	27 (9.7)	0.26 (0.11, 0.58)	<0.001
Pneumonia	11 (4.0)	14 (5.1)	0.78 (0.36, 1.69)	0.533
Prostate Cancer (*M)	2 (0.7)	4 (1.4)	0.50 (0.09, 2.70)	0.409
Angina Pectoris	3 (1.1)	0 (0.0)		0.083
Nausea	3 (1.1)	0 (0.0)		0.083
Atrial Fibrillation	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Bronchitis	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Dyspnoea	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Pulmonary Embolism	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Septic Shock	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Cardiac Failure Congestive	2 (0.7)	0 (0.0)		0.157
Rib Fracture	2 (0.7)	0 (0.0)		0.157
Aortic Aneurysm	0 (0.0)	2 (0.7)	0.00	0.156
Gastroenteritis Viral	0 (0.0)	2 (0.7)	0.00	0.156

Each patient is counted only once for each preferred term. For terms followed by (*M), percentages are based on the number of males within each treatment group. Preferred terms are listed in order of decreasing frequency in the total study population.
TE SAE = treatment-emergent serious adverse events, defined as occurring after the first dose and within 28 days after the last dose of study treatment.

Table 3: Summary of All Treatment-Emergent Serious Adverse Events Reported by ≥2 Patients in Any Treatment Group (All Randomised Patients) from CAPACITY 1

Preferred Term	Number of Patients, n (%)		Rate ratio (95% CI)	Pr>chi2
	Pirfenidone 2403 mg/d	Placebo		
	(N=171)	(N=173)		
Patients With Any TE SAE	**53 (31.0%)**	**51 (29.5%)**	**1.05 (0.76, 1.45)**	**0.760**
Idiopathic Pulmonary Fibrosis	13 (7.6%)	17 (9.8%)	0.77 (0.39, 1.54)	0.465
Pneumonia	7 (4.1%)	7 (4.0%)	1.01 (0.36, 2.82)	0.982
Coronary Artery Disease	6 (3.5%)	0 (0.0%)	0.00 (0.00, 0.00)	0.013
Respiratory Failure	4 (2.3%)	6 (3.5%)	0.67 (0.19, 2.35)	0.533
Acute Respiratory Failure	2 (1.2%)	3 (1.7%)	0.67 (0.11, 3.99)	0.662
Renal Failure Acute	2 (1.2%)	2 (1.2%)	1.01 (0.14, 7.10)	0.991
Atrial Fibrillation	2 (1.2%)	1 (0.6%)	2.02 (0.19, 22.11)	0.555
Fall	2 (1.2%)	1 (0.6%)	2.02 (0.19, 22.11)	0.555
Hypotension	2 (1.2%)	1 (0.6%)	2.02 (0.19, 22.11)	0.555
Colitis	2 (1.2%)	0 (0.0%)		0.154
Hip Fracture	2 (1.2%)	0 (0.0%)		0.154
Intervertebral Disc Protrusion	2 (1.2%)	0 (0.0%)		0.154
Liver Function Test Abnormal	2 (1.2%)	0 (0.0%)		0.154
Nephrolithiasis	2 (1.2%)	0 (0.0%)		0.154
Sick Sinus Syndrome	2 (1.2%)	0 (0.0%)		0.154
Prostate Cancer (*M)	2 (1.2%)	0 (0.0%)		0.154
Angina Pectoris	1 (0.6%)	1 (0.6%)	1.01 (0.06, 16.04)	0.993
Bladder Cancer	1 (0.6%)	1 (0.6%)	1.01 (0.06, 16.04)	0.993
Cardiac Failure Congestive	1 (0.6%)	1 (0.6%)	1.01 (0.06, 16.04)	0.993
Carotid Artery Stenosis	1 (0.6%)	1 (0.6%)	1.01 (0.06, 16.04)	0.993
Diverticular Perforation	1 (0.6%)	1 (0.6%)	1.01 (0.06, 16.04)	0.993
Dyspnoea	1 (0.6%)	1 (0.6%)	1.01 (0.06, 16.04)	0.993
Pneumothorax	1 (0.6%)	1 (0.6%)	1.01 (0.06, 16.04)	0.993
Small Cell Lung Cancer Stage Unspecified	1 (0.6%)	1 (0.6%)	1.01 (0.06, 16.04)	0.993
Urinary Tract Infection	1 (0.6%)	1 (0.6%)	1.01 (0.06, 16.04)	0.993
Bronchitis	0 (0.0%)	5 (2.9%)	0.00	0.025
Hypertension	0 (0.0%)	2 (1.2%)	0.00	0.159
Hypoxia	0 (0.0%)	2 (1.2%)	0.00	0.159
Transitional Cell Carcinoma	0 (0.0%)	2 (1.2%)	0.00	0.159

Includes all TE SAEs reported after or reported before but worsened after administration of study treatment until 28 days from last dose of study treatment. Each patient is counted only once for each preferred term. For terms followed by (*M) or (*F), percentages are based on the number of males or females with each treatment group.

Table 4: Summary of All Treatment-Emergent Serious Adverse Events Reported by ≥2 Patients in Any Treatment Group (All Randomised Patients) from CAPACITY 2

Preferred Term	Number of Patients, n (%)		Rate ratio (95% CI)	Pr>chi2
	Pirfenidone 2403 mg/d (N=174)	Placebo (N=174)		
Patients with Any TE SAE	**60 (34.5%)**	**58 (33.3%)**	**1.03 (0.77, 1.39)**	**0.821**
Idiopathic pulmonary fibrosis	13 (7.5%)	14 (8.0%)	0.93 (0.45, 1.92)	0.841
Pneumonia	4 (2.3%)	6 (3.4%)	0.67 (0.19, 2.32)	0.521
Syncope	3 (1.7%)	1 (0.6%)	3.00 (0.32, 28.56)	0.315
Pneumothorax	3 (1.7%)	0		0.082
Chest pain	3 (1.7%)	0		0.082
Acute respiratory failure	2 (1.1%)	3 (1.7%)	0.67 (0.11, 3.94)	0.652
Respiratory failure	2 (1.1%)	2 (1.1%)	1.00 (0.14, 7.02)	1.000
Bronchitis	2 (1.1%)	2 (1.1%)	1.00 (0.14, 7.02)	1.000
Lobar pneumonia	2 (1.1%)	2 (1.1%)	1.00 (0.14, 7.02)	1.000
Noncardiac chest pain	2 (1.1%)	2 (1.1%)	1.00 (0.14, 7.02)	1.000
Angina pectoris	2 (1.1%)	1 (0.6%)	2.00 (0.18, 21.85)	0.562
Aortic aneurysm	2 (1.1%)	0		0.156
GERD	2 (1.1%)	0		0.156
Bladder cancer	2 (1.1%)	0		0.156
Atrial fibrillation	1 (0.6%)	1 (0.6%)	1.00 (0.06, 15.86)	1.000
Pulmonary embolism	1 (0.6%)	1 (0.6%)	1.00 (0.06, 15.86)	1.000
Colon cancer	1 (0.6%)	1 (0.6%)	1.00 (0.06, 15.86)	1.000
Renal failure acute	1 (0.6%)	0		0.317
Back pain	1 (0.6%)	0		0.317
Pleural effusion	1 (0.6%)	0		0.317
Vertigo	1 (0.6%)	0		0.317
Myocardial infarction	0 (0.0%)	4 (2.3%)	0.00 (0.00, 0.00)	0.044
Coronary artery disease	0 (0.0%)	2 (1.1%)	0.00 (0.00, 0.00)	0.156
Rectal cancer	0 (0.0%)	0	NA	NA

GERD = Gastroesophageal reflux disease
Includes all TE SAEs reported after or reported before but worsened after administration of study treatment until 28 days from last dose of study treatment. Each patient is counted only once for each preferred term.

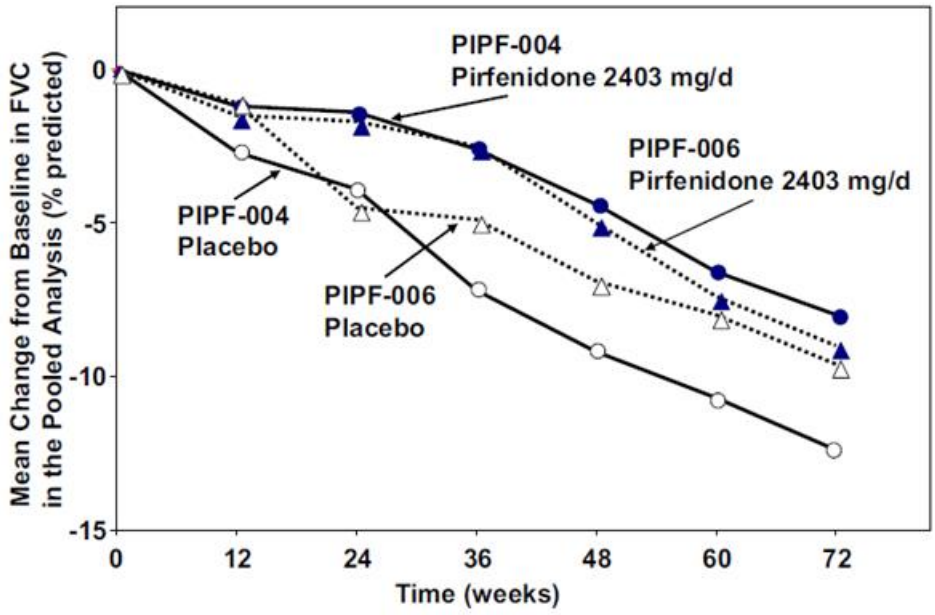
A26. ***Priority question:*** *Section 1.3, p.22: An imbalance in baseline characteristics in CAPACITY 1 is given as a possible reason for a failure to demonstrate a significant treatment effect on the trial's primary outcome. The company suggests that an example of this imbalance is that "numerically more patients in CAPACITY 1 had been diagnosed with IPF ≥1 year". This example appears to be a comparison between CAPACTIY 1 and another trial rather than a description of a baseline imbalance within CAPACITY 1. Please provide further clarification on how this statement relates to baseline imbalances. If instead it relates to differences between the trial population between CAPACITY 1 and CAPACITY 2, please explain how this would lead to differences in the primary outcomes of these trials and which direction the difference would be expected to go based on current evidence.*

We agree with the ERG that the statements provided on page 22 of the manufacturer submission relate to differences in baseline characteristics between the CAPACITY 1 and CAPACITY 2 studies.

A rationale for CAPACITY 1 not demonstrating a statistically significant difference in its primary outcome was discussed as part of the earlier NICE review of pirfenidone. The Committee's considerations on this were summarised in paragraph 4.10 of the TA282 guidance document [NICE 2013a]: "*It accepted that there was no clear explanation for the different results between PIPF-004 and PIPF-006, but was aware that the placebo group in PIPF-006 experienced less decline in FVC than expected, which could potentially be explained by the higher proportion of patients with borderline obstructive disease (FEV1/FVC less than 0.8) in this study. The Committee concluded that although it had concerns about the differences in the results of the 2 studies (particularly in relation to the outcomes considered important to people with idiopathic pulmonary fibrosis), the pooled analysis of the primary end point of PIPF-004 and PIPF-006 provided acceptable evidence of pirfenidone's overall modest treatment effect over a short duration*."

The differences in the rate of decline in FVC percent-predicted between CAPACITY 1 and CAPACITY 2 was also discussed on page 183 (and Figure 38 – reproduced below) of the submission. This outlined that the different FVC outcomes across studies may be a product of the natural variability in rates of FVC percent predicted decline in this heterogeneous disease, as supported by the difference in behaviour of the placebo groups between the two studies.

Figure 38: Mean Change from Baseline in Percent Predicted FVC in CAPACITY 2 (PIPF-004) and CAPACITY 1 (PIPF-006) (ITT population)



Further analyses (provided by InterMune as part of the original NICE appraisal, and accounted for within paragraph 4.10 above) investigated factors which may have contributed to the attenuated FVC decline in the placebo arm of CAPACITY 1.

A comparison of baseline characteristics found that the placebo group of CAPACITY 2 had a similar proportion of patients with more obstructive physiology (FEV1/FVC < 0.8) to patients randomised to pirfenidone (Table 5). Placebo patients in CAPACITY 1, however, had a higher proportion of patients with a FEV1/FVC ratio < 0.8. This is also supported by the higher portion of patients using salbutamol in the placebo arm of CAPACITY 1, relative both to patients receiving pirfenidone in CAPACITY 1, and their counterparts in CAPACITY 2.

Obstructive physiology and the presence of emphysema are associated with less decline in lung volume despite similar progression of fibrosis in the lungs [Akagi 2009]. Therefore, these imbalances stemming from the heterogeneous nature of IPF likely underlie the attenuated rate of decline in FVC in the placebo group and help to explain the different FVC outcomes across the two studies.

Table 5: Potential factors contributing to differences in FVC outcomes at Week 72 in the CAPACITY studies

	CAPACITY 2 % of patients		CAPACITY 1 % of patients	
Baseline variable	Pirfenidone	Placebo	Pirfenidone	Placebo
FEV1/FVC ratio <0.8	25	24	19	34
Salbutamol use	26	31	28	41

A27. *Section 4.7, p.104: Please clarify the consistency of the following statements: on p.104 the company states, "In SP2, the incidence of acute exacerbation of IPF was 14% (n=5) in the placebo group and was none in the pirfenidone group during the 9 months (p=0.0031) (Azuma 2005)." However, on p.105 the company states, "Whilst the incidence of hospitalisation was similar, the duration of these hospital stays was consistently numerically longer in the placebo arms. In SP2, five patients in the pirfenidone treatment arm were hospitalised due to exacerbations (Azuma 2005)."*

We would like to clarify that the statement "*whilst the incidence of hospitalisation was similar, the duration of these hospital stays was consistently numerically longer in the placebo arms*" (p105) was not related to SP2; rather, it was highlighting the numerical difference in the mean length of stay for respiratory hospitalisations for placebo vs. pirfenidone across the CAPACITY trials, which was in relation to the paragraph directly above.

A28. *Section 4.12, p.171: Please provide data on serious adverse events for SP3.*

REVISED QUESTION FROM NICE: *Please provide the numbers of serious adverse events, and p values to indicate whether there was any statistically significant difference between arms. Please see the data presented for CAPACITY 1 and 2 in the webappendix (pp.8-9) of the Noble 2011 publication and present the equivalent data for ASCEND, and SP2 and SP3.*

Summary tables for ASCEND, CAPACITY 1 and CAPACITY 2 are provided in our response to A25.

As clarified with the NICE Technical Lead on 1 March, a tabulated summary of treatment-emergent serious adverse events from the SP3 study will be sent to NICE as soon as possible, and no later than 18 March.

Subgroup analysis

A29. ***Priority question:*** *Section 4.8, p.113-117: For each subgroup analyses, please provide detailed results for each outcome (including event rates, hazard ratios, confidence intervals and p values). Please include enough information to support the statement that exploratory findings confirm the robustness and consistency of the findings across the study population. Please state whether these analyses were considered a priori or post hoc.*

Point estimates, confidence intervals and p-values for Figure 16 of the submission (p113) are presented in Table 6 below.

Figure 16 was provided in the submission based on its inclusion in the FDA briefing document [FDA 2010]. Although this analysis was pre-specified in the subgroup analyses section of the SAP for analysis for the FDA, it is different to the analysis of ranks suggested for the analysis of primary and secondary endpoints in the SAP of the clinical trial itself. Compared to the analysis of ranks, differences in observed mean effects is less robust to deviations from the normality assumption. The SAP specified that data for change-from-baseline outcomes were not expected to be normally distributed, and data were to be

analysed using a rank analysis of covariance (ANCOVA) model with a standardised rank change-from-baseline as the outcome and standardised rank baseline value as a covariate.

As results reported in Figure 16 deviated from this more robust approach for the primary outcome, we believe they should not be further used for assessment of robustness and consistency of results in subpopulations. This view is supported by the pre-specified approach used to assess subgroups in the pooled ASCEND and CAPACITY studies (ANCOVA), as presented in Figure 17 of submission.

Table 6: Statistical analyses supporting forest plots presented in Figure 16 in submission

Stratification	Absolute difference	Lower 95% CI*	Lower 95% CI*	p-value	Interaction test p-value
Sex: Male	1.1	-2.28	4.48	0.5293	0.2625
Sex: Female	5.9	1.13	10.67	0.0142	
Age: <65	3.5	-1.06	8.06	0.1288	0.8642
Age: 65 to 74	2.3	-1.66	6.26	0.2531	
Age: >=75	1.2	-5.68	8.08	0.7428	
Race: White	2.6	-0.29	5.49	0.0739	0.8070
Race: Non white	1.7	-7.30	10.70	0.7215	
Region: USA	2.9	-0.28	6.08	0.0707	0.3585
Region: ROW	1.1	-4.44	6.64	0.7071	
Time since diagnosis: <1 year	0.6	-3.26	4.46	0.7707	0.0212
Time since diagnosis: >=1 year	4.7	0.75	8.65	0.0184	
Baseline Severity FVC: <70%	3.8	-0.47	8.07	0.0773	0.3524
Baseline Severity FVC: 70 to 80%	4.8	-0.99	10.59	0.1003	
Baseline Severity FVC: >=80%	-1.2	-6.04	3.64	1.2864	
Baseline Oxygen Use: Yes	5.6	-0.77	11.97	0.0812	0.3314
Baseline Oxygen Use: No	1.7	-1.35	4.75	0.2730	
Baseline 6MWT O2 Use: Yes	9.9	0.55	19.25	0.0357	0.0036
Baseline 6MWT O2 Use: No	1.4	-1.43	4.23	0.3322	
* Confidence Intervals and p-value are calculated according to the Satterthwaite method					

Figure 17 of the submission was the presentation by Albera [2015]. This analysis is in line with the analysis of ranks that was pre-specified for the primary endpoint, however, the choice of subgroup was non pre-specified *a priori*. As per the request, analyses have been provided for the pooled population (Table 7), and for each trial (ASCEND, CAPACITY 1, CAPACITY 2 [Table 8 to Table 10]).

These exploratory findings confirm the robustness of the results in favour of pirfenidone. Figure 17 of the submission shows that the standardised treatment effect of pirfenidone is consistent across subgroups and multiple endpoints (in particular, FVC decline and 6MWT). Moreover, as shown in the Tables below, while the uncertainty of the inferences increases as we look at study-specific effects (due to smaller groups of individuals being analysed), the

results remain consistent across studies, endpoints and subgroups. The exception is the CAPACITY 1 study, where differences in the magnitude of point estimates were observed for the mild and moderate population, nevertheless this effect was not statistically significant (p-value for treatment by subgroup interaction >0.05).

Table 7: Rank ANCOVA analysis and derivation of standardised treatment effect – pooled data

Endpoint	Subgroup	Pirfenidone - Average Standardised Rank*	95% CI	Placebo - Average Standardised Rank*	(95% CI)	Difference in standardised ranks*	(95% CI)	Standardised Treatment Effect	(95% CI)	Interaction Test p-value
Decline in FVC, result at 12 months	FVC <80%	0.55	(0.52;0.58)	0.45	(0.42;0.48)	0.1	(0.06;0.13)	0.38	(0.24;0.53)	0.397
	FVC >=80%	0.56	(0.51;0.62)	0.44	(0.38;0.49)	0.13	(0.06;0.19)	0.49	(0.25;0.74)	
	GAP stage I	0.56	(0.52;0.59)	0.46	(0.42;0.5)	0.1	(0.05;0.15)	0.4	(0.2;0.6)	0.8153
	GAP stage II-III	0.55	(0.52;0.58)	0.44	(0.41;0.47)	0.11	(0.07;0.15)	0.42	(0.26;0.58)	
Decline in 6MWD ability, results at 12 months	FVC <80%	0.51	(0.48;0.54)	0.45	(0.42;0.48)	0.06	(0.02;0.1)	0.23	(0.09;0.38)	0.9583
	FVC >=80%	0.58	(0.54;0.63)	0.52	(0.48;0.57)	0.06	(0;0.12)	0.25	(-0.02;0.51)	
	GAP stage I	0.57	(0.53;0.61)	0.51	(0.48;0.55)	0.06	(0.01;0.11)	0.23	(0.03;0.44)	0.9328
	GAP stage II-III	0.5	(0.47;0.53)	0.44	(0.41;0.47)	0.05	(0.01;0.1)	0.21	(0.05;0.37)	
Change in SOBQ, results at 12 months	FVC <80%	0.5	(0.47;0.53)	0.55	(0.52;0.58)	-0.05	(-0.09;-0.01)	0.19	(0.33;0.05)	0.1957
	FVC >=80%	0.43	(0.39;0.48)	0.43	(0.39;0.48)	0	(-0.06;0.06)	0	(0.27;-0.26)	
	GAP stage I	0.45	(0.41;0.48)	0.44	(0.41;0.48)	0	(-0.05;0.05)	-0.01	(0.19;-0.22)	0.0804
	GAP stage II-III	0.51	(0.48;0.54)	0.57	(0.53;0.6)	-0.05	(-0.1;-0.01)	0.21	(0.37;0.06)	

* From the Rank ANCOVA stratified by study and region of the world and adjusted for baseline

Table 8: Rank ANCOVA analysis and derivation of standardised treatment effect – ASCEND study

Endpoint	Subgroup	Pirfenidone - Average Standardised Rank*	95% CI	Placebo - Average Standardised Rank*	(95% CI)	Difference in standardised ranks*	(95% CI)	Standardised Treatment Effect	(95% CI)	Interaction Test p-value
Decline in FVC, result at 12 months	FVC <80%	0.56	(0.52;0.6)	0.44	(0.4;0.48)	0.12	(0.06;0.17)	0.47	(0.26;0.68)	0.7817
	FVC >=80%	0.56	(0.47;0.65)	0.43	(0.34;0.51)	0.13	(0.02;0.24)	0.52	(0.09;0.95)	
	GAP stage I	0.59	(0.52;0.65)	0.44	(0.38;0.5)	0.15	(0.06;0.24)	0.59	(0.23;0.95)	0.4873
	GAP stage II-III	0.55	(0.51;0.59)	0.44	(0.4;0.48)	0.11	(0.05;0.16)	0.44	(0.22;0.65)	
Decline in 6MWD ability, results at 12 months	FVC <80%	0.51	(0.48;0.55)	0.46	(0.42;0.5)	0.06	(0;0.11)	0.22	(0.01;0.42)	0.7244
	FVC >=80%	0.57	(0.49;0.66)	0.54	(0.47;0.62)	0.03	(-0.08;0.15)	0.13	(-0.32;0.59)	
	GAP stage I	0.57	(0.51;0.64)	0.54	(0.47;0.6)	0.04	(-0.05;0.13)	0.15	(-0.22;0.53)	0.7474
	GAP stage II-III	0.5	(0.46;0.55)	0.45	(0.41;0.49)	0.05	(0;0.11)	0.21	(-0.01;0.43)	
Change in SOBQ, results at 12 months	FVC <80%	0.51	(0.47;0.55)	0.55	(0.51;0.59)	-0.04	(-0.09;0.01)	0.15	(0.35;-0.05)	0.9355
	FVC >=80%	0.36	(0.28;0.45)	0.41	(0.33;0.48)	-0.04	(-0.16;0.07)	0.19	(0.67;-0.29)	
	GAP stage I	0.37	(0.3;0.43)	0.42	(0.36;0.48)	-0.05	(-0.14;0.04)	0.22	(0.59;-0.16)	0.6879
	GAP stage II-III	0.53	(0.49;0.57)	0.56	(0.52;0.6)	-0.03	(-0.09;0.02)	0.12	(0.33;-0.09)	

* From the Rank ANCOVA stratified by study and region of the world and adjusted for baseline

Table 9: Rank ANCOVA analysis and derivation of standardised treatment effect – CAPACITY 2 study

Endpoint	Subgroup	Pirfenidone - Average Standardised Rank*	95% CI	Placebo - Average Standardised Rank*	(95% CI)	Difference in standardised ranks*	(95% CI)	Standardised Treatment Effect	(95% CI)	Interaction Test p-value
Decline in FVC, result at 12 months	FVC <80%	0.51	(0.46;0.57)	0.41	(0.35;0.48)	0.1	(0.03;0.17)	0.4	(0.11;0.69)	0.7339
	FVC >=80%	0.64	(0.54;0.74)	0.52	(0.43;0.61)	0.12	(0.02;0.23)	0.48	(0.07;0.89)	
	GAP stage I	0.56	(0.5;0.62)	0.45	(0.39;0.52)	0.1	(0.02;0.19)	0.41	(0.07;0.76)	0.9109
	GAP stage II-III	0.54	(0.48;0.6)	0.45	(0.39;0.5)	0.1	(0.01;0.18)	0.38	(0.05;0.7)	
Decline in 6MWD ability, results at 12 months	FVC <80%	0.51	(0.46;0.56)	0.43	(0.38;0.49)	0.08	(0;0.15)	0.29	(0.02;0.57)	0.5611
	FVC >=80%	0.57	(0.49;0.66)	0.53	(0.46;0.6)	0.04	(-0.07;0.15)	0.17	(-0.3;0.65)	
	GAP stage I	0.59	(0.53;0.65)	0.51	(0.45;0.58)	0.08	(-0.01;0.17)	0.32	(-0.03;0.67)	0.3443
	GAP stage II-III	0.46	(0.4;0.52)	0.44	(0.38;0.5)	0.02	(-0.06;0.11)	0.08	(-0.24;0.41)	
Change in SOBQ, results at 12 months	FVC <80%	0.5	(0.45;0.55)	0.56	(0.5;0.61)	-0.06	(-0.14;0.01)	0.23	(0.52;-0.05)	0.4773
	FVC >=80%	0.44	(0.36;0.53)	0.46	(0.39;0.53)	-0.01	(-0.12;0.09)	0.06	(0.51;-0.4)	
	GAP stage I	0.47	(0.41;0.53)	0.49	(0.42;0.55)	-0.02	(-0.11;0.07)	0.08	(0.44;-0.28)	0.6582
	GAP stage II-III	0.5	(0.43;0.56)	0.54	(0.48;0.6)	-0.05	(-0.13;0.04)	0.18	(0.51;-0.14)	

* From the Rank ANCOVA stratified by study and region of the world and adjusted for baseline

Table 10: Rank ANCOVA analysis and derivation of standardised treatment effect – CAPACITY 1 study

Endpoint	Subgroup	Pirfenidone - Average Standardised Rank*	95% CI	Placebo - Average Standardised Rank*	(95% CI)	Difference in standardised ranks*	(95% CI)	Standardised Treatment Effect	(95% CI)	Interaction Test p-value
Decline in FVC, result at 12 months	FVC <80%	0.55	(0.49;0.62)	0.49	(0.43;0.55)	0.06	(-0.01;0.13)	0.25	(-0.04;0.53)	0.2041
	FVC >=80%	0.51	(0.4;0.61)	0.36	(0.25;0.47)	0.15	(0.04;0.26)	0.58	(0.14;1.02)	
	GAP stage I	0.53	(0.46;0.6)	0.48	(0.4;0.55)	0.06	(-0.03;0.15)	0.23	(-0.12;0.58)	0.3352
	GAP stage II-III	0.56	(0.49;0.63)	0.44	(0.37;0.51)	0.12	(0.03;0.2)	0.47	(0.14;0.79)	
Decline in 6MWD ability, results at 12 months	FVC <80%	0.51	(0.45;0.57)	0.45	(0.4;0.51)	0.06	(-0.02;0.13)	0.21	(-0.07;0.49)	0.482
	FVC >=80%	0.6	(0.51;0.68)	0.49	(0.4;0.58)	0.1	(-0.01;0.22)	0.44	(-0.04;0.93)	
	GAP stage I	0.54	(0.47;0.61)	0.49	(0.42;0.57)	0.05	(-0.04;0.14)	0.2	(-0.16;0.55)	0.513
	GAP stage II-III	0.53	(0.46;0.6)	0.44	(0.37;0.51)	0.09	(0.01;0.17)	0.35	(0.02;0.68)	
Change in SOBQ, results at 12 months	FVC <80%	0.48	(0.42;0.54)	0.54	(0.48;0.6)	-0.06	(-0.13;0.02)	0.22	(0.5;-0.06)	0.1041
	FVC >=80%	0.48	(0.4;0.57)	0.43	(0.34;0.52)	0.05	(-0.06;0.17)	-0.22	(0.24;-0.68)	
	GAP stage I	0.51	(0.44;0.57)	0.43	(0.35;0.5)	0.08	(-0.01;0.17)	-0.31	(0.03;-0.66)	0.0013
	GAP stage II-III	0.46	(0.4;0.53)	0.58	(0.52;0.65)	-0.12	(-0.2;-0.04)	0.48	(0.81;0.15)	

* From the Rank ANCOVA stratified by study and region of the world and adjusted for baseline

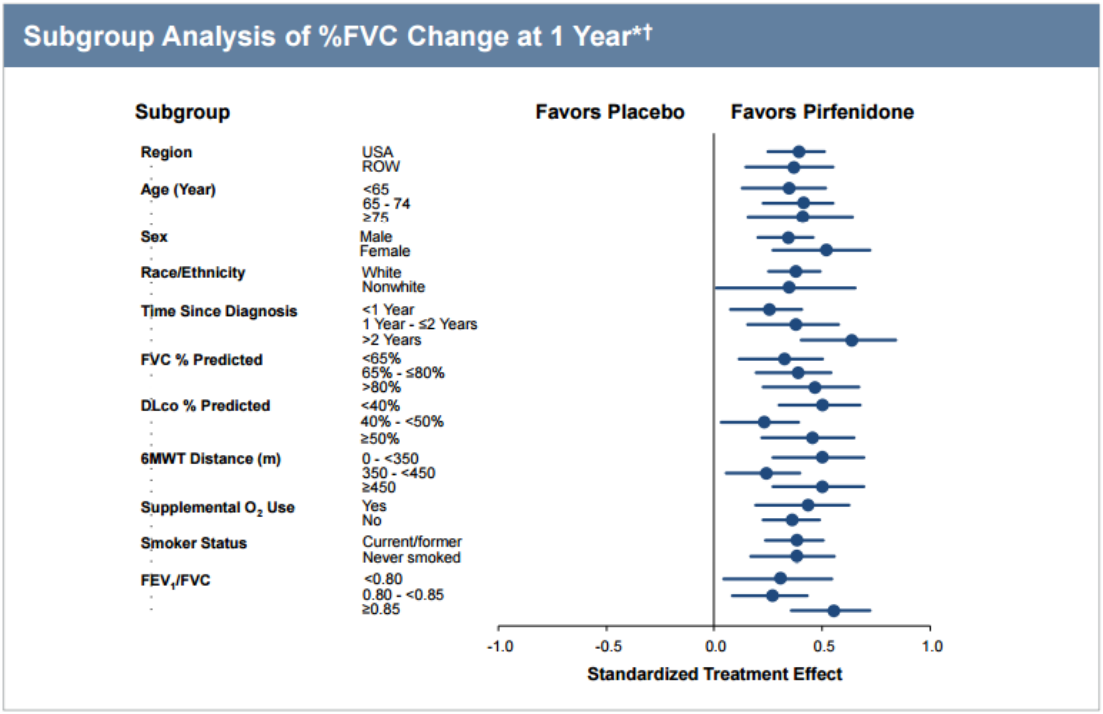
A30. ***Priority question:*** *Section 4.8, p.113: The company present a subgroup analysis of the CAPACITY trials using 3 categories of baseline percent predicted FVC: <70%, 70-80% and ≥80%. Please provide a rationale for investigating these 3 subgroups.*

Roche were not involved in the design of the CAPACITY trials, and the clinical rationale for the pre-specification of these subgroups is not known.

Pirfenidone was the first treatment to be licensed in the management of IPF: the design of the CAPACITY programme, therefore, took place at a time where knowledge of the condition and its management were still growing. On this basis, it is possible that the specific cut-offs were based on clinical rationale, as opposed to published evidence.

With the emergence of published evidence showing the significant mortality risk for patients below an FVC of 80% [du Bois 2011], the ASCEND study considered alternative cut-offs of FVC % predicted, which included a cut-point of 80%. Based on pooled subgroup analyses of the ASCEND and CAPACITY studies, the therapeutic benefit of pirfenidone in reducing the decline in lung function in patients with IPF was consistent across a wide range of patient subgroups and baseline disease severities, with no evidence of a patient subgroup in which the treatment effect is unfavourable, as shown in Figure 1 [Noble 2014].

Figure 1: Forest plot of subgroups in pooled trial data (ASCEND & CAPACITY trials)



* Rank ANCOVA model with standardized treatment effects
† Statistical test for interaction provides no evidence that treatment effect is different at different levels of any covariate, except time since IPF diagnosis (p=0.034)

A31. ***Priority question****: Section 4.8, p.114-5: Please provide additional results from the subgroup analysis stratified by baseline percent predicted FVC (<80% and ≥80%):*

a. *Provide the results in figure 17 separately for each of the 3 individual trials (ASCEND, CAPACITY 1 and CAPACITY 2).*

b. *In addition, provide OS and PFS results for both subgroups at both 52 and 72 weeks. Please provide these for the 3 individual trials (ASCEND, CAPACITY 1 and CAPACITY 2) and for the pooled dataset.*

Part a: The requested analyses can be found in the answer provided for question A29.

Part b: PFS results for the patients with baseline FVC <80% predicted and ≥80% predicted across the three trials (and pooled results) are presented in Table 11. OS results are presented in Table 12.

Table 11: PFS results by baseline FVC percent predicted subgroup at 52 and 72 weeks

Study & time point	Baseline FVC ≤80% predicted			Baseline FVC >80% predicted		
	Adjusted HR	95% CI	p-value	Adjusted HR	95% CI	p-value
CAPACITY 1						
52 weeks	0.84	0.53-1.32	0.4438	0.63	0.29-1.41	0.2571
72 weeks	0.85	0.58-1.26	0.4128	0.56	0.28-1.11	0.0919
CAPACITY 2						
52 weeks	0.60	0.40-0.92	0.0159	0.40	0.18-0.89	0.0193
72 weeks	0.58	0.39-0.86	0.0590	0.48	0.25-0.92	0.0233
ASCEND						
52 weeks	0.56	0.41-0.76	0.0002	0.64	0.30-1.40	0.2584
72 weeks	N/A	N/A	N/A	N/A	N/A	N/A
Pooled trials						
52 weeks	0.62	0.52-0.78	<0.0001	0.54	0.35-0.75	0.0069
72 weeks	0.64	0.52-0.79	<0.0001	0.53	0.35-0.79	0.0017

Table 12: OS results by baseline FVC percent predicted subgroup at 52 and 72 weeks

Study & time point	Baseline FVC ≤80% predicted			Baseline FVC >80% predicted		
	Adjusted HR	95% CI	p-value	Adjusted HR	95% CI	p-value
CAPACITY 1						
52 weeks	0.6	0.17-2.04	0.4051	0.77*	0.11-5.59	0.7932
72 weeks	0.89	0.40-1.99	0.7763	0.77	0.11-5.59	0.7932
CAPACITY 2						
52 weeks	0.25	0.08-0.76	0.0080	NE**	**	**
72 weeks	0.29	0.10-0.79	0.0102	4.04***	0.42-38.87***	0.1900***
ASCEND						
52 weeks	0.63	0.29-1.34	0.2215	<0.01	0.00-NE	0.1231
72 weeks	N/A	N/A	N/A	N/A	N/A	N/A
Pooled trials						
52 weeks	0.48	0.27-0.83	0.0071	0.59	0.14-2.51	0.4682
72 weeks	0.58	0.36-0.94	0.0240	0.90	0.27-2.99	0.8610

NE: not evaluable
* Only two deaths occurred in CAPACITY 1 before 52 weeks
** There were no additional deaths observed in either arm of CAPACITY 2 between 52 and 72 weeks in patients with FVC >80% predicted

*** Low number of events

Network meta-analysis (NMA)

A32. ***Priority question:*** *Section 4.10 p.125. A network diagram is provided for all studies that contribute to the NMA, however different trials are included in the NMA for each outcome. Please present separate network diagrams for each outcome.*

The same evidence network was assessed for all outcomes, although not all trials provided evidence for each outcome. Where data on a particular outcome were available from one of the included trials, this was incorporated into the NMA.

For space considerations, the evidence networks for each outcome were not included in the submission, although these are presented below. We start by providing a table which summarises which data were available in each trial for each endpoint (in the base case network, Table 13). The network diagrams are provided in Appendix B.

Table 13: Summary of evidence considered in NMAs

	Study duration (weeks)	All-cause mortality	IPF-related mortality	Progression-Free Survival	Exacerbations	10% categorical decline FVC% pred [d]	FVC% pred	FVC Litres	6MWD	SGRQ	UCSD SOBQ
CAPACITY1 and 2	72	√	√	√	√	√	√	√	√	√	√
ASCEND	52	√	√	√	√	√	√	√	√		√
SP3	52	√[a]		√	√	[e]	√	√			
INPULSIS1 and 2	52	√	√*	√*	√	√	√	√		√	
TOMORROW	52	√[a]	√[b]	[f]	√	[g]	√	√	√	√	
PANTHER (NAC)	60	√[b]	√[a]	√[a]	√	√	√	√	√	√	√
PANTHER (Triple)	32	√	√[a]	√	[h]	[h]		√[c]	√[c]	√[c]	√[c]

[a] HRs were unavailable: number of events and number of patients were used as an alternative (via the Woods model)
[b] HR was calculated from other available data using the methods of Parmar
[c] For FVC (L), 6MWD, SGRQ and UCSD SOBQ, results could be included in the NMAs, as the publication presented estimated changes over 60 weeks (based on a repeated measures model)
[d] The NMA for this outcome assumes that all patients with missing values are non-responders (i.e. have a decline of more than 10%)
[e] Taniguchi 2010 reported some results for FVC 10% (Table E2 in the supplementary appendix) however there was insufficient information to calculate FVC 10% in line with the above definition
[f] only reported the proportion of patients who progressed, rather than the proportion of patients who either progressed or died. Although the number of deaths was reported, it was unclear how many patients progressed before they died and therefore PFS cannot be calculated.
[g] The outcome is not clearly defined in the nintedanib manufacturer submission to NICE. Based on the manufacturer's response to clarification questions, the submission may be measuring any decline up to 52 weeks, whereas the other studies are measuring declines at exactly 48/52 weeks.
[h] Results were reported but the time point was not comparable.
* Our review – including the direction provided by NICE on 26 February 2016 over email – did not identify individual HRs for the INPULSIS trials for IPF-related mortality and PFS [1]. Based on this, we used the pooled HR. [1] Post hoc analysis only supplied as part of the submission to NICE by Boehringer Ingelheim.

A33. ***Priority question:*** *Section 4.7, p.99, Section 4.10, p.143: For CAPACITY 1 & 2, PFS was reanalysed using the definition used in the ASCEND trial. For other trials (INPULSIS, SP3, PANTHER), it is "assumed that they will lead to similar hazard ratios and odds ratios between a given pair of treatments, and thus that it is appropriate to combine them in an NMA" (p.143). Please validate this assumption by reanalysing the individual patient data from CAPACITY 1 & 2 and ASCEND, using the PFS definition(s) used in INPULSIS, SP3, and PANTHER trials.*

The rationale for performing the PFS NMA included in our submission is described in our response to A14.

Reanalysis of INPULSIS data was provided in the nintedanib manufacturer's evidence submission [Boehringer Ingelheim 2015] to match the pirfenidone NDA 2009 definition (which includes FVC decline, death and/or DLCO decline), as published in Noble [2011].

The SP3 and PANTHER definitions are the same, and limited to FVC decline and death – i.e. without a DLCO or 6MWDT decline component.

Table 14 provides a comparison of the PFS hazard ratios included in our PFS NMAs at 52 and 72 weeks (using the ASCEND study definition for pirfenidone trials, as presented in the submission) with PFS hazards ratios resulting from reanalysis of individual patient data from CAPACITY 1 & 2 and ASCEND according to the NDA 2009 and 2014 definitions.

Results demonstrate that the hazard ratios used in our NMA for PFS at 52 weeks and PFS at 72 weeks are more conservative at comparable time points than the comparable hazard ratio estimates for ASCEND, CAPACITY 1 and CAPACITY 2 by the PANTHER/SP3 PFS definition. They are comparable to the respective NDA 2009 definition PFS hazard ratios.

The NMAs we have run are the most robust way we have of including all the relevant information (Figure 2) and provide a conservative estimate of comparative effectiveness (Table 14).

Table 14: Reanalysis of the individual patient data from CAPACITY 1 & 2 and ASCEND, using the PFS definition(s) used in CAPACITY and SP3/PANTHER trials – HRs

	ASCEND	CAPACITY 2			CAPACITY 1			Pooled CAPACITY studies			Pooled ASCEND & CAPACITY studies		
Censoring time point (in weeks)	**52**	**52**	**72**	**End of study**	**52**	**72**	**End of study**	**52**	**72**	**End of study**	**52**	**72**	**End of study**
InterMune ASCEND definition [a]	0.57 (0.42, 0.77)	0.58 (0.40, 0.83)	0.57 (0.41, 0.80)	0.63 (0.46, 0.87)	0.78 (0.52, 1.15)	0.75 (0.54, 1.06)	0.79 (0.58, 1.08)				0.62 (0.50, 0.76)	0.62 (0.51, 0.75)	0.65 (0.55, 0.78)
Definition used in SP3/PANTHER [c]	0.36 (0.23, 0.56)	0.43 (0.24, 0.76)	0.45 (0.28, 0.72)	0.60 (0.40, 0.90)	0.77 (0.44, 1.35)	0.73 (0.46, 1.15)	0.80 (0.54, 1.18)				0.46 (0.34, 0.62)	0.48 (0.37, 0.63)	0.57 (0.45, 0.72)
NDA 2009 definition (Noble publication, incl DLco) [b]	not possible	0.41 (0.24, 0.70)	0.49 (0.31, 0.75)	0.64 (0.44, 0.95)*	0.86 (0.51, 1.43)	0.82 (0.54, 1.25)	0.84 (0.58, 1.22)*	0.59 (0.41, 0.85)	0.64 (0.47, 0.86)	0.74 (0.57, 0.96)*			

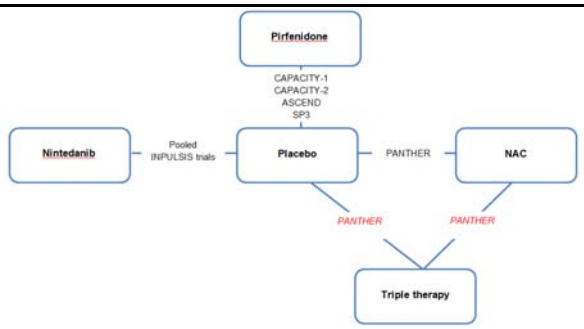
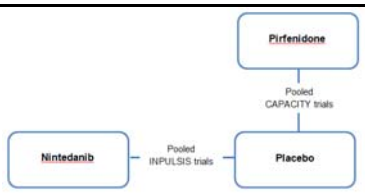
[a] confirmed ≥10% decline from baseline in %FVC, confirmed ≥50 m decline from baseline in 6MWT distance, or death

[b] confirmed ≥10% decline in % predicted FVC, ≥15% decline in % predicted DLco or death

[c] decline of 10% or more in VC/FVC or death

*available in the Noble [2011]

Figure 2: Network diagrams for a) pirfenidone PFS NMA [NICE 2016] compared to b) nintedanib PFS NMA [NICE 2015]

a) PFS NMA in our submission [NICE 2016]	b) PFS NMA in nintedanib submission [NICE 2016]
	
All trials are included in the base-case network *Red italicised text indicates trials excluded from the restricted network*	

A34. ***Priority question:*** *Section 4.10 p.133-155. Please provide the following additional information for the NMA results:*

a. *Estimates of the between-study heterogeneity for all random effects models.*

b. *The 95% predictive intervals (PrI) in addition to the credible intervals (CrI) that are currently presented. The NICE Decision Support Unit (in Technical Support Document 2) recommends that the predictive distribution, rather than the distribution of the mean treatment effect, better represents uncertainty about comparative effectiveness in the presence of heterogeneity.*

c. *Model fit statistics (total residual deviance and deviance information criteria [DIC]) to allow a comparison of the random and fixed-effects analyses.*

The models had to be re-run to derive the predictive intervals and the total residual deviance. Based on the time available, models were re-run for the following key outcomes informing the economic model: all-cause mortality at 52 weeks; all-cause mortality at 72 weeks; PFS at 52 weeks; PFS at 72 weeks; IPF-related mortality at 52 weeks; IPF-related mortality at 72 weeks, and acute exacerbations.

Posterior summaries of the between-study standard deviation for all random effects models, the 95% predictive intervals for all random effects models, and model fit statistics (total residual deviance and DIC) are provided in Appendix C for these key outcomes.

We do not observe any meaningful differences in DIC between random effects and fixed effect models. All DICs comparisons were below 3.

In addition, we derived the posterior probability of a treatment being better than another treatment for each pairwise comparison.

A35. ***Priority question****: Section 4.10. Please re-run the NMA with the changes below.*

a. *The evidence network includes one 3 arm trial (PANTHER), however no correction for multi-arm trials was implemented. The ERG considers that although the correlation between arms may be reduced, the assumption of zero correlation is not appropriate. Please implement correction for multi-arm trials for the NMA where required.*

b. *The NMAs of survival outcomes used a pooled hazard ratio for the INPULSIS trials. The ERG notes that results from the individual trials are available and should be used to inform the NMA. Please repeat the analyses with the trial level estimates of treatment effect.*

Part a: The systematic review identified one multi-arm trial for inclusion in the NMA. PANTHER included three arms: placebo, NAC and triple therapy. However, the triple therapy arm was terminated early. The first PANTHER publication reported results for triple therapy vs. placebo only [Raghu 2012]. Comparisons were based on 78 patients in the placebo arm and 77 patients in the triple-therapy arm. Comparisons were reported up to a mean follow-up of approximately 32 weeks (of a planned 60 weeks). After termination of the triple therapy arm, recruitment continued for the placebo and NAC arms. The second PANTHER publication reported results for NAC vs. placebo only, based on a total of 131 patients in the placebo arm and 131 patients in the NAC arm [Martinez 2014]. Comparisons were reported up to 60 weeks as planned.

Hence, PANTHER is an atypical multi-arm trial. The comparisons of placebo with triple-therapy involve only a subset of the placebo patients included in the comparisons of placebo with NAC. The two comparisons were also conducted at different time points. Correlations between the arms will be less than those in a regular multi-arm trial.

Furthermore, triple therapy is only included in the NMAs performed via a sensitivity analysis, and so correlations do not need to be accounted for in the base case analysis. Hence, for the sensitivity analysis, it is assumed that the correlation between the active arms in PANTHER is negligible and it is not explicitly accounted for in the NMA models, but this assumption *does not* apply to the base case analyses.

On this basis, we do not believe that it is necessary to rerun the NMAs correcting for multi-arm trials.

Part b: In line with discussion over email with the NICE Project Manager and Associate Director (25-26 February), we aimed to include separate study inputs over pooled estimates wherever available.

We did not pick up on the individual estimates for OS in the individual INPULSIS trials, as reported in Table 19 of the nintedanib manufacturer's evidence submission, as the submission used a pooled estimate in their overall survival NMA (Table 29 of their submission dossier [Boehringer Ingelheim 2015]). We apologise for this oversight. As part of our response to A34, we have now rerun the all-cause mortality NMAs, using individual data inputs for the INPULSIS trials.

Neither our systematic review, nor our review of the nintedanib manufacturer submission or SmPC, identified individual hazard ratios for the INPULSIS trials for IPF-related mortality or PFS. On this basis, the pooled hazard ratio for these studies was used.

A36. ***Priority question**: Section 4.10, p.125: The individual trials report outcomes at different time points. These are synthesised under the assumption that the treatment effects are constant over time.*

a. *Please provide evidence to justify this assumption, by considering the effect of including a covariate for trial duration through meta-regression.*

b. *If time allows, for binary outcomes please consider the use of a complementary log-log (cloglog) link function (as described in the NICE Decision Support Unit Technical Support Document 2) or the piece-wise constant hazard model of Lu et al (2007). These approaches do not rely on the assumption of constant treatment effects.*

Part a: As discussed with the NICE Project Manager and Associate Director over email on 25-26 February, these analyses will be supplied as soon as possible, and by 18 March.

Part b: Unfortunately, it is not possible to provide this analysis within the requested timeframes.

A37. ***Priority question:** Section 4.10, p.126: For the NMA of the survival outcomes, the principal analysis assumed that the "proportional hazards is an acceptable assumption up to 52 weeks, but not beyond". A sensitivity analysis explored the assumption that proportional hazards apply until week 72. Please justify choosing the weaker assumption for the base case analyses. Is there reason to believe that the hazards may not be proportional beyond 52 weeks?*

For the base case analyses we choose the 52 weeks endpoint in order to optimize comparability with other therapies (i.e. nintedanib). Moreover, beyond 52 weeks more than half of the patients contributing to the pooled pirfenidone dataset were no longer followed-up (patients enrolled into ASCEND, as presented in Figure 3 of the submission [p70]). In survival analyses, termination patterns can affect the estimate of effect: when few patients are at risk, large steps in the survival curve are to be expected that impact on the quality of the estimates.

The choice of assessing survival outcomes at 52 weeks was, therefore, based on following factors:

- Full follow up data were available for the majority of the 1247 patients enrolled across the 3 trials;
- The running of the ASCEND trial was specified by the FDA as a requirement in order to confirm the OS benefit of pirfenidone; analysis at 52 weeks allows inclusion of data from this trial with equal weighting to the other trials. When looking at longer time horizons, data are not available from this trial;

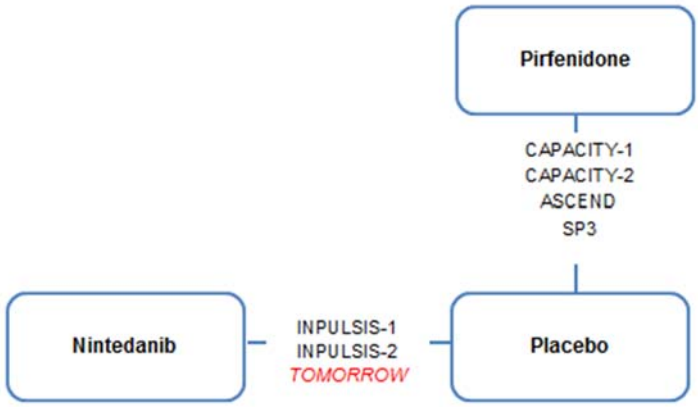
- Clinical trial data from ASCEND and CAPACITY studies were pre-specified to be pooled at the 52 week time point;
- The period of observation was similar to that assessed with comparator treatments included in the NMA and appraisal scope allowing comparison of data across similar timeframes, and;
- There are no data available to support an assumption of proportional hazards in the longer term for nintedanib versus placebo.

As noted within Appendix 20 of the original submission (and demonstrated by the answers supplied in response to question A14) there is no evidence to indicate an assumption of proportional hazards for longer time periods for the comparison of pirfenidone to placebo is not appropriate within the data being analysed. The issue is that such evidence is not available for nintedanib.

A38. Section 4.10, p.124: Please provide a sensitivity analysis, excluding PANTHER and SP3 from the NMA. The ERG notes that there are questions over the generalisability of the population in the SP3 trial, and over the inclusion of the PANTHER trial because the active treatments are not listed as comparators in the scope.

Revised results, excluding results of the PANTHER study (Figure 3) have been provided as part of our response to A34/A35 (Appendix C) for all-cause mortality at 52 weeks, as proof of concept that in such a star-shaped network, excluding PANTHER does not change the comparative results of pirfenidone, nintedanib and placebo (Table 15).

Figure 3: Sensitivity analysis for exclusion of PANTHER results from NMAs of key outcomes



All trials are included in the base-case network
Red italicised text indicates trials excluded from the restricted network

Table 15: Comparison of ACM 52 weeks main NMA results, with and without the PANTHER trial (random effect and fixed effects model)*

	Original expanded network				Excluding the PANTHER trial			
	HR estimate (95%CrI)				HR estimate (95%CrI)			
	Pirf vs placebo	Nin vs placebo	Pirf vs Nin	Pirf vs NAC	Pirf vs placebo	Nin vs placebo	Pirf vs Nin	Pirf vs NAC
RE model	0.52 (0.30, 0.88)**	0.71 (0.43, 1.16)	0.73 (0.35, 1.50)	0.26 (0.06, 1.12)	0.52 (0.30, 0.88)**	0.71 (0.43, 1.16)	0.73 (0.35, 1.51)	NA
FE model	0.52 (0.32, 0.84)**	0.71 (0.46, 1.09)	0.73 (0.38, 1.40)	0.26 (0.06, 1.04)	0.52 (0.32, 0.84)**	0.71 (0.46, 1.09)	0.73 (0.38, 1.40)	NA

* Amendments related to question A35b have been incorporated in the generation of these estimates; for full detail of these estimates see Appendix C
** HRs do not cross 1

We do not agree that there are concerns over the generalisability of the population in the SP3 trial. Although SP2 was recognised as an outlier as part of the nintedanib appraisal [NICE 2016], and has accordingly not been included in the networks for this submission, SP3 has been recognised as providing valuable evidence in several reviews: the initial NICE technology appraisal of pirfenidone [Intermune 2011, NICE 2013]; the nintedanib appraisal [Boehringer Ingelheim 2015, NICE 2016], and; as part of the EMA's review of the marketing authorisation application for pirfenidone [EMA 2015]. Indeed, data from SP3 are incorporated into the text of the SmPC for pirfenidone.

A39. *Section 4.10 p.146: The company states, "To mitigate the differences in definitions, we reanalysed our IPD to match BI's definition, adjusted for different base case by meta-regression, and corrected actual data based on the baseline prevalence of adverse events as an additional sensitivity analysis." Please provide further details on how this analysis was conducted.*

This sentence is incorrect, and we apologise for this mistake.

We did not "reanalyse our IPD to match BI's definition", or "[adjust] for different base case by meta-regression". We did, however, "[correct] actual data based on the baseline prevalence of adverse events as an additional sensitivity analysis" for the acute exacerbations NMA, as presented in Appendix 14 of our submission (p166-168).

In order to provide a more robust estimate of comparative safety an additional NMA was run, based on outcomes reported in Phase III studies for adverse events. Results of this NMA were also incorporated into the economic model. Running the NMA allows a more thorough exploration of how adverse event rates differ between treatments. Full details of the NMA, along with base case results and sensitivity analyses, can be found in Appendix D. A summary of key results are presented in Table 16.

Table 16: Key outcomes of the adverse event NMA

AE of interest*	Measure	Estimate** (95%CrI)		
		Pirfenidone vs placebo	Nintedanib vs placebo	Pirfenidone vs nintedanib
Diarrhoea	Odds ratio	1.39 (0.94, 2.11)	7.32 (4.82, 11.13)	0.19 (0.11, 0.35)
	Relative risks***	1.30 (0.95, 1.76)	3.34 (2.73, 4.02)	0.39 (0.28, 0.55)
Rash	Odds ratio	3.85 (2.38, 6.29)	1.29 (0.49, 3.35)	2.99 (1.03, 8.88)
	Relative risks***	3.16 (2.13, 4.61)	1.26 (0.51, 2.87)	2.51 (1.03, 6.59)
Discontinuation of treatment due to AE	Odds ratio	1.58 (1.04, 2.39)	1.52 (1.01, 2.29)	1.04 (0.58, 1.85)
	Relative risks***	1.45 (1.03, 1.99)	1.41 (1.01, 1.94)	1.03 (0.65, 1.61)
Serious cardiac AE	Odds ratio	1.36 (0.54, 3.46)	0.64 (0.17, 1.49)	2.11 (0.65, 11.34)
	Relative risks***	1.33 (0.56, 3.07)	0.66 (0.18, 1.46)	2.03 (0.67, 9.85)

* We also attempted to extract AE data on photosensitivity. However, such data were identified only for trials comparing pirfenidone vs placebo (ASCEND, CAPACITY 1 & 2, and SP3). Therefore, an NMA was not possible for this endpoint.
** Base case results (random effects model)
*** The relative risk estimates were derived from the odds ratio estimates and the placebo rate, estimated as the average rate from all the placebo arms in the analysis.

Running the NMA did not have a substantial impact on model results, but was deemed necessary in the interest of demonstrating a clear comparison between all treatments, utilising as much available data as possible. Results using the NMA can be found in the revised model base case results presented in Section B below.

Section B: Clarification on cost-effectiveness data

Revised base case model results are given in Table 17 to Table 19, following the changes made in response to the clarification questions. Changes were made for the revised model results in response to the following specific questions:

- A39: Incorporation of the AE NMA
- B8: Revision of tunnel state methodology
- B9/B10: Revision of IPF-related mortality figures used in the model
- B23: Revision of oxygen cost application in the model

Additionally, model results after re-running the PFS parametric curves using the same structure as used for OS are presented as a discrete scenario analysis in response to question B20.

The results are presented at the list prices for both pirfenidone and nintedanib. Full model results (including sensitivity analyses, disaggregated outcomes etc.) are given in Appendix E.

Table 17: ITT population – List prices

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	5.38	3.80				
PFN	[redacted]	8.67	5.67	[redacted]	3.29	1.87	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years.

Table 18: Moderate population – List prices

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)	Versus baseline (QALYs)
BSC	24,868	4.80	3.44					
NTB	65,065	6.06	4.23	40,197	1.26	0.78	51,331	51,331
PFN	[redacted]	7.67	5.14	[redacted]	1.61	0.91	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years.

Table 19: Mild population – List prices

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	7.11	4.82				
PFN	[redacted]	11.26	6.99	[redacted]	4.15	2.17	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years.

Non-randomised trial evidence

B1. Section 4.11, p.158: Please provide the actual numbers of people who entered the RECAP study from each of the trials: CAPACITY 1, CAPACITY 2 and ASCEND.

The numbers of patients who entered the RECAP study from each of the ASCEND and CAPACITY trials are outlined in Table 20 [Roche 2016a].

Table 20: Summary of Patient Counts into RECAP by Study Source – all treated patients

Study, n (%)	Pirfenidone 1197 mg/day (N=68)	Pirfenidone 2403 mg/day (N=484)	Placebo (N=506)	Total (N=1058)
CAPACITY 2	68 (100)	130 (26.9)	137 (27.1)	335 (31.7)
CAPACITY 1	NA	131 (27.1)	137 (27.1)	268 (25.3)
ASCEND	NA	223 (46.1)	232 (45.8)	455 (43.0)

Table includes data as of 30 June 2015 clinical data cut-off.
Percentages are based on number of patients in each dose group.

B2. According to the study design summarised in Table 56, patients who did not take >80% of study drug in ASCEND and CAPACITY were excluded from RECAP. Please provide the rationale for excluding these patients from RECAP.

Patients using less than 80% of drug are considered to be non-compliant (a standard cut-off being used in many trials), and for this reason were not included in RECAP. Although RECAP was an open-label extension study, the standard compliance considerations were still applied.

B3. Section 4.12, p.172-174: Please explain and justify the inclusion of the PIPF-002 trial, which includes a population and pirfenidone dosing regimen outside of the scope of this appraisal.

In this section, we are providing an overview of the safety profile of pirfenidone, so whilst the trial population in PIPF-002 is not the same as the scope of this appraisal, the safety data is included for completeness to provide as much information as is available.

Systematic literature review

B4. Section 5.1, p.190: The combined searches were run without a date limit. However, the submission states that only data published after 2010 was screened. A large number of the sources (including NHS EED and the Cost-Effectiveness Analysis registry) were not searched in the original submission. Please explain how the company ensured key data were not missed by limiting the results by date.

The ERG report produced as part of the earlier NICE appraisal of pirfenidone critically appraised the search methodology, and also noted that the NHS EED had not been included in the manufacturer's search strategy, although no additional relevant studies were identified by the ERG in their own search: "*The ERG checked the search strategy used for the cost effectiveness searches and considered them reasonably comprehensive, fit for purpose and reproducible (Section 3.1.1 of this report). An additional search of NHSEED has been run by the ERG and has not found any cost effectiveness studies for pirfenidone*" (page 53) [Cooper 2012].

On this basis, the screening of publications was limited to those made available after the cut-off date of the original searches.

B5. Appendix 17 (pp.195-201): The company does not provide a reference to any published filter that has been used; however, the utilities search filter appears to have been directly derived (with no variation) from Arber et al (2015) (http://www.yhec.co.uk/yhec-content/uploads/2015/06/Poster-374-Sensitivity-Of-A-Search-Filter.pdf). The cost-effectiveness filter appears to be a modified version of the NHS EED search filter (line 56 has been amended and lines 25-27 have been omitted). Please confirm the sources of published search filters and explain why they have been modified.

We note that NICE do not usually require the use of any specified filter, or require that filters are justified. As identified by the ERG, the HRQoL searches run to support Section 5.4 of the submission has used a filter developed by York Health Economics Consortium, which was presented at the 2015 HTAi Conference and 2015 Cochrane Colloquium [Arber 2015]. This filter was used without modification.

The ERG are also correct that an amended NHS EED filter has been used for the identification of cost-effectiveness studies. The rationale for the amendments is to allow an improvement in sensitivity:

- line 56: 'value for money' has been replaced with 'value adj2 money' which replaces a single word with the possibility of many words in addition to 'for' being interposed between value and money.

- lines 25-27: these lines are omitted, as these were part of the NHS EED processing activity. The NHS EED team use these lines to remove journals which they are hand-searching to avoid duplication of effort. As we did not want to potentially miss these publications, the lines were removed from the strategy.

Both modifications to the NHS EED filter led to an enhancement of sensitivity, and reduced the chance of missing relevant studies.

Model structure and key assumptions

B6. ***Priority question:*** *Given the company's modelling approach (partitioned-survival) OS, PFS and discontinuations are modelled independently of each other. The effect of this is that increasing the discontinuation rate (for example, by applying the stopping rule) only affects costs – it does not impact the effectiveness in terms of PFS and OS. Please comment on whether this is appropriate and provide evidence to support the assumption that OS and PFS are independent of the time on treatment.*

The ERG is correct that the model was constructed utilising the simplifying assumption that time on treatment, OS and PFS are independent of each other. This is a common practice in NICE submissions using time to event data (such as oncology submissions where disease is similar in severity and impact to IPF). To accurately quantify the relationship between time on treatment, OS and PFS, additional data would be required which are not publically available for nintedanib.

Recent studies comparing the state-transition method (i.e. modelling time on treatment, PFS and OS separately) and area-under-the-curve (AUC) partitioned survival models show that the two methods produce similar results, and that either approach may be considered appropriate to a given decision problem, depending on the available data and scope of the evaluation [Briggs 2015]. We consider our approach the most appropriate given the data available.

B7. ***Priority question****: Following pirfenidone discontinuation, the company assumes that patients receive best supportive care even if the patient still has moderate IPF (percent predicted FVC 50–80%). Nintedanib has recently been recommended by NICE for use in people with moderate disease. Please comment on whether nintedanib could be used following pirfenidone, and whether pirfenidone could be used following discontinuation of nintedanib (in patients who still have moderate IPF).*

As discussed within the appraisal for nintedanib it is possible that clinicians may sequence pirfenidone and nintedanib within the moderate population. However, in line with the approach accepted within the submission for nintedanib, the use of nintedanib following discontinuation of pirfenidone or vice versa has not been explored in our modelling approach.

This is due to the lack of clinical or safety evidence available on the impact of sequencing pirfenidone and nintedanib and uncertainty regarding the exact mechanism of action of both nintedanib and pirfenidone in the treatment of IPF. There is an ongoing safety and tolerability study which seeks to assess the impact of the addition of nintedanib to pirfenidone in patients with IPF. Results from this study are anticipated in late 2017 / early 2018. Efficacy data from sequencing studies are not anticipated to become available in the near future, with no such trials currently listed on clinicaltrials.gov.

The inclusion of treatment sequencing based on the prevailing clinical data would require a major assumption regarding the treatment effect of both therapies in patients who are not treatment naive. It is plausible that either or both of these treatments may be less efficacious when given as 2nd line treatment options, although – based on the differing mechanisms of action – it is also possible that they may be equally effective; any analysis would be pure speculative in design.

B8. Section 5.3, p.218: The company states, "the stopping rule was applied in the model using tunnel states to estimate the proportion of patients who progress within 12 months", but no further details are provided regarding the tunnel states. Please describe in more detail how the tunnel states operate within the model.

Tunnel states were included in the model in order to incorporate the stopping rule according to currently recommended practice with nintedanib. The stopping rule states: "*Treatment should be stopped if a patient's predicted FVC declines by 10% or more in any 12 month period*" [NICE 2016].

To incorporate this stopping rule into the Markov health state structure used in the model, tunnel states were used due to the requirement to track when patients enter the progressive disease state in order to implement the stopping rule.

Incorporating the tunnel states was done using the following steps:

1) OS and PFS are read from the relevant patient flow sheet for all model cycles.

2) Using the OS data, transition probabilities for survival were derived between cycles for the model time horizon using the formula:

   $= EXP(-(-LN(S_{t+1}/S_t)))$

3) Up until 12 months in the model, the proportion of patients in the progression-free health state separated by model cycle was estimated using the OS transition probabilities. This was done to approximate how long each proportion of patients have been on treatment allowing us to estimate the proportion of patients who have progressed and remain alive. An example is shown in Figure 4.

Figure 4: Incorporation of tunnel states in the economic model

Tunnel states - Esbriet

			8.71		13.53					
	OS	Trans prob	PFS	Progressed	Still on tx	cycle 1	cycle 2	cycle 3	cycle 4	cycle 5
0	100.0%		100.0%	0.00%	100.00%					
1	99.3%	99.3%	96.1%	3.16%	100.00%	3.16%				
2	98.2%	98.9%	89.7%	8.53%	100.00%	5.39%	3.14%			
3	96.9%	98.7%	82.1%	14.82%	100.00%	6.39%	5.33%	3.11%		
4	95.5%	98.5%	74.0%	21.46%	100.00%	6.83%	6.31%	5.26%	3.07%	
5	93.9%	98.4%	65.9%	28.04%	96.78%	6.90%	6.73%	6.21%	5.18%	3.02%
6	92.3%	98.2%	58.0%	34.31%	91.26%	6.72%	6.79%	6.62%	6.11%	5.09%
7	90.5%	98.1%	50.5%	40.08%	84.62%	6.37%	6.61%	6.67%	6.50%	6.00%

The transition probability of OS between cycle 0 and cycle 1 was used in the submission document to estimate the proportion of patients who progressed at cycle 1 and survived up until cycle 2. This was done using the formula: 3.16% * 99.3%, which gives 3.14%.

Following a review of the implementation of this step as part of the response to these clarification questions, an error was identified in the submitted model, as the transition probability between cycle 1 and 2 should have been used to estimate the proportion of patients who progressed at cycle 1 and survived up until cycle 2. This has been addressed in the revised model base case results.

4) For newly-progressed patients, the proportion of patients in cycle 1 of post-progression was estimated as the total number of progressed patients minus those who have previously progressed and survived.

5) Using these formulae, we obtained an estimate of the number of progressed patients based on the time of progression (to the nearest cycle). The proportion of patients who have progressed, and would still be on treatment according to the stopping rule, was derived using the following formula:

= (Patients in cycles 1-4 of post-progression + Progression-free patients)/ Alive patients

This proportion was then applied to the number of patients still on treatment in the relevant patient flow sheets, to obtain an estimate of those who would still be on treatment in accordance to the stopping rule.

Mortality rates

B9. Section 5.3, p.215: In the base case, the company assumes that 57.89% of deaths are attributed to IPF for patients receiving pirfenidone. Please clarify the data source used to inform this estimate, including information on the trials and time points used.

The proportion of deaths related to IPF for pirfenidone and BSC were captured from pooled analysis of ASCEND/CAPACITY data. The value above (57.89%) is calculated as the number of IPF deaths recorded in patients receiving pirfenidone, divided by total deaths across CAPACITY and ASCEND trials (i.e. 22/38).

The equivalent value for BSC patients is 72.22% (39/54).

The model has since been updated to use the data extracted from the NMA (see answer to question B10, Table 21).

B10. Section 5.3, p.215 (table 74): Please explain how the proportions of deaths that are IPF-related for best supportive care and nintedanib were calculated using the data from the NMA. The ERG cannot see how the hazard ratios from the NMA for OS and IPF-related survival can be used to estimate the proportion of deaths which are IPF-related.

The NMA was used to determine the proportion of deaths for patients on each treatment attributable directly to IPF. The model utilises odds ratios to determine the equivalent proportion of IPF-related deaths for patients on nintedanib and BSC using the NMA.

The use of odds ratios to determine the equivalent risk of mortality-related outcomes has been used in other diseases areas in published literature.

As an alternative, the figures in Table 21 have been taken from Appendix 11 of the submission document. These figures utilise the consistent week 52 time point, and have been used to provide an alternative estimate of IPF-related mortality across the different treatments arms, without using the odds ratios presented in the previous model.

Table 21: Revised IPF-related mortality figures

Intervention	Time point	n of IPF-related deaths	N of all-cause deaths
Pirfenidone	52 weeks	17	32
Placebo		35	50
Nintedanib		26	42

Survival modelling

B11. **Priority question**: In Table 70 (p.209) it is stated that individual patient data (IPD) *from CAPACITY, ASCEND and RECAP were used to fit the OS curve applied in the model. Please clarify which patients from RECAP were included. In particular:*

a. Were patients enrolled in RECAP from the placebo arms of CAPACITY included in the IPD?

b. Were patients enrolled in RECAP from the non-licensed dose arm of CAPACITY 2 included in the IPD?

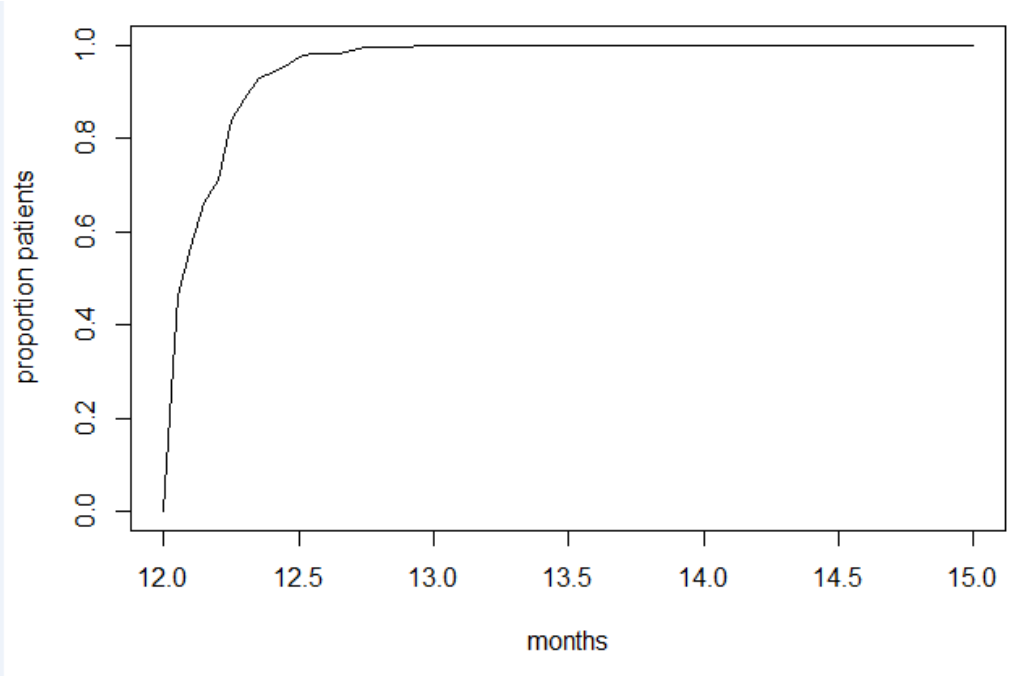
c. Were all patients enrolled in ASCEND censored at 52 weeks in this analysis?

Part a: No – only patients from the licensed dose arm of the two CAPACITY trials were used to fit the OS curve

Part b: No – only patients from the licensed dose arm of the two CAPACITY trials were used to fit the OS curve

Part c: Figure 5 shows the cumulative time of censoring for patients from the ASCEND study. All patients were censored between 12 and 13 months: as Figure 5 shows, the majority of this censoring occurred within the first two weeks.

Figure 5: Cumulative censoring of patients in the ASCEND trial



B12. ***Priority question**: Justification of the assumption of proportional hazards is provided in Appendix 20, however it is not clear which data were used for the test of interaction.*

a. *Please confirm which data were used for the test of interaction and for the test for the proportionality assumption for both OS and PFS.*

b. *Please clarify whether the analysis is based on data from 52 or 72 weeks.*

c. *If the analyses presented are based on data pooled across multiple trials, please repeat the analysis for each separate trial (using 72 week data, where available).*

Part a: The test for interaction between time and treatment presented in Appendix 20, along with the graphs of scaled Schoenfeld residuals, are based on pooled data from ASCEND, CAPACITY 1 and CAPACITY 2 studies (collected until the end of the study for the OS endpoint). For PFS analyses, these are based on data up to 72 weeks.

In response to the ERG's request in B12c, we have provided test results for CAPACITY 1, CAPACITY 2, ASCEND and updated test results for pooled trial data (see below). Consistent with the results provided in Appendix 20 of the submission, the p-values for all tests of interaction between treatment and time is not significant, and Schoenfeld residual plots for both OS and PFS indicate there is no substantial deviation from proportional hazards (across all trials and pooled data).

Part b: OS data are assessed using data collected until the end of study period; PFS data are assessed through to 72 weeks, as PFS was not recorded after this point.

Part c: Tests for the interaction between time and treatment, along with plots of scaled Schoenfeld residuals have been re-presented for all trials, in addition to the pooled data, for OS and PFS. Of note, the figures of scaled Schoenfeld residuals provided in Appendix 20 of the submission presented ranks of time (as opposed to time) on the x-axis. For consistency with the test, scaled Schoenfeld residuals over (linear) time have been presented. Full results of these analyses are provided in Appendix F. Consistent with the results provided in Appendix 20 of the submission, the p-values for all tests of interaction between treatment and time are not significant, and Schoenfeld residual plots for both OS and PFS indicate there is no substantial deviation from proportional hazards.

Calculation of costs

B13. ***Priority question****: In Table 92 of the company submission, the mean actual pirfenidone dose received in the ASCEND and CAPACITY trials is reported. The company states that pirfenidone needs to be titrated in the first 2 weeks. Please provide the mean (and standard error) for the actual dose received:*

a. estimated for the first 3 months of treatment only (to represent the first cycle of the model)

b. excluding the first 3 months of treatment (to represent subsequent cycles in the model).

Part a: As discussed with the NICE Project Manager and Associate Director over email on 25-26 February, these analyses will be supplied as soon as possible, and by 18 March.

Part b: As discussed with the NICE Project Manager and Associate Director over email on 25-26 February, these analyses will be supplied as soon as possible, and by 18 March.

B14. End of life costs are included for IPF-related deaths but not for deaths from other causes. It is stated on p.246, "This is because costs associated with IPF are greatly increased in the last year of life due to increased resource use, home care and length of stay in hospital". Please provide evidence which demonstrates that the costs in the last year of life for IPF-related deaths are higher than the costs in the last year of life for deaths from other causes, to support this assumption.

The purpose of the submission was to assess the cost and clinical implications associated with pirfenidone for the treatment of IPF compared with current care. As a result of this, only costs borne by the condition have been considered in the analysis, with other costs deemed out of the scope of this analysis and unrelated to the decision problem. Consequently for end of life care, costs attributable to death from other causes than IPF are not included.

B15. The company uses data on hospitalisations from the trials to inform the costs associated with acute exacerbations (Table 99) but uses estimates from the published literature combined with data from the NMA to inform the disutility associated with acute exacerbation. Please explain and justify why the same data for incidence of exacerbation is not used to inform the costs and disutilities from exacerbations.

In line with the methodology used within the nintedanib submission, the incidence of acute exacerbations was derived using the NMA, which considered all available trial evidence for all comparators in estimating the occurrence of acute exacerbation. These acute exacerbation rates were used as they facilitated comparison across different treatment groups. It should be noted, as discussed within the NMA, that the ability to accurately compare acute exacerbation rates across trials is, however, limited due to differences in the definitions used in the relevant trials.

The cost of acute exacerbations in the nintedanib NICE company submission considered a range of medical resource use items, including hospitalisations, emergency room visits, general practitioner visits and specialist visits [Boehringer Ingelheim 2015]. The majority of the cost was accounted for by the hospitalisation cost (98.8% of the total cost for acute exacerbation). Therefore, the cost of acute exacerbation in our modelling approach was assumed to be approximately equal to the cost of hospitalisation, in consideration of the small proportion of other costs applied for acute exacerbation.

These incidence rates informed the HRQL of patients experiencing an acute exacerbation. As per the health state utility values, disutilities associated with acute exacerbation were not available in many of the studies. Therefore, the disutility associated with acute exacerbation was taken from an external source. For consistency with the previous nintedanib NICE company submission, the same disutility for acute exacerbation was assumed.

It is expected that the vast majority of acute exacerbations recorded within the clinical trials available would result in hospitalisation, however, not all relevant hospitalisations may have been recorded within the clinical trial as acute exacerbation. The use of hospitalisation data, rather than using data only for acute exacerbations as derived from the clinical trials to inform costs allowed us to include the costs for all relevant hospitalisations and to accurately account for the differences in length of stay observed between pirfenidone and nintedanib – it can clearly be seen that in addition to reducing the rate of hospitalisation, pirfenidone reduces the severity of hospitalisation (i.e. length of stay) when hospitalisation does occur. The cost of acute exacerbations was therefore included in the model within the cost of hospitalisation for respiratory failure. The cost of a bed day was taken from the NHS reference costs 2014-15 (as per the NICE reference case).

For example, assuming the largest occurrence of acute exacerbation in the model (for BSC patients) of 1.46% per cycle and the total cost for a hospitalisation of £4,339.37 gives an estimated cost for acute exacerbation-related hospitalisation of £63.23 per model cycle, as opposed to the cost of overall hospitalisation per model cycle of £226.34 for BSC patients.

A broader approach was not taken to estimate utilities in the same way as for hospitalisation as it was not possible to find disutilities specifically relating to hospitalisation for IPF (rather than acute exacerbation). This is conservative as inclusion of disutility for a larger proportion of patients receiving BSC would have made the cost-effectiveness case for pirfenidone more favourable. Additionally, our methodology is consistent with the previous nintedanib NICE company submission which also did not incorporate a disutility for hospitalisation for IPF [Boehringer Ingelheim 2015].

B16. Section 5.5, p.239: The company states that resource use was based upon advice on UK clinical practice in IPF from a panel of UK clinicians, stratified by treatment type and progression status. Please provide any materials used to elicit this information from the panel, and the analysis of these data.

One-to-one telephone interviews were conducted with the panel of UK clinical experts. Content of the earlier NICE manufacturer submission was discussed, along with how the approach employed to assess resource use in the earlier submission matched current clinical practice in IPF. Discussions accounted for the revised descriptions of the NHS Reference Cost list for 2014-15 compared to earlier years (e.g. revision of 'simple lung exercise function test' to 'field exercise test').

Utility estimates

B17. Please explain how the Freemantle (2015) and Starkie (2011) mapping studies, which were used to map from the St George's Respiratory Questionnaire (SGRQ) to EQ-5D, were identified. In particular, was a systematic search conducted to identify all relevant mapping algorithms? How were these 2 studies selected from those identified in the search?

No additional mapping studies were identified as part of the systematic search for utility information conducted within this submission (see Section 5.4 of the submission for details for this search).

The Starkie [2011] mapping algorithm was used in the previous pirfenidone NICE STA submission, and was included in this submission for completeness. This algorithm was identified as part of a literature search conducted within the previous submission to identify studies which mapped either the SGRQ or WHO-QOL onto the EQ-5D [InterMune 2011].

This mapping study was used in the previous pirfenidone NICE STA submission as "No EQ-5D data was available, and the mapping study was based on a patient population of COPD patients. However, a mapping study in an orphan disease such as IPF would be very rare, and the mapping of COPD patients is likely the next best estimate." [InterMune 2011].

Due to the lack of alternative evidence, InterMune utilised SGRQ and EQ-5D from a double-blind multicentre study conducted in England and Wales in IPF patients in order to generate a mapping algorithm between the two aforementioned instruments. This mapping was also used in the previous manufacturer submission of pirfenidone to NICE (unpublished at the time) and was accepted in the base case estimate of cost-effectiveness; this has since been published by Freemantle [2015].

As discussed in this submission document, the study by Starkie [2011] was conducted in patients with COPD (as opposed to IPF patients), and was associated with larger RMSE (0.1723 versus 0.1391) when compared to the mapping algorithm by Freemantle [2015] therefore the Freemantle algorithm was selected as the preferred algorithm to map utilities for this submission.

The range of sources used within the model produced similar estimates of health state utilities, and therefore the choice of mapping algorithm is not extensively influential on model results.

B18. *In the base case, the company estimates the mean SGRQ for patients in the progression-free and progression health states using a generalised estimating equation (GEE) model; it estimates the mean EQ-5D score based on a linear mapping algorithm estimated in IPF patients (Freemantle, 2015). However, the mapping algorithm used in the base case (EQ-5D = 1.3246 – 0.01276*SGRQ) predicts an EQ-5D value over 1 when the SGRQ becomes below 25.*

a. *Please provide the proportion of patients with a SGRQ score below 25 at baseline and last follow-up in the CAPACITY trials.*

b. *Please provide more information on how the GEE model was estimated, in particular please comment on whether each patient contributed data at multiple time points. If so, how those time points were selected?*

c. *Finally, please provide the following analyses:*

i. *Apply the unconstrained mapping algorithm (EQ-5D = 1.3246 – 0.01276*SGRQ) to the individual patient level data from the CAPACITY trials, then estimate the mean EQ-5D score using a GEE model (similar to the approach used for SGRQ in the original submission).*

ii. *Apply the mapping algorithm (EQ-5D = 1.3246 – 0.01276*SGRQ) to the individual patient level data from the CAPACITY trials and truncate the predicted EQ-5D to a maximum of 1, then estimate the mean EQ-5D score using a GEE model (similar approach used for SGRQ in the original submission).*

Part a: The proportion of patients with a SGRQ score below 25 was 25.5% at baseline and 19.5% based upon the last observation.

Part b: The GEE model was estimated using an exchangeable correlation matrix. The total St George Respiratory Questionnaire score was regressed on progression status. All visits were included in the model with patients contributing between 1 and 11 observations each. Observations occurring after progression (using RISE definition including 6MWT and FVC) were classified as being progressed for this model.

The mapping algorithm was performed using the proc glimmix in SAS. The data analysed were from a double-blind multicentre study conducted in England and Wales of which 181 IPF patients received either co-trimoxazole or placebo. In total, 202 pairs of data were collected recording both SGRQ and the EQ-5D-3L. Hence some patients contributed with more than one data point to the regression. All pairs of data for all time points from the study were used for the mapping.

Part c: Table 22 below displays the utility estimates based upon applying the mapping algorithm between SGRQ and EQ-5D.

Table 22: Estimated utility values from mapping algorithm

Model state	Previous submission	Not truncated	Truncated
Progression free	0.8485	0.8485	0.8185
Progressed	0.7835	0.7835	0.7597

Subgroup analyses

B19. ***Priority question**: In addition to the intention to treat population, the company reports results for two subgroups; mild (percent predicted FVC ≥80%) and moderate (percent predicted FVC 50–80%). Whilst the ERG understands these analyses are post-hoc (not pre-specified), could the company provide the following:*

a. *Kaplan–Meier curves for OS for best supportive care from the pooled ASCEND and CAPACITY trials. Provide these separately for patients with mild and moderate IPF (as defined in the model) at baseline.*

b. *Evidence that the proportional hazard assumption holds for patients with moderate and mild IPF separately (at 72 weeks) for PFS and OS.*

Part a: Kaplan-Meier curves for OS in the placebo arms of the pooled ASCEND and CAPACITY trials are presented in Figure 6. Separate curves are presented by baseline FVC % predicted (50-80% and 80%).

Figure 6: Kaplan-Meier curves for OS in placebo patients in pooled trial population, by baseline FVC % predicted (academic in confidence)

Part b: Tests for the interaction between time and treatment, along with plots of scaled Schoenfeld residuals have been generated from the pooled trials (ASCEND, CAPACITY 1, CAPACITY 2) at 72 weeks for PFS and OS. Both outcomes are assessed for baseline FVC 50-80% predicted, and FVC ≥80% predicted. Full results of these analyses are provided in Appendix G. The p-values for all tests of interaction between treatment and time are not significant, and Schoenfeld residual plots for both OS and PFS indicate there is no substantial deviation from proportional hazards in either disease severity subgroup.

B20. *For the subgroup analyses (mild and moderate IPF), the company fit a series of parametric curves to the OS and time to discontinuation for the pirfenidone arms using percent predicted FVC at baseline as covariates (FVC<50%, FVC 50–80% and FVC≥80%). However, for PFS, the company appears to use a different approach, fitting parametric curves to the 2 subgroups separately. Please justify the use of different approaches for OS and PFS?*

The fitting of parametric functions i.e. PFS and OS were done at different time points, and this was the reason for the different approaches taken. The parametric fit for PFS was re-run in order to be consistent with the OS approach. The impact of this was explored in the model, where the change was found to have minimal impact on results (Table 23). As a result of this limited impact, this change has not been included in the updated base-case results or scenario analyses (Appendix E).

Table 23: Impact on ICER of using a common approach to fitting parametric curves for PFS & OS

Population	ICERs prior to adjustment		ICERs after adjustment	
	vs. BSC	vs.nintedanib	vs. BSC	vs.nintedanib
ITT	[redacted]		[redacted]	
Mild	[redacted]		[redacted]	
Moderate	[redacted]	£28,893	[redacted]	£29,078

Executable model

B21. ***Priority question**: In the economic model, in sheets "KM TTOT", "KM PFS" and "KM OS" the number of patients at risk in the Kaplan–Meier plots for pirfenidone appear to be different depending on the outcome (PFS n=618, OS n=623 and TTOT n=490). Could the company:*

a. *Clarify the source of data used for these Kaplan–Meier plots (ASCEND/CAPACITY/RECAP?)*

b. *Confirm whether only patients receiving 2,403 mg daily for pirfenidone are included for each of the outcomes.*

c. *Explain the reasons for the different numbers of patients at risk for these outcomes.*

Part a: Sources have been added to all the KM sheets in the revised model ("KM TTOT", "KM PFS" and "KM OS"). The previous model had placeholder data for the "KM TTOT", this has been corrected and updated with the correct data. Data are from ASCEND, CAPACITY 1 and CAPACITY 2 for each of the three outcomes. RECAP data are only available for TTOT and OS (Table 24).

Table 24: Source of data informing KM curves for OS, PFS and TTOT

KM data	CAPACITY 1 & 2 (13 Jan 2009)*	ASCEND (14 Feb 2014)*	RECAP (30 June 2015)*
OS			
Pirfenidone – all	√	√	√
Pirfenidone - FVC >=50% & <80%	√	√	√
Pirfenidone - FVC >=80%	√	√	√
BSC	√	√	
PFS			
Pirfenidone – all	√	√	
Pirfenidone - FVC >=50% & <80%	√	√	
Pirfenidone - FVC >=80%	√	√	
BSC	√	√	
TTOT			
Pirfenidone - all	√	√	√
Pirfenidone - FVC >=50% & <80%	√	√	√
Pirfenidone - FVC >=80%	√	√	√

* Date of data-cut

Part b: It is correct that we have only used data from patients receiving the 2403 mg daily dose of pirfenidone.

Part c: The OS and TTOT outcomes are based upon all patients who received the 2403 mg dose (n=623); the corresponding number for placebo is 624. For PFS this number is 618 in both groups: the difference is driven by the fact that some patients did not have the assessment required for the 'progressed' definition.

B22. In the economic model, in sheet "KM TTOT", please:

a. Clarify whether the time to discontinuation (column B-I) is for the pooled intention to treat population from the ASCEND/CAPACITY/RECAP trial in patients receiving 2,403 mg daily for pirfenidone.

b. Provide correct label for the subsequent Kaplan–Meier curves (columns K-BK)

Part a: The correct KM data have now also been added to the "KM TTOT" sheet, before this was only available on the "KM for plotting" sheet. It is correct that time to treatment discontinuation is the pooled estimate of all patients 2403 mg, from CAPACITY and ASCEND.

Part b: These were placeholders and have been removed.

B23. The submission states that costs for supplemental oxygen are applied to patients with a percent predicted FVC <80%. There appears to be some data in the model to facilitate this calculation but the costs applied beyond the first cycle always remain at zero due to a sumproduct calculation (in G63:H63 of Sheet named 'costs') that refers to a set of blank cells. Please clarify why costs for supplemental oxygen are not implemented in the model.

This is an error. The calculation previously referred to a set of cells which ensured the model only considered patients with a percent predicted FVC <80%.

Figure 7: Costs of supplemental oxygen previously used in economic model

		Subsequent cycles	
	First cycle	PF	PP
Liver function test*	Please see liver function test section below		
Gas transfer	£202.08	£151.56	£151.56
Lung volume study	£170.54	£0.00	£0.00
Full pulmonary (covers spirometry)	£165.85	£124.39	£124.39
Field exercise test	£177.13	£88.57	£88.57
Oxygen	£0.00	£0.00	£0.00
Healthcare professional visit	£248.17	£142.08	£154.48
GP visit	£0.00	£6.62	£6.45

Figure 8: Revised costs of supplemental oxygen used in economic model

		Subsequent cycles	
	First cycle	PF	PP
Liver function test*	Please see liver function test section below		
Gas transfer	£202.08	£151.56	£151.56
Lung volume study	£170.54	£0.00	£0.00
Full pulmonary (covers spirometry)	£165.85	£124.39	£124.39
Field exercise test	£177.13	£88.57	£88.57
Oxygen	£0.00	£154.39	£188.68
Healthcare professional visit	£248.17	£142.08	£154.48
GP visit	£0.00	£6.62	£6.45

As the larger cost for oxygen is incurred in the post-progression model health state (since FVC percent predicted is generally lower for progressed patients), the base case model ICER increases slightly after amending this modelling error as shown in the revised model base case results.

Section C: Textual clarifications and additional points

C1. Section 6.1, p.283: The company states, "An estimated 30.5% of patients with prevalent IPF have mild IPF, 54.0% have moderate IPF, and 15.5% have severe IPF [Roche 2016a]." Please provide more information regarding the source of data for these proportions and discuss whether they are representative of the distribution of severity within the population of England.

These data are based on interviews using quantitative questionnaire, conducted among a group of 58 Respiratory/Chest Physicians across the UK, who treat patients with IPF. Interviews were conducted in two waves: during February and April-May 2015. Physicians were identified on the basis of being with a Consultant or senior grade physician, with 3-20 years' experience in clinical practice, and spending >70% of their time in clinical practice, seeing a minimum of 3 patients with IPF per month. It was also a requirement for physicians to be personally responsible for prescribing drug treatment for patients with IPF for inclusion in the interviews, and they were not permitted to be affiliated with any pharmaceutical company.

C2. The baseline characteristics for 6MWD and SGRQ in Table 67 do not look realistic. Please confirm if these figures have been mislabelled

These figures have been mislabelled. The correct baseline characteristics should be as follows: SGRQ: 37; 6MWD: 418.

C3. Different ICERs for pirfenidone compared with BSC (base case analysis, list prices) are reported in sections 5.7 (£38,779) and 5.11 (£38,644). Please confirm that £38,644 is a typographical error.

£38,644 is a typographical error.

C4. The penultimate bullet point in the summary box on page 48 of the submission appears to be incomplete. Please provide the complete sentence.

This is a transcription error and the final word should have been deleted. The bullet should read: "*In a subgroup analysis from the pooled data from CAPACITY 1 & 2, those patients diagnosed more than a year before randomisation experienced a significantly greater treatment effect*".

C5. The company submission refers to a "decline in FVC", when it would be more appropriate to refer to a "decline in percent predicted FVC", for example on pages 60 and 89 (which report the primary outcome of the ASCEND trial). Please confirm or clarify.

Agree: "decline in FVC" has occasionally been stated in lieu of the preferred terminology "decline in percent predicted FVC".

C6. Please clarify the statement in the summary box on page 89, because this analysis does not appear to be described elsewhere in the submission: "When considering patients with earlier (FVC ≥ 80% predicted) vs. later (FVC <80% predicted) disease s…. There was also a numerically lower risk of FVC decline ≥10% or death in those

with FVC ≥ 80% predicted, although this was not statistically significant.(p= 0.2403)". Does the p value for the lower risk of FVC decline refer to the comparison between patients with earlier disease versus later disease, regardless of treatment received? That is, is the analysis unrelated to the effect of pirfenidone on FVC decline?

This is a typographical error in the submission: the p=0.2403 referred to the between group (early [FVC ≥80% predicted] vs. later [FVC <80% predicted]) difference for patients treated with placebo, and should not have been included in the summary.

The statement is based on a post-hoc analysis of pooled placebo population from ASCEND and CAPACITY trials [Albera 2015]. Analyses were conducted by stratifying placebo patients by baseline disease severity.

The analysis for pirfenidone treatment effect was presented as Figure 17, on p115 of the submission.

References

Albera 2015. Albera C. Pirfenidone Is Efficacious in Patients With Idiopathic Pulmonary Fibrosis (IPF) and Mild Restrictive Disease: Benefit of Early Intervention. Oral presentation presented at ATS 2015 annual congress; 15-20 May, Denver, USA

Akagi 2009. Akagi T, Matsumoto T, Harada T et al. Coexistent emphysema delays the decrease of vital capacity in idiopathic pulmonary fibrosis. Respiratory Medicine, 2009. 103, 1209-1215

Arber 2015. Arber M., Garcia S., Veale T. et al. Sensitivity of a search filter designed to identify studies reporting health state utility values. Poster presented at HTAi conference and the Cochrane Colloquium 2015. http://www.yhec.co.uk/yhec-content/uploads/2015/06/Poster-374-Sensitivity-Of-A-Search-Filter.pdf

Boehringer Ingelheim 2015. Boehringer Ingelheim Ltd. Single technology appraisal: Nintedanib for the treatment of adults with idiopathic pulmonary fibrosis: Company evidence submission. Bracknell: Boehringer Ingelheim Ltd; 2015.

Briggs 2015. Briggs A, Baker T, Gilloteau I, et al. (2015) Partitioned Survival Versus State Transition Modeling in Oncology: a Case Study with Nivolumab in Advanced Melanoma. Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research. 18(7): A338.

Cooper 2012. Cooper K, Mendes D, Picot, J, Loveman E. Pirfenidone for the treatment of idiopathic pulmonary fibrosis: A Single Technology Appraisal. SHTAC, 2012.

Costabel 2016 (Academic in Confidence). Costabel U, Albera C, Kirchgassler KU et al. Analysis of patients with idiopathic pulmonary fibrosis (IPF) with percent predicted forced vital capacity (FVC)<50% treated with pirfenidone (PFD) in RECAP. To be presented at ERS 2016 [AiC]

du Bois 2011. du Bois, R. M., Weycker, D., Albera, C. et al. Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference (a). Am.J.Respir.Crit Care Med. 2011c; 183(9): 1231-1237.

du Bois 2014. du Bois RM, Albera C, Bradford WZ, et al. Idiopathic pulmonary fibrosis: present understanding and future options. Eur Respir J. 2014 May;43(5):1421-9.

EMA 2015b. European Medicines Agency. EMA/PRAC/799709/2015. Pharmacovigilance risk assessment committee minutes. 3 Dec 2015.

EMC 2015. Esbriet Summary of Product Characteristics. Available on: www.medicines.org.uk [date of revision of text: 08 Sept 2015]

FDA 2010. U.S. Food and Drug Administration. Pirfenidone Capsules [NDA 22-535]. Sponsor's Briefing Document for the FDA Pulmonary-Allergy Drugs Advisory Committee Meeting Scheduled for 09 March 2010. Available on: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/Pulmonary-AllergyDrugsAdvisoryCommittee/UCM203083.pdf

Flaherty 2003. Flaherty, K. R., Mumford, J. A., Murray, S. et al. Prognostic implications of physiologic and radiographic changes in idiopathic interstitial pneumonia. Am.J.Respir.Crit Care Med. 2003; 168(5): 543-548.

Freemantle 2015. Freemantle N, Wilson A, Fisher M. Mapping The St George's Respiratory Questionnaire To The Euroqol 5 Dimensions: A Study In Patients With Idiopathic Pulmonary Fibrosis. Value in Health. 2015;18(7):A503.

Higgins 2008. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions. Chichester: Wiley-Blackwell; 2008.

InterMune 2006a. InterMune. CAPACITY 1 PIPF-006 study protocol. Date of original protocol: 3 February 2006

InterMune 2006b. InterMune. CAPACITY 2 PIPF-004 study protocol. Date of original protocol: 27 Jan 2006

InterMune 2009a. InterMune. CAPACITY 2 PIPF-004 Final Clinical Study Report. Date of final report: 26 June 2009

InterMune 2009b . InterMune. CAPACITY 1 PIPF-006 Final Clinical Study Report. Date of final report: 30 June 2009

InterMune 2011. InterMune. Single technology appraisal: Pirfenidone for the treatment of idiopathic pulmonary fibrosis: Manufacturer's submission. Brisbane, CA: InterMune; 2011

InterMune 2012. InterMune. ASCEND PIPF-016 study protocol. Date of protocol: 25 April 2012

InterMune 2014 . InterMune. ASCEND PIPF-016 Final Clinical Study Report. Date of final report: 22 May 2014.

King 2005. King, T. E., Safrin, S., Starko, K. M. et al. Analyses of efficacy endpoints in a controlled trial of interferon-gamma1b for IPF. Chest. 2005; 127(1): 171-177.

King 2014. King TE, Bradford WZ, Castro-Bernardini S. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. N Engl J Med 2014; 370: 2083-2092

Martinez 2014. The Idiopathic Pulmonary Fibrosis Clinical Research Network. Randomised trial of acetylcysteine in idiopathic pulmonary fibrosis. N Engl J Med 2014; 370: 2093-2101

NICE 2012. Single technology appraisal (STA). Specification for manufacturer/sponsor submission of evidence. London: NICE

NICE 2013. National Institute for Health and Care Excellence. Pirfenidone for treating idiopathic pulmonary fibrosis [TA282]. Published: 24 April 2013

NICE 2015. National Institute for Health and Care Excellence. Nintedanib for treating idiopathic pulmonary fibrosis: Appraisal consultation document Committee Papers: Company's response to NICE's resquest for clarification. 2015

NICE 2016. National Institute for Health and Care Excellence. Nintedanib for treating idiopathic pulmonary fibrosis [TA379]

Noble 2011. Noble PW, Albera C, Bradford WZ et al. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. Lancet 2011; 377: 1760–69

Noble 2014. Noble PW, Albera C, Bradford WZ et al. Analysis of Pooled Data from Three Phase 3, Multinational, Randomized, Double-blind, Placebo Controlled Trials Evaluating Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis. Presentation to ATS 2014.

Raghu 2011. Raghu G, Collard HR, Egan JJ, et al. An Official ATS/ERS/JRS/ALAT Statement: Idiopathic Pulmonary Fibrosis: Evidence-based Guidelines for Diagnosis and Management. Issued 2011 (link: http://www.ncbi.nlm.nih.gov/pubmed/21471066)

Raghu 2012. The Idiopathic Pulmonary Fibrosis Clinical Research Network. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. N Engl J Med 2012; 366: 1968-1977

Raghu 2012 Suppl. The Idiopathic Pulmonary Fibrosis Clinical Research Network. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. N Engl J Med 2012; 366: 1968-1977 [Supplementary Appendix]

Richeldi 2011. Richeldi L, Costabel U, Selman M et al. Efficacy of a Tyrosine Kinase Inhibitor in Idiopathic Pulmonary Fibrosis. N Engl J Med 2011; 365: 1079-1087

Richeldi 2011 Suppl. Richeldi L, Costabel U, Selman M et al. Efficacy of a Tyrosine Kinase Inhibitor in Idiopathic Pulmonary Fibrosis. N Engl J Med 2011; 365: 1079-1087 [Supplementary Appendix]

Richeldi 2014. Richeldi L, du Bois R, Raghu G et al. Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis. N Engl J Med 2014; 370: 2071-82

Richeldi 2014 Suppl. Richeldi L, du Bois R, Raghu G et al. Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis. N Engl J Med 2014; 370: 2071-82 [Supplementary Appendix]

Roche 2016. Data on file

Shinogi 2001. Shinogi. SP2 study protocol. Date of preparation: 9 October 2001

Shinogi 2006. Shinogi. SP3 study protocol. Date of preparation: 16 October 2006

Spagnolo 2010. Spagnolo, P., et al. Non-steroid agents for idiopathic pulmonary fibrosis. 2010, The Cochrane Library.

Starkie 2011. Starkie HJ, Briggs AH, Chambers MG, et al. Predicting EQ-5D Values Using the SGRQ. Value in Health Volume 14, Issue 2, March–April 2011, Pages 354–360 doi:10.1016/j.jval.2010.09.011

Swigris 2010. Swigris JJ, Brown KK, Behr J, du Bois RM, King TE, Raghu G, et al. The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. Respiratory Medicine. 2010;104(2):296-304.

Swigris 2012. Swigris JJ, Han M, Vij R et al. The UCSD Shortness of Breath Questionnaire has Longitudinal Construct Validity in Idiopathic Pulmonary Fibrosis. Respir Med 2012; 106(10): 1447–1455

Taniguchi 2010. Taniguchi H, Ebina M, Kondoh Y et al. Pirfenidone in idiopathic pulmonary fibrosis. Eur Respir J 2010; 35: 821–829

Turner 2012. Turner RM, Davey J, Clarke MJ, et al. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews, Int. J. Epidemiol. (2012), doi: 10.1093/ije/dys041, First published online: March 29, 2012

.

Appendix A: Searches of clinical trial registers

1. Searches

We searched clinical trial registers as described in Table 25.

Table 25: Details of the trial register searches

Register	URL	Search date	Search terms	Number of records downloaded
ClinicalTrials.gov	https://clinicaltrials.gov/	24/2/16	pirfenidone OR esbriet OR pirespa OR pirfenex OR 53179-13-8 OR "53179138" OR "amr 69" OR amr69 OR deskar	42
ICTRP	http://www.who.int/ictrp/en/	24/2/16	pirfenidone OR esbriet OR pirespa OR pirfenex OR 53179-13-8 OR 53179138 OR amr 69 OR amr69 OR deskar	89 records for 58 trials
PharmNet Bund	https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html	24/2/16	Pirfenidone in title Esbriet in title Pirespa in title Pirfenex in title 53179-13-8 in title Amr 69 in title Amr69 in title Deskar in title	6 records

2. The record selection process

Of the 106 records identified, 44 were duplicates and were not reviewed further and 32 were rejected as being:

- Ineligible condition (n=28);
- Ineligible intervention (n=2)
- Studies in healthy people (n=2)

30 records were considered relevant and were categorised as follows:

- Ongoing trials (n=28) – some of these are for combined therapies
- Completed studies (n=11) –some of these are for combined therapies
- Terminated early (n=1)

For the completed studies, all of the trials with publications were identified in the searches. Some completed studies have not yet published. The studies are listed in Table 26 to Table 28.

Table 26: Completed trials (Pirfenidone)

Title	Trial id	Intervention	Link to full record	Publication	Status in the SR
Clinical Efficacy and Safety of Pirfenidone in Patients With Idiopathic Pulmonary Fibrosis (IPF) (Phase II) trial	NCT02136992	A Multicenter, Randomized, Double-blind, Placebo-controlled Trial for Clinical Efficacy and Safety of Pirfenidone in Patients With Idiopathic Pulmonary Fibrosis	https://clinicaltrials.gov/show/NCT02136992	Study complete but no publications	No study report identified
Efficacy and Safety of Pirfenidone in Patients With Idiopathic Pulmonary Fibrosis (IPF) (ASCEND)	NCT01366209	PIPF-016 (ASCEND) is a Randomized, Double-Blind, Placebo Controlled, Phase 3 Study of the Efficacy and Safety of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis. The study objectives are to confirm the treatment effect of pirfenidone compared with placebo on change in percent predicted forced vital capacity (%FVC) in patients with idiopathic pulmonary fibrosis (IPF), and to confirm the safety of treatment with pirfenidone compared	https://clinicaltrials.gov/show/NCT01366209	ederer DJ, Bradford WZ, Fagan EA, Glaspole I, Glassberg MK, Glasscock KF, Kardatzke D, King TE Jr, Lancaster LH, Nathan SD, Pereira CA, Sahn SA, Swigris JJ, Noble PW. Sensitivity Analyses of the Change in FVC in a Phase 3 Trial of Pirfenidone for Idiopathic Pulmonary Fibrosis. Chest. 2015 Jul;148(1):196-201. doi: 10.1378/chest.14-2817. King TE Jr, Bradford WZ, Castro-Bernardini S, Fagan EA, Glaspole I, Glassberg MK, Gorina E, Hopkins PM, Kardatzke D, Lancaster	Included in the SR

Title	Trial id	Intervention	Link to full record	Publication	Status in the SR
		with placebo in patients with IPF.		L, Lederer DJ, Nathan SD, Pereira CA, Sahn SA, Sussman R, Swigris JJ, Noble PW; ASCEND Study Group. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. N Engl J Med. 2014 May 29;370(22):2083-92. doi: 10.1056/NEJMoa1402582. Epub 2014 May 18. Erratum in: N Engl J Med. 2014 Sep 18;371(12):1172.	
Safety Study of Oral Pirfenidone in Patients With Pulmonary Fibrosis/Idiopathic Pulmonary Fibrosis	NCT00080223	An Open-Label, Phase 2 Study	https://clinicaltrials.gov/show/NCT00080223	No publications identified	No publications identified
Safety and Efficacy Study of Pirfenidone to Treat Idiopathic Pulmonary Fibrosis(IPF)	NCT01504334	A Multicenter, Randomized, Double-blind, Placebo-controlled Trial	https://clinicaltrials.gov/show/NCT01504334	Huang H, Dai HP, Kang J, Chen BY, Sun TY, Xu ZJ. Double-Blind Randomized Trial of Pirfenidone in Chinese Idiopathic Pulmonary Fibrosis Patients. Medicine (Baltimore). 2015 Oct;94(42):e1600. doi: 10.1097/MD.0000000000	Identified in our searches but it is Pirfenidone+NAC

Title	Trial id	Intervention	Link to full record	Publication	Status in the SR
				0001600.	
Three-Arm Study of the Safety and Efficacy of Pirfenidone in Patients With Idiopathic Pulmonary Fibrosis (CAPACITY 2 trial)	NCT00287716 EUDRACT 2006-000252-41 PIPF-004	The objectives of this study are to assess the safety and efficacy of treatment with pirfenidone 2403 milligrams per day (mg/d) compared with placebo in patients with idiopathic pulmonary fibrosis (IPF), to assess the safety and efficacy of treatment with pirfenidone 1197 mg/d in patients with idiopathic pulmonary fibrosis and to characterize the pharmacokinetic disposition of pirfenidone in patients with idiopathic pulmonary fibrosis.	https://clinicaltrials.gov/show/NCT00287716	Noble PW, Albera C, Bradford WZ, Costabel U, Glassberg MK, Kardatzke D, King TE Jr, Lancaster L, Sahn SA, Szwarcberg J, Valeyre D, du Bois RM; CAPACITY Study Group. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. Lancet. 2011 May 21;377(9779):1760-9. doi: 10.1016/S0140-6736(11)60405-4. Epub 2011 May 13.	Included in the SR
Safety and Efficacy of Pirfenidone in Patients With Idiopathic Pulmonary Fibrosis (CAPACITY 1)	NCT00287729 EUDRACT 2006-000138-11 PIPF-006	The purposes of this study are to assess the efficacy of treatment with pirfenidone 2403 milligrams per day compared with placebo in patients with idiopathic pulmonary fibrosis (IPF)and to assess the safety of treatment with pirfenidone 2403 milligrams per day compared with placebo in patients with	https://clinicaltrials.gov/show/NCT00287729	Noble PW, Albera C, Bradford WZ, Costabel U, Glassberg MK, Kardatzke D, King TE Jr, Lancaster L, Sahn SA, Szwarcberg J, Valeyre D, du Bois RM; CAPACITY Study Group. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. Lancet. 2011 May 21;377(9779):1760-9.	Included in the SR

Title	Trial id	Intervention	Link to full record	Publication	Status in the SR
		idiopathic pulmonary fibrosis.		doi: 10.1016/S0140-6736(11)60405-4. Epub 2011 May 13.	
An observational, PractIce based, Open label, Non-comparative, multicEnter study to Evaluate the efficacy, toleRability and safety of pirfenidone in idiopathic pulmonary fibrosis [PIONEER]	CTRI/2012/05/002707	This study is an observational practice based study to monitor the safety, tolerability and efficacy of Pirfenidone in Idiopathic Pulmonary Fibrosis. Total 150 patients will be enrolled in the study from 8 centres across India. The patient will be prescribed Pirfenidone for treatment duration of 48 weeks. The patients will be called for follow up visits at week 4, week 12, week 24 and week 48. Adverse event monitoring will be done at all visits.	http://www.ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=4689	Dhar, R., et al. (2014). Clinical profile and early follow-up of patients receiving pirfenidone in the PIONEER observational study. International Congress 2014. Munich European Respiratory Society: Abstract 3792. This is an interim analysis	Excluded from SR since not an RCT

Table 27: Completed trials (Pirfenidone combined with other treatments)

Title	Trial id	Intervention	Link to full record	Publication	Status in the SR
A case-control study evaluating the efficacy and safety of combined therapy with pirfenidone and inhaled N-acetylcysteine for	JPRN-UMIN000016045	Using Micro Air nebulizers and vibration mesh technology (NE-U07, Omron, Tokyo, Japan), patients receiving NAC combined with pirfenidone were treated twice daily with 352.4 mg of inhaled NAC, which was diluted with saline to a total	https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr.cgi?function=brows&action=brows&type=summary&language=E&recptno=R000018649	Sakamoto, S., et al. (2015). "Effectiveness of combined therapy with pirfenidone and inhaled N-acetylcysteine for advanced idiopathic pulmonary fibrosis: A case-control study." Respirology **20**(3): 445-452.	Excluded as not an RCT & use of pirfenidone with NAC

Title	Trial id	Intervention	Link to full record	Publication	Status in the SR
idiopathic pulmonary fibrosis		volume of 6 mL for 48 weeks. Pirfenidone 1200-1800mg (control group)			
Safety and PK Study of BIBF 1120 in Japanese Patients With IPF: Follow up Study From 1199.31	NCT01136174	long-term tolerability and safety profile of BIBF 1120 on top of pirfenidone treatment in patients with Idiopathic Pulmonary Fibrosis who have completed a prior clinical trial of BIBF 1120 (1199.31).	https://clinicaltrials.gov/show/NCT01417156	No publication	No publication
Safety and PK Study of BIBF 1120 in Japanese Patients With IPF	NCT01136174	To investigate safety of BIBF 1120 in Japanese patients with idiopathic pulmonary fibrosis (IPF), with and without pirfenidone background treatment.	https://clinicaltrials.gov/show/NCT01136174	Ogura T, Taniguchi H, Azuma A, Inoue Y, Kondoh Y, Hasegawa Y, Bando M, Abe S, Mochizuki Y, Chida K, Klüglich M, Fujimoto T, Okazaki K, Tadayasu Y, Sakamoto W, Sugiyama Y. Safety and pharmacokinetics of nintedanib and pirfenidone in idiopathic pulmonary fibrosis. Eur Respir J. 2015 May;45(5):1382-92. doi: 10.1183/09031936.00198013. Epub 2014 Dec 10. PubMed PMID: 25504994.	Identified in our searches
A Randomized, Double-Blind, Placebo-Controlled, Phase 2 Study of the Safety and Tolerability of N–Acetylcysteine in Patients with Idiopathic Pulmonary Fibrosis with Background	EUDRACT 2012-000564-14	PIPF-023	https://www.clinicaltrialsregister.eu/ctr-search/trial/2012-000564-14/DE	Behr, J., et al. (2015). "Safety and tolerability of N-acetylcysteine (NAC) with pirfenidone in idiopathic pulmonary fibrosis (IPF): PANORAMA." European Respiratory Journal. Conference: European Respiratory Society Annual Congress 46(no pagination).	Identified in our searches

Title	Trial id	Intervention	Link to full record	Publication	Status in the SR
Treatment of Pirfenidone (PANORAMA)					

Table 28: Ongoing trials (pirfenidone as a single treatment or as a combined treatment)

Title	Trial id	Intervention	Link to full record
Clinical Progression of Mild to Moderate Idiopathic Pulmonary Fibrosis (IPF) Under a Therapy With Esbriet® (Pirfenidone)	NCT02622477	Observational study of treatment with pirfenidone	https://clinicaltrials.gov/show/NCT02622477
The comparison of the efficacy and safety of pirfenidone and nintedanib in patients with idiopathic pulmonary fibrosis	UMIN000020682	We compare the efficacy and safety of pirfenidone and nintedanib to identify the background and characteristics of the responders to each drug and discover appropriate use of these antifibrotic agents based on each individual patients. Parallel open label non-randomized trial	https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr.cgi?function=brows&action=brows&type=summary&language=E&recptno=R000022819
The Effect of Pirfenidone on Cough in Patients With Idiopathic Pulmonary Fibrosis (Cough-IPF)	NCT02009293	The effect of Pirfenidone on cough and quality of life in patients with idiopathic pulmonary fibrosis (IPF) that are treated with Pirfenidone in daily practice. The hypothesis is that Pirfenidone will decrease cough and increase quality of life. Observational study	https://clinicaltrials.gov/show/NCT02009293
The evaluation of oxidative stress and the effect of pirfenidone in patients with idiopathic pulmonary fibrosis.	UMIN000016826	To investigate the association between oxidative stress and disease severity in patients with IPF. And, to evaluate the effects of pirfenidone on oxidative stress in patients with IPF. Parallel open non-randomised study	https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr.cgi?function=brows&action=brows&type=summary&language=E&recptno=R000019528
Expanded Access Program (EAP): Allow Patients in the US With Idiopathic Pulmonary Fibrosis Access to	NCT02141087	A Treatment Protocol to Allow Patients in the US With Idiopathic Pulmonary Fibrosis Access to Pirfenidone	https://clinicaltrials.gov/show/NCT02141087

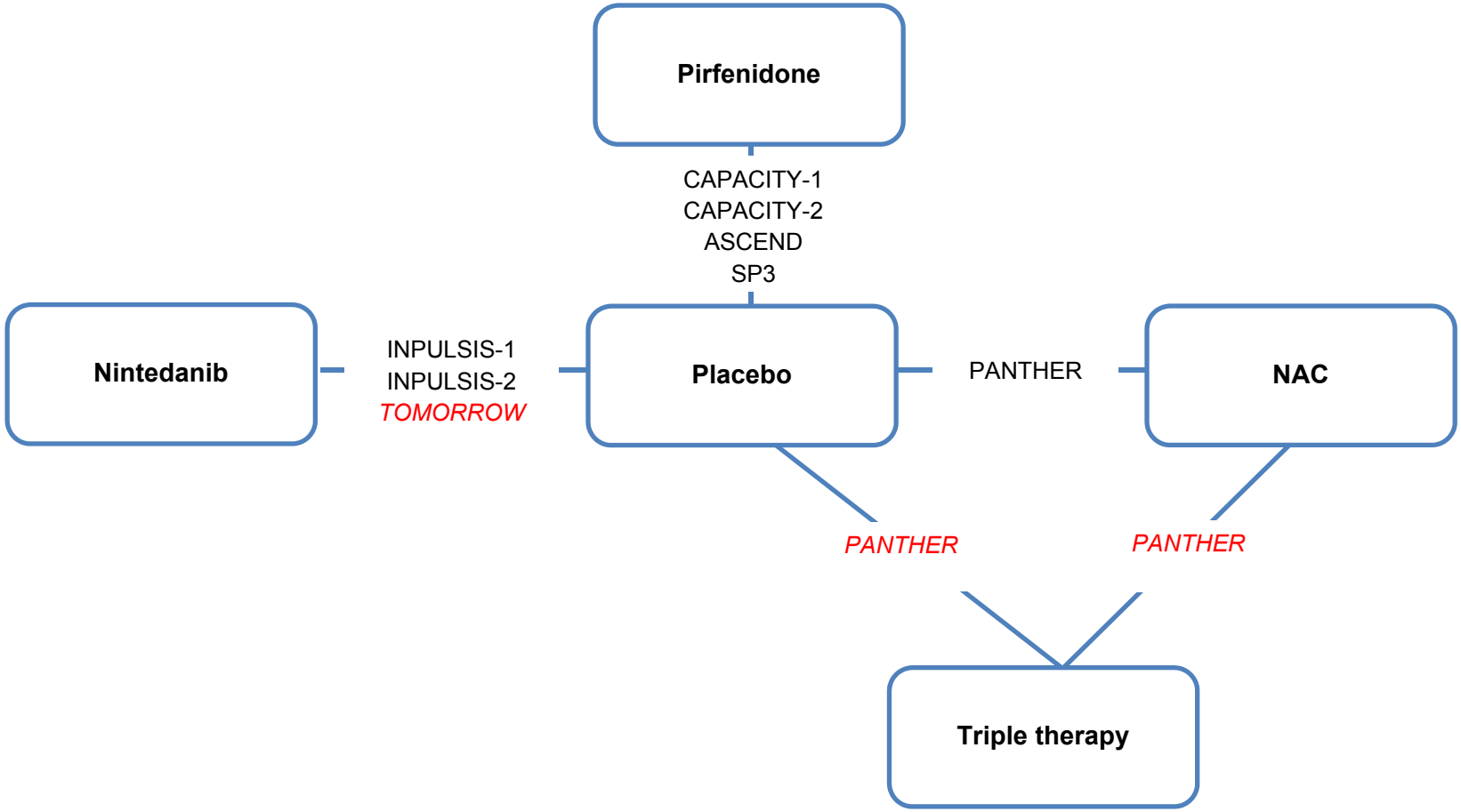
Title	Trial id	Intervention	Link to full record
Pirfenidone			
Investigating Significant Health Trends in Idiopathic Pulmonary Fibrosis (INSIGHTS-IPF)	NCT01695408	As in Europe data are limited on the characteristics and management of such patients, INSIGHTS-IPF was initiated as a new registry that documents newly diagnosed (incident) and prevalent patients with confirmed IPF diagnosis prospectively.The registry will contribute to the optimization of the management of IPF patients in the long term. Observational cohort	https://clinicaltrials.gov/show/NCT01695408
Non-interventional study of the clinical course of mild to moderate IPF under therapy with Esbriet(R)	DRKS00006040	Observational study	http://drks-neu.uniklinik-freiburg.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00006040
Observational Study to Evaluate Disease Course and Outcomes in Patients Treated With Esbriet (Pirfenidone) for Idiopathic Pulmonary Fibrosis (IPF) in Canada	NCT02552849	Observational study	https://clinicaltrials.gov/show/NCT02552849
An Open-Label Extension Study of the Long-Term Safety of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis (IPF) Who Complete the CAPACITY Studies - Open-Label Extension of Pirfenidone CAPACITY Studies	EUCTR2007-007800-13-IE NCT00662038	Open label extension study of CAPACITY studies (PIPF-012)	https://www.clinicaltrialsregister.eu/ctr-search/trial/2007-007800-13/GB#P
A Safety and Tolerability Study of Oral Vismodegib in Combination With Pirfenidone in	NCT02648048	Phase 1 trial	https://clinicaltrials.gov/show/NCT02648048

Title	Trial id	Intervention	Link to full record
Participants With Idiopathic Pulmonary Fibrosis.			
Safety and Tolerability Study of Pirfenidone in Combination With Nintedanib in Participants With Idiopathic Pulmonary Fibrosis (IPF)	NCT02598193	Exploratory Multicenter, Open-Label, Single Arm Study	https://clinicaltrials.gov/show/NCT02598193
Safety, Tolerability and PK of Nintedanib in Combination With Pirfenidone in IPF	NCT02579603	phase IV, twelve week, open label, randomized, parallel group study	https://clinicaltrials.gov/show/NCT02579603
A Study of Lebrikizumab in Patients With Idiopathic Pulmonary Fibrosis	NCT01872689	randomized, multicenter, double-blind, placebo-controlled, parallel-group study will evaluate the efficacy and safety of lebrikizumab as monotherapy in the absence of background IPF therapy or as combination therapy with pirfenidone background therapy in patients with idiopathic pulmonary fibrosis. Patients will be randomized to receive either lebrikizumab or placebo subcutaneously (SC) every 4 weeks. There is an arm receiving pirfenidone+lebrikizumab and an arm receiving pirfenidone+placebo	https://clinicaltrials.gov/show/NCT01872689
Randomized Phase 2 study of Nintedanib and Pirfenidone versus Nintedanib following a clinically meaningful decline in forced vital capacity in patients with idiopathic pulmonary fibrosis administering pirfenidone	UMIN000019436	Phase II trial, Open label	https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr.cgi?function=brows&action=brows&type=summary&language=E&recptno=R000022471
A prospective randomized, multicenter trial evaluating the	UMIN000015508	Combined therapy	https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr.cgi?function=brows&action=brows&type=summary&a

Title	Trial id	Intervention	Link to full record
efficacy and safety of combined therapy with pirfenidone and inhaled N-acetylcysteine for idiopathic pulmonary fibrosis.			mp;language=E&recptno=R000018019
Open-label study of tiotropium/pirfenidone for idiopathic pulmonary fibrosis with emphysema	UMIN000005793	Combined therapy	https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr.cgi?function=brows&action=brows&type=summary&language=E&recptno=R000006225
A twelve week, open-label, randomised, parallel-group study evaluating safety, tolerability and pharmacokinetics (PK) of oral nintedanib in combination with oral pirfenidone, compared to treatment ...	EUDRACT **2015-000640-42**	Combined nintedanib+pirfenidone vs nintedanib alone	https://www.clinicaltrialsregister.eu/ctr-search/trial/2015-000640-42/FR
AN EXPLORATORY MULTICENTER, OPEN-LABEL, SINGLE ARM STUDY OF THE SAFETY AND TOLERABILITY OF PIRFENIDONE (ESBRIET®) IN COMBINATION WITH NINTEDANIB (OFEV®) IN PATIENTS WITH IDIOPATHIC PULMONARY FIBROSIS	EUDRACT 2015-003280-11	Combined therapy	https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003280-11

Appendix B: Evidence networks for each outcome assessed in A32

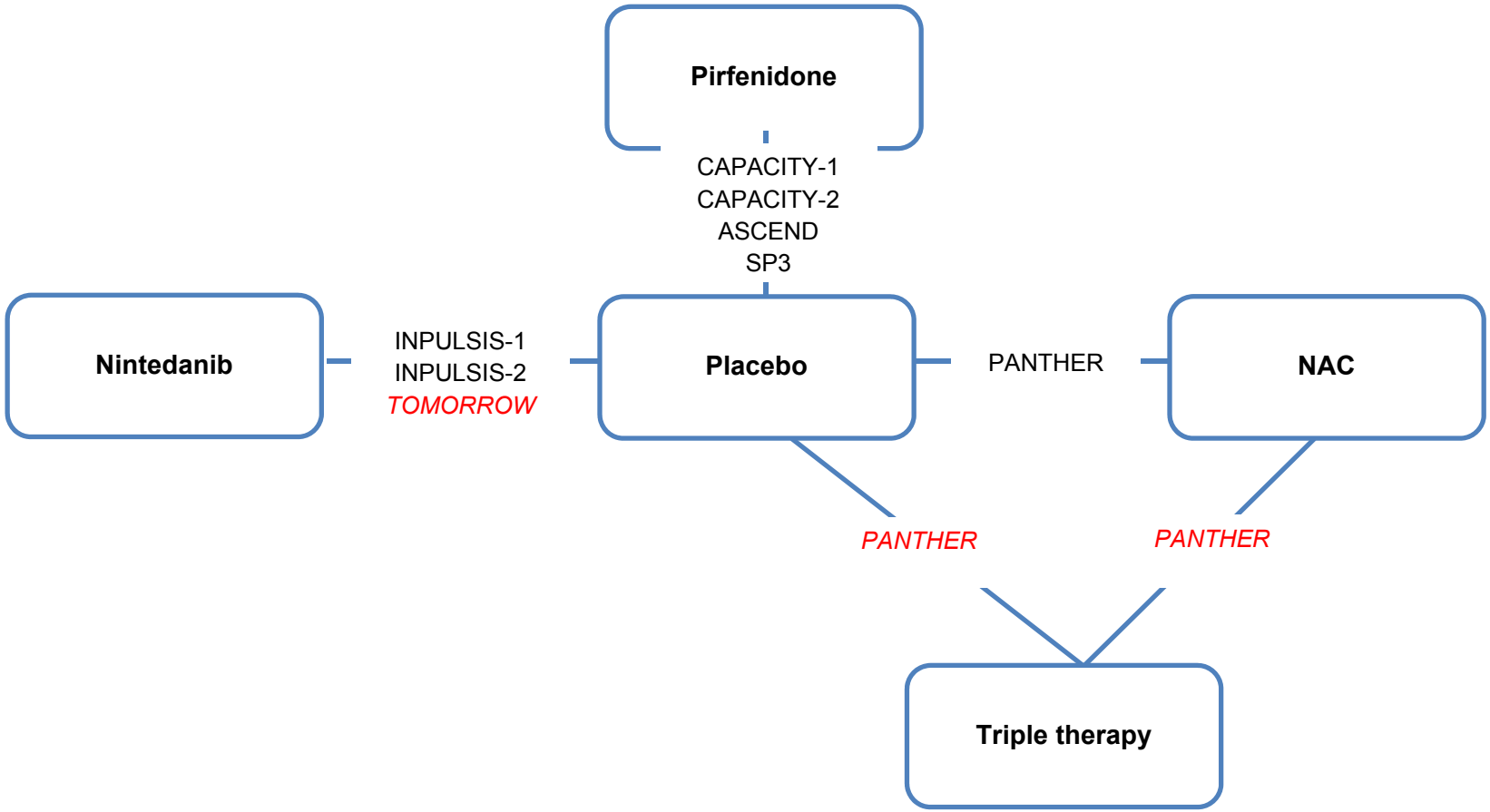
1. Network diagram including all trials



All trials are included in the base-case network

Red italicised text indicates trials excluded from the restricted network

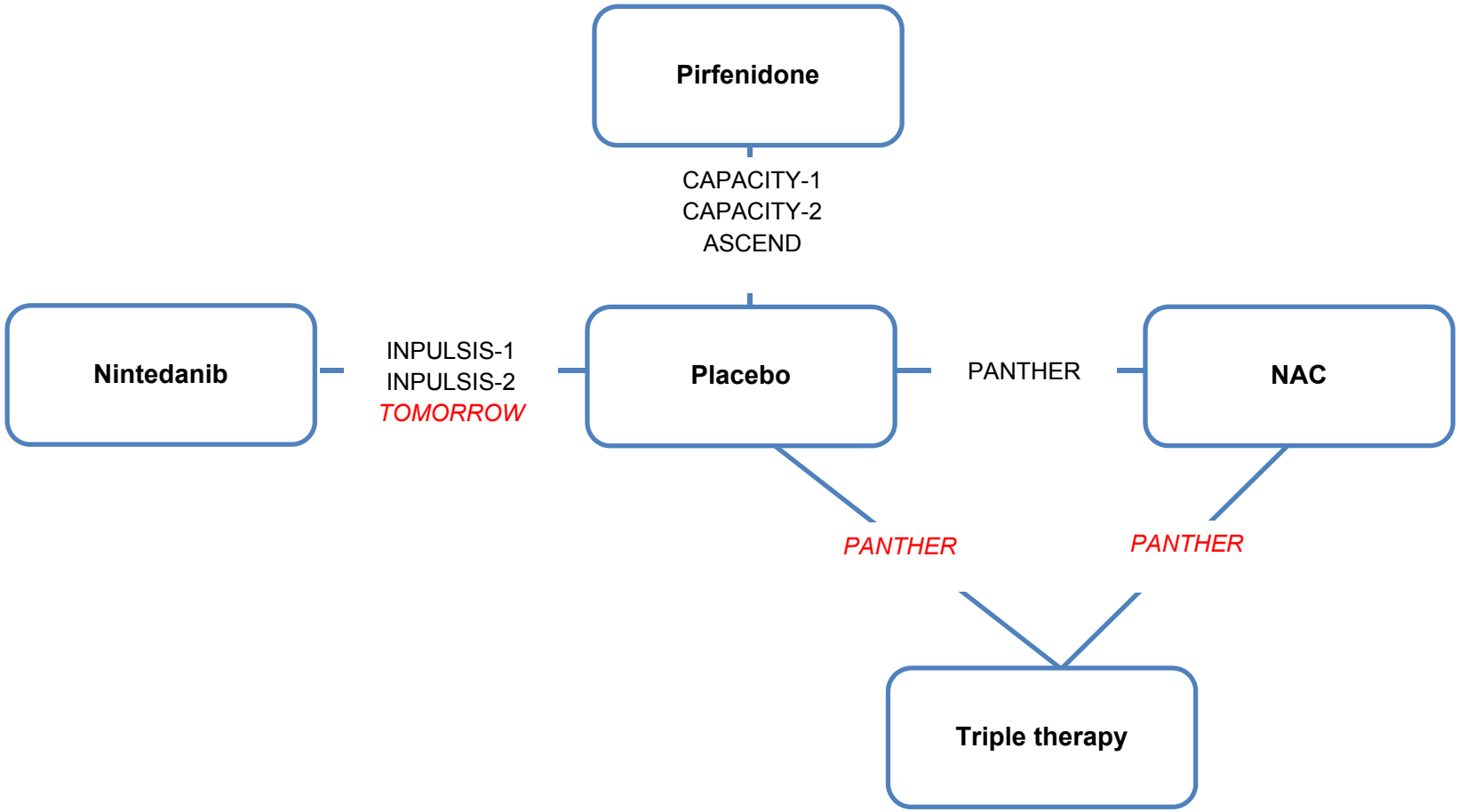
2. Network diagrams for each endpoint: All-cause mortality



All trials are included in the base-case network

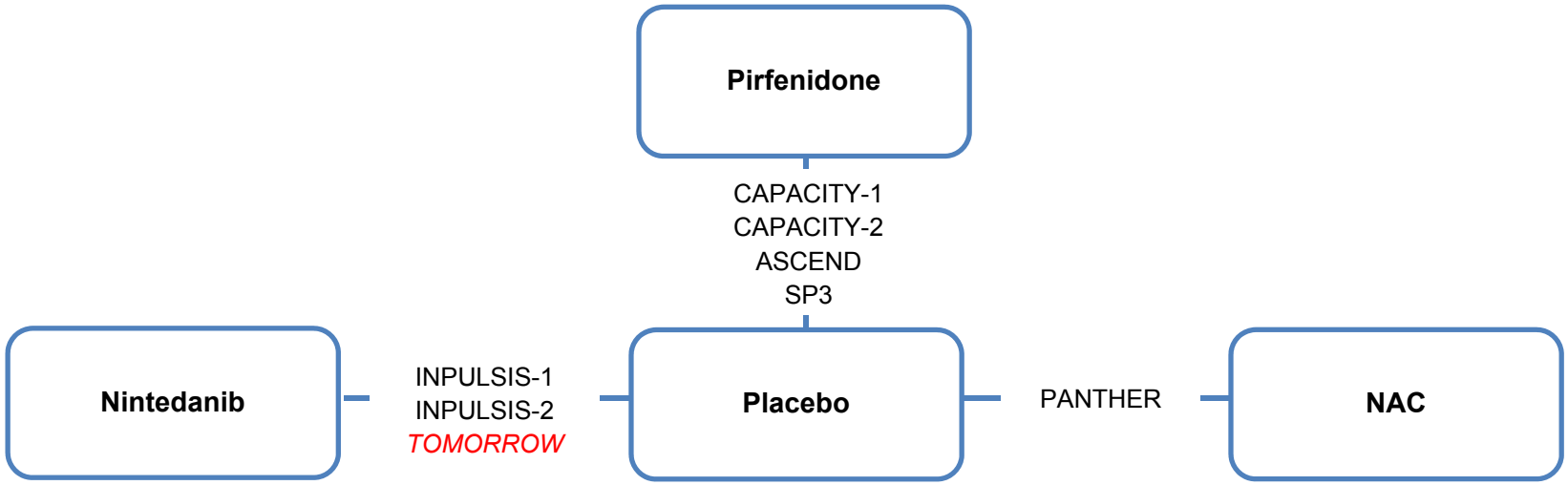
Red italicised text indicates trials excluded from the restricted network

3. Network diagrams for each endpoint: IPF-related mortality



All trials are included in the base-case network
Red italicised text indicates trials excluded from the restricted network

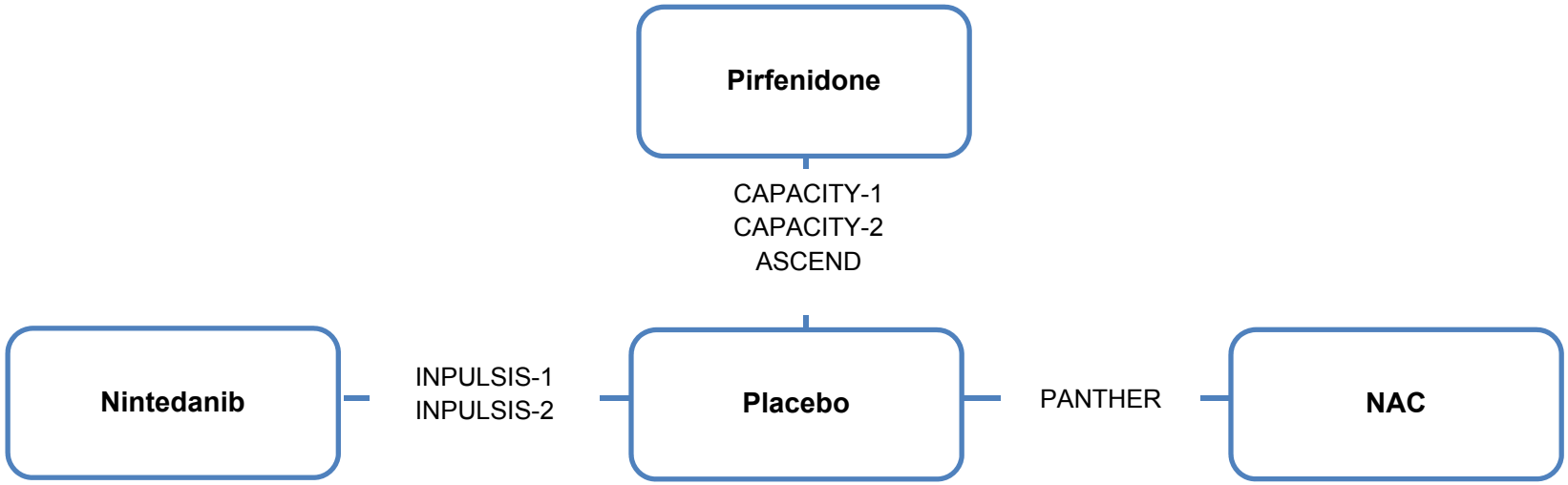
4. Network diagrams for each endpoint: Acute exacerbations



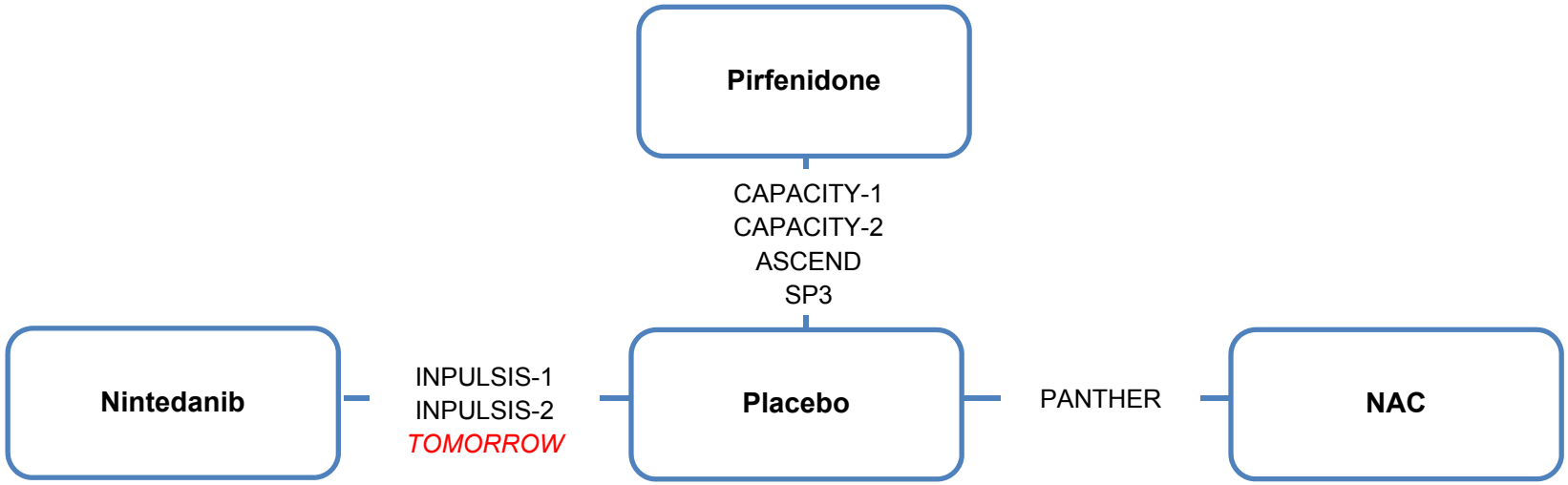
All trials are included in the base-case network

Red italicised text indicates trials excluded from the restricted network

5. Network diagrams for each endpoint: 10 % categorical decline in FVC % predicted



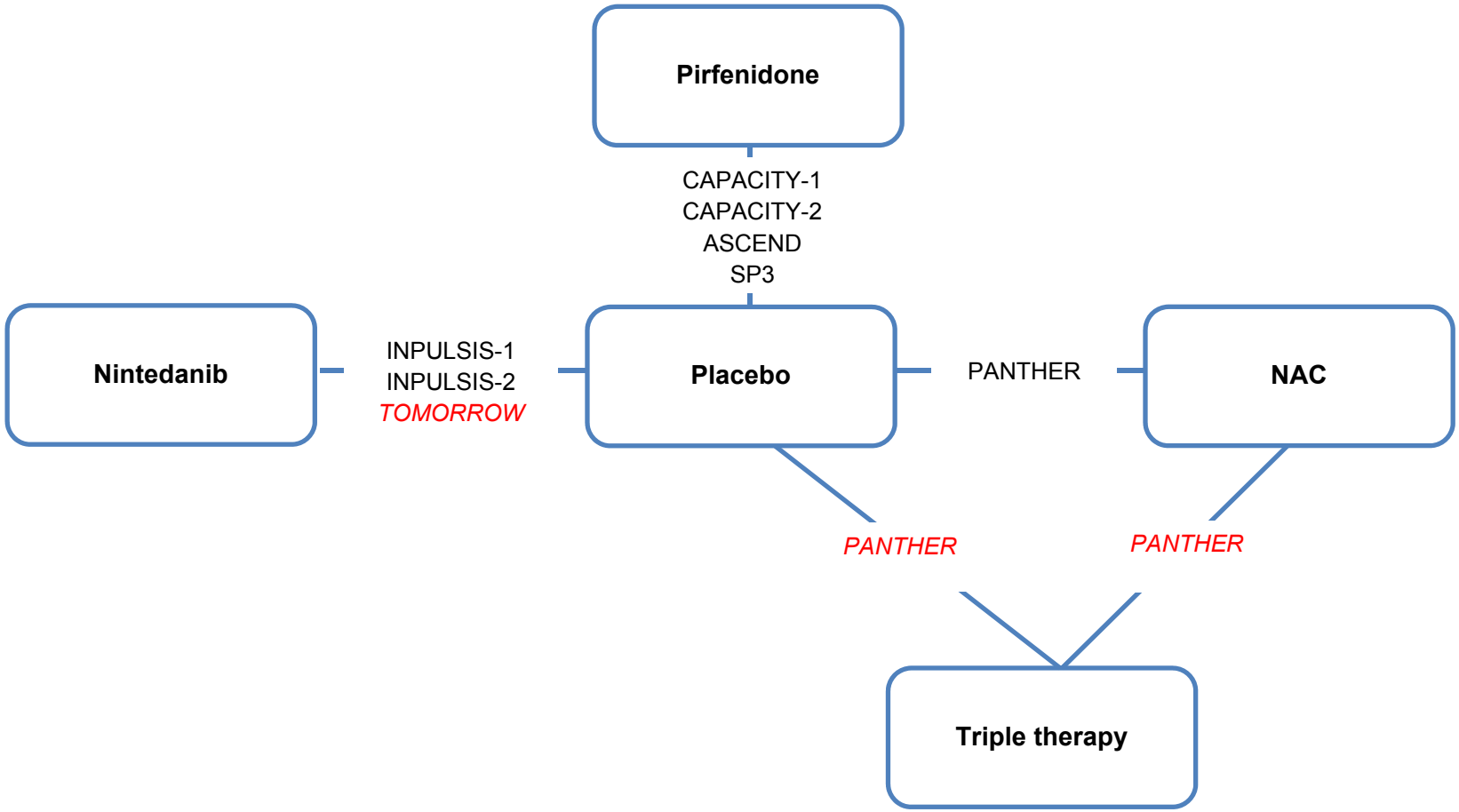
6. Network diagrams for each endpoint: FVC% predicted



All trials are included in the base-case network

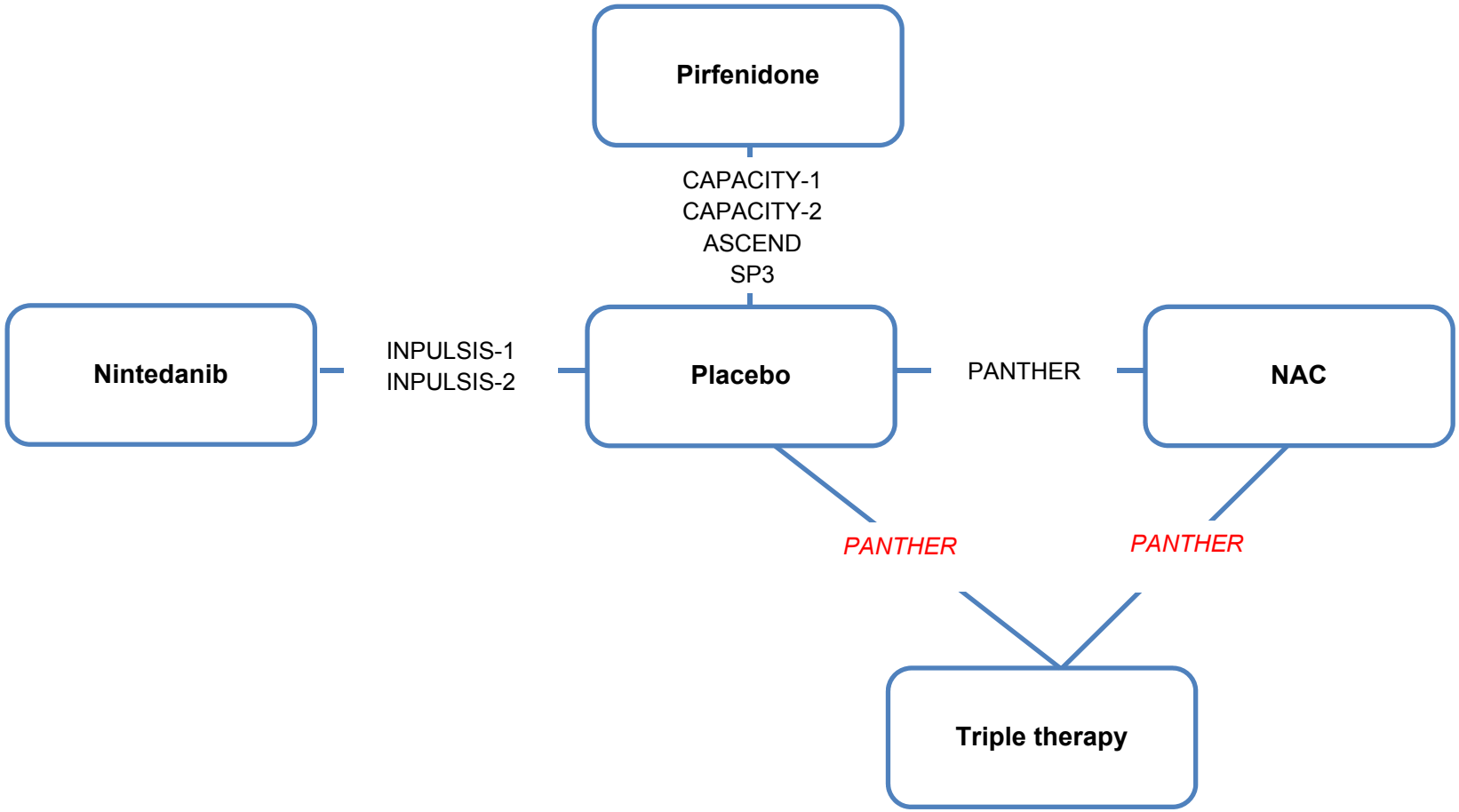
Red italicised text indicates trials excluded from the restricted network

7. Network diagrams for each endpoint: FVC (L)



All trials are included in the base-case network
Red italicised text indicates trials excluded from the restricted network

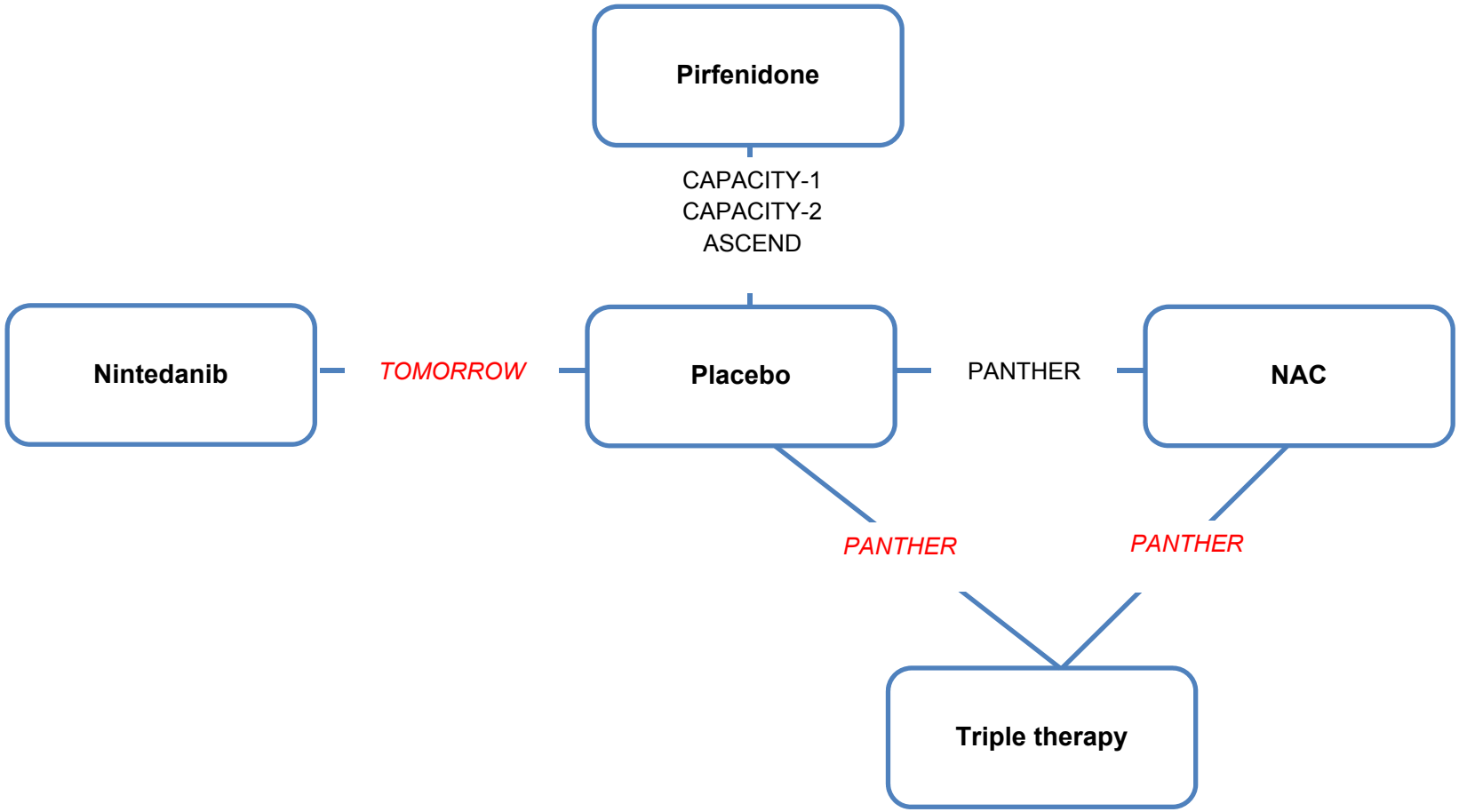
8. Network diagrams for each endpoint: PFS



All trials are included in the base-case network

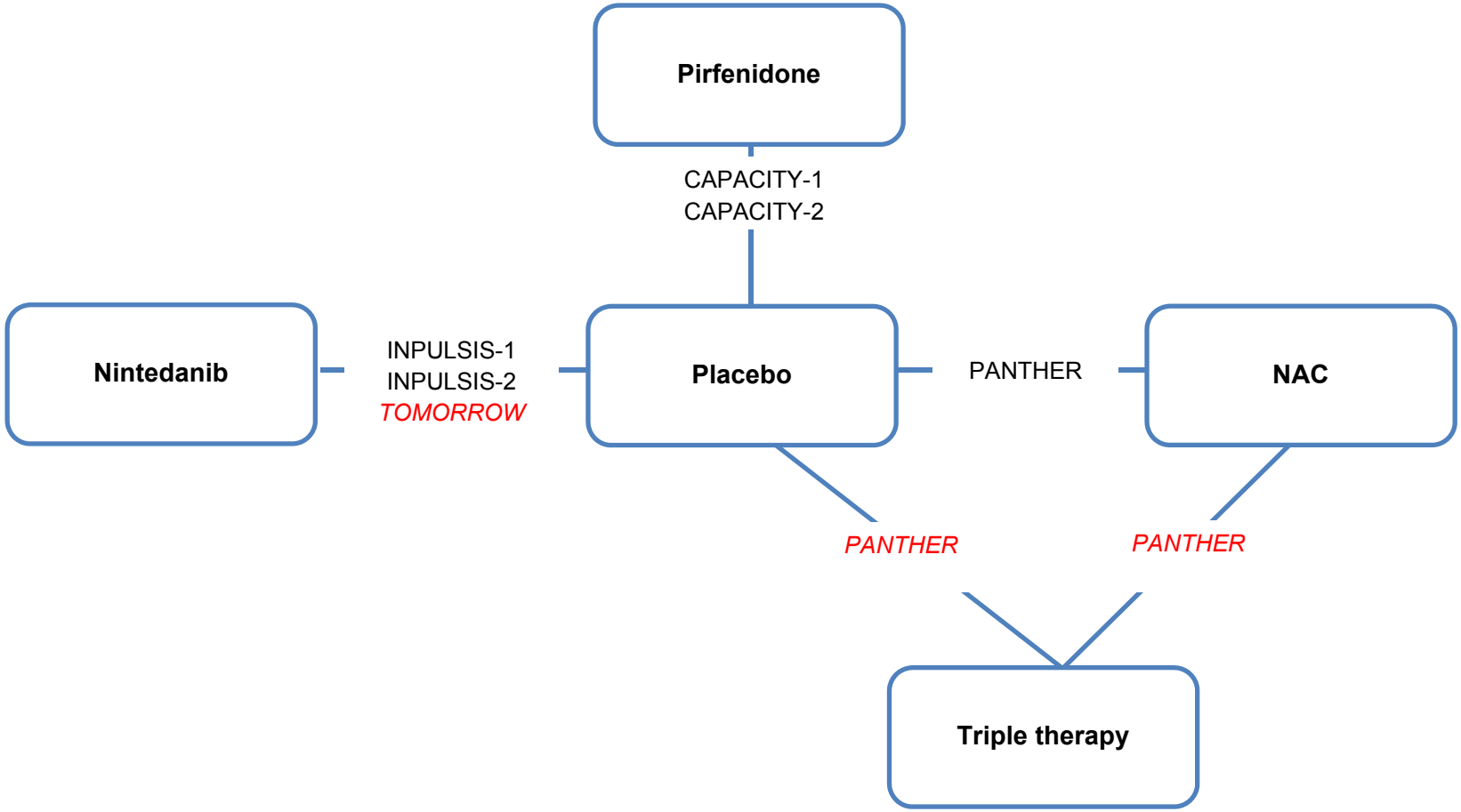
Red italicised text indicates trials excluded from the restricted network

9. Network diagrams for each endpoint: 6MWD



All trials are included in the base-case network
Red italicised text indicates trials excluded from the restricted network

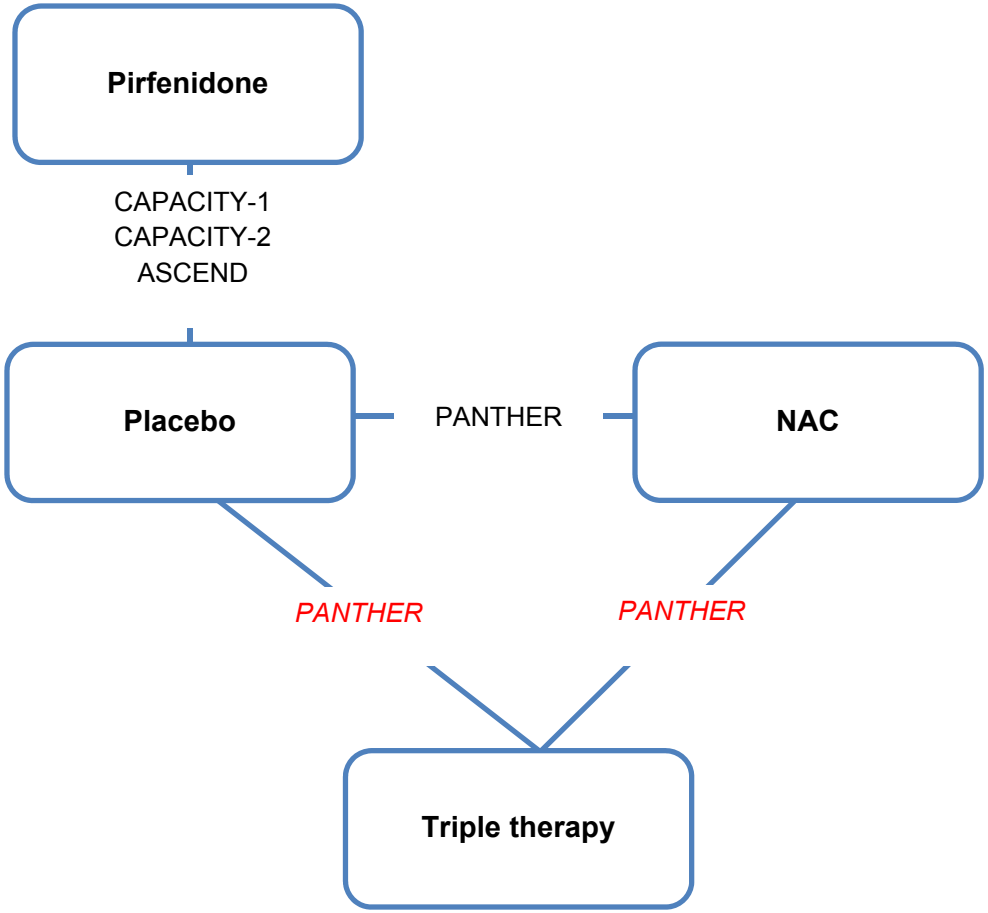
10. Network diagrams for each endpoint: SGRQ



All trials are included in the base-case network

Red italicised text indicates trials excluded from the restricted network

11. Network diagrams for each endpoint: UCSD SOBQ



All trials are included in the base-case network

Red italicised text indicates trials excluded from the restricted network

Appendix C: Analyses supporting response to A34

a) All-cause mortality (HR 52 weeks)

1) Introduction

New input: separate data for the two INPULSIS studies (instead of pooled). The following NMA results are presented:

(a) Base case network, RE model

(b) Restricted network, RE model

(c) Restricted network, FE model

(d) Base case network, FE model

(e) Base case network without PANTHER, RE model

(f) Base case network without PANTHER, FE model

(g) Model comparison

Given the star-shaped network, the scenarios with all trials except PANTHER must lead to the same results (up to Monte-Carlo error) as the extended network. The last two scenarios are run for confirmation only.

The primary RE models used an Inv-Gamma(0.001, 0.001) prior for the between study variance.

Table 29: The (full) analysis data set used in the JAGS fits:

Study	Treatment	Comparator	HR	logHR	SE	N	n	t	b
PIPF-004 CAPACITY 2	Pirfenidone 2403mg/day (*)	Placebo	0.370000	-0.9942523	0.5304795	NA	NA	2	1
PIPF-006 CAPACITY 1	Pirfenidone 2403mg/day (*)	Placebo	0.660000	-0.4155154	0.5237481	NA	NA	2	1
ASCEND	Pirfenidone 2403mg/day (*)	Placebo	0.550000	-0.5978370	0.3793018	NA	NA	2	1
INPULSIS 1	**Nintedanib 300mg/day**	**Placebo**	**0.630000**	**-0.4620355**	**0.3942242**	**NA**	**NA**	**3**	**1**
INPULSIS 2	**Nintedanib 300mg/day**	**Placebo**	**0.740000**	**-0.3011051**	**0.3103049**	**NA**	**NA**	**3**	**1**
PANTHER (Triple therapy)	Triple therapy	Placebo	9.260000	2.2257040	1.0604775	NA	NA	5	1
PANTHER (NAC)	NAC	Placebo	1.995622	0.6909556	0.6666667	NA	NA	4	1
SP3	Pirfenidone 2403mg/day (*)	Placebo	NA	NA	NA	110	3	2	1
SP3	Placebo	Placebo	NA	NA	NA	109	6	1	1
TOMORROW	Nintedanib 300mg/day	Placebo	NA	NA	NA	85	7	3	1
TOMORROW	Placebo	Placebo	NA	NA	NA	85	9	1	1

2) Base case network, RE model

Total residual deviance (posterior mean): 7.25
DIC: 31.46
pD: 6.51

Table 30: Hazard ratios and 95% credible intervals (CrI): ACM at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.52 (0.3, 0.88)	0.71 (0.43, 1.16)	2 (0.51, 7.84)	9.26 (1.11, 77.24)
Pirfenidone 2403mg/day (*)	1.94 (1.14, 3.32)		1.37 (0.66, 2.84)	3.87 (0.89, 16.82)	17.96 (2.01, 160.45)
Nintedanib 300mg/day	1.41 (0.86, 2.32)	0.73 (0.35, 1.5)		2.82 (0.66, 12.08)	13.07 (1.48, 115.32)
NAC	0.5 (0.13, 1.97)	0.26 (0.06, 1.12)	0.35 (0.08, 1.52)		4.64 (0.37, 57.92)
Triple therapy	0.11 (0.01, 0.9)	0.06 (0.01, 0.5)	0.08 (0.01, 0.68)	0.22 (0.02, 2.7)	

[For SP3 pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. For all tables hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.]*

Table 31: Hazard ratios and 95% predictive intervals (PrI): ACM at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo	1 (0.65, 1.54)	0.52 (0.26, 1)	0.71 (0.38, 1.33)	2 (0.48, 8.32)	9.26 (1.07, 80.13)
Pirfenidone 2403mg/day (*)	1.94 (1, 3.78)	1 (0.65, 1.53)	1.37 (0.6, 3.14)	3.87 (0.84, 17.76)	17.95 (1.93, 166.44)
Nintedanib 300mg/day	1.41 (0.75, 2.67)	0.73 (0.32, 1.66)	1 (0.65, 1.53)	2.82 (0.62, 12.77)	13.06 (1.42, 120.13)
NAC	0.5 (0.12, 2.09)	0.26 (0.06, 1.19)	0.35 (0.08, 1.61)	1 (0.65, 1.54)	4.64 (0.36, 59.96)
Triple therapy	0.11 (0.01, 0.94)	0.06 (0.01, 0.52)	0.08 (0.01, 0.7)	0.22 (0.02, 2.79)	1 (0.65, 1.53)

Table 32: Posterior summaries of between-trial standard deviation: ACM at 52 weeks (base case network, RE model)

mean	50%	2.5%	97.5%
0.153	0.107	0.025	0.542

Table 33: Probability of column treatment being better than row treatment: ACM at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.992	0.92	0.16	0.02
Pirfenidone 2403mg/day (*)	0.008		0.19	0.035	0.005
Nintedanib 300mg/day	0.08	0.81		0.08	0.011
NAC	0.84	0.965	0.92		0.116
Triple therapy	0.98	0.995	0.989	0.884	

3) Restricted network, RE model

Total residual deviance (posterior mean): 5.15
DIC: 17.7
pD: 4.47

Table 34: Hazard ratios and 95% credible intervals (CrI): ACM at 52 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.52 (0.3, 0.89)	0.69 (0.39, 1.24)	1.99 (0.5, 8.06)
Pirfenidone 2403mg/day (*)	1.94 (1.12, 3.37)		1.35 (0.61, 3)	3.88 (0.87, 17.37)
Nintedanib 300mg/day	1.44 (0.81, 2.59)	0.74 (0.33, 1.65)		2.88 (0.63, 13.09)
NAC	0.5 (0.12, 2.02)	0.26 (0.06, 1.15)	0.35 (0.08, 1.58)	

Table 35: Hazard ratios and 95% predictive intervals (PrI): ACM at 52 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo	1 (0.61, 1.65)	0.52 (0.25, 1.05)	0.69 (0.33, 1.46)	2 (0.46, 8.67)
Pirfenidone 2403mg/day (*)	1.94 (0.95, 4)	1 (0.6, 1.65)	1.35 (0.54, 3.39)	3.88 (0.81, 18.68)
Nintedanib 300mg/day	1.44 (0.68, 3.05)	0.74 (0.3, 1.86)	1 (0.61, 1.64)	2.88 (0.59, 14.06)
NAC	0.5 (0.11, 2.18)	0.26 (0.05, 1.23)	0.35 (0.07, 1.69)	1 (0.61, 1.65)

Table 36: Posterior summaries of between-trial standard deviation: ACM at 52 weeks (restricted network, RE model)

mean	50%	2.5%	97.5%
0.174	0.115	0.026	0.66

Table 37: Probability of column treatment being better than row treatment: ACM at 52 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.99	0.903	0.163
Pirfenidone 2403mg/day (*)	0.01		0.222	0.037
Nintedanib 300mg/day	0.097	0.778		0.082
NAC	0.837	0.963	0.918	

4) Restricted network, FE model

Total residual deviance (posterior mean): 4.77
DIC: 16.87
pD: 4.03

Table 38: Hazard ratios and 95% credible intervals (CrI): ACM at 52 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.52 (0.32, 0.84)	0.7 (0.43, 1.12)	1.99 (0.54, 7.37)
Pirfenidone 2403mg/day (*)	1.93 (1.19, 3.15)		1.35 (0.68, 2.66)	3.86 (0.96, 15.53)
Nintedanib 300mg/day	1.44 (0.89, 2.32)	0.74 (0.38, 1.47)		2.87 (0.71, 11.52)
NAC	0.5 (0.14, 1.85)	0.26 (0.06, 1.04)	0.35 (0.09, 1.4)	

Table 39: Probability of column treatment being better than row treatment: ACM at 52 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.996	0.932	0.15
Pirfenidone 2403mg/day (*)	0.004		0.197	0.029
Nintedanib 300mg/day	0.068	0.803		0.069
NAC	0.85	0.971	0.931	

5) Base case network, FE model

Total residual deviance (posterior mean): 6.82
DIC: 30.58
pD: 6.07

Table 40: Hazard ratios and 95% credible intervals (CrI): ACM at 52 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.52 (0.32, 0.84)	0.71 (0.46, 1.09)	2 (0.54, 7.36)	9.27 (1.15, 74.16)
Pirfenidone 2403mg/day (*)	1.93 (1.19, 3.16)		1.37 (0.71, 2.64)	3.87 (0.96, 15.52)	17.91 (2.1, 151.25)
Nintedanib 300mg/day	1.41 (0.92, 2.18)	0.73 (0.38, 1.4)		2.82 (0.71, 11.12)	13.07 (1.55, 109.74)
NAC	0.5 (0.14, 1.85)	0.26 (0.06, 1.04)	0.35 (0.09, 1.4)		4.64 (0.4, 54.24)
Triple therapy	0.11 (0.01, 0.87)	0.06 (0.01, 0.48)	0.08 (0.01, 0.64)	0.22 (0.02, 2.52)	

Table 41: Probability of column treatment being better than row treatment: ACM at 52 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.996	0.942	0.15	0.018
Pirfenidone 2403mg/day (*)	0.004		0.171	0.029	0.004
Nintedanib 300mg/day	0.058	0.829		0.07	0.009
NAC	0.85	0.971	0.93		0.11
Triple therapy	0.982	0.996	0.991	0.89	

6) Base case network without PANTHER, RE model

Total residual deviance (posterior mean): 5.25
DIC: 24.48
pD: 4.52

Table 42: Hazard ratios and 95% credible intervals (CrI): ACM at 52 weeks (base case network without PANTHER, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day
Placebo		0.52 (0.3, 0.88)	0.71 (0.43, 1.16)
Pirfenidone 2403mg/day (*)	1.94 (1.14, 3.32)		1.37 (0.66, 2.85)
Nintedanib 300mg/day	1.41 (0.86, 2.32)	0.73 (0.35, 1.51)	

Table 43: Hazard ratios and 95% predictive intervals (PrI): ACM at 52 weeks (base case network without PANTHER, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day
Placebo	1 (0.65, 1.54)	0.52 (0.26, 1)	0.71 (0.37, 1.33)
Pirfenidone 2403mg/day (*)	1.94 (1, 3.78)	1 (0.65, 1.53)	1.37 (0.6, 3.15)
Nintedanib 300mg/day	1.41 (0.75, 2.67)	0.73 (0.32, 1.67)	1 (0.65, 1.53)

Table 44: Posterior summaries of between-trial standard deviation: ACM at 52 weeks (base case network without PANTHER, RE model)

mean	50%	2.5%	97.5%
0.153	0.107	0.025	0.54

Table 45: Probability of column treatment being better than row treatment: ACM at 52 weeks (base case network without PANTHER, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day
Placebo		0.992	0.92
Pirfenidone 2403mg/day (*)	0.008		0.19
Nintedanib 300mg/day	0.08	0.81	

7) Base case network without PANTHER, FE model

Total residual deviance (posterior mean): 4.82
DIC: 23.6
pD: 4.07

Table 46: Hazard ratios and 95% credible intervals (CrI): ACM at 52 weeks (base case network without PANTHER, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day
Placebo		0.52 (0.32, 0.84)	0.71 (0.46, 1.09)
Pirfenidone 2403mg/day (*)	1.93 (1.19, 3.15)		1.37 (0.71, 2.62)
Nintedanib 300mg/day	1.41 (0.92, 2.18)	0.73 (0.38, 1.4)	

Table 47: Probability of column treatment being better than row treatment: ACM at 52 weeks (base case network without PANTHER, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day
Placebo		0.996	0.942
Pirfenidone 2403mg/day (*)	0.004		0.172
Nintedanib 300mg/day	0.058	0.828	

8) Model comparison (ACM at 52 weeks)

Note that fit statistics can be compared only for models fitted to the same data scenario. DIC differences of less than 5 are not considered meaningful (DSU TSD 2).

Table 48: Model comparison: ACM at 52 weeks (base case network)

description	Residual deviance	DIC	pD
Base case network, RE model	7.25	31.46	6.51
Base case network, FE model	6.82	30.58	6.07

Table 49: Model comparison: ACM at 52 weeks (restricted network)

description	Residual deviance	DIC	pD
Restricted network, RE model	5.15	17.70	4.47
Restricted network, FE model	4.77	16.87	4.03

Table 50: Model comparison: ACM at 52 weeks (base case network without PANTHER)

description	Residual deviance	DIC	pD
Base case network without PANTHER, RE model	5.25	24.48	4.52
Base case network without PANTHER, FE model	4.82	23.60	4.07

b) All-cause mortality (HR at 72 weeks)

1) Introduction

New input: separate data for the two INPULSIS studies (instead of pooled). The following NMA results are presented:

(a) Base case network, RE model

(b) Restricted network, RE model

(c) Restricted network, FE model

(d) Base case network, FE model

(e) Model comparison

The primary RE models used an Inv-Gamma(0.001, 0.001) prior for the between study variance.

Table 51: The (full) analysis data set used in the JAGS fits:

Study	Treatment	Comparator	HR	logHR	SE	N	n	t	b
PIPF-004 CAPACITY 2	Pirfenidone 2403mg/day (*)	Placebo	0.508910	-0.6754841	0.4378553	NA	NA	2	1
PIPF-006 CAPACITY 1	Pirfenidone 2403mg/day (*)	Placebo	0.866370	-0.1434432	0.3789628	NA	NA	2	1
ASCEND	Pirfenidone 2403mg/day (*)	Placebo	0.550000	-0.5978370	0.3793018	NA	NA	2	1
INPULSIS 1	**Nintedanib 300mg/day**	**Placebo**	**0.630000**	**-0.4620355**	**0.3942242**	**NA**	**NA**	**3**	**1**
INPULSIS 2	**Nintedanib 300mg/day**	**Placebo**	**0.740000**	**-0.3011051**	**0.3103049**	**NA**	**NA**	**3**	**1**
PANTHER (Triple therapy)	Triple therapy	Placebo	9.260000	2.2257040	1.0604775	NA	NA	5	1
PANTHER (NAC)	NAC	Placebo	1.995622	0.6909556	0.6666667	NA	NA	4	1
SP3	Pirfenidone 2403mg/day (*)	Placebo	NA	NA	NA	110	3	2	1
SP3	Placebo	Placebo	NA	NA	NA	109	6	1	1
TOMORROW	Nintedanib 300mg/day	Placebo	NA	NA	NA	85	7	3	1
TOMORROW	Placebo	Placebo	NA	NA	NA	85	9	1	1

2) Base case network, RE model

Total residual deviance (posterior mean): 7.69
DIC: 30.91
pD: 6.55

Table 52: Hazard ratios and 95% credible intervals (CrI): ACM at 72 weeks (Base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.62 (0.38, 0.99)	0.71 (0.43, 1.16)	2 (0.51, 7.87)	9.27 (1.12, 77.26)
Pirfenidone 2403mg/day (*)	1.62 (1.01, 2.63)		1.15 (0.58, 2.29)	3.24 (0.76, 13.86)	15.06 (1.71, 132.22)
Nintedanib 300mg/day	1.41 (0.86, 2.31)	0.87 (0.44, 1.72)		2.82 (0.66, 12.07)	13.09 (1.48, 115.45)
NAC	0.5 (0.13, 1.97)	0.31 (0.07, 1.31)	0.35 (0.08, 1.52)		4.65 (0.37, 57.71)
Triple therapy	0.11 (0.01, 0.9)	0.07 (0.01, 0.58)	0.08 (0.01, 0.67)	0.21 (0.02, 2.68)	

[For SP3 pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. For all tables hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.]*

Table 53: Hazard ratios and 95% predictive intervals (Prl): ACM at 72 weeks (Base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo	1 (0.65, 1.53)	0.62 (0.33, 1.13)	0.71 (0.38, 1.33)	2 (0.48, 8.33)	9.26 (1.07, 80.09)
Pirfenidone 2403mg/day (*)	1.62 (0.88, 3.03)	1 (0.65, 1.53)	1.15 (0.53, 2.55)	3.24 (0.72, 14.67)	15.05 (1.65, 137.11)
Nintedanib 300mg/day	1.41 (0.75, 2.65)	0.87 (0.39, 1.91)	1 (0.66, 1.53)	2.82 (0.62, 12.73)	13.09 (1.43, 119.73)
NAC	0.5 (0.12, 2.08)	0.31 (0.07, 1.39)	0.35 (0.08, 1.6)	1 (0.66, 1.52)	4.65 (0.36, 59.68)
Triple therapy	0.11 (0.01, 0.93)	0.07 (0.01, 0.6)	0.08 (0.01, 0.7)	0.22 (0.02, 2.77)	1 (0.65, 1.53)

Table 54: Posterior summaries of between-trial standard deviation: ACM at 72 weeks (Base case network, RE model)

mean	50%	2.5%	97.5%
0.152	0.106	0.025	0.532

Table 55: Probability of column treatment being better than row treatment: ACM at 72 weeks (Base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.978	0.92	0.159	0.02
Pirfenidone 2403mg/day (*)	0.022		0.34	0.055	0.007
Nintedanib 300mg/day	0.08	0.66		0.08	0.01
NAC	0.841	0.945	0.92		0.116
Triple therapy	0.98	0.993	0.99	0.884	

3) Restricted network, RE model

Total residual deviance (posterior mean): 5.58
DIC: 17.16
pD: 4.53

Table 56: Hazard ratios and 95% credible intervals (Crl): ACM at 72 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.62 (0.37, 1)	0.69 (0.39, 1.24)	1.99 (0.5, 8)
Pirfenidone 2403mg/day (*)	1.62 (1, 2.67)		1.13 (0.53, 2.42)	3.24 (0.74, 14.24)
Nintedanib 300mg/day	1.44 (0.81, 2.58)	0.89 (0.41, 1.89)		2.88 (0.64, 12.99)
NAC	0.5 (0.12, 2.02)	0.31 (0.07, 1.35)	0.35 (0.08, 1.57)	

Table 57: Hazard ratios and 95% predictive intervals (Prl): ACM at 72 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo	1 (0.61, 1.63)	0.62 (0.31, 1.2)	0.69 (0.33, 1.45)	2 (0.46, 8.63)
Pirfenidone 2403mg/day (*)	1.62 (0.84, 3.21)	1 (0.61, 1.64)	1.13 (0.47, 2.75)	3.24 (0.69, 15.24)
Nintedanib 300mg/day	1.44 (0.69, 3.03)	0.89 (0.37, 2.14)	1 (0.61, 1.63)	2.87 (0.59, 13.89)
NAC	0.5 (0.12, 2.18)	0.31 (0.07, 1.45)	0.35 (0.07, 1.68)	1 (0.61, 1.63)

Table 58: Posterior summaries of between-trial standard deviation: ACM at 72 weeks (restricted network, RE model)

mean	50%	2.5%	97.5%
0.173	0.115	0.026	0.647

Table 59: Probability of column treatment being better than row treatment: ACM at 72 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.975	0.903	0.162
Pirfenidone 2403mg/day (*)	0.025		0.373	0.057
Nintedanib 300mg/day	0.097	0.627		0.082
NAC	0.838	0.943	0.918	

4) Restricted network, FE model

Total residual deviance (posterior mean): 5.32
DIC: 16.4
pD: 4.03

Table 60: Hazard ratios and 95% credible intervals (Crl): ACM at 72 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.62 (0.4, 0.95)	0.7 (0.43, 1.12)	2 (0.54, 7.39)
Pirfenidone 2403mg/day (*)	1.62 (1.06, 2.47)		1.12 (0.59, 2.13)	3.23 (0.82, 12.79)
Nintedanib 300mg/day	1.44 (0.89, 2.32)	0.89 (0.47, 1.68)		2.87 (0.71, 11.55)
NAC	0.5 (0.14, 1.85)	0.31 (0.08, 1.22)	0.35 (0.09, 1.4)	

Table 61: Probability of column treatment being better than row treatment: ACM at 72 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.987	0.931	0.15
Pirfenidone 2403mg/day (*)	0.013		0.36	0.048
Nintedanib 300mg/day	0.069	0.64		0.069
NAC	0.85	0.952	0.931	

5) Base case network, FE model

Total residual deviance (posterior mean): 7.36
DIC: 30.11
pD: 6.08

Table 62: Hazard ratios and 95% credible intervals (CrI): ACM at 72 weeks (Base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.62 (0.4, 0.95)	0.71 (0.46, 1.09)	1.99 (0.54, 7.37)	9.26 (1.16, 73.77)
Pirfenidone 2403mg/day (*)	1.62 (1.06, 2.47)		1.14 (0.62, 2.1)	3.22 (0.81, 12.74)	14.97 (1.79, 124.19)
Nintedanib 300mg/day	1.41 (0.92, 2.18)	0.87 (0.48, 1.6)		2.81 (0.71, 11.16)	13.06 (1.56, 108.45)
NAC	0.5 (0.14, 1.86)	0.31 (0.08, 1.23)	0.36 (0.09, 1.41)		4.64 (0.4, 53.74)
Triple therapy	0.11 (0.01, 0.86)	0.07 (0.01, 0.56)	0.08 (0.01, 0.64)	0.22 (0.02, 2.5)	

Table 63: Probability of column treatment being better than row treatment: ACM at 72 weeks (Base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.986	0.942	0.15	0.018
Pirfenidone 2403mg/day (*)	0.014		0.331	0.048	0.006
Nintedanib 300mg/day	0.058	0.669		0.071	0.009
NAC	0.85	0.952	0.929		0.11
Triple therapy	0.982	0.994	0.991	0.89	

6) Model comparison (ACM at 72 weeks)

Note that fit statistics can be compared only for models fitted to the same data scenario. DIC differences of less than 5 are not considered meaningful (DSU TSD 2).

Table 64: Model comparison: ACM at 72 weeks (base case network)

description	Residual deviance	DIC	pD
Base case network, RE model	7.69	30.91	6.55
Base case network, FE model	7.36	30.11	6.08

Table 65: Model comparison: ACM at 72 weeks (restricted network)

description	Residual deviance	DIC	pD
Restricted network, RE model	5.58	17.16	4.53
Restricted network, FE model	5.32	16.40	4.03

c) Progression free survival (HR at 52 weeks)

1) Introduction

The following NMA results are presented:

(a) Base case network, RE model

(b) Restricted network, RE model

(c) Restricted network, FE model

(d) Base case network, FE model

(e) Model comparison

The primary RE models used an Inv-Gamma(0.001, 0.001) prior for the between study variance.

2) Base case network, RE model

Total residual deviance (posterior mean): 6.94
DIC: 13.34
pD: 5.8

Table 66: HR estimates and 95% credible intervals (CrI): PFS at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.63 (0.5, 0.8)	0.74 (0.5, 1.08)	1.01 (0.57, 1.8)	1.46 (0.65, 3.28)
Pirfenidone 2403mg/day (*)	1.59 (1.25, 1.99)		1.17 (0.75, 1.82)	1.61 (0.87, 2.97)	2.31 (1, 5.36)
Nintedanib 300mg/day	1.35 (0.92, 1.98)	0.85 (0.55, 1.34)		1.37 (0.69, 2.72)	1.97 (0.81, 4.83)
NAC	0.99 (0.56, 1.75)	0.62 (0.34, 1.16)	0.73 (0.37, 1.45)		1.44 (0.53, 3.88)
Triple therapy	0.68 (0.31, 1.54)	0.43 (0.19, 1)	0.51 (0.21, 1.24)	0.7 (0.26, 1.87)	

[For SP3 pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. For all tables hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.]*

Table 67: HR estimates and 95% predictive intervals (PrI): PFS at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo	1 (0.71, 1.41)	0.63 (0.42, 0.95)	0.74 (0.45, 1.23)	1.01 (0.53, 1.95)	1.46 (0.61, 3.48)
Pirfenidone 2403mg/day (*)	1.59 (1.05, 2.36)	1 (0.71, 1.41)	1.17 (0.67, 2.04)	1.61 (0.8, 3.2)	2.32 (0.94, 5.67)
Nintedanib 300mg/day	1.35 (0.81, 2.25)	0.85 (0.49, 1.5)	1 (0.71, 1.4)	1.37 (0.65, 2.92)	1.98 (0.77, 5.1)
NAC	0.99 (0.51, 1.89)	0.62 (0.31, 1.25)	0.73 (0.34, 1.55)	1 (0.71, 1.4)	1.44 (0.51, 4.07)
Triple therapy	0.68 (0.29, 1.63)	0.43 (0.18, 1.06)	0.51 (0.2, 1.3)	0.7 (0.25, 1.97)	1 (0.71, 1.4)

Table 68: Posterior summaries of between-trial standard deviation: PFS at 52 weeks (base case network, RE model)

mean	50%	2.5%	97.5%
0.126	0.088	0.024	0.449

Table 69: Probability of column treatment being better than row treatment: PFS at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.997	0.956	0.479	0.174
Pirfenidone 2403mg/day (*)	0.003		0.193	0.059	0.026
Nintedanib 300mg/day	0.044	0.807		0.159	0.063
NAC	0.521	0.941	0.841		0.227
Triple therapy	0.826	0.974	0.937	0.773	

3) Restricted network, RE model

Total residual deviance (posterior mean): 5.95
DIC: 11.47
pD: 4.8

Table 70: HR estimates and 95% credible intervals (CrI): PFS at 52 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.63 (0.5, 0.8)	0.74 (0.5, 1.08)	1.02 (0.57, 1.8)
Pirfenidone 2403mg/day (*)	1.59 (1.25, 1.99)		1.17 (0.75, 1.82)	1.61 (0.87, 2.97)
Nintedanib 300mg/day	1.35 (0.92, 1.98)	0.85 (0.55, 1.34)		1.37 (0.69, 2.72)
NAC	0.99 (0.56, 1.74)	0.62 (0.34, 1.16)	0.73 (0.37, 1.45)	

Table 71: HR estimates and 95% predictive intervals (Prl): PFS at 52 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo	1 (0.71, 1.41)	0.63 (0.42, 0.95)	0.74 (0.44, 1.23)	1.02 (0.53, 1.95)
Pirfenidone 2403mg/day (*)	1.59 (1.05, 2.37)	1 (0.71, 1.4)	1.17 (0.67, 2.04)	1.61 (0.8, 3.21)
Nintedanib 300mg/day	1.35 (0.81, 2.25)	0.85 (0.49, 1.5)	1 (0.71, 1.41)	1.37 (0.65, 2.92)
NAC	0.99 (0.51, 1.89)	0.62 (0.31, 1.25)	0.73 (0.34, 1.55)	1 (0.71, 1.41)

Table 72: Posterior summaries of between-trial standard deviation: PFS at 52 weeks (restricted network, RE model)

mean	50%	2.5%	97.5%
0.126	0.088	0.024	0.449

Table 73: Probability of column treatment being better than row treatment: PFS at 52 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.997	0.956	0.478
Pirfenidone 2403mg/day (*)	0.003		0.193	0.059
Nintedanib 300mg/day	0.044	0.807		0.159
NAC	0.522	0.941	0.841	

4) Restricted network, FE model

Total residual deviance (posterior mean): 5.81
DIC: 10.57
pD: 4.03

Table 74: HR estimates and 95% credible intervals (Crl): PFS at 52 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.63 (0.53, 0.74)	0.74 (0.61, 0.9)	1.02 (0.63, 1.63)
Pirfenidone 2403mg/day (*)	1.59 (1.35, 1.88)		1.18 (0.91, 1.53)	1.62 (0.98, 2.67)
Nintedanib 300mg/day	1.35 (1.11, 1.65)	0.85 (0.65, 1.1)		1.37 (0.82, 2.29)
NAC	0.98 (0.61, 1.58)	0.62 (0.38, 1.02)	0.73 (0.44, 1.21)	

Table 75: Probability of column treatment being better than row treatment: PFS at 52 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		1	0.998	0.474
Pirfenidone 2403mg/day (*)	0		0.108	0.029
Nintedanib 300mg/day	0.002	0.892		0.112
NAC	0.526	0.971	0.888	

5) Base case network, FE model

Total residual deviance (posterior mean): 6.81
DIC: 12.4
pD: 5

Table 76: HR estimates and 95% credible intervals (CrI): PFS at 52 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.63 (0.53, 0.74)	0.74 (0.61, 0.9)	1.02 (0.63, 1.63)	1.46 (0.7, 3.05)
Pirfenidone 2403mg/day (*)	1.59 (1.35, 1.88)		1.18 (0.91, 1.53)	1.62 (0.98, 2.66)	2.32 (1.09, 4.94)
Nintedanib 300mg/day	1.35 (1.11, 1.65)	0.85 (0.66, 1.1)		1.37 (0.82, 2.29)	1.97 (0.92, 4.23)
NAC	0.99 (0.61, 1.58)	0.62 (0.38, 1.02)	0.73 (0.44, 1.22)		1.44 (0.6, 3.44)
Triple therapy	0.68 (0.33, 1.43)	0.43 (0.2, 0.92)	0.51 (0.24, 1.09)	0.7 (0.29, 1.67)	

Table 77: Probability of column treatment being better than row treatment: PFS at 52 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1	0.999	0.475	0.157
Pirfenidone 2403mg/day (*)	0		0.108	0.03	0.014
Nintedanib 300mg/day	0.001	0.892		0.112	0.04
NAC	0.525	0.97	0.888		0.208
Triple therapy	0.843	0.986	0.96	0.792	

6) Model comparison (PFS at 52 weeks)

Note that fit statistics can be compared only for models fitted to the same data scenario. DIC differences of less than 5 are not considered meaningful (DSU TSD 2).

Table 78: Model comparison: PFS at 52 weeks (base case network)

description	Residual deviance	DIC	pD
Base case network, RE model	6.94	13.34	5.8
Base case network, FE model	6.81	12.40	5.0

Table 79: Model comparison: PFS at 52 weeks (restricted network)

description	Residual deviance	DIC	pD
Restricted network, RE model	5.95	11.47	4.80
Restricted network, FE model	5.81	10.57	4.03

d) Progression free survival (HR at 72 weeks)

1) Introduction

The following NMA results are presented:

(a) Base case network, RE model

(b) Restricted network, RE model

(c) Restricted network, FE model

(d) Base case network, FE model

(e) Model comparison

The primary RE models used an Inv-Gamma(0.001, 0.001) prior for the between study variance.

2) Base case network, RE model

Total residual deviance (posterior mean): 6.96
DIC: 12.89
pD: 5.84

Table 80: HR estimates and 95% credible intervals (CrI): PFS at 72 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.63 (0.5, 0.78)	0.74 (0.51, 1.07)	1.02 (0.58, 1.79)	1.46 (0.65, 3.26)
Pirfenidone 2403mg/day (*)	1.59 (1.28, 1.98)		1.18 (0.76, 1.8)	1.62 (0.88, 2.97)	2.32 (1.01, 5.34)
Nintedanib 300mg/day	1.35 (0.93, 1.96)	0.85 (0.55, 1.31)		1.37 (0.7, 2.7)	1.97 (0.82, 4.77)
NAC	0.98 (0.56, 1.73)	0.62 (0.34, 1.14)	0.73 (0.37, 1.43)		1.44 (0.54, 3.84)
Triple therapy	0.69 (0.31, 1.53)	0.43 (0.19, 0.99)	0.51 (0.21, 1.23)	0.7 (0.26, 1.86)	

*[*For SP3 pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. For all tables hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.]*

Table 81: HR estimates and 95% predictive intervals (PrI): PFS at 72 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo	1 (0.72, 1.39)	0.63 (0.43, 0.93)	0.74 (0.45, 1.21)	1.02 (0.53, 1.93)	1.46 (0.62, 3.45)
Pirfenidone 2403mg/day (*)	1.59 (1.08, 2.35)	1 (0.72, 1.39)	1.18 (0.68, 2.01)	1.62 (0.82, 3.19)	2.32 (0.96, 5.64)
Nintedanib 300mg/day	1.35 (0.83, 2.21)	0.85 (0.5, 1.46)	1 (0.72, 1.39)	1.37 (0.65, 2.88)	1.97 (0.77, 5.03)
NAC	0.98 (0.52, 1.87)	0.62 (0.31, 1.22)	0.73 (0.35, 1.52)	1 (0.72, 1.39)	1.44 (0.51, 4.02)
Triple therapy	0.69 (0.29, 1.62)	0.43 (0.18, 1.05)	0.51 (0.2, 1.29)	0.7 (0.25, 1.95)	1 (0.72, 1.39)

Table 82: Posterior summaries of between-trial standard deviation: PFS at 72 weeks (base case network, RE model)

mean	50%	2.5%	97.5%
0.122	0.086	0.024	0.426

Table 83: Probability of column treatment being better than row treatment: PFS at 72 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.998	0.959	0.477	0.173
Pirfenidone 2403mg/day (*)	0.002		0.182	0.054	0.024
Nintedanib 300mg/day	0.041	0.818		0.156	0.062
NAC	0.523	0.946	0.844		0.228
Triple therapy	0.827	0.976	0.938	0.772	

3) Restricted network, RE model

Total residual deviance (posterior mean): 5.95

DIC: 10.99

pD: 4.82

Table 84: HR estimates and 95% credible intervals (CrI): PFS at 72 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.63 (0.5, 0.78)	0.74 (0.51, 1.07)	1.02 (0.58, 1.79)
Pirfenidone 2403mg/day (*)	1.59 (1.28, 1.98)		1.18 (0.77, 1.81)	1.62 (0.88, 2.97)
Nintedanib 300mg/day	1.35 (0.93, 1.96)	0.85 (0.55, 1.31)		1.37 (0.7, 2.69)
NAC	0.98 (0.56, 1.73)	0.62 (0.34, 1.14)	0.73 (0.37, 1.43)	

Table 85: HR estimates and 95% predictive intervals (PrI): PFS at 72 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo	1 (0.72, 1.39)	0.63 (0.43, 0.93)	0.74 (0.45, 1.21)	1.02 (0.53, 1.93)
Pirfenidone 2403mg/day (*)	1.59 (1.08, 2.34)	1 (0.72, 1.39)	1.18 (0.69, 2.02)	1.62 (0.82, 3.19)
Nintedanib 300mg/day	1.35 (0.83, 2.22)	0.85 (0.5, 1.46)	1 (0.72, 1.39)	1.37 (0.65, 2.88)
NAC	0.98 (0.52, 1.87)	0.62 (0.31, 1.22)	0.73 (0.35, 1.52)	1 (0.72, 1.39)

Table 86: Posterior summaries of between-trial standard deviation: PFS at 72 weeks (restricted network, RE model)

mean	50%	2.5%	97.5%
0.124	0.087	0.024	0.431

Table 87: Probability of column treatment being better than row treatment: PFS at 72 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.998	0.958	0.476
Pirfenidone 2403mg/day (*)	0.002		0.182	0.055
Nintedanib 300mg/day	0.042	0.818		0.157
NAC	0.524	0.945	0.843	

4) Restricted network, FE model

Total residual deviance (posterior mean): 5.87
DIC: 10.12
pD: 4.03

Table 88: HR estimates and 95% credible intervals (CrI): PFS at 72 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.63 (0.54, 0.73)	0.74 (0.61, 0.9)	1.02 (0.63, 1.63)
Pirfenidone 2403mg/day (*)	1.6 (1.36, 1.87)		1.18 (0.92, 1.52)	1.62 (0.99, 2.66)
Nintedanib 300mg/day	1.35 (1.11, 1.65)	0.85 (0.66, 1.09)		1.37 (0.82, 2.3)
NAC	0.98 (0.61, 1.58)	0.62 (0.38, 1.01)	0.73 (0.44, 1.21)	

Table 89: Probability of column treatment being better than row treatment: PFS at 72 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		1	0.998	0.474
Pirfenidone 2403mg/day (*)	0		0.101	0.028
Nintedanib 300mg/day	0.002	0.899		0.112
NAC	0.526	0.972	0.888	

5) Base case network, FE model

Total residual deviance (posterior mean): 6.87
DIC: 11.98
pD: 5.01

Table 90: HR estimates and 95% credible intervals (CrI): PFS at 72 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.63 (0.54, 0.73)	0.74 (0.61, 0.9)	1.02 (0.63, 1.63)	1.46 (0.7, 3.04)
Pirfenidone 2403mg/day (*)	1.6 (1.36, 1.87)		1.18 (0.92, 1.52)	1.62 (0.99, 2.66)	2.33 (1.1, 4.94)
Nintedanib 300mg/day	1.35 (1.11, 1.65)	0.85 (0.66, 1.09)		1.37 (0.82, 2.29)	1.97 (0.92, 4.23)
NAC	0.99 (0.61, 1.58)	0.62 (0.38, 1.01)	0.73 (0.44, 1.22)		1.44 (0.6, 3.44)
Triple therapy	0.68 (0.33, 1.43)	0.43 (0.2, 0.91)	0.51 (0.24, 1.08)	0.7 (0.29, 1.66)	

Table 91: Probability of column treatment being better than row treatment: PFS at 72 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1	0.998	0.475	0.156
Pirfenidone 2403mg/day (*)	0		0.1	0.028	0.014
Nintedanib 300mg/day	0.002	0.9		0.112	0.04
NAC	0.525	0.972	0.888		0.207
Triple therapy	0.844	0.986	0.96	0.793	

6) Model comparison (PFS at 72 weeks)

Note that fit statistics can be compared only for models fitted to the same data scenario. DIC differences of less than 5 are not considered meaningful (DSU TSD 2).

Table 92: Model comparison: PFS at 72 weeks (base case network)

description	Residual deviance	DIC	pD
Base case network, RE model	6.96	12.89	5.84
Base case network, FE model	6.87	11.98	5.01

Table 93: Model comparison: PFS at 72 weeks (restricted network)

description	Residual deviance	DIC	pD
Restricted network, RE model	5.95	10.99	4.82
Restricted network, FE model	5.87	10.12	4.03

e) IPF-related mortality (HR at 52 weeks)

1) Introduction

The following NMA results are presented:

(a) Base case network, RE model

(b) Restricted network, RE model

(c) Restricted network, FE model

(d) Base case network, FE model

(e) Model comparison

The primary RE models used an Inv-Gamma(0.001, 0.001) prior for the between study variance.

2) Base case network, RE model

Total residual deviance (posterior mean): 8.81
DIC: 31.51
pD: 7.25

Table 94: Hazard ratios and 95% credible intervals (CrI): IPF-related mortality at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.36 (0.14, 0.9)	0.6 (0.22, 1.32)	1.74 (0.31, 12.05)	10.27 (1.21, 343.41)
Pirfenidone 2403mg/day	2.74 (1.11, 6.97)		1.63 (0.43, 5.51)	4.8 (0.67, 41.03)	28.7 (2.75, 1064.04)
Nintedanib 300mg/day	1.67 (0.76, 4.61)	0.61 (0.18, 2.33)		2.94 (0.44, 26.06)	17.61 (1.77, 677.5)
NAC	0.57 (0.08, 3.27)	0.21 (0.02, 1.48)	0.34 (0.04, 2.26)		6.1 (0.33, 291.99)
Triple therapy	0.1 (0, 0.82)	0.03 (0, 0.36)	0.06 (0, 0.56)	0.16 (0, 3.04)	

*[*For all tables hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.]*

Table 95: Hazard ratios and 95% predictive intervals (PrI): IPF-related mortality at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo	1 (0.36, 2.76)	0.37 (0.09, 1.33)	0.6 (0.14, 2)	1.74 (0.24, 14.94)	10.34 (1.01, 397.75)
Pirfenidone 2403mg/day	2.74 (0.75, 10.63)	1 (0.36, 2.76)	1.64 (0.3, 7.4)	4.8 (0.55, 49.78)	28.81 (2.34, 1237.34)
Nintedanib 300mg/day	1.66 (0.5, 7.38)	0.61 (0.13, 3.29)	1 (0.36, 2.76)	2.93 (0.37, 32.51)	17.66 (1.53, 786.46)
NAC	0.57 (0.07, 4.12)	0.21 (0.02, 1.8)	0.34 (0.03, 2.75)	1 (0.37, 2.76)	6.13 (0.29, 329.38)
Triple therapy	0.1 (0, 0.98)	0.03 (0, 0.43)	0.06 (0, 0.65)	0.16 (0, 3.51)	1 (0.36, 2.77)

Table 96: Posterior summaries of between-trial standard deviation: IPF-related mortality at 52 weeks (base case network, RE model)

mean	50%	2.5%	97.5%
0.333	0.187	0.028	1.44

Table 97: Probability of column treatment being better than row treatment: IPF-related mortality at 52 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.983	0.921	0.254	0.017
Pirfenidone 2403mg/day	0.017		0.195	0.053	0.005
Nintedanib 300mg/day	0.079	0.805		0.119	0.009
NAC	0.746	0.947	0.881		0.109
Triple therapy	0.983	0.995	0.991	0.891	

3) Restricted network, RE model

Total residual deviance (posterior mean): 5.36
DIC: 18.76
pD: 4.57

Table 98: Hazard ratios and 95% credible intervals (CrI): IPF-related mortality at 52 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.36 (0.13, 0.99)	0.74 (0.18, 3.05)	1.74 (0.26, 14.77)
Pirfenidone 2403mg/day	2.75 (1.01, 7.82)		2.03 (0.37, 11.93)	4.8 (0.57, 51.99)
Nintedanib 300mg/day	1.35 (0.33, 5.57)	0.49 (0.08, 2.68)		2.36 (0.23, 29.77)
NAC	0.57 (0.07, 3.82)	0.21 (0.02, 1.75)	0.42 (0.03, 4.35)	

Table 99: Hazard ratios and 95% predictive intervals (PrI): IPF-related mortality at 52 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo	1 (0.26, 3.77)	0.37 (0.07, 1.81)	0.74 (0.11, 5.25)	1.74 (0.19, 20.78)
Pirfenidone 2403mg/day	2.74 (0.55, 14.91)	1 (0.26, 3.8)	2.03 (0.24, 18.66)	4.8 (0.42, 70.4)
Nintedanib 300mg/day	1.35 (0.19, 9.53)	0.49 (0.05, 4.19)	1 (0.26, 3.8)	2.36 (0.17, 40.33)
NAC	0.57 (0.05, 5.39)	0.21 (0.01, 2.37)	0.42 (0.03, 5.95)	1 (0.26, 3.83)

Table 100: Posterior summaries of between-trial standard deviation: IPF-related mortality at 52 weeks (restricted network, RE model)

mean	50%	2.5%	97.5%
0.48	0.197	0.029	2.207

Table 101: Probability of column treatment being better than row treatment: IPF-related mortality at 52 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.975	0.759	0.258
Pirfenidone 2403mg/day	0.025		0.134	0.06
Nintedanib 300mg/day	0.241	0.866		0.189
NAC	0.742	0.94	0.811	

4) Restricted network, FE model

Total residual deviance (posterior mean): 5.28
DIC: 18.28
pD: 4.17

Table 102: Hazard ratios and 95% credible intervals (CrI): IPF-related mortality at 52 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.37 (0.18, 0.76)	0.74 (0.41, 1.34)	1.73 (0.4, 9.04)
Pirfenidone 2403mg/day	2.72 (1.31, 5.63)		2.01 (0.79, 5.14)	4.73 (0.93, 28.58)
Nintedanib 300mg/day	1.35 (0.75, 2.44)	0.5 (0.19, 1.27)		2.35 (0.49, 13.43)
NAC	0.58 (0.11, 2.47)	0.21 (0.03, 1.08)	0.43 (0.07, 2.06)	

Table 103: Probability of column treatment being better than row treatment: IPF-related mortality at 52 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.996	0.841	0.23
Pirfenidone 2403mg/day	0.004		0.072	0.031
Nintedanib 300mg/day	0.159	0.928		0.144
NAC	0.77	0.969	0.856	

5) Base case network, FE model

Total residual deviance (posterior mean): 9.08
DIC: 31.06
pD: 6.53

Table 104: Hazard ratios and 95% credible intervals (CrI): IPF-related mortality at 52 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.37 (0.18, 0.76)	0.63 (0.37, 1.07)	1.73 (0.4, 9.05)	10.2 (1.53, 298.14)
Pirfenidone 2403mg/day	2.72 (1.31, 5.64)		1.7 (0.69, 4.21)	4.73 (0.92, 28.51)	28.11 (3.57, 862.42)
Nintedanib 300mg/day	1.59 (0.93, 2.72)	0.59 (0.24, 1.45)		2.77 (0.59, 15.66)	16.38 (2.24, 490.56)
NAC	0.58 (0.11, 2.47)	0.21 (0.04, 1.08)	0.36 (0.06, 1.71)		6.06 (0.47, 220.8)
Triple therapy	0.1 (0, 0.66)	0.04 (0, 0.28)	0.06 (0, 0.45)	0.16 (0, 2.14)	

Table 105: Probability of column treatment being better than row treatment: IPF-related mortality at 52 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.996	0.956	0.23	0.006
Pirfenidone 2403mg/day	0.004		0.124	0.031	0
Nintedanib 300mg/day	0.044	0.876		0.099	0.002
NAC	0.77	0.969	0.901		0.088
Triple therapy	0.994	1	0.998	0.912	

6) Model comparison (IPF-related mortality at 52 weeks)

Note that fit statistics can be compared only for models fitted to the same data scenario. DIC differences of less than 5 are not considered meaningful (DSU TSD 2).

Table 106: Model comparison: IPF-related mortality at 52 weeks (base case network)

description	Residual deviance	DIC	pD
Base case network, RE model	8.81	31.51	7.25
Base case network, FE model	9.08	31.06	6.53

Table 107: Model comparison: IPF-related mortality at 52 weeks (restricted network)

description	Residual deviance	DIC	pD
Restricted network, RE model	5.36	18.76	4.57
Restricted network, FE model	5.28	18.28	4.17

f) IPF-related mortality (HR at 72 weeks)

1) Introduction

The following NMA results are presented:

(a) Base case network, RE model

(b) Restricted network, RE model

(c) Restricted network, FE model

(d) Base case network, FE model

(e) Model comparison

The primary RE models used an Inv-Gamma(0.001, 0.001) prior for the between study variance.

2) Base case network, RE model

Total residual deviance (posterior mean): 8.8

DIC: 30.2
pD: 7.27

Table 108: Hazard ratios and 95% credible intervals (CrI): IPF-related mortality at 72 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.48 (0.22, 1.01)	0.6 (0.23, 1.28)	1.74 (0.32, 11.3)	10.27 (1.23, 334.52)
Pirfenidone 2403mg/day	2.09 (0.99, 4.63)		1.25 (0.38, 3.71)	3.65 (0.57, 27.99)	21.78 (2.26, 764.79)
Nintedanib 300mg/day	1.67 (0.78, 4.32)	0.8 (0.27, 2.65)		2.93 (0.46, 23.97)	17.5 (1.82, 644.67)
NAC	0.58 (0.09, 3.16)	0.27 (0.04, 1.75)	0.34 (0.04, 2.18)		6.11 (0.35, 279.98)
Triple therapy	0.1 (0, 0.81)	0.05 (0, 0.44)	0.06 (0, 0.55)	0.16 (0, 2.88)	

*[*For all tables hazard ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.]*

Table 109: Estimates and 95% predictive intervals (PrI): IPF-related mortality at 72 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo	1 (0.4, 2.51)	0.48 (0.14, 1.49)	0.6 (0.15, 1.83)	1.74 (0.26, 13.65)	10.34 (1.06, 380.49)
Pirfenidone 2403mg/day	2.08 (0.67, 7.02)	1 (0.4, 2.53)	1.25 (0.27, 4.94)	3.65 (0.48, 33.3)	21.86 (1.97, 869.99)
Nintedanib 300mg/day	1.66 (0.55, 6.54)	0.8 (0.2, 3.63)	1 (0.4, 2.53)	2.93 (0.39, 28.74)	17.59 (1.6, 733.08)
NAC	0.57 (0.07, 3.86)	0.27 (0.03, 2.09)	0.34 (0.03, 2.58)	1 (0.39, 2.52)	6.12 (0.31, 312.84)
Triple therapy	0.1 (0, 0.95)	0.05 (0, 0.51)	0.06 (0, 0.63)	0.16 (0, 3.25)	1 (0.4, 2.53)

Table 110: Posterior summaries of between-trial standard deviation: IPF-related mortality at 72 weeks (base case network, RE model)

mean	50%	2.5%	97.5%
0.306	0.179	0.028	1.293

Table 111: Probability of column treatment being better than row treatment: IPF-related mortality at 72 weeks (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.974	0.925	0.251	0.016
Pirfenidone 2403mg/day	0.026		0.327	0.077	0.006
Nintedanib 300mg/day	0.075	0.673		0.116	0.008
NAC	0.749	0.923	0.884		0.107
Triple therapy	0.984	0.994	0.992	0.893	

3) Restricted network, RE model

Total residual deviance (posterior mean): 5.35
DIC: 17.46
pD: 4.61

Table 112: Estimates and 95% credible intervals (CrI): IPF-related mortality at 72 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.48 (0.2, 1.08)	0.74 (0.22, 2.5)	1.73 (0.28, 12.6)
Pirfenidone 2403mg/day	2.09 (0.93, 5.04)		1.54 (0.37, 7.13)	3.64 (0.5, 31.75)
Nintedanib 300mg/day	1.35 (0.4, 4.59)	0.65 (0.14, 2.73)		2.35 (0.27, 23.5)
NAC	0.58 (0.08, 3.57)	0.27 (0.03, 2.01)	0.42 (0.04, 3.75)	

Table 113: Estimates and 95% predictive intervals (PrI): IPF-related mortality at 72 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo	1 (0.32, 3.09)	0.48 (0.11, 1.81)	0.74 (0.14, 3.91)	1.73 (0.21, 16.05)
Pirfenidone 2403mg/day	2.08 (0.55, 8.9)	1 (0.32, 3.08)	1.54 (0.25, 10.41)	3.64 (0.39, 40.16)
Nintedanib 300mg/day	1.35 (0.26, 7.11)	0.65 (0.1, 3.93)	1 (0.32, 3.11)	2.35 (0.21, 29.23)
NAC	0.58 (0.06, 4.68)	0.27 (0.02, 2.55)	0.42 (0.03, 4.74)	1 (0.32, 3.08)

Table 114: Posterior summaries of between-trial standard deviation: IPF-related mortality at 72 weeks (restricted network, RE model)

mean	50%	2.5%	97.5%
0.389	0.182	0.028	1.824

Table 115: Probability of column treatment being better than row treatment: IPF-related mortality at 72 weeks (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.967	0.769	0.256
Pirfenidone 2403mg/day	0.033		0.215	0.084
Nintedanib 300mg/day	0.231	0.785		0.185
NAC	0.744	0.916	0.815	

4) Restricted network, FE model

Total residual deviance (posterior mean): 5.34
DIC: 17.01
pD: 4.17

Table 116: Estimates and 95% credible intervals (CrI): IPF-related mortality at 72 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.49 (0.27, 0.87)	0.74 (0.41, 1.34)	1.73 (0.41, 9.07)
Pirfenidone 2403mg/day	2.06 (1.15, 3.67)		1.52 (0.66, 3.48)	3.57 (0.75, 20.43)
Nintedanib 300mg/day	1.35 (0.75, 2.44)	0.66 (0.29, 1.5)		2.35 (0.49, 13.54)
NAC	0.58 (0.11, 2.46)	0.28 (0.05, 1.34)	0.43 (0.07, 2.05)	

Table 117: Probability of column treatment being better than row treatment: IPF-related mortality at 72 weeks (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.993	0.841	0.23
Pirfenidone 2403mg/day	0.007		0.161	0.055
Nintedanib 300mg/day	0.159	0.839		0.144
NAC	0.77	0.945	0.856	

5) Base case network, FE model

Total residual deviance (posterior mean): 9.13
DIC: 29.77
pD: 6.52

Table 118: Estimates and 95% credible intervals (CrI): IPF-related mortality at 72 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.49 (0.27, 0.87)	0.63 (0.37, 1.07)	1.73 (0.4, 9.04)	10.13 (1.52, 293.32)
Pirfenidone 2403mg/day	2.05 (1.15, 3.67)		1.29 (0.59, 2.84)	3.56 (0.74, 20.37)	21.04 (2.84, 629.57)
Nintedanib 300mg/day	1.59 (0.93, 2.72)	0.78 (0.35, 1.71)		2.76 (0.58, 15.63)	16.31 (2.23, 485.66)
NAC	0.58 (0.11, 2.47)	0.28 (0.05, 1.35)	0.36 (0.06, 1.71)		6.05 (0.46, 219.26)
Triple therapy	0.1 (0, 0.66)	0.05 (0, 0.35)	0.06 (0, 0.45)	0.17 (0, 2.15)	

Table 119: Probability of column treatment being better than row treatment: IPF-related mortality at 72 weeks (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.992	0.956	0.23	0.006
Pirfenidone 2403mg/day	0.008		0.264	0.056	0.001
Nintedanib 300mg/day	0.044	0.736		0.1	0.002
NAC	0.77	0.944	0.9		0.088
Triple therapy	0.994	0.999	0.998	0.912	

6) Model comparison (IPF-related mortality at 72 weeks)

Note that fit statistics can be compared only for models fitted to the same data scenario. DIC differences of less than 5 are not considered meaningful (DSU TSD 2).

Table 120: Model comparison: IPF-related mortality at 72 weeks (base case network)

description	Residual deviance	DIC	pD
Base case network, RE model	8.80	30.20	7.27
Base case network, FE model	9.13	29.77	6.52

Table 121: Model comparison: IPF-related mortality at 72 weeks (restricted network)

description	Residual deviance	DIC	pD
Restricted network, RE model	5.35	17.46	4.61
Restricted network, FE model	5.34	17.01	4.17

g) Acute exacerbations

1) Introduction

The following NMA results are presented:

(a) Base case network, RE model

(b) Restricted network, RE model

(c) Restricted network, FE model

(d) Base case network, FE model

(e) Model comparison

The primary RE models used a LogNorml(-3.02, 1.85^2) prior based on Turner et al. (2012) for the between study variance. This prior has median at 0.05 and 95% interval [0.001, 1.83].

2) Base case network, RE model

Total residual deviance (posterior mean): 20.33
DIC: 89.58
pD: 15.66

Table 122: Odds ratios and 95% credible intervals (CrI): Acute exacerbations (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.62 (0.29, 1.39)	0.55 (0.26, 1.08)	0.99 (0.13, 7.3)
Pirfenidone 2403mg/day (*)	1.6 (0.72, 3.46)		0.88 (0.29, 2.43)	1.58 (0.18, 13.38)
Nintedanib 300mg/day	1.83 (0.92, 3.83)	1.14 (0.41, 3.43)		1.81 (0.22, 15.32)
NAC	1.01 (0.14, 7.47)	0.63 (0.07, 5.41)	0.55 (0.07, 4.55)	

[For SP3 pirfenidone 1800mg/day is assumed to be equivalent to pirfenidone 2403mg/day. For all tables odds ratios are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment.]*

Table 123: Odds ratios and 95% predictive intervals (PrI): Acute exacerbations (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo	1 (0.39, 2.6)	0.62 (0.19, 2.17)	0.55 (0.16, 1.73)	0.99 (0.11, 8.84)
Pirfenidone 2403mg/day (*)	1.61 (0.46, 5.32)	1 (0.39, 2.59)	0.88 (0.2, 3.43)	1.58 (0.15, 15.94)
Nintedanib 300mg/day	1.82 (0.58, 6.15)	1.14 (0.29, 4.95)	1 (0.39, 2.59)	1.8 (0.19, 18.46)
NAC	1.01 (0.11, 9.06)	0.63 (0.06, 6.54)	0.56 (0.05, 5.46)	1 (0.39, 2.59)

Table 124: Posterior summaries of between-trial standard deviation: Acute exacerbations (base case network, RE model)

mean	50%	2.5%	97.5%
0.357	0.289	0.044	1.069

Table 125: Probability of column treatment being better than row treatment: Acute exacerbations (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.902	0.962	0.506
Pirfenidone 2403mg/day (*)	0.098		0.612	0.328
Nintedanib 300mg/day	0.038	0.388		0.28
NAC	0.494	0.672	0.72	

3) Restricted network, RE model

Total residual deviance (posterior mean): 18.32
DIC: 78.91
pD: 14.34

Table 126: Odds ratios and 95% credible intervals (CrI): Acute exacerbations (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.63 (0.28, 1.47)	0.63 (0.26, 1.54)	0.98 (0.13, 7.73)
Pirfenidone 2403mg/day (*)	1.6 (0.68, 3.61)		1.01 (0.3, 3.36)	1.56 (0.17, 14.31)
Nintedanib 300mg/day	1.59 (0.65, 3.81)	0.99 (0.3, 3.39)		1.55 (0.16, 14.58)
NAC	1.02 (0.13, 8)	0.64 (0.07, 5.92)	0.64 (0.07, 6.06)	

Table 127: Odds ratios and 95% predictive intervals (PrI): Acute exacerbations (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo	1 (0.35, 2.87)	0.62 (0.17, 2.47)	0.63 (0.16, 2.51)	0.98 (0.1, 9.74)
Pirfenidone 2403mg/day (*)	1.61 (0.41, 6.01)	1 (0.35, 2.88)	1.01 (0.2, 4.91)	1.56 (0.13, 17.55)
Nintedanib 300mg/day	1.59 (0.4, 6.22)	0.99 (0.2, 5.04)	1 (0.34, 2.89)	1.55 (0.13, 18.17)
NAC	1.02 (0.1, 10.06)	0.64 (0.06, 7.36)	0.65 (0.06, 7.54)	1 (0.35, 2.9)

Table 128: Posterior summaries of between-trial standard deviation: Acute exacerbations (restricted network, RE model)

mean	50%	2.5%	97.5%
0.395	0.311	0.045	1.241

Table 129: Probability of column treatment being better than row treatment: Acute exacerbations (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.892	0.883	0.51
Pirfenidone 2403mg/day (*)	0.108		0.494	0.335
Nintedanib 300mg/day	0.117	0.506		0.338
NAC	0.49	0.665	0.662	

4) Restricted network, FE model

Total residual deviance (posterior mean): 21.34
DIC: 78.6
pD: 11.01

Table 130: Odds ratios and 95% credible intervals (CrI): Acute exacerbations (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.61 (0.37, 0.98)	0.62 (0.37, 1.04)	0.99 (0.17, 5.9)
Pirfenidone 2403mg/day (*)	1.65 (1.03, 2.67)		1.03 (0.51, 2.08)	1.63 (0.26, 10.4)
Nintedanib 300mg/day	1.61 (0.96, 2.68)	0.97 (0.48, 1.96)		1.58 (0.25, 10.16)
NAC	1.01 (0.17, 5.99)	0.61 (0.1, 3.86)	0.63 (0.1, 4)	

Table 131: Probability of column treatment being better than row treatment: Acute exacerbations (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.98	0.965	0.506
Pirfenidone 2403mg/day (*)	0.02		0.471	0.295
Nintedanib 300mg/day	0.035	0.529		0.306
NAC	0.494	0.705	0.694	

5) Base case network, FE model

Total residual deviance (posterior mean): 23.23
DIC: 88.63
pD: 11.81

Table 132: Odds ratios and 95% credible intervals (CrI): Acute exacerbations (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		0.61 (0.37, 0.98)	0.56 (0.35, 0.89)	0.98 (0.17, 5.8)
Pirfenidone 2403mg/day (*)	1.65 (1.03, 2.68)		0.92 (0.47, 1.8)	1.62 (0.26, 10.23)
Nintedanib 300mg/day	1.79 (1.12, 2.86)	1.09 (0.56, 2.12)		1.76 (0.28, 11.05)
NAC	1.02 (0.17, 6)	0.62 (0.1, 3.87)	0.57 (0.09, 3.56)	

Table 133: Probability of column treatment being better than row treatment: Acute exacerbations (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day (*)	Nintedanib 300mg/day	NAC
Placebo		0.981	0.993	0.508
Pirfenidone 2403mg/day (*)	0.019		0.595	0.295
Nintedanib 300mg/day	0.007	0.405		0.265
NAC	0.492	0.705	0.735	

6) Model comparison (acute exacerbations)

Note that fit statistics can be compared only for models fitted to the same data scenario. DIC differences of less than 5 are not considered meaningful (DSU TSD 2).

Table 134: Model comparison: Acute exacerbations (base case network)

description	Residual deviance	DIC	pD
Base case network, RE model	20.33	89.58	15.66
Base case network, FE model	23.23	88.63	11.81

Table 135: Model comparison: Acute exacerbations (restricted network)

description	Residual deviance	DIC	pD
Restricted network, RE model	18.32	78.91	14.34
Restricted network, FE model	21.34	78.60	11.01

Appendix D: Adverse event NMAs

a) Introduction

In line with our response to A39, in order to allow for a more thorough exploration of how adverse event (AE) rates differ between treatments, additional NMAs were run. These were also used to inform results from the updated economic model. Inputs into the NMA are based on outcomes reported in the Phase III studies, as described below.

Network meta-analyses were run for the following key AEs of interest, in line with the approach used by the manufacturer of nintedanib in their submission to NICE and the recommendation in this same submission that AEs of particular focus to clinicians with respect to pirfenidone are photosensitivity and rash (p.97) [Boehringer Ingelheim 2015]:

1) Diarrhoea

2) Rash

3) Discontinuation due to adverse event

4) Serious cardiac adverse events

With specific regard to the NMA for discontinuation due to AE, there were two potential sources of data for nintedanib: the NICE manufacturer submission document or published data. To maintain a consistent approach with that used in the submission, data from the manufacturer submission were used as the primary source [Boehringer Ingelheim 2015]. In the case of any discrepancy between the literature and the manufacturer submission, the submission data were used.

For each NMA the fit was performed on the log-OR scale. We then also calculated relative risks (RRs). The reference AE rate used to calculate the RRs was the average placebo AE rate from all studies in the NMA which had placebo as baseline arm. All studies providing data on each endpoint/outcome were placebo controlled. To maintain consistency with Section 4.10 of the submission document, the NMAs presented below use the same base case evidence network (and random effects [RE] models) to generate results. Sensitivity analyses are also presented for each NMA, again in line with the approach used in the submission (Appendix 14 of the submission). [As a reminder, the base case network contains one study more than the restricted network (TOMORROW), along with a comparison to triple therapy (from PANTHER)].

The following NMA results are presented for each AE of interest:

- Base case network, RE models
- Restricted network, RE models
- Restricted network, fixed effect (FE) models
- Base case network, FE models

Common to each NMA, the prior for between-trial-variance was based on the publication by Turner et al [2012].

For the NMAs for diarrhoea, rash and serious cardiac AEs, we classified the events as subjective endpoints. This means we used the Log-normal(-2.13, 1.58^2) as prior distribution for the between-study-variance in each of these NMAs. Recognising treatment discontinuation due to AEs as semi-objective endpoint, we used the Log-normal(-3.02, 1.85^2) as prior distribution for the between-study-variance.

The (full) analysis data set used in the JAGS fits for each AE of interest are shown in Table 136.

Table 136: Data inputs for each AE NMA

Study	Treatment	Baseline	N	Number of events:			
				Diarrhoea	Rash	Discontinuation due to AE	Serious cardiac AE
ASCEND	Pirfenidone 2403mg/day (*)	Placebo	278	62	78	40	16
ASCEND	Placebo	Placebo	277	60	24	30	11
CAPACITY 1	Pirfenidone 2403mg/day (*)	Placebo	171	56	58	NR	NR
CAPACITY 1	Placebo	Placebo	173	37	22	NR	NR
CAPACITY 2	Pirfenidone 2403mg/day (*)	Placebo	174	43	53	NR	NR
CAPACITY 2	Placebo	Placebo	174	30	18	NR	NR
CAPACITY 1 & CAPACITY 2*	Pirfenidone 2403mg/day (*)	Placebo	345	-	-	51	21
CAPACITY 1 & CAPACITY 2*	Placebo	Placebo	347	-	-	30	17
SP3	Pirfenidone 2403mg/day (*)	Placebo	108	NR	NR	20	NR
SP3	Placebo	Placebo	104	NR	NR	14	NR
INPULSIS 1	Nintedanib 300mg/day	Placebo	309	190	16	65	NR
INPULSIS 1	Placebo	Placebo	204	38	6	22	NR
INPULSIS 2	Nintedanib 300mg/day	Placebo	329	208	NR	58	NR
INPULSIS 2	Placebo	Placebo	219	40	NR	33	NR
INPULSIS 1 & INPULSIS 2*	Nintedanib 300mg/day	Placebo	638	-	-	NR	32
INPULSIS 1 & INPULSIS 2*	Placebo	Placebo	423	-	-	NR	23
TOMORROW	Nintedanib 300mg/day	Placebo	85	47	2	26	1
TOMORROW	Placebo	Placebo	85	13	4	22	7
PANTHER (NAC)	NAC	Placebo	133	18	NR	1	9
PANTHER (NAC)	Placebo	Placebo	131	15	NR	4	2
PANTHER (Triple therapy)	Triple therapy	Placebo	77	6	13	NR	3
PANTHER (Triple therapy)	Placebo	Placebo	78	7	4	NR	0

* Results from individual studies are used where available; pooled study data are used where results from individual studies are not available
NR: not reported
Sources: *Diarrhoea*: main publications for ASCEND, CAPACITY 1, CAPACITY 2, INPULSIS 1, INPULSIS 2 and TOMORROW [King 2014, Noble 2011, Richeldi 2011 and Richeldi 2014]. PANTHER (NAC and triple therapy arms) data extracted from ClinicalTrials.gov [NCT00650091]
Rash: main publications for ASCEND and CAPACITY [King 2014, Noble 2011]. CAPACITY data as presented in Table 119 of manufacturer's NICE submission [Boehringer Ingelheim 2015]. INPULSIS I (NCT01335464) and TOMORROW (NCT00514683) data from ClinicalTrials.gov. For PANTHER (triple therapy arm), specific data for rash were not reported – we have extracted data for 'Skin' AEs as an alternative (table S3 in Raghu 2012 Suppl).

Study	Treatment	Baseline	N	Number of events:			
				Diarrhoea	Rash	Discontinuation due to AE	Serious cardiac AE
Discontinuation due to AE: Whenever possible discontinuation of treatment due to AE was used. ASCEND: discontinuation of treatment due to AE [King 2014]. CAPACITY: discontinuation of treatment due to AE [Noble 2011]. Note, ASCEND & CAPACITY data taken from publication text vs. patient flow diagrams, consistent with approach used in NICE nintedanib manufacturer submission[1]. SP3: discontinuation of treatment due to AE [Taniguchi 2010]. INPULSIS 1 & INPULSIS 2, TOMORROW: discontinuation of treatment due to AE used as per NICE manufacturer submission for nintedanib [Boehringer Ingelheim 2015]. PANTHER (NAC): the nintedanib NICE manufacturer submission is based on Appendix table S3. However, the clinical trial record reported the 1 / 133 for NAC and 4 / 131 for placebo under the amended study design & the latter has been used (https://clinicaltrials.gov/ct2/show/results/NCT00650091?term=NCT00650091&rank=1&sect=X30156#evnt). *Serious cardiac AE*: ASCEND; clinicaltrials.org [NCT01366209: Serious AE section of the Study Results tab]. Please note that the manufacturer of nintedanib did not extract serious cardiac events for ASCEND in their NICE submission [Boehringer Ingelheim 2015]. Pooled CAPACITY; InterMune submission[2] (Table B5.9.2 and B5.9 [InterMune 2011]). Pooled INPULSIS; Appendix Table S10 [Richeldi 2014 Suppl]. PANTHER (NAC); Table 3 [Martinez 2014]. PANTHER (triple therapy); Table 3 reports Cardiac as serious AE [Raghu 2012]. TOMORROW; [Richeldi 2011].							

[1] Discrepancy due to recording of "primary" vs "not primary" reason for treatment discontinuation. To summarise: ASCEND [King 2014] flow diagram pirfenidone: 39 (=35+4); placebo: 29 (=24+5). Text: pirfenidone: 40; placebo: 30. CAPACITY 1 & 2 [Noble 2011] flow diagram: pirfenidone: 26 (=21+5) + 25 (=24+1) [total=51]; placebo: 23 (=14+9) + 25 (=14+11) [total=48]. Text: pirfenidone: 51; placebo: 30

[2] Discrepancy between clinicaltrials.gov data and the InterMune STA data. Data from InterMune STA submission used [InterMune 2011], consistent with nintedanib NICE manufacturer submission [Boehringer Ingelheim 2015]

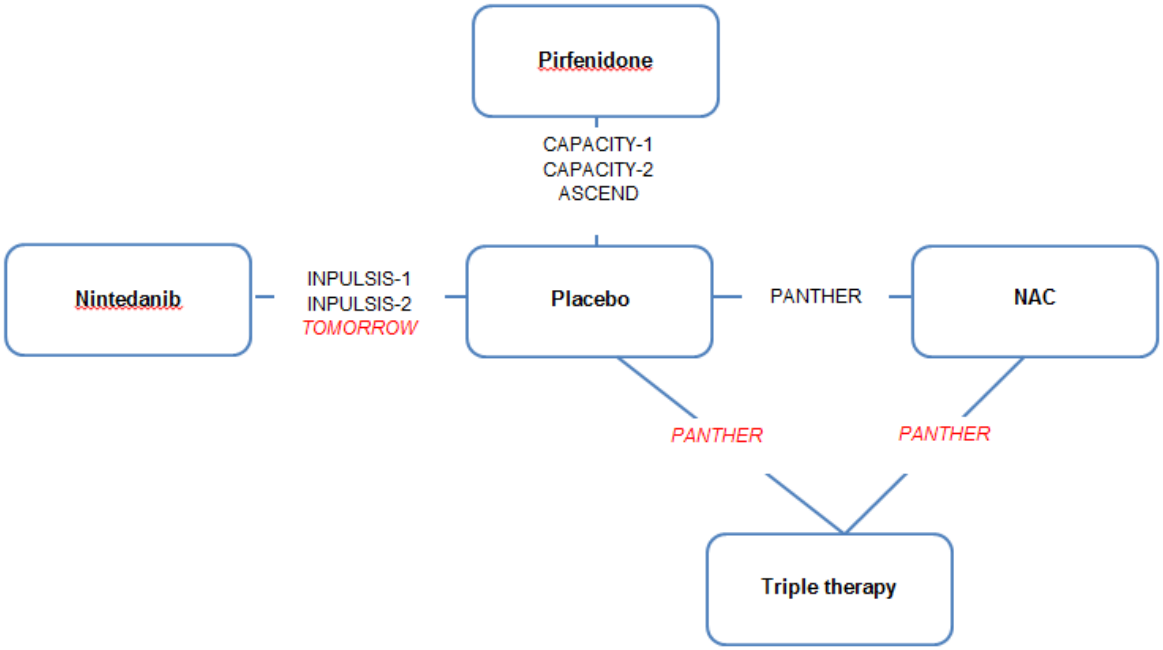
Table 137: Summary of evidence considered in the adverse events NMAs

Trials	Study duration (weeks)	Diarrhoea	Rash	Discontinuation of treatment due to AE	Serious cardiac events
CAPACITY1 and CAPACITY 2	72	√	√	√*	√*
ASCEND	52	√	√	√	√
SP3	52			√	
PANTHER (NAC)	60	√		√	√
PANTHER (Triple)	32	√	√		√
INPULSIS1 and INPULSIS2	52	√	√ (only INPULSIS 1)	√	√*
TOMORROW	52	√	√	√	√
* Pooled trials					

b) Network diagrams

1) Diarrhoea:

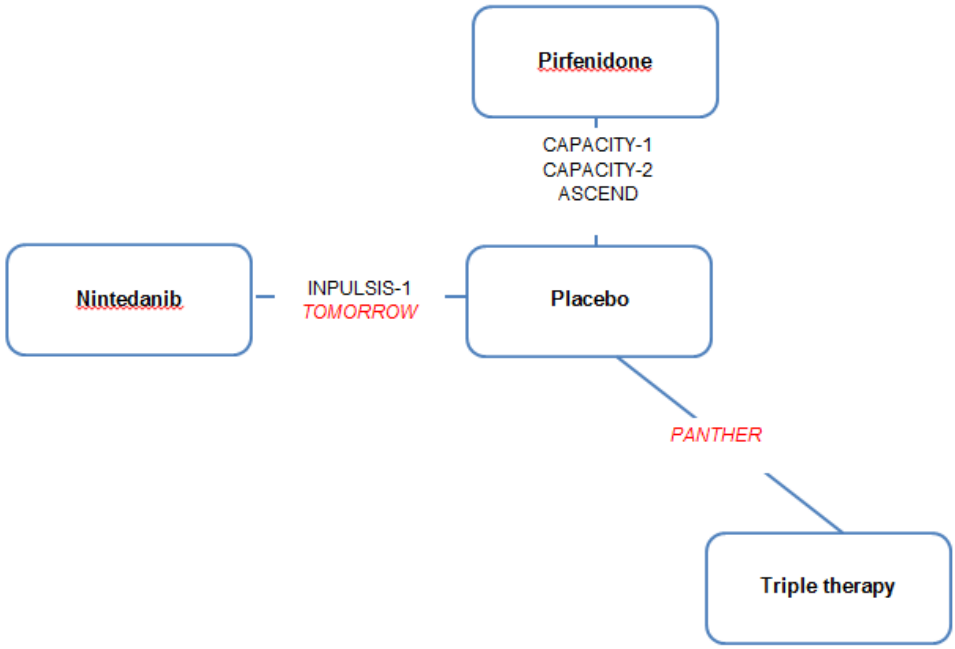
Figure 9: Network diagram for diarrhoea as an adverse event



All trials are included in the base-case network
Red italicised text indicates trials excluded in the restricted network

2) Rash:

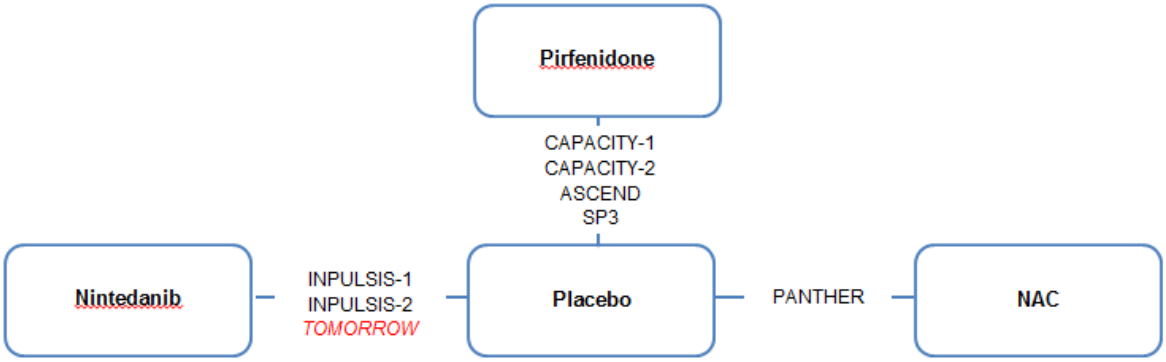
Figure 10: Network diagram for rash as an adverse event



All trials are included in the base-case network
Red italicised text indicates trials excluded in the restricted network

3) Discontinuation due to adverse event:

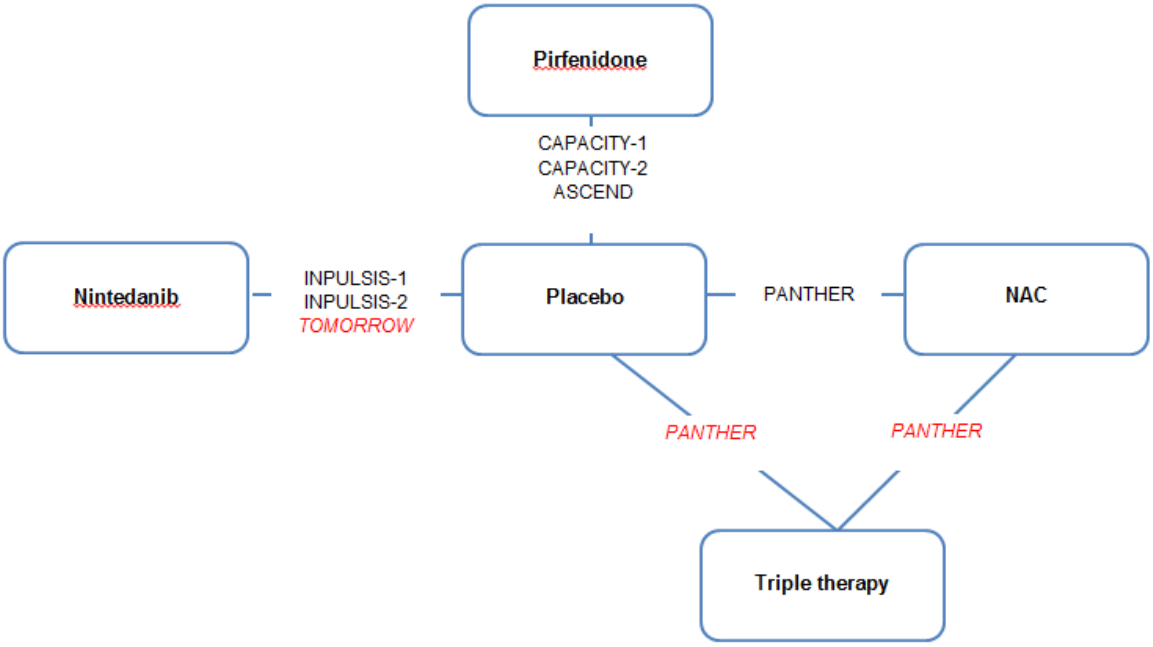
Figure 11: Network diagram for discontinuation due to an adverse event



All trials are included in the base-case network
Red italicised text indicates trials excluded in the restricted network

4) Serious cardiac adverse event:

Figure 12: Network diagram for serious cardiac adverse events



All trials are included in the base-case network
Red italicised text indicates trials excluded in the restricted network

c) Base case NMA results

For all tables, odds ratios/relative risks are provided for the column treatment relative to the row treatment, a result less than one is favourable to the column treatment

1) Diarrhoea:

Table 138: OR estimates and 95% credible intervals: diarrhoea (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1.39 (0.94, 2.11)	7.32 (4.82, 11.13)	1.22 (0.49, 3.03)	0.85 (0.23, 3.1)
Pirfenidone 2403mg/day	0.72 (0.47, 1.07)		5.25 (2.9, 9.32)	0.87 (0.32, 2.35)	0.61 (0.15, 2.35)
Nintedanib 300mg/day	0.14 (0.09, 0.21)	0.19 (0.11, 0.35)		0.17 (0.06, 0.45)	0.12 (0.03, 0.45)
NAC	0.82 (0.33, 2.03)	1.15 (0.42, 3.11)	6.03 (2.2, 16.24)		0.7 (0.14, 3.39)
Triple therapy	1.17 (0.32, 4.36)	1.64 (0.43, 6.48)	8.6 (2.21, 34.12)	1.43 (0.3, 7.04)	

Table 139: RR estimates and 95% credible intervals: diarrhoea (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1.3 (0.95, 1.76)	3.34 (2.73, 4.02)	1.17 (0.55, 2.19)	0.88 (0.27, 2.22)
Pirfenidone 2403mg/day	0.77 (0.57, 1.05)		2.57 (1.82, 3.61)	0.9 (0.4, 1.81)	0.67 (0.2, 1.8)
Nintedanib 300mg/day	0.3 (0.25, 0.37)	0.39 (0.28, 0.55)		0.35 (0.16, 0.68)	0.26 (0.08, 0.68)
NAC	0.86 (0.46, 1.83)	1.11 (0.55, 2.52)	2.85 (1.48, 6.25)		0.75 (0.2, 2.51)
Triple therapy	1.14 (0.45, 3.73)	1.48 (0.56, 5.02)	3.8 (1.48, 12.62)	1.33 (0.4, 5.08)	

2) Rash:

Table 140: OR estimates and 95% credible intervals: rash (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	Triple therapy
Placebo		3.85 (2.38, 6.29)	1.29 (0.49, 3.35)	4.01 (1.1, 17.79)
Pirfenidone 2403mg/day	0.26 (0.16, 0.42)		0.33 (0.11, 0.97)	1.04 (0.26, 4.98)
Nintedanib 300mg/day	0.78 (0.3, 2.05)	2.99 (1.03, 8.88)		3.13 (0.63, 18.56)
Triple therapy	0.25 (0.06, 0.91)	0.96 (0.2, 3.85)	0.32 (0.05, 1.59)	

Table 141: RR estimates and 95% credible intervals: rash (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	Triple therapy
Placebo		3.16 (2.13, 4.61)	1.26 (0.51, 2.87)	3.25 (1.09, 7.82)
Pirfenidone 2403mg/day (*)	0.32 (0.22, 0.47)		0.4 (0.15, 0.97)	1.03 (0.32, 2.64)
Nintedanib 300mg/day	0.79 (0.35, 1.96)	2.51 (1.03, 6.59)		2.57 (0.67, 9.02)
Triple therapy	0.31 (0.13, 0.92)	0.97 (0.38, 3.08)	0.39 (0.11, 1.49)	

3) Discontinuation due to adverse event:

Table 142: OR estimates and 95% credible intervals: discontinuation due to AE (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.58 (1.04, 2.39)	1.52 (1.01, 2.29)	0.18 (0.01, 1.58)
Pirfenidone 2403mg/day	0.63 (0.42, 0.96)		0.96 (0.54, 1.73)	0.11 (0, 1.04)
Nintedanib 300mg/day	0.66 (0.44, 0.99)	1.04 (0.58, 1.85)		0.12 (0, 1.08)
NAC	5.54 (0.63, 172.65)	8.79 (0.96, 279.56)	8.46 (0.92, 269.39)	

Table 143: RR estimates and 95% credible intervals: discontinuation due to AE (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.45 (1.03, 1.99)	1.41 (1.01, 1.94)	0.21 (0.01, 1.45)
Pirfenidone 2403mg/day	0.69 (0.5, 0.97)		0.97 (0.62, 1.54)	0.14 (0, 1.03)
Nintedanib 300mg/day	0.71 (0.51, 0.99)	1.03 (0.65, 1.61)		0.15 (0, 1.07)
NAC	4.87 (0.69, 147.37)	7.08 (0.97, 216.81)	6.88 (0.94, 211.45)	

4) Serious cardiac adverse events:

Table 144: OR estimates and 95% credible intervals: serious cardiac AEs (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1.36 (0.54, 3.46)	0.64 (0.17, 1.49)	5.45 (0.9, 54.4)	*
Pirfenidone 2403mg/day	0.74 (0.29, 1.84)		0.47 (0.09, 1.53)	4.04 (0.52, 47.38)	*
Nintedanib 300mg/day	1.55 (0.67, 5.82)	2.11 (0.65, 11.34)		8.75 (1.22, 123.94)	*
NAC	0.18 (0.02, 1.12)	0.25 (0.02, 1.91)	0.11 (0.01, 0.82)		*
Triple therapy	*	*	*	*	

* As no patients in the placebo arm of the PANTHER trial were reported to experience a serious cardiac AE (Table 136), a meaningful comparison with triple therapy would require additional information/assumptions on the baseline risk; therefore, results for triple therapy are not shown

Table 145: RR estimates and 95% credible intervals: serious cardiac AEs (base case network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1.33 (0.56, 3.07)	0.66 (0.18, 1.46)	4.43 (0.9, 14.97)	*
Pirfenidone 2403mg/day	0.75 (0.33, 1.79)		0.49 (0.1, 1.5)	3.31 (0.55, 14.56)	*
Nintedanib 300mg/day	1.52 (0.69, 5.53)	2.03 (0.67, 9.85)		6.88 (1.21, 37.91)	*
NAC	0.23 (0.07, 1.11)	0.3 (0.07, 1.8)	0.15 (0.03, 0.83)		*
Triple therapy	*	*	*	*	

* As no patients in the placebo arm of the PANTHER trial were reported to experience a serious cardiac AE (Table 136), a meaningful comparison with triple therapy would require additional information/assumptions on the baseline risk; therefore, results for triple therapy are not shown

d) Interpretation of NMA results

Common to all results described above, it may be important to note that – prior to commencement of therapy – some treatments are known to have different AE profiles; patients are sometimes even forewarned about these possible side effects. As a result, those AEs are likely to be subject to differential scrutiny in different trials which may hinder the comparability of trials on AEs.

1) Diarrhoea

The results of the base case random effects model reveal a difference between the rate of diarrhoea in the nintedanib (300mg/day) group compared to placebo. The odds ratio estimate for nintedanib versus placebo was 7.32, with a 95% CrI of (4.82, 11.13). This is

consistent with the AE profile for nintedanib, as raised during the NICE assessment of this treatment [NICE 2016].

There is no conclusive evidence of a difference in odds of getting diarrhoea in the pirfenidone (2403mg/day) group compared to placebo, though the point estimate favours placebo. The odds ratio estimate for pirfenidone versus placebo was 1.39, with a 95% CrI of (0.94, 2.11).

There is strong evidence of a difference between pirfenidone and nintedanib on this adverse event in favour of pirfenidone. The odds ratio estimate for pirfenidone versus nintedanib is 0.19, with a 95% CrI of (0.11, 0.35).

Note that CAPACITY trial data are collected at 72 weeks, that PANTHER trial data (NAC arm) are collected at 60 weeks and that PANTHER trial data (triple therapy) are collected at 32 weeks; all other data points (for ASCEND, INPULSIS I & II, and the TOMORROW trials) come from 52 weeks trials. This difference in follow-up time may add some bias to the results.

2) Rash

The results of the base case random effects model reveal a difference between the rate of rash in the pirfenidone group compared to placebo. The odds ratio estimate for pirfenidone versus placebo was 3.85, with a 95% CrI of (2.38, 6.29).

There is no conclusive evidence of a difference in odds of getting a rash in the nintedanib group compared to placebo, though the point estimate favours placebo. The odds ratio estimate for nintedanib versus placebo was 1.29, with a 95% CrI of (0.49, 3.35).

There is evidence of a difference between pirfenidone and nintedanib on this adverse event in favour of nintedanib. The odds ratio estimate for pirfenidone versus nintedanib is 2.99, with a 95% CrI of (1.03, 8.88).

Note that CAPACITY trial data are collected at 72 weeks, and that PANTHER trial data (triple therapy) are collected at 32 weeks; all other data points (for ASCEND, INPULSIS I, and the TOMORROW trials) come from 52 weeks trials. This difference in follow-up time may add some bias to the results.

3) Discontinuation due to adverse event

In contrast to the approach taken for diarrhoea and rash, the NICE manufacturer submission for nintedanib presented a NMA for discontinuation due to AE [Boehringer Ingelheim 2015]. We have produced NMAs for discontinuation of treatment due to AEs (rather than discontinuation of study data) due to better data availability on this specific outcome.

We find that our results are quite consistent with those presented by Boehringer Ingelheim, in that there is no evidence of a difference between nintedanib and pirfenidone on rate of discontinuations of treatment due to AEs collected during the trial periods.

Comparison with NICE manufacturer submission for nintedanib [Boehringer Ingelheim 2015]: The key difference between NMAs produced by the manufacturer of nintedanib and the manufacturer of pirfenidone is a difference in data inputs. For the nintedanib trials (TOMORROW and INPULSIS 1 & 2) there were some discrepancies in the available data, both within the nintedanib submission and between the submission and the published literature (p14 of Boehringer Ingelheim's response to NICE clarification questions provides an explanation of some of the discrepancies within the submission [NICE 2015]). For SP3 and PANTHER there were also some differences between the data extracted for the nintedanib submission and the data extracted for this analysis.

4) Serious cardiac adverse events

EMA has recently requested Boehringer Ingelheim to update the nintedanib SmPC with a warning on the risk of haemorrhage and epistaxis and new data on mild/moderate hepatic impairment [EMA, 2015b]. The data which led to this request are not yet publically available, and so not included in the NMA reported above. Such warnings do not appear in the SmPC for pirfenidone [EMC 2015].

As with discontinuation of treatment due to AE, but unlike the outcomes for diarrhoea and rash, Boehringer Ingelheim presented an NMA for serious cardiac events as part of their evidence submission [Boehringer Ingelheim 2015]. We find that our results are quite consistent with those from the nintedanib submission, in that there is no evidence of a difference between nintedanib and pirfenidone on risk of serious cardiac events during the trial periods.

Comparison with NICE manufacturer submission for nintedanib [Boehringer Ingelheim 2015]: The NMA presented in the nintedanib submission suggests there is no evidence of a difference between pirfenidone and nintedanib in the risk of serious cardiac events, though the point estimate favours nintedanib.

All of the results presented above also suggest no evidence of a difference between pirfenidone and nintedanib, with point estimates favouring nintedanib. Note that CAPACITY trial data are collected at 72 weeks, PANTHER trial data (NAC arm) are collected at 60 weeks and PANTHER trial data (triple therapy) are collected at 32 weeks, while other data points (for ASCEND, INPULSIS I & II, and the TOMORROW trials) come from 52 weeks trials.

e) NMA sensitivity analysis: Restricted network, RE model

1) Diarrhoea:

Table 146: OR estimates and 95% credible intervals: diarrhoea (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.4 (0.91, 2.19)	7.39 (4.43, 12.42)	1.21 (0.47, 3.17)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Pirfenidone 2403mg/day	0.72 (0.46, 1.1)		5.3 (2.66, 10.33)	0.87 (0.3, 2.47)
Nintedanib 300mg/day	0.14 (0.08, 0.23)	0.19 (0.1, 0.38)		0.16 (0.06, 0.49)
NAC	0.82 (0.32, 2.12)	1.15 (0.41, 3.28)	6.09 (2.06, 17.86)	

Table 147: RR estimates and 95% credible intervals: diarrhoea (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.3 (0.93, 1.79)	3.28 (2.6, 4)	1.17 (0.53, 2.22)
Pirfenidone 2403mg/day	0.77 (0.56, 1.08)		2.53 (1.73, 3.64)	0.9 (0.38, 1.85)
Nintedanib 300mg/day	0.3 (0.25, 0.38)	0.4 (0.28, 0.58)		0.36 (0.16, 0.71)
NAC	0.86 (0.45, 1.9)	1.11 (0.54, 2.63)	2.81 (1.42, 6.37)	

2) Rash:

Table 148: OR estimates and 95% credible intervals: rash (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day
Placebo		3.86 (2.39, 6.26)	1.88 (0.61, 6.45)
Pirfenidone 2403mg/day	0.26 (0.16, 0.42)		0.49 (0.14, 1.83)
Nintedanib 300mg/day	0.53 (0.16, 1.63)	2.05 (0.55, 6.97)	

Table 149: RR estimates and 95% credible intervals: rash (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day
Placebo		3.09 (2.11, 4.48)	1.75 (0.63, 4.46)
Pirfenidone 2403mg/day (*)	0.32 (0.22, 0.47)		0.57 (0.19, 1.53)
Nintedanib 300mg/day	0.57 (0.22, 1.58)	1.77 (0.66, 5.14)	

3) Discontinuation due to adverse event:

Table 150: OR estimates and 95% credible intervals: discontinuation due to AE (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.58 (1.02, 2.44)	1.61 (0.99, 2.69)	0.18 (0.01, 1.59)
Pirfenidone 2403mg/day	0.63 (0.41, 0.98)		1.02 (0.53, 2.01)	0.12 (0, 1.06)
Nintedanib 300mg/day	0.62 (0.37, 1.01)	0.98 (0.5, 1.88)		0.11 (0, 1.04)
NAC	5.44 (0.63, 149.1)	8.62 (0.95, 241)	8.85 (0.96, 248.14)	

Table 151: RR estimates and 95% credible intervals: discontinuation due to AE (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.47 (1.01, 2.09)	1.5 (0.99, 2.25)	0.2 (0.01, 1.48)
Pirfenidone 2403mg/day	0.68 (0.48, 0.99)		1.02 (0.6, 1.76)	0.14 (0.01, 1.05)
Nintedanib 300mg/day	0.67 (0.44, 1.01)	0.98 (0.57, 1.68)		0.14 (0.01, 1.04)
NAC	4.89 (0.68, 130.87)	7.22 (0.96, 195.55)	7.37 (0.97, 199.72)	

4) Serious cardiac adverse events:

Table 152: OR estimates and 95% credible intervals: serious cardiac AEs (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.35 (0.61, 3.04)	0.92 (0.32, 2.62)	5.41 (0.99, 48.9)
Pirfenidone 2403mg/day	0.74 (0.33, 1.63)		0.68 (0.18, 2.51)	4.02 (0.61, 41.11)
Nintedanib 300mg/day	1.09 (0.38, 3.08)	1.47 (0.4, 5.5)		5.93 (0.8, 66.78)
NAC	0.18 (0.02, 1.01)	0.25 (0.02, 1.65)	0.17 (0.01, 1.24)	

Table 153: RR estimates and 95% credible intervals: serious cardiac AEs (restricted network, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.34 (0.62, 2.85)	0.92 (0.33, 2.45)	4.57 (0.99, 17.05)
Pirfenidone 2403mg/day	0.75 (0.35, 1.6)		0.69 (0.19, 2.36)	3.43 (0.63, 15.01)
Nintedanib 300mg/day	1.08 (0.41, 3)	1.44 (0.42, 5.13)		4.92 (0.82, 26.19)
NAC	0.22 (0.06, 1.01)	0.29 (0.07, 1.6)	0.2 (0.04, 1.22)	

f) NMA sensitivity analysis: Restricted network, FE model

1) Diarrhoea

Table 154: OR estimates and 95% credible intervals: diarrhoea (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.37 (1.05, 1.79)	7.39 (5.53, 9.95)	1.21 (0.58, 2.58)
Pirfenidone 2403mg/day	0.73 (0.56, 0.95)		5.4 (3.64, 8.05)	0.89 (0.41, 1.97)
Nintedanib 300mg/day	0.14 (0.1, 0.18)	0.19 (0.12, 0.27)		0.16 (0.07, 0.37)
NAC	0.82 (0.39, 1.72)	1.13 (0.51, 2.47)	6.09 (2.72, 13.45)	

Table 155: RR estimates and 95% credible intervals: diarrhoea (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.28 (1.04, 1.56)	3.27 (2.81, 3.81)	1.17 (0.63, 1.97)
Pirfenidone 2403mg/day	0.78 (0.64, 0.96)		2.56 (2.06, 3.18)	0.91 (0.48, 1.6)
Nintedanib 300mg/day	0.31 (0.26, 0.36)	0.39 (0.31, 0.48)		0.36 (0.19, 0.61)
NAC	0.86 (0.51, 1.58)	1.09 (0.63, 2.08)	2.81 (1.63, 5.26)	

2) Rash

Table 156: OR estimates and 95% credible intervals: rash (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day
Placebo		3.85 (2.84, 5.28)	1.87 (0.75, 5.39)
Pirfenidone 2403mg/day	0.26 (0.19, 0.35)		0.49 (0.18, 1.46)
Nintedanib 300mg/day	0.54 (0.19, 1.34)	2.06 (0.69, 5.45)	

Table 157: RR estimates and 95% credible intervals: rash (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day
Placebo		3.09 (2.39, 4.06)	1.74 (0.76, 3.97)
Pirfenidone 2403mg/day (*)	0.32 (0.25, 0.42)		0.56 (0.24, 1.31)
Nintedanib 300mg/day	0.58 (0.25, 1.31)	1.78 (0.76, 4.18)	

3) Discontinuation due to adverse event

Table 158: OR estimates and 95% credible intervals: discontinuation due to AE (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.58 (1.16, 2.18)	1.6 (1.14, 2.28)	0.18 (0.01, 1.46)
Pirfenidone 2403mg/day	0.63 (0.46, 0.86)		1.01 (0.64, 1.62)	0.11 (0, 0.95)
Nintedanib 300mg/day	0.62 (0.44, 0.88)	0.99 (0.62, 1.57)		0.11 (0, 0.94)
NAC	5.61 (0.68, 192.15)	8.92 (1.05, 308.63)	9.05 (1.06, 314.1)	

Table 159: RR estimates and 95% credible intervals: discontinuation due to AE (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.48 (1.13, 1.92)	1.49 (1.12, 2)	0.2 (0.01, 1.38)
Pirfenidone 2403mg/day	0.68 (0.52, 0.88)		1.01 (0.69, 1.48)	0.13 (0, 0.96)
Nintedanib	0.67 (0.5,	0.99 (0.67,		0.13 (0,

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
300mg/day	0.89)	1.45)		0.95)
NAC	5.05 (0.72, 168.77)	7.47 (1.04, 251.96)	7.56 (1.05, 255.45)	

4) Serious cardiac adverse events

Table 160: OR estimates and 95% credible intervals: serious cardiac AEs (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.35 (0.82, 2.26)	0.92 (0.53, 1.62)	5.38 (1.27, 42.23)
Pirfenidone 2403mg/day	0.74 (0.44, 1.23)		0.68 (0.32, 1.46)	4.01 (0.86, 33.26)
Nintedanib 300mg/day	1.08 (0.62, 1.87)	1.46 (0.69, 3.1)		5.88 (1.23, 48.6)
NAC	0.19 (0.02, 0.79)	0.25 (0.03, 1.16)	0.17 (0.02, 0.81)	

Table 161: RR estimates and 95% credible intervals: serious cardiac AEs (restricted network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.33 (0.82, 2.17)	0.93 (0.55, 1.58)	4.54 (1.26, 15.89)
Pirfenidone 2403mg/day	0.75 (0.46, 1.21)		0.7 (0.34, 1.43)	3.43 (0.87, 12.53)
Nintedanib 300mg/day	1.08 (0.63, 1.83)	1.44 (0.7, 2.97)		4.92 (1.21, 19.11)
NAC	0.22 (0.06, 0.8)	0.29 (0.08, 1.15)	0.2 (0.05, 0.82)	

g) NMA sensitivity analysis: Base case network, FE model

1) Diarrhoea

Table 162: OR estimates and 95% credible intervals: diarrhoea (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1.37 (1.05, 1.79)	7.33 (5.6, 9.66)	1.22 (0.58, 2.57)	0.85 (0.26, 2.75)
Pirfenidone 2403mg/day	0.73 (0.56, 0.95)		5.36 (3.67, 7.86)	0.89 (0.41, 1.97)	0.62 (0.18, 2.07)
Nintedanib	0.14 (0.1,	0.19 (0.13,		0.17 (0.08, 0.37)	0.12 (0.03, 0.39)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
300mg/day	0.18)	0.27)			
NAC	0.82 (0.39, 1.71)	1.12 (0.51, 2.46)	6.03 (2.72, 13.18)		0.7 (0.17, 2.8)
Triple therapy	1.18 (0.36, 3.91)	1.61 (0.48, 5.5)	8.63 (2.58, 29.5)	1.43 (0.36, 5.87)	

Table 163: RR estimates and 95% credible intervals: diarrhoea (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1.28 (1.04, 1.57)	3.33 (2.87, 3.87)	1.17 (0.63, 1.98)	0.88 (0.3, 2.07)
Pirfenidone 2403mg/day	0.78 (0.64, 0.96)		2.61 (2.09, 3.26)	0.91 (0.48, 1.61)	0.68 (0.23, 1.65)
Nintedanib 300mg/day	0.3 (0.26, 0.35)	0.38 (0.31, 0.48)		0.35 (0.19, 0.6)	0.26 (0.09, 0.63)
NAC	0.86 (0.5, 1.58)	1.09 (0.62, 2.09)	2.85 (1.65, 5.36)		0.75 (0.23, 2.16)
Triple therapy	1.14 (0.48, 3.36)	1.46 (0.6, 4.38)	3.81 (1.6, 11.32)	1.33 (0.46, 4.42)	

2) Rash

Table 164: OR estimates and 95% credible intervals: rash (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	Triple therapy
Placebo		3.85 (2.83, 5.29)	1.34 (0.61, 3.09)	4 (1.3, 15.36)
Pirfenidone 2403mg/day	0.26 (0.19, 0.35)		0.35 (0.15, 0.85)	1.04 (0.32, 4.12)
Nintedanib 300mg/day	0.75 (0.32, 1.63)	2.88 (1.18, 6.68)		3 (0.73, 14.02)
Triple therapy	0.25 (0.07, 0.77)	0.96 (0.24, 3.09)	0.33 (0.07, 1.36)	

Table 165: RR estimates and 95% credible intervals: rash (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	Triple therapy
Placebo		3.16 (2.44, 4.15)	1.31 (0.63, 2.7)	3.25 (1.27, 7.37)
Pirfenidone 2403mg/day (*)	0.32 (0.24, 0.41)		0.41 (0.19, 0.88)	1.03 (0.39, 2.38)
Nintedanib 300mg/day	0.77 (0.37, 1.58)	2.42 (1.14, 5.17)		2.48 (0.77, 7.32)
Triple therapy	0.31 (0.14,	0.97 (0.42,	0.4 (0.14,	

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	Triple therapy
	0.79)	2.57)	1.31)	

3) Discontinuation due to adverse event

Table 166: OR estimates and 95% credible intervals: discontinuation due to AE (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.58 (1.16, 2.18)	1.53 (1.13, 2.08)	0.18 (0.01, 1.45)
Pirfenidone 2403mg/day	0.63 (0.46, 0.86)		0.96 (0.62, 1.5)	0.11 (0, 0.94)
Nintedanib 300mg/day	0.66 (0.48, 0.89)	1.04 (0.67, 1.61)		0.12 (0, 0.98)
NAC	5.54 (0.69, 169.41)	8.8 (1.06, 271.45)	8.48 (1.02, 261.11)	

Table 167: RR estimates and 95% credible intervals: discontinuation due to AE (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC
Placebo		1.46 (1.13, 1.87)	1.42 (1.11, 1.82)	0.21 (0.01, 1.36)
Pirfenidone 2403mg/day	0.69 (0.53, 0.88)		0.97 (0.69, 1.38)	0.14 (0, 0.96)
Nintedanib 300mg/day	0.71 (0.55, 0.9)	1.03 (0.73, 1.45)		0.14 (0, 0.98)
NAC	4.87 (0.73, 144.74)	7.11 (1.05, 212.07)	6.91 (1.02, 206.51)	

4) Serious cardiac adverse events

Table 168: OR estimates and 95% credible intervals: serious cardiac AEs (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1.35 (0.81, 2.25)	0.76 (0.45, 1.27)	5.36 (1.27, 40.22)	*
Pirfenidone 2403mg/day	0.74 (0.44, 1.23)		0.56 (0.27, 1.16)	4.01 (0.86, 31.46)	*
Nintedanib 300mg/day	1.32 (0.79, 2.21)	1.78 (0.86, 3.67)		7.12 (1.53, 56.28)	*
NAC	0.19 (0.02, 0.79)	0.25 (0.03, 1.16)	0.14 (0.02, 0.66)		*
Triple therapy	0 (0, 0.06)	0 (0, 0.09)	0 (0, 0.05)	0 (0, 0.39)	

* As no patients in the placebo arm of the PANTHER trial were reported to experience a serious cardiac AE (Table 136), a meaningful comparison with triple therapy would require additional information/assumptions on the

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
baseline risk; therefore, results for triple therapy are not shown					

Table 169: RR estimates and 95% credible intervals: serious cardiac AEs (base case network, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		1.32 (0.82, 2.14)	0.77 (0.47, 1.25)	4.44 (1.26, 14.43)	*
Pirfenidone 2403mg/day	0.75 (0.47, 1.22)		0.58 (0.29, 1.15)	3.36 (0.87, 11.69)	*
Nintedanib 300mg/day	1.3 (0.8, 2.14)	1.73 (0.87, 3.44)		5.81 (1.49, 20.45)	*
NAC	0.23 (0.07, 0.8)	0.3 (0.09, 1.15)	0.17 (0.05, 0.67)		*
Triple therapy	0.05 (0.03, 0.11)	0.06 (0.04, 0.15)	0.04 (0.02, 0.09)	0.22 (0.06, 0.79)	
* As no patients in the placebo arm of the PANTHER trial were reported to experience a serious cardiac AE (Table 136), a meaningful comparison with triple therapy would require additional information/assumptions on the baseline risk; therefore, results for triple therapy are not shown					

Appendix E: Cost-effectiveness analyses: Updated base case results – list price

This document presents revised base case model results at the list prices for both pirfenidone and nintedanib. Revisions are based on the updates outlined in the introduction to our responses to Section B.

Results are presented for the ITT, mild and moderate populations.

a) ITT population – List price for pirfenidone

Table 170: Discounted base case model results – ITT population – List price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	5.38	3.80				
PFN	[redacted]	8.67	5.67	[redacted]	3.29	1.87	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 171: Undiscounted base case model results – ITT population – List price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	7.11	5.63				
PFN	[redacted]	11.26	[redacted]	[redacted]	4.15	3.28	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 172: LY gain by health state – ITT population – List price

Health state	LY PFN	LY BSC	Increment	Absolute increment	% increment
Progression-free	2.61	1.90	0.71	0.71	17%
Progressed	8.64	5.21	3.44	3.44	83%
Total LYs	**11.26**	**7.11**	**4.15**	**4.15**	**100%**

Key: BSC, best supportive care; LY, life year; PFN, pirfenidone.

Table 173: QALY gain by health state – discounted – ITT population – List price

Health state	QALY PFN	QALY BSC	Increment	Absolute increment	% increment
Progression-free	1.68	1.24	0.44	0.44	23%
Progressed	4.03	2.59	1.44	1.44	76%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.03	-0.03	-0.01	0.01	0%
Total QALYs	**5.67**	**3.80**	**1.87**	**1.89**	**100%**

Key: BSC, best supportive care; QALY, quality-adjusted life year; PFN, pirfenidone.

Table 174: QALY gain by health state – undiscounted – ITT population – List price

Health state	QALY PFN	QALY BSC	Increment	Absolute increment	% increment
Progression-free	1.74	1.27	0.47	0.47	18%
Progressed	5.18	3.04	2.14	2.14	82%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.04	-0.03	-0.01	0.01	1%
Total QALYs	**6.86**	**4.27**	**2.60**	**2.62**	**100%**

Key: BSC, best supportive care; QALY, quality-adjusted life year; PFN, pirfenidone.

Table 175: Costs by health state – discounted – ITT population – List price

Health state	Cost PFN	Cost BSC	Increment	Absolute increment	% increment
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£3,256	£4,324	-£1,068	£1,068	1%
End of life	£4,063	£5,874	-£1,811	£1,811	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Table 176: Costs by resource use category – discounted – ITT population – List price

Item	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hospitalisation	£3,256	£4,324	-£1,068	£1,068	1%
Disease management costs	£20,055	£14,339	£5,716	£5,716	7%
Terminal care	£4,063	£5,874	-£1,811	£1,811	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Table 177: Costs by health state – undiscounted – ITT population – List price

Health state	Cost PFN	Cost BSC	Increment	Absolute increment	% increment
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£3,958	£4,873	-£915	£915	1%
End of life	£5,246	£6,878	-£1,632	£1,632	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Table 178: Costs by resource use category – undiscounted – ITT population – List price

Item	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hospitalisation	£3,958	£4,873	-£915	£915	1%
Disease management costs	£24,166	£16,065	£8,102	£8,102	9%
Terminal care	£5,246	£6,878	-£1,632	£1,632	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Figure 13: PSA scatterplot – ITT population – List price

[redacted]

Figure 14: CEAC – ITT population – List price

[redacted]

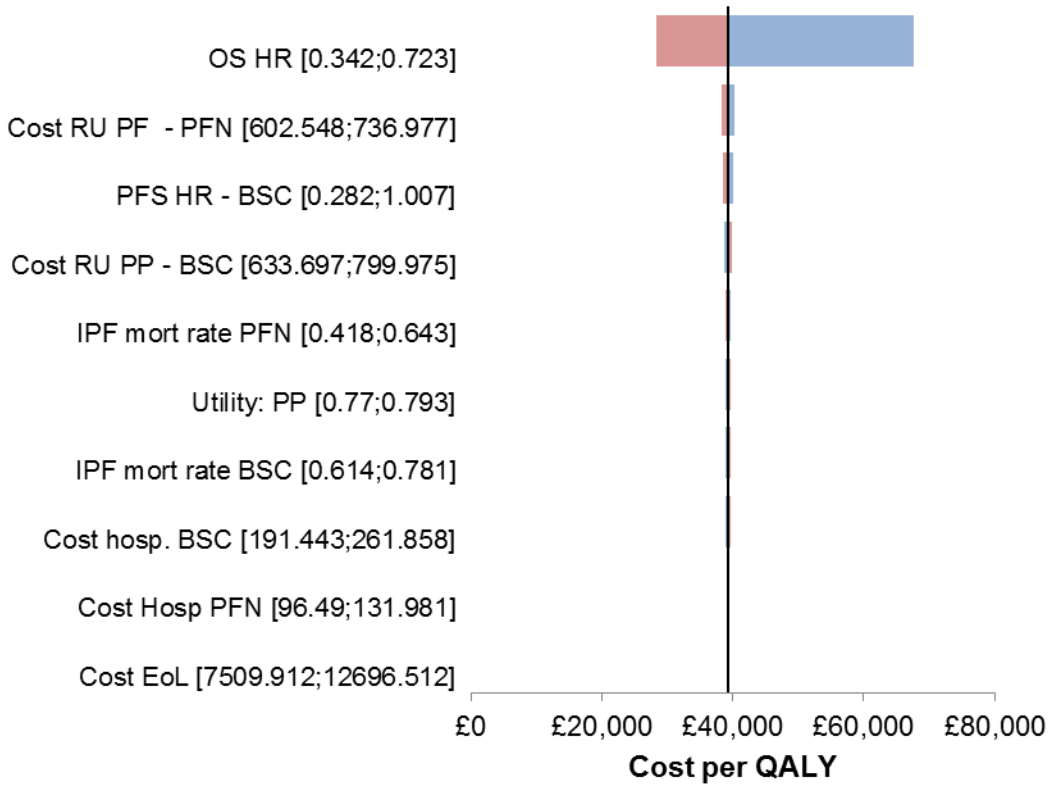
Table 179: PSA results – ITT population – List price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Vs.baseline (QALYs)
Deterministic model results							
BSC	[redacted]	5.38	3.80				
PFN	[redacted]	8.67	5.67	[redacted]	3.29	1.87	[redacted]
Mean probabilistic model results							
BSC 95% CI	[redacted] [redacted]	5.37 (3.36; 7.94)	3.765 (2.51; 5.24)				
PFN 95% CI	[redacted] [redacted]	8.71 (7.38; 10.16)	5.68 (5.02; 6.36)	[redacted]	3.34	1.91	[redacted]

% chance of being cost effective	£20,000 per QALY	£25,000 per QALY	£30,000 per QALY
BSC	100%	97%	82%
PFN	0%	3%	18%

Key: BSC, best supportive care; CI, confidence interval; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Figure 15: OWSA – ITT population – List price



Key: BSC, best supportive care; EoL, end of life; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; mort, mortality; PF, progression-free; PFN, pirfenidone; PP, post-progression; QALY, quality-adjusted life year; RU, resource use.

Table 180: Scenario analysis – ITT population – List price

	Category	Base case setting	Model change	PFN			BSC			ICER vs. BSC (£)
				Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Base Case				[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
1	Time horizon	Lifetime (34 years)	10 years	[redacted]	6.67	4.72	[redacted]	4.97	3.59	[redacted]
2			15 years	[redacted]	7.86	5.34	[redacted]	5.30	3.76	[redacted]
3			20 years	[redacted]	8.38	5.56	[redacted]	5.37	3.79	[redacted]
4			25 years	[redacted]	8.58	5.64	[redacted]	5.38	3.80	[redacted]
5			30 years	[redacted]	8.65	5.66	[redacted]	5.38	3.80	[redacted]
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	[redacted]	8.67	5.40	[redacted]	5.38	3.62	[redacted]
7			Nintedanib NICE company submission	[redacted]	8.67	5.33	[redacted]	5.38	3.57	[redacted]
8			Starkie et al. (2012) mapping algorithm	[redacted]	8.67	5.36	[redacted]	5.38	3.59	[redacted]
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	[redacted]	8.67	5.67	[redacted]	6.22	4.19	[redacted]
10			Treatment effect applied for up to 10 years	[redacted]	8.67	5.67	[redacted]	5.74	3.95	[redacted]
11			Treatment effect applied for up to 14 years	[redacted]	8.67	5.67	[redacted]	5.48	3.84	[redacted]
12	OS	Weibull	Exponential	[redacted]	11.91	6.99	[redacted]	6.62	4.36	[redacted]
13			Log-normal	[redacted]	12.71	7.24	[redacted]	6.89	4.45	[redacted]
14			Gamma	[redacted]	9.49	6.00	[redacted]	5.59	3.89	[redacted]
15			Log-Logistic	[redacted]	11.19	6.63	[redacted]	6.12	4.11	[redacted]

16			Gompertz	[redacted]	7.60	5.20	[redacted]	5.16	3.69	[redacted]
17		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	8.67	5.67	[redacted]	5.45	3.84	[redacted]
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	[redacted]	8.67	5.67	[redacted]	5.75	3.94	[redacted]
19			Edinburgh registry used for BSC OS	[redacted]	8.67	5.67	[redacted]	5.03	3.58	[redacted]
20	PFS	Weibull	Exponential	[redacted]	8.67	5.72	[redacted]	5.38	3.82	[redacted]
21			Log-normal	[redacted]	8.67	5.72	[redacted]	5.38	3.82	[redacted]
22			Gamma	[redacted]	8.67	5.75	[redacted]	5.38	3.82	[redacted]
23			Log-Logistic	[redacted]	8.67	5.72	[redacted]	5.38	3.81	[redacted]
24			Gompertz	[redacted]	8.67	5.65	[redacted]	5.38	3.79	[redacted]
25		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
26	TTD	Weibull	Exponential	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
27			Log-normal	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
28			Gamma	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
29			Log-Logistic	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]

30			Gompertz	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
31	Stopping rule	Not applied	Applied for pirfenidone patients	[redacted]	8.67	5.66	[redacted]	5.38	3.80	[redacted]
32										
33	OS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.31	3.75	[redacted]
34			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
35			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
36			Phase III trials only, random effects, 72 weeks cut off	[redacted]	8.67	5.67	[redacted]	6.12	4.23	[redacted]
37			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	8.67	5.67	[redacted]	6.12	4.23	[redacted]
38			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	8.67	5.67	[redacted]	6.12	4.23	[redacted]
39			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	8.67	5.67	[redacted]	6.12	4.23	[redacted]
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
41			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
42			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]

43			Phase III trials only, random effects, 72 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
44			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
45			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
46			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
47	AEs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
48			Phase III trials only, fixed effects	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
49			Phase II and III trials, fixed effects	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
50			Phase II and III trials, random effects with adjustments in data for differences in end point	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
51			Phase III trials only, fixed effects with adjustments in data for differences in end point	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
52	Lung transplant	Not applied	Applied	[redacted]	8.63	5.67	[redacted]	5.40	3.81	[redacted]
53	Gas Transfer	Test administered every 4 months	Test administered every 6 months	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
54	Full pulmonary	Test administered	Test administered every 6 months	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]

		every 4 months								
55	Lung Volume Transfer	No subsequent tests administered	Test administered every 4 months	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
56	Field exercise test	Test administered every 6 months	Test administered every 12 months	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
57			Test administered every 3 months if on oxygen	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]
58	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	[redacted]	8.67	5.67	[redacted]	5.38	3.80	[redacted]

Key: BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use; NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

* Result driven by low number of patients with mild disease in the INOVA patient registry: numbers at risk: 53 at baseline; 24 at 50 months; 16 at 100 months; 2 at 150 months.

b) Mild population – List price for pirfenidone

Table 181: Discounted base case model results – Mild population – List price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	7.11	4.82				
PFN	[redacted]	11.26	6.99	[redacted]	4.15	2.17	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 182: Undiscounted base case model results – Mild population – List price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	7.11	5.65				
PFN	[redacted]	11.26	8.92	[redacted]	4.15	3.28	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 183: LY gain by health state – Mild population – List price

Health state	LY PFN	LY BSC	Increment	Absolute increment	% increment
Progression-free	2.61	1.90	0.71	0.71	17%
Progressed	8.64	5.21	3.44	3.44	83%
Total LYs	**11.26**	**7.11**	**4.15**	**4.15**	**100%**

Key: BSC, best supportive care; LY, life year; PFN, pirfenidone.

Table 184: QALY gain by health state – discounted – Mild population – List price

Health state	QALY PFN	QALY BSC	Increment	Absolute increment	% increment
Progression-free	2.11	1.56	0.55	0.55	25%
Progressed	4.94	3.31	1.63	1.63	74%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.04	-0.03	-0.01	0.01	0%
Total QALYs	**6.99**	**4.82**	**2.17**	**2.19**	**100%**

Key: BSC, best supportive care; QALY, quality-adjusted life year; PFN, pirfenidone.

Table 185: QALY gain by health state – undiscounted – Mild population – List price

Health state	QALY PFN	QALY BSC	Increment	Absolute increment	% increment
Progression-free	2.21	1.61	0.60	0.60	18%
Progressed	6.76	4.07	2.69	2.69	81%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.05	-0.04	-0.02	0.02	0%
Total QALYs	**8.91**	**5.63**	**3.28**	**3.31**	**100%**

Key: BSC, best supportive care; QALY, quality-adjusted life year; PFN, pirfenidone.

Table 186: Costs by health state – discounted – Mild population – List price

Health state	Cost PFN	Cost BSC	Increment	Absolute increment	% increment
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£4,016	£5,496	-£1,480	£1,480	1%
End of life	£3,752	£5,605	-£1,852	£1,852	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Table 187: Costs by resource use category – discounted – Mild population – List price

Item	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hospitalisation	£4,016	£5,496	-£1,480	£1,480	1%
Disease management costs	£24,512	£17,971	£6,541	£6,541	6%
Terminal care	£3,752	£5,605	-£1,852	£1,852	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Table 188: Costs by health state – undiscounted – Mild population – List price

Health state	Cost PFN	Cost BSC	Increment	Absolute increment	% increment
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£5,138	£6,435	-£1,297	£1,297	1%
End of life	£5,191	£6,920	-£1,728	£1,728	1%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Table 189: Costs by resource use category – undiscounted – Mild population – List price

Item	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hospitalisation	£5,138	£6,435	-£1,297	£1,297	1%
Disease management costs	£31,086	£20,927	£10,159	£10,159	9%
Terminal care	£5,191	£6,920	-£1,728	£1,728	1%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Figure 16: PSA scatterplot – Mild population – List price

[redacted]

Figure 17: CEAC – Mild population – List price

[redacted]

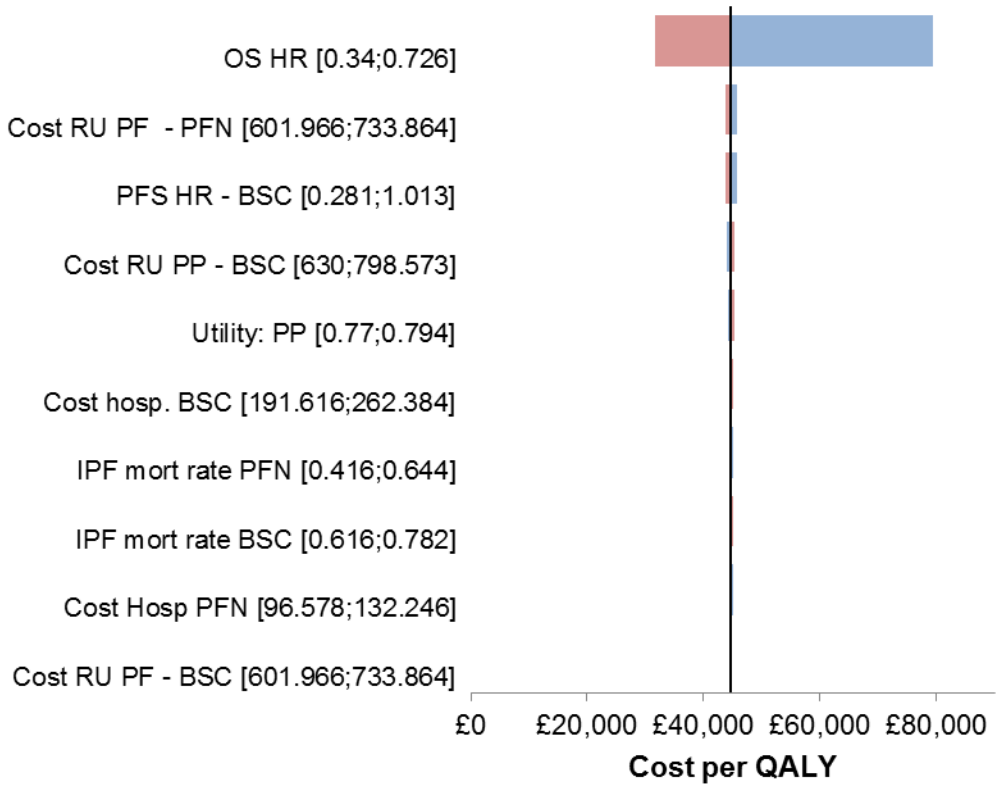
Table 190: PSA results – Mild population – List price

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Vs.baseline (QALYs)
Deterministic model results							
BSC	[redacted]	7.11	4.82				
PFN	[redacted]	11.26	6.99	[redacted]	4.15	2.17	[redacted]
Mean probabilistic model results							
BSC	[redacted]	7.14	4.799				
95% CI	[redacted]	(4.20; 11.03)	(3.11; 6.85)				
PFN	[redacted]	11.31	7.00	[redacted]	4.18	2.21	[redacted]
95% CI	[redacted]	(8.64; 14.39)	(5.72; 8.41)				

% chance of being cost effective	£20,000 per QALY	£25,000 per QALY	£30,000 per QALY
BSC	99%	99%	92%
PFN	1%	1%	8%

Key: BSC, best supportive care; CI, confidence interval; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Figure 18: OWSA – Mild population – List price



Key: BSC, best supportive care; EoL, end of life; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; mort, mortality; PF, progression-free; PFN, pirfenidone; PP, post-progression; QALY, quality-adjusted life year; RU, resource use.

Table 191: Scenario analysis – Mild population – List price

	Category	Base case setting	Model change	PFN			BSC			ICER vs. BSC (£)
				Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Base Case				[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
1	Time horizon	Lifetime (34 years)	10 years	[redacted]	7.53	5.30	[redacted]	6.03	4.30	[redacted]
2			15 years	[redacted]	9.38	6.26	[redacted]	6.78	4.69	[redacted]
3			20 years	[redacted]	10.39	6.70	[redacted]	7.02	4.79	[redacted]
4			25 years	[redacted]	10.91	6.89	[redacted]	7.09	4.82	[redacted]
5			30 years	[redacted]	11.16	6.97	[redacted]	7.10	4.82	[redacted]
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	[redacted]	11.26	6.66	[redacted]	7.11	4.60	[redacted]
7			Nintedanib NICE company submission	[redacted]	11.26	6.58	[redacted]	7.11	4.53	[redacted]
8			Starkie et al. (2012) mapping algorithm	[redacted]	11.26	6.61	[redacted]	7.11	4.56	[redacted]
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	[redacted]	11.26	6.99	[redacted]	8.70	5.53	[redacted]
10			Treatment effect applied for up to 10 years	[redacted]	11.26	6.99	[redacted]	7.96	5.17	[redacted]
11			Treatment effect applied for up to 14 years	[redacted]	11.26	6.99	[redacted]	7.44	4.95	[redacted]
12	OS	Weibull	Exponential	[redacted]	15.16	8.47	[redacted]	9.02	5.61	[redacted]
13			Log-normal	[redacted]	15.11	8.40	[redacted]	8.87	5.52	[redacted]
14			Gamma	[redacted]	11.93	7.24	[redacted]	7.33	4.92	[redacted]

15			Log-Logistic	[redacted]	13.64	7.85	[redacted]	8.01	5.19	[redacted]
16			Gompertz	[redacted]	9.09	6.09	[redacted]	6.46	4.52	[redacted]
17		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	11.26	6.99	[redacted]	7.05	4.79	[redacted]
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	[redacted]	11.26	6.99	[redacted]	9.27	5.81	[redacted]
19			Edinburgh registry used for BSC OS	[redacted]	11.26	6.99	[redacted]	6.18	4.30	[redacted]
20	PFS	Weibull	Exponential	[redacted]	11.26	7.07	[redacted]	7.11	4.86	[redacted]
21			Log-normal	[redacted]	11.26	7.07	[redacted]	7.11	4.85	[redacted]
22			Gamma	[redacted]	11.26	7.09	[redacted]	7.11	4.86	[redacted]
23			Log-Logistic	[redacted]	11.26	7.06	[redacted]	7.11	4.85	[redacted]
24			Gompertz	[redacted]	11.26	6.95	[redacted]	7.11	4.79	[redacted]
25		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	11.26	6.99	[redacted]	7.11	4.81	[redacted]
26	TTD	Weibull	Exponential	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
27			Log-normal	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
28			Gamma	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]

29			Log-Logistic	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
30			Gompertz	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
31	Stopping rule	Not applied	Applied for pirfenidone patients	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
32										
33	OS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	11.26	6.99	[redacted]	7.01	4.77	[redacted]
34			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
35			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
36			Phase III trials only, random effects, 72 weeks cut off	[redacted]	11.26	6.99	[redacted]	8.06	5.34	[redacted]
37			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	11.26	6.99	[redacted]	8.06	5.34	[redacted]
38			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	11.26	6.99	[redacted]	8.06	5.34	[redacted]
39			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	11.26	6.99	[redacted]	8.06	5.34	[redacted]
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
41			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
42			Phase II and III trials, fixed effects, 52 weeks	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]

			cut off							
43			Phase III trials only, random effects, 72 weeks cut off	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
44			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
45			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
46			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
47	AEs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
48			Phase III trials only, fixed effects	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
49			Phase II and III trials, fixed effects	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
50			Phase II and III trials, random effects with adjustments in data for differences in end point	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
51			Phase III trials only, fixed effects with adjustments in data for differences in end point	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
52	Lung transplant	Not applied	Applied	[redacted]	11.20	6.99	[redacted]	7.12	4.83	[redacted]
53	Gas Transfer	Test administered every 4 months	Test administered every 6 months	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]

54	Full pulmonary	Test administered every 4 months	Test administered every 6 months	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
55	Lung Volume Transfer	No subsequent tests administered	Test administered every 4 months	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
56	Field exercise test	Test administered every 6 months	Test administered every 12 months	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
57			Test administered every 3 months if on oxygen	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]
58	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	[redacted]	11.26	6.99	[redacted]	7.11	4.82	[redacted]

Key: BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use; NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

* Result driven by low number of patients with mild disease in the INOVA patient registry: numbers at risk: 53 at baseline; 24 at 50 months; 16 at 100 months; 2 at 150 months.

c) Moderate population – List price for pirfenidone and nintedanib

Table 192: Discounted base case model results – Moderate population – List price

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)	Incremental (QALYs)
BSC	24,868	4.80	3.44					
NTB	65,065	6.06	4.23	40,197	1.26	0.78	51,331	51,331
PFN	[redacted]	7.67	5.14	[redacted]	1.61	0.91	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 193: Undiscounted base case model results – Moderate population – List price

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)	Incremental (QALYs)
BSC	27,753	4.80	3.81					
NTB	69,955	6.06	4.82	42,202	1.26	1.01	41,986	41,986
PFN	[redacted]	7.67	6.07	[redacted]	1.61	1.26	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Table 194: LY gain by health state (vs. BSC) – Moderate population – List price

Health state	LY PFN	LY BSC	Increment	Absolute increment	% increment
Progression-free	1.88	1.38	0.50	0.50	18%
Progressed	5.79	3.42	2.36	2.36	82%
Total LYs	**7.67**	**4.80**	**2.87**	**2.87**	**100%**

Key: BSC, best supportive care; LY, life year; PFN, pirfenidone.

Table 195: LY gain by health state (vs. NTB) – Moderate population – List price

Health state	LY PFN	LY NTB	Increment	Absolute increment	% increment
Progression-free	1.88	1.69	0.19	0.19	12%
Progressed	5.79	4.38	1.41	1.41	88%
Total LYs	**7.67**	**6.06**	**1.61**	**1.61**	**100%**

Key: LY, life year; NTB, nintedanib; PFN, pirfenidone.

Table 196: QALY gain by health state (vs. BSC) – discounted – Moderate population – List price

Health state	QALY PFN	QALY BSC	Increment	Absolute increment	% increment
Progression-free	1.55	1.15	0.40	0.40	23%
Progressed	3.63	2.33	1.30	1.30	76%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.03	-0.02	-0.01	0.01	0%
Total QALYs	**5.14**	**3.44**	**1.70**	**1.71**	**100%**

Key: BSC, best supportive care; QALY, quality-adjusted life year; PFN, pirfenidone.

Table 197: QALY gain by health state (vs. NTB) – discounted – Moderate population – List price

Health state	QALY PFN	QALY NTB	Increment	Absolute increment	% increment
Progression-free	1.55	1.39	0.15	0.15	16%
Progressed	3.63	2.86	0.77	0.77	82%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.03	-0.02	-0.01	0.01	1%
Total QALYs	**5.14**	**4.23**	**0.91**	**0.93**	**100%**

Key: NTB, nintedanib; QALY, quality-adjusted life year; PFN, pirfenidone.

Table 198: QALY gain by health state (vs. BSC) – undiscounted – Moderate population – List price

Health state	QALY PFN	QALY BSC	Increment	Absolute increment	% increment
Progression-free	1.60	1.17	0.43	0.43	19%
Progressed	4.52	2.68	1.85	1.85	81%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.04	-0.03	-0.01	0.01	1%
Total QALYs	**6.07**	**3.81**	**2.26**	**2.29**	**100%**

Key: BSC, best supportive care; QALY, quality-adjusted life year; PFN, pirfenidone.

Table 199: QALY gain by health state (vs. NTB) – undiscounted – Moderate population – List price

Health state	QALY PFN	QALY NTB	Increment	Absolute increment	% increment
Progression-free	1.60	1.43	0.16	0.16	13%
Progressed	4.52	3.42	1.10	1.10	86%
Acute exacerbations	-0.01	-0.01	0.00	0.00	0%
Adverse events	-0.04	-0.03	-0.01	0.01	1%
Total QALYs	**6.07**	**4.82**	**1.26**	**1.28**	**100%**

Key: NTB, nintedanib; QALY, quality-adjusted life year; PFN, pirfenidone.

Table 200: Costs by health state (vs. BSC) – discounted – Moderate population – List price

Health state	Cost PFN	Cost BSC	Increment	Absolute increment	% increment
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£2,951	£3,919	-£968	£968	1%
End of life	£4,186	£5,972	-£1,786	£1,786	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Table 201: Costs by health state (vs. NTB) – discounted – Moderate population – List price

Health state	Cost PFN	Cost NTB	Increment	Absolute increment	% increment
Progression-free	[redacted]	£32,617	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	£24,923	[redacted]	[redacted]	[redacted]
Acute exacerbations	£2,951	£2,421	£530	£530	2%
End of life	£4,186	£5,104	-£919	£919	3%
Total costs	[redacted]	**£65,065**	[redacted]	[redacted]	**100%**
Key: NTB, nintedanib; PFN, pirfenidone					

Table 202: Costs by resource use category (vs. BSC) – discounted – Moderate population – List price

Item	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hospitalisation	£2,951	£3,919	-£968	£968	1%
Disease management costs	£18,266	£13,081	£5,186	£5,186	7%
Terminal care	£4,186	£5,972	-£1,786	£1,786	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**
Key: BSC, best supportive care; PFN, pirfenidone					

Table 203: Costs by resource use category (vs. NTB) – discounted – Moderate population – List price

Item	Cost PFN	Cost NTB	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	£39,921	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	£1,785	[redacted]	[redacted]	[redacted]
Hospitalisation	£2,951	£2,421	£530	£530	2%
Disease management costs	£18,266	£15,834	£2,432	£2,432	9%
Terminal care	£4,186	£5,104	-£919	£919	3%
Total costs	[redacted]	**£65,065**	[redacted]	[redacted]	**100%**
Key: NTB, nintedanib; PFN, pirfenidone					

Table 204: Costs by health state (vs. BSC) – undiscounted – Moderate population – List price

Health state	Cost PFN	Cost BSC	Increment	Absolute increment	% increment
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Acute exacerbations	£3,500	£4,349	-£849	£849	1%
End of life	£5,255	£6,869	-£1,614	£1,614	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Table 205: Costs by health state (vs. NTB) – undiscounted – Moderate population – List price

Health state	Cost PFN	Cost NTB	Increment	Absolute increment	% increment
Progression-free	[redacted]	£33,137	[redacted]	[redacted]	[redacted]
Progressed	[redacted]	£27,945	[redacted]	[redacted]	[redacted]
Acute exacerbations	£3,500	£2,767	£733	£733	2%
End of life	£5,255	£6,105	-£851	£851	3%
Total costs	[redacted]	**£69,955**	[redacted]	[redacted]	**100%**

Key: NTB, nintedanib; PFN, pirfenidone

Table 206: Costs by resource use category (vs. BSC) – undiscounted – Moderate population – List price

Item	Cost PFN	Cost BSC	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Hospitalisation	£3,500	£4,349	-£849	£849	1%
Disease management costs	£21,483	£14,432	£7,050	£7,050	9%
Terminal care	£5,255	£6,869	-£1,614	£1,614	2%
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	**100%**

Key: BSC, best supportive care; PFN, pirfenidone

Table 207: Costs by resource use category (vs. NTB) – undiscounted – Moderate population – List price

Item	Cost PFN	Cost NTB	Increment	Absolute increment	% absolute increment
Therapy cost	[redacted]	£40,983	[redacted]	[redacted]	[redacted]
Adverse event	[redacted]	£2,103	[redacted]	[redacted]	[redacted]
Hospitalisation	£3,500	£2,767	£733	£733	2%
Disease management costs	£21,483	£17,997	£3,486	£3,486	11%
Terminal care	£5,255	£6,105	-£851	£851	3%
Total costs	[redacted]	**£69,955**	[redacted]	[redacted]	**100%**
Key: NTB, nintedanib; PFN, pirfenidone					

Figure 19: PSA scatterplot – Moderate population – List price

[redacted]

Figure 20: CEAC – Moderate population – List price

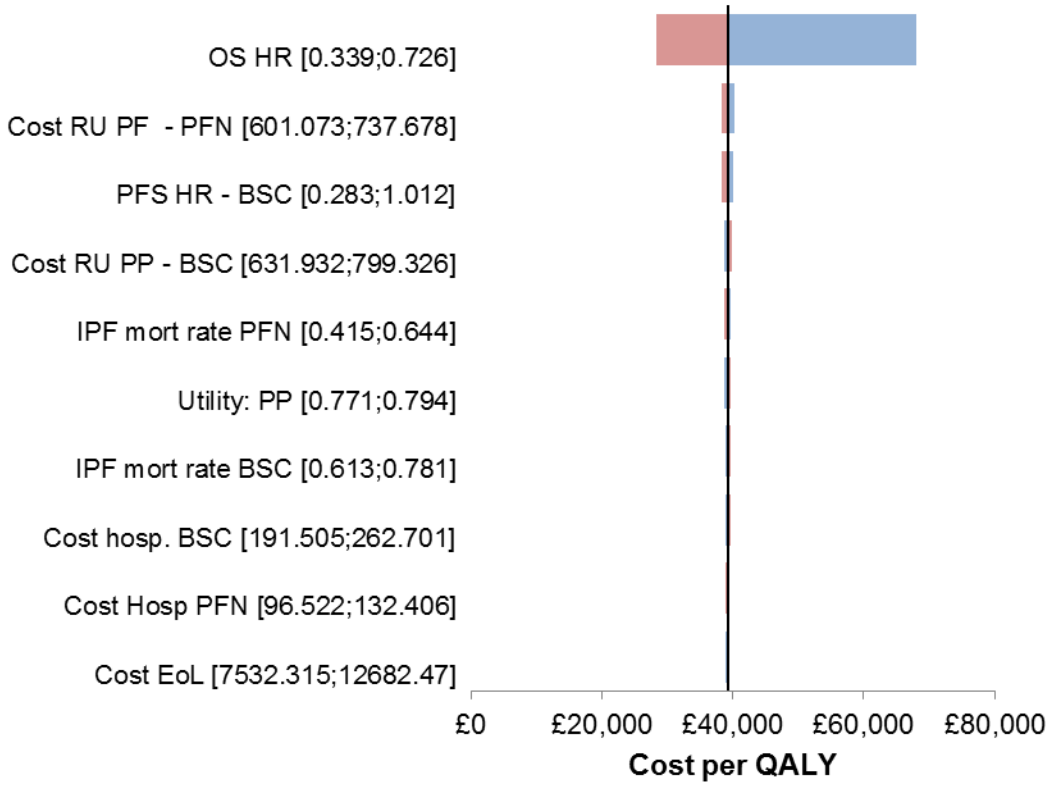
Table 208: PSA results – Moderate population – List price

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)	Incremental (QALYs)
Deterministic model results								
BSC	24,868	4.80	3.44					
NTB	65,065	6.06	4.23	40,197	1.26	0.78	51,331	51,331
PFN	[redacted]	7.67	[redacted]	[redacted]	1.61	0.91	[redacted]	[redacted]
Mean probabilistic model results								
BSC	24,755	4.80	3.42					
95% CI	(17,875; 32,799)	(2.98; 7.15)	(2.26; 4.81)					
NTB	63,792	5.95	4.10	39,038	1.15	0.69	56,976	56,976
95% CI	(48,499; 78,555)	(2.63; 10.49)	(2.06; 6.56)					
PFN	[redacted]	7.71	5.15	[redacted]	1.76	1.05	[redacted]	[redacted]
95% CI	[redacted]	(6.53; 9.05)	(4.53; 5.83)					

% chance of being cost effective	£20,000 per QALY	£25,000 per QALY	£30,000 per QALY
BSC	91%	81%	65%
NTB	9%	18%	24%
PFN	0%	2%	11%

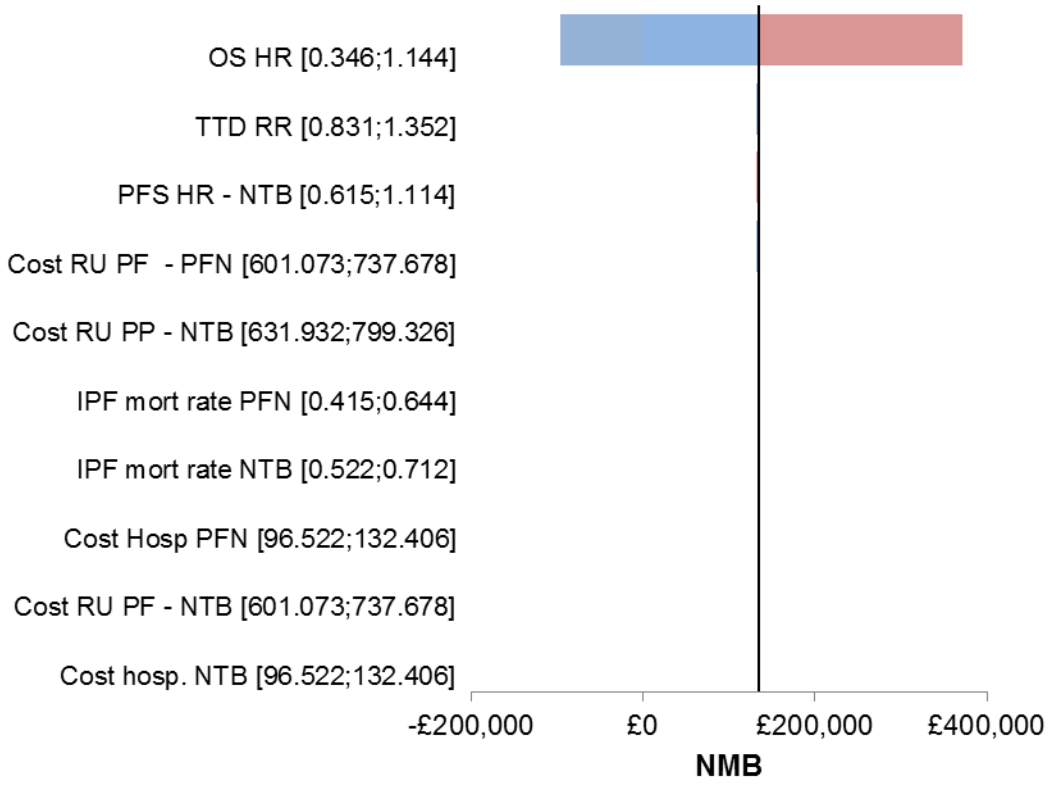
Key: BSC, best supportive care; CI, confidence interval; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NTB, nintedanib; PFN, pirfenidone; QALYs, quality-adjusted life years; TRT, treatment.

Figure 21: OWSA (vs. BSC; Cost per QALY) – Moderate population – List price



Key: BSC, best supportive care; EoL, end of life; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; mort, mortality; PF, progression-free; PFN, pirfenidone; PP, post-progression; QALY, quality-adjusted life year; RU, resource use.

Figure 22: OWSA (vs. NTB; net monetary benefit) – Moderate population – List price



Key: BSC, best supportive care; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; mort, mortality; PF, progression-free; PFN, pirfenidone; PP, post-progression; QALY, quality-adjusted life year; RU, resource use; TTD, time to treatment discontinuation.

Table 209: Scenario analysis – Moderate population – List price

	Category	Base case setting	Model change	PFN			BSC			ICER vs. BSC (£)	NTB			ICER vs. NTB (£)
				Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs		Costs (£)	LYs	QALYs	
Base Case				[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
1	Time horizon	Lifetime (34 years)	10 years	[redacted]	6.30	4.47	23,836	4.57	3.32	[redacted]	62,904	5.44	3.92	[redacted]
2			15 years	[redacted]	7.21	4.95	24,742	4.77	3.43	[redacted]	64,625	5.92	4.17	[redacted]
3			20 years	[redacted]	7.53	5.09	24,856	4.80	3.44	[redacted]	64,990	6.04	4.22	[redacted]
4			25 years	[redacted]	7.63	5.13	24,867	4.80	3.44	[redacted]	65,054	6.06	4.22	[redacted]
5			30 years	[redacted]	7.66	5.14	24,868	4.80	3.44	[redacted]	65,064	6.06	4.23	[redacted]
6	Utilities	Freemantle et al. (2015) mapping algorithm	PANTHER and ACE	[redacted]	7.67	4.89	24,868	4.80	3.28	[redacted]	65,065	6.06	4.03	[redacted]
7			Nintedanib NICE company submission	[redacted]	7.67	4.83	24,868	4.80	3.23	[redacted]	65,065	6.06	3.97	[redacted]
8			Starkie et al. (2012) mapping algorithm	[redacted]	7.67	4.86	24,868	4.80	3.25	[redacted]	65,065	6.06	4.00	[redacted]
9	Treatment effect	Treatment effect applied for the lifetime horizon	Treatment effect applied for up to 7 years	[redacted]	7.67	5.14	26,324	5.38	3.73	[redacted]	66,134	6.55	4.46	[redacted]
10			Treatment effect applied for up to 10 years	[redacted]	7.67	5.14	25,349	5.01	3.54	[redacted]	65,527	6.29	4.33	[redacted]
11			Treatment effect applied for up to 14 years	[redacted]	7.67	5.14	24,959	4.85	3.46	[redacted]	65,197	6.14	4.25	[redacted]

12	OS	Weibull	Exponential	[redacted]	10.83	6.48	27,374	5.92	3.97	[redacted]	68,280	8.09	5.14	[redacted]
13			Log-normal	[redacted]	11.68	6.76	28,052	6.18	4.06	[redacted]	69,741	8.62	5.31	[redacted]
14			Gamma	[redacted]	11.93	7.20	32,309	7.33	4.89	[redacted]	74,236	9.38	5.97	[redacted]
15			Log-Logistic	[redacted]	10.14	6.11	26,337	5.44	3.73	[redacted]	67,636	7.45	4.81	[redacted]
16			Gompertz	[redacted]	6.87	4.78	24,430	4.67	3.37	[redacted]	64,021	5.70	4.05	[redacted]
17		BSC trial data not applied for the first 52 weeks of the model.	BSC trial data applied for the first 52 weeks of the model.	[redacted]	7.67	5.14	25,215	4.89	3.50	[redacted]	65,065	6.06	4.23	[redacted]
18		Real-world data not applied for BSC OS	INOVA registry used for BSC OS	[redacted]	7.67	5.14	26,064	5.35	3.70	[redacted]	65,065	6.06	4.23	[redacted]
19			Edinburgh registry used for BSC OS	[redacted]	7.67	5.14	20,493	3.48	2.60	[redacted]	65,065	6.06	4.23	[redacted]
20	PFS	Weibull	Exponential	[redacted]	7.67	5.19	24,806	4.80	3.46	[redacted]	67,255	6.06	4.26	[redacted]
21			Log-normal	[redacted]	7.67	5.19	24,820	4.80	3.46	[redacted]	66,766	6.06	4.26	[redacted]
22			Gamma	[redacted]	7.67	5.07	25,041	4.80	3.38	[redacted]	53,870	6.06	4.16	[redacted]
23			Log-Logistic	[redacted]	7.67	5.18	24,827	4.80	3.46	[redacted]	66,483	6.06	4.26	[redacted]
24			Gompertz	[redacted]	7.67	5.13	24,865	4.80	3.44	[redacted]	65,170	6.06	4.22	[redacted]
25		BSC trial data not applied for the first 52 weeks of the	BSC trial data applied for the first 52 weeks of the model.	[redacted]	7.67	5.14	24,865	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]

		model.												
26	TTD	Weibull	Exponential	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	63,656	6.06	4.23	[redacted]
27			Log-normal	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	63,572	6.06	4.23	[redacted]
28			Gamma	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	82,439	6.06	4.23	[redacted]
29			Log-Logistic	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	64,649	6.06	4.23	[redacted]
30			Gompertz	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	64,634	6.06	4.23	[redacted]
31	Stopping rule	Applied for nintedanib patients only	Not applied	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	81,462	6.06	4.23	[redacted]
32			Applied for pirfenidone and nintedanib patients	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
33	OS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	7.67	5.14	24,655	4.74	3.40	[redacted]	65,458	6.18	4.30	[redacted]
34			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,458	6.18	4.30	[redacted]
35			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,263	6.12	4.26	[redacted]
36			Phase III trials only, random effects, 72 weeks cut off	[redacted]	7.67	5.14	26,896	5.45	3.84	[redacted]	68,009	7.00	4.76	[redacted]
37			Phase II and III trials, random effects, 72	[redacted]	7.67	5.14	26,896	5.45	3.84	[redacted]	67,664	6.89	4.70	[redacted]

			weeks cut off											
38			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	7.67	5.14	26,896	5.45	3.84	[redacted]	68,009	7.00	4.76	[redacted]
39			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	7.67	5.14	26,896	5.45	3.84	[redacted]	67,837	6.94	4.73	[redacted]
40	PFS - NMA	Phase II and III trials, random effects, 52 weeks cut off	Phase III trials only, random effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
41			Phase III trials only, fixed effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
42			Phase II and III trials, fixed effects, 52 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
43			Phase III trials only, random effects, 72 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
44			Phase II and III trials, random effects, 72 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
45			Phase III trials only, fixed effects, 72 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]

46			Phase II and III trials, fixed effects, 72 weeks cut off	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
47	TTD - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,170	6.06	4.23	[redacted]
48			Phase III trials only, fixed effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,275	6.06	4.23	[redacted]
49			Phase II and III trials, fixed effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,170	6.06	4.23	[redacted]
50	AEs - NMA	Phase II and III trials, random effects	Phase III trials only, random effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
51			Phase III trials only, fixed effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
52			Phase II and III trials, fixed effects	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
53			Phase II and III trials, random effects with adjustments in data for differences in end point	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
54			Phase III trials only, fixed effects with adjustments in data for differences in end point	[redacted]	7.67	5.14	24,868	4.80	3.44	[redacted]	65,065	6.06	4.23	[redacted]
55	Lung	Not applied	Applied	[redacted]	7.64	5.15	25,053	4.82	3.46	[redacted]	65,199	6.08	4.24	[redacted]

	transplant													
56	Gas Transfer	Test administered every 4 months	Test administered every 6 months	[redacted]	7.67	5.14	23,993	4.80	3.44	[redacted]	63,993	6.06	4.23	[redacted]
57	Full pulmonary	Test administered every 4 months	Test administered every 6 months	[redacted]	7.67	5.14	24,150	4.80	3.44	[redacted]	64,185	6.06	4.23	[redacted]
58	Lung Volume Transfer	No subsequent tests administered	Test administered every 4 months	[redacted]	7.67	5.14	27,082	4.80	3.44	[redacted]	67,779	6.06	4.23	[redacted]
59	Field exercise test	Test administered every 6 months	Test administered every 12 months	[redacted]	7.67	5.14	24,101	4.80	3.44	[redacted]	64,125	6.06	4.23	[redacted]
57			Test administered every 3 months if on oxygen	[redacted]	7.67	5.14	25,983	4.80	3.44	[redacted]	66,433	6.06	4.23	[redacted]
58	Healthcare professional visit	Test administered every 4 months if FVC >60%, and every 3 months if FVC<60%	Healthcare professional visit - Sub MRU = every 6 months if FVC >60%, every 3 months if FVC<60%	[redacted]	7.67	5.14	24,401	4.80	3.44	[redacted]	64,494	6.06	4.23	[redacted]

Key: BSC, best supportive care; FVC, forced vital capacity; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; MRU, medical resource use;

NMA network meta-analysis; NTB, nintedanib; OS, overall survival; PFN, pirfenidone; PFS, progression-free survival; Ph2, Phase II; Ph3, Phase III; QALY, quality-adjusted life year; RCT, randomised controlled trial; SA, sensitivity analysis; STA, single technology appraisal; Sub, subsequent; TTD, time to treatment discontinuation.

Appendix F: Response to B12c

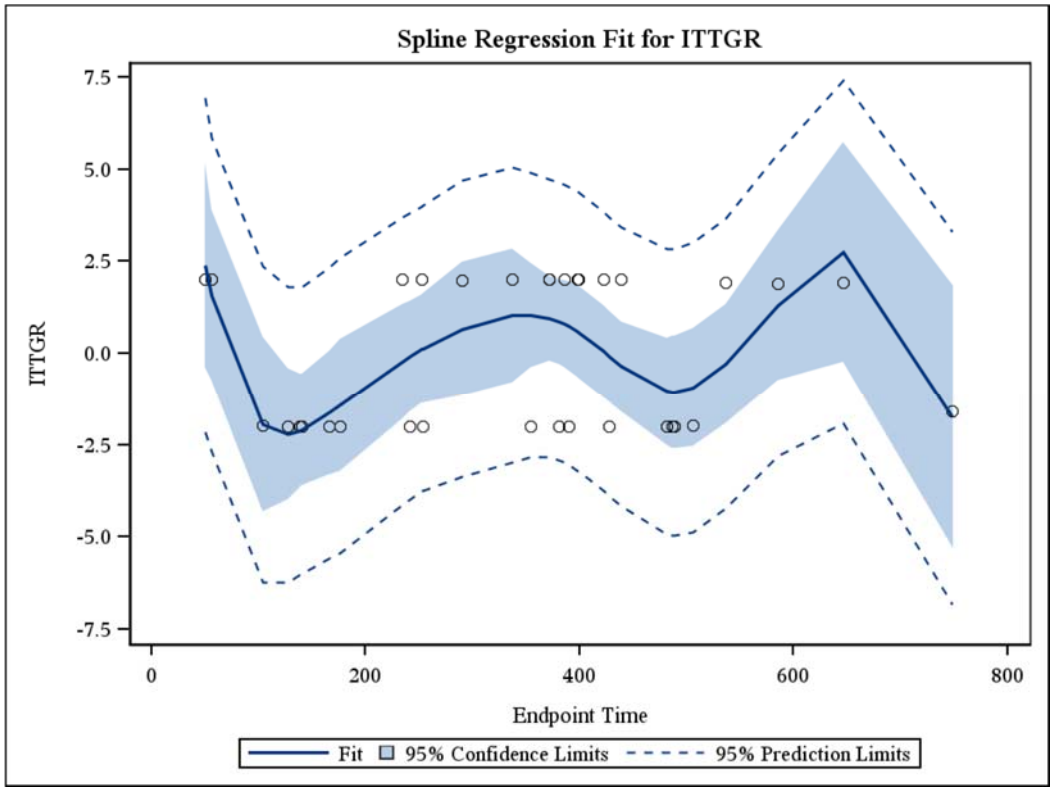
a) CAPACITY 1 analysis

1) OS endpoint

Table 210: CAPACITY 1 analysis, OS endpoint – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	-1.11708	3.10164	0.1297	0.7187	0.327
treat_time_int	1	0.18719	0.54243	0.1191	0.7300	1.206

Figure 23: CAPACITY 1 analysis, OS endpoint – scaled Schoenfeld residuals over time (in days)

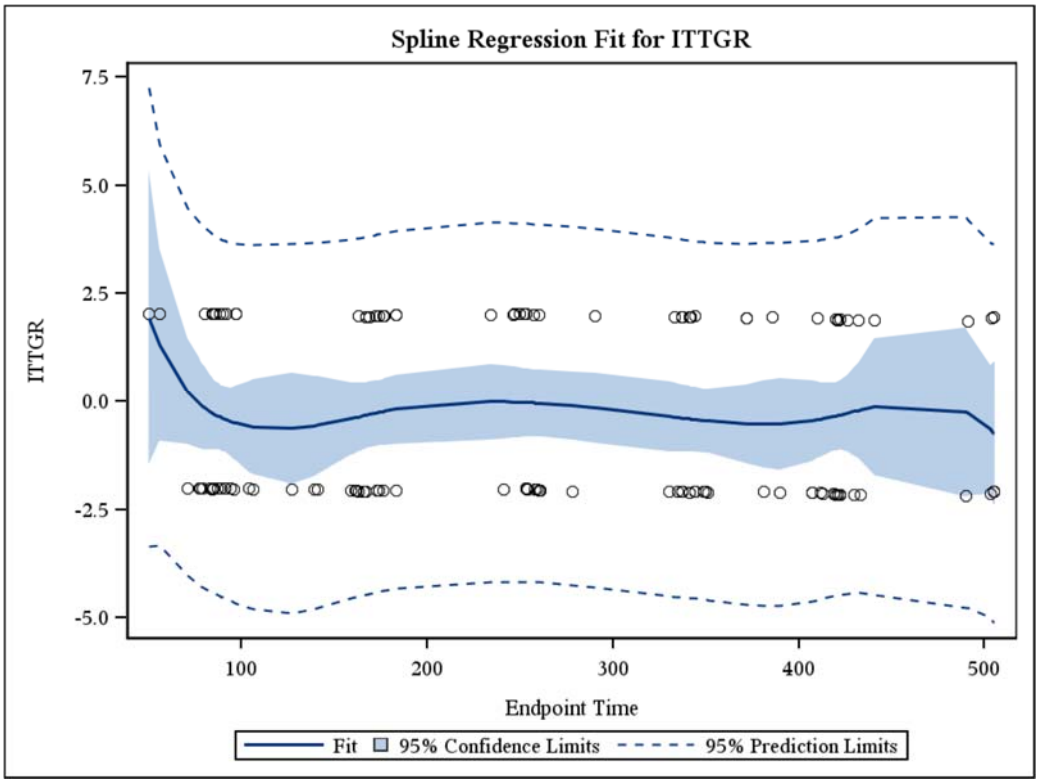


2) PFS endpoint

Table 211: CAPACITY 1 analysis, PFS endpoint – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	0.58117	1.51148	0.1478	0.7006	1.788
treat_time_int	1	-0.16013	0.27778	0.3323	0.5643	0.852

Figure 24: CAPACITY 1 analysis, PFS endpoint – scaled Schoenfeld residuals over time (in days)



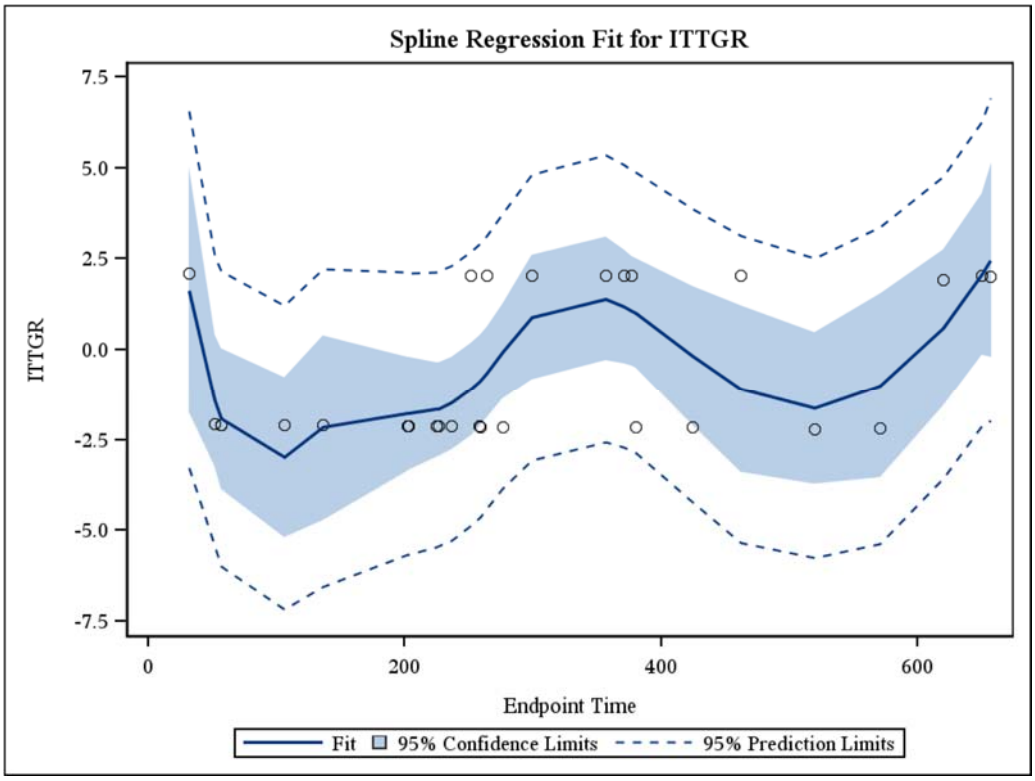
b) CAPACITY 2 analysis

1) OS endpoint

Table 212: CAPACITY 2 analysis, OS endpoint – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	-5.53701	3.60231	2.3626	0.1243	0.004
treat_time_int	1	0.90637	0.63400	2.0437	0.1528	2.475

Figure 25: CAPACITY 2 analysis, OS endpoint – scaled Schoenfeld residuals over time (in days)

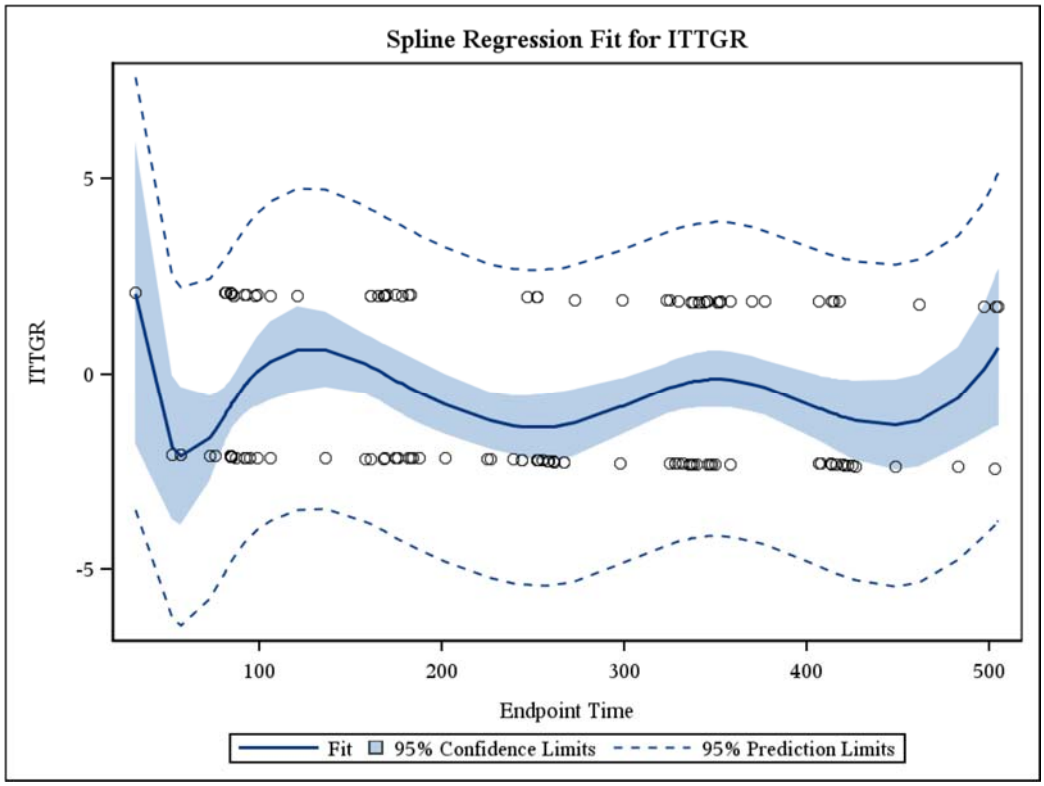


2) PFS endpoint

Table 213: CAPACITY 2 analysis, PFS endpoint – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	-0.53325	1.43265	0.1385	0.7097	0.587
treat_time_int	1	-0.00424	0.26790	0.0003	0.9874	0.996

Figure 26: CAPACITY 2 analysis, PFS endpoint – scaled Schoenfeld residuals over time (in days)



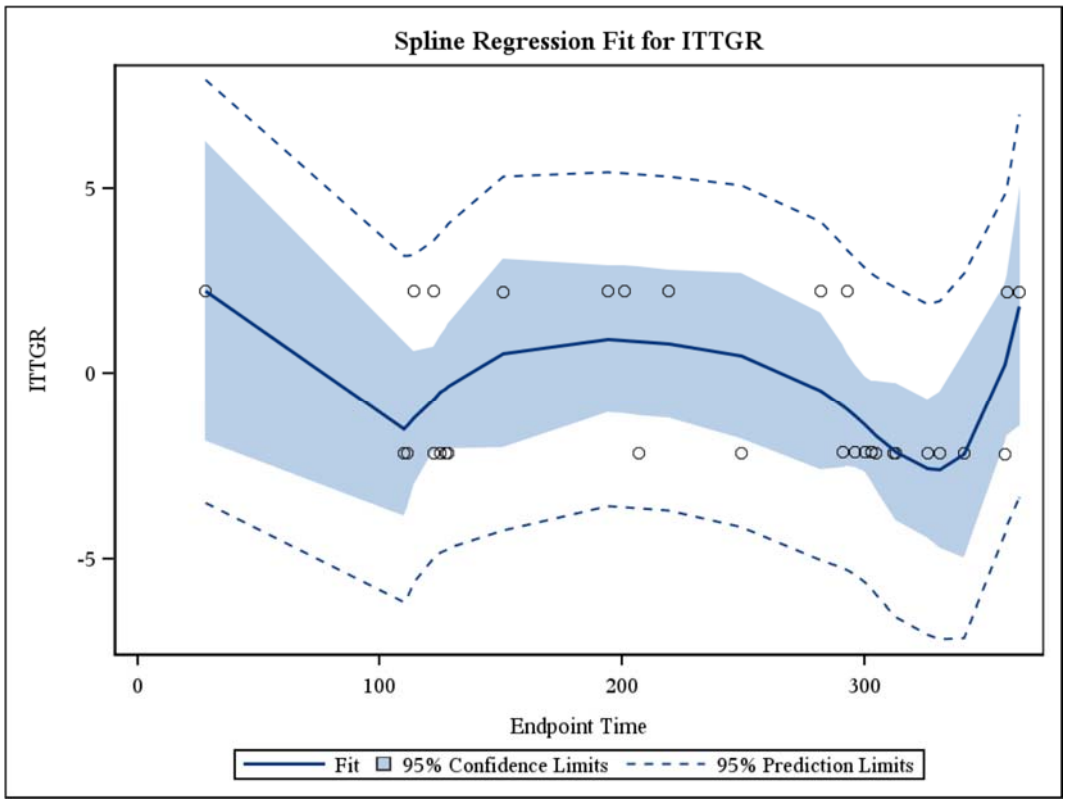
c) ASCEND analysis

1) OS endpoint

Table 214: ASCEND analysis, OS endpoint – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	3.74355	3.77147	0.9852	0.3209	42.248
treat_time_int	1	-0.81690	0.70684	1.3356	0.2478	0.442

Figure 27: ASCEND analysis, OS endpoint – scaled Schoenfeld residuals over time (in days)

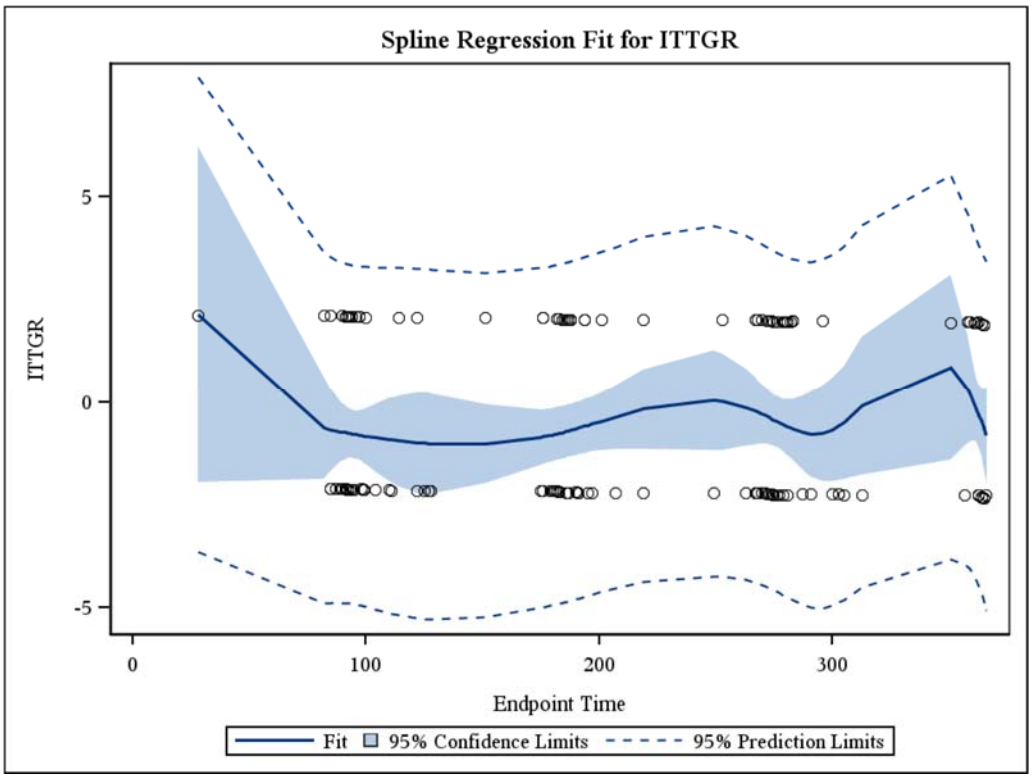


2) PFS endpoint

Table 215: ASCEND analysis, PFS endpoint – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	-1.65336	1.60339	1.0633	0.3025	0.191
treat_time_int	1	0.20576	0.30326	0.4603	0.4975	1.228

Figure 28: ASCEND analysis, PFS endpoint – scaled Schoenfeld residuals over time (in days)



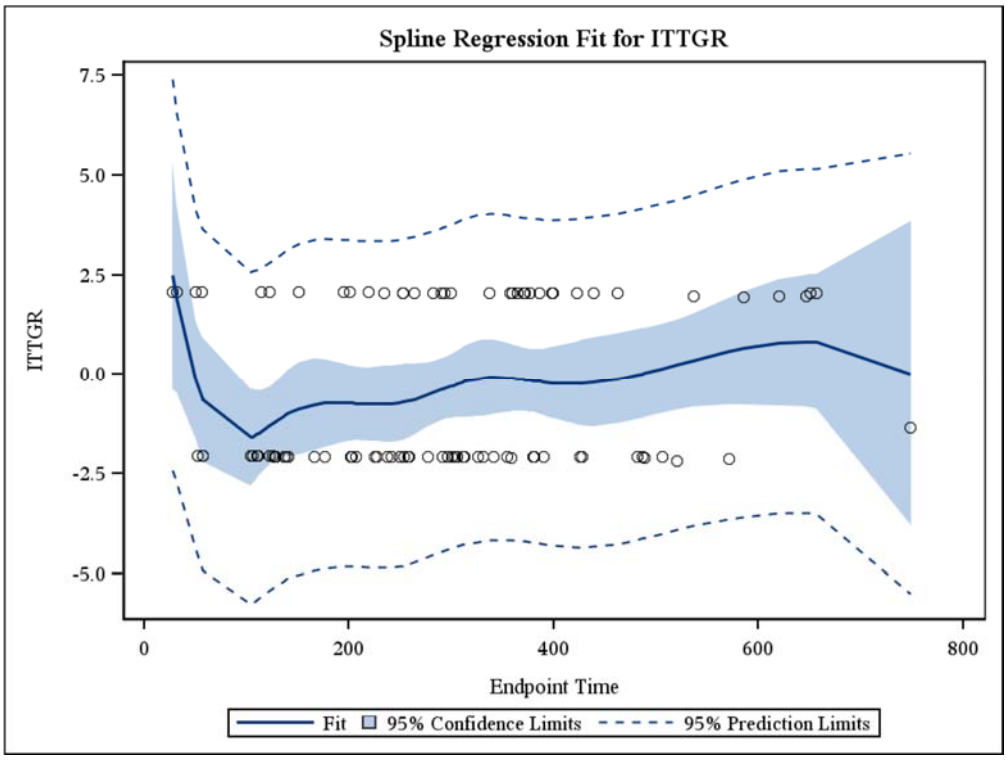
d) Pooled study analysis

1) OS endpoint

Table 216: Pooled analysis, OS endpoint – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	-1.88346	1.79623	1.0995	0.2944	0.152
treat_time_int	1	0.27513	0.32254	0.7277	0.3936	1.317

Figure 29: Pooled analysis, OS endpoint – scaled Schoenfeld residuals over time (in days)

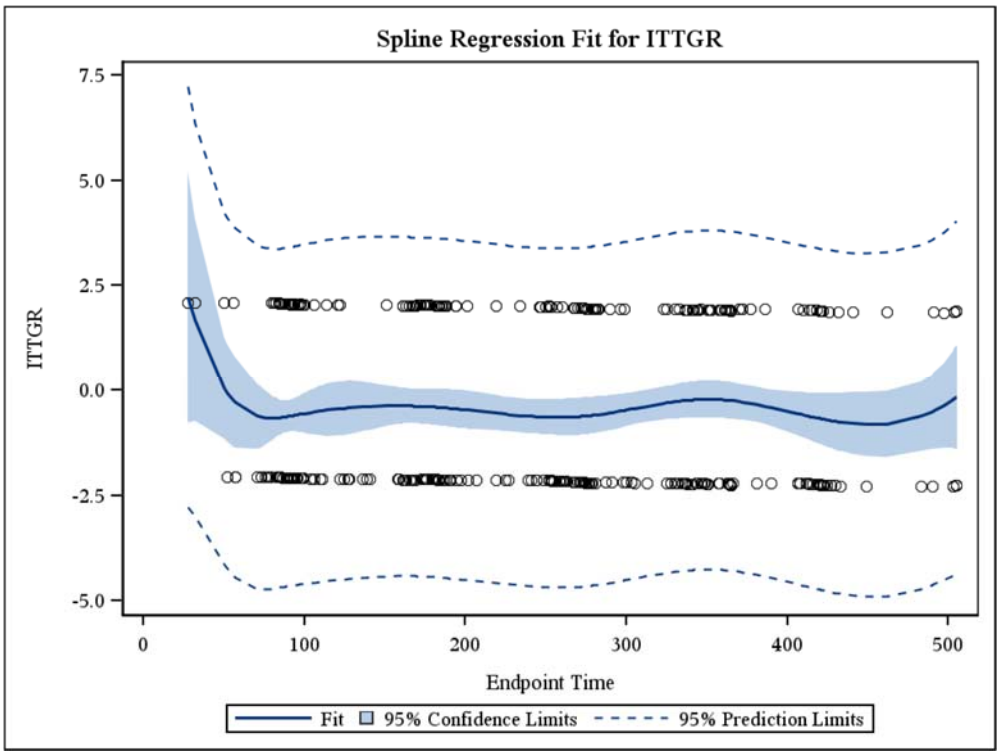


2) PFS endpoint

Table 217: Pooled analysis, PFS endpoint – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	-0.62000	0.86294	0.5162	0.4725	0.538
treat_time_int	1	0.02661	0.16126	0.0272	0.8689	1.027

Figure 30: Pooled analysis, PFS endpoint – scaled Schoenfeld residuals over time (in days)



Appendix G: Response to B19b

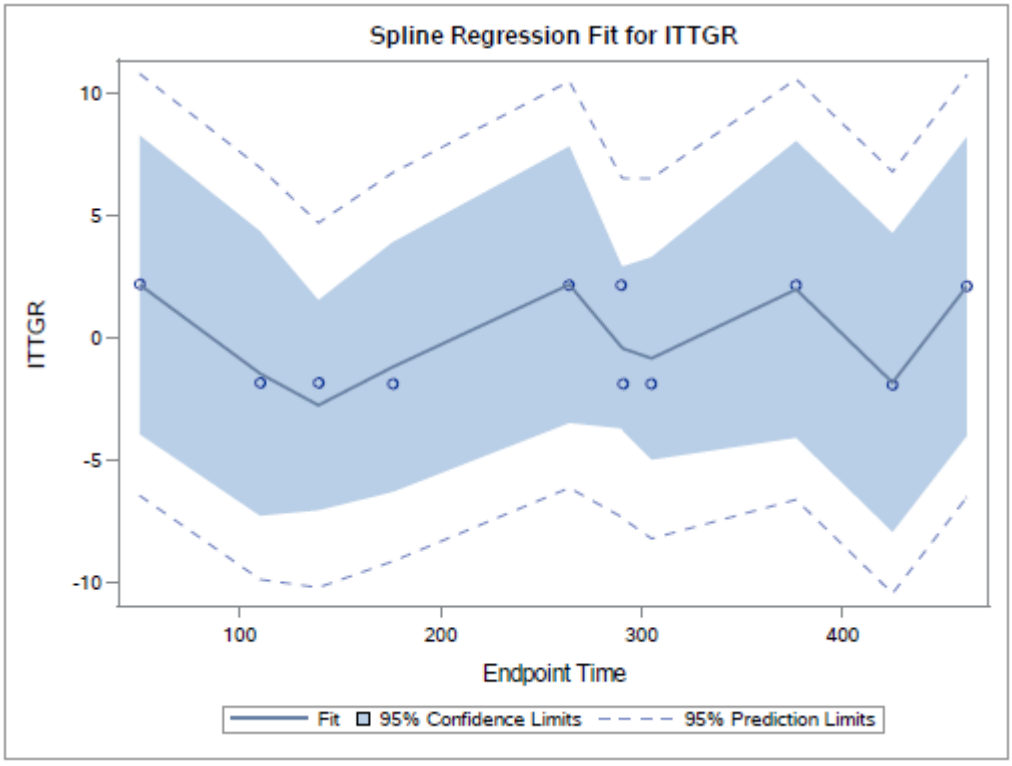
a) Patients with mild IPF at baseline (FVC ≥80% predicted)

1) OS endpoint

Table 218: Pooled trials, OS endpoint at 72 weeks – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	-0.84640	5.20796	0.0264	0.8709	0.429
treat_time_int	1	0.14611	0.95522	0.0234	0.8784	1.157

Figure 31: Pooled trials, OS endpoint at 72 weeks – scaled Schoenfeld residuals over time (in days)

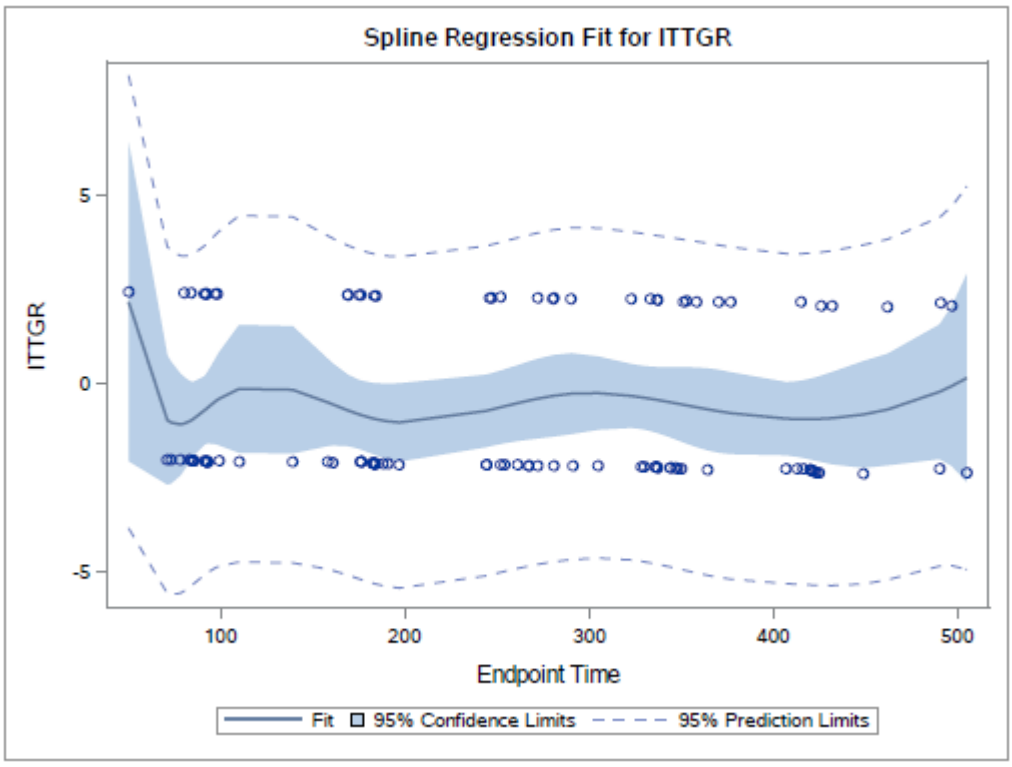


2) PFS endpoint

Table 219: Pooled trials, PFS endpoint at 72 weeks – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	- 0.76908	1.87403	0.1684	0.6815	0.463
treat_time_int	1	0.02075	0.34544	0.0036	0.9521	1.021

Figure 32: Pooled trials, PFS endpoint at 72 weeks – scaled Schoenfeld residuals over time (in days)



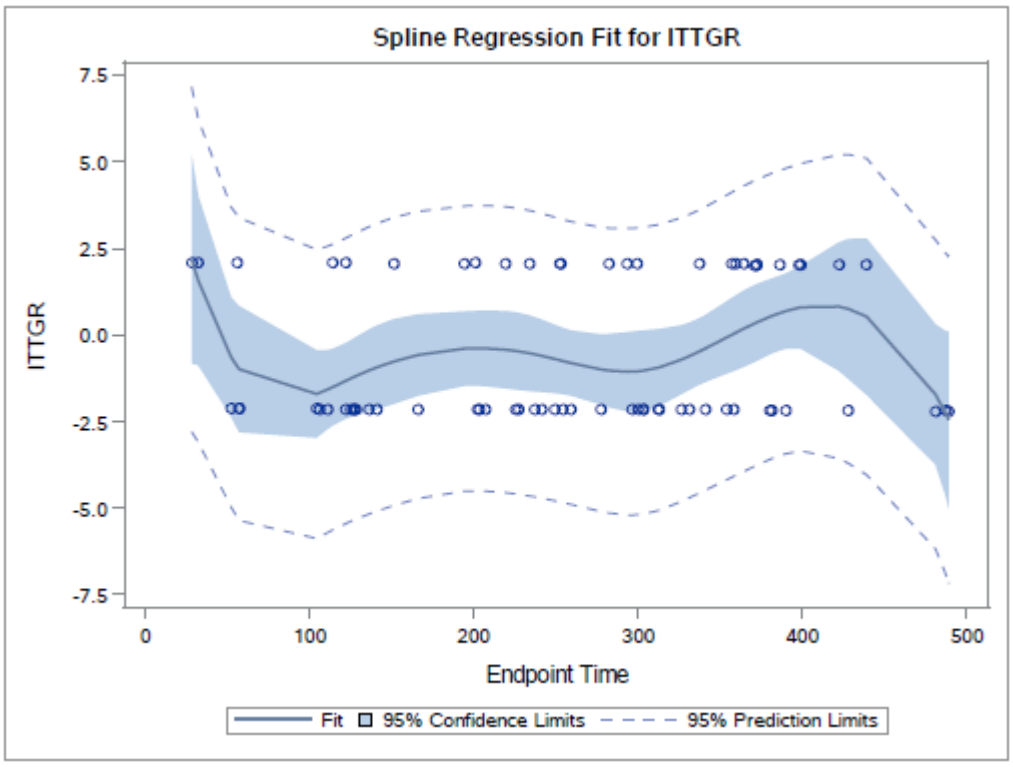
b) Patients with moderate IPF at baseline (FVC 50-80% predicted)

1) OS endpoint

Table 220: Pooled trials, OS endpoint at 72 weeks – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	- 0.91557	2.09007	0.1919	0.6613	0.400
treat_time_int	1	0.07211	0.38441	0.0352	0.8512	1.075

Figure 33: Pooled trials, OS endpoint at 72 weeks – scaled Schoenfeld residuals over time (in days)



2) PFS endpoint

Table 221: Pooled trials, PFS endpoint at 72 weeks – test for interaction between time and treatment

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Parameter estimate**	**Standard error**	**Chi-square**	**Pr > ChiSq**	**Hazard ratio**
ITTGR	1	- 0.64500	0.98511	0.4287	0.5126	0.525
treat_time_int	1	0.04103	0.18473	0.0493	0.8242	1.042

Figure 34: Pooled trials, PFS endpoint at 72 weeks – scaled Schoenfeld residuals over time (in days)

ID837 – Addendum to Roche responses to clarification questions (18 March)

A21. ***Priority question:*** *Section 4.4, p.69 and Section 4.7, p.90, Table 18: Please refer to the NICE guide to the methods of technology appraisal and provide additional detail on the method of pooling data from ASCEND, CAPACITY 1 and CAPACITY 2 (including data synthesis methods and the rationale, software and models used, and any methodological limitations).*

Pooling data from the ASCEND and CAPACITY studies was requested by the FDA and consequently pre-specified in the statistical analysis plan of ASCEND and the Integrated Summary of Efficacy (ISE) submitted to the EMA. The pooling of trial data is justified when considering the similarity in:

1. The design of the three studies;
2. Patient characteristics between the studies, and;
3. Outcomes at 12 months

Individual patient data from ASCEND, CAPACITY 1 and CAPACITY 2 were pooled and analysed in a model, including study and region of the world (US vs rest of the world) as stratification factors. Region of the world was a pre-specified variable in the Statistical Analysis Plans of the CAPACITY studies, with randomisation also accounting for this factor. These studies were designed and conducted prior to Roche acquiring the rights to pirfenidone.

Details of the methods of pooling and model used are provided below by type of endpoint. Analyses were run on SAS software. While the InterMune analyses were run on a Windows platform, the Roche analyses have been performed on a Unix system.

Pooling of mortality endpoints: The pooled hazard ratio was estimated by stratifying by study and region of the world. This approach assumes an underlying common effect across strata. The strata specific estimates are combined in a single measure of effect.

Frequentist approaches did not detect any heterogeneity in mortality endpoints (heterogeneity equal to zero) at 12 months [Nathan 2016].

Pooling of lung function decline data (Rank ANCOVA): The primary efficacy outcome of the pooled analysis was the absolute change in percent predicted FVC from Baseline to Month 12, which was analysed using a rank ANCOVA model. Use of ranks as opposed to crude values is advantageous when the outcome of interest is not normally distributed.

For the analysis of ranks patients who died were ranked according to time from randomization until death. Patients who did not die were ranked according to their change in percent predicted FVC from Baseline.

The ranks were standardised: i.e. ranks (1,2,...,n) are transformed into numbers which are distributed between 0 and 1. This standardisation can aid in the interpretation of results: 0.5

corresponds to the median, the smallest value closest to 0 represents the worst outcome and the largest value closest to 1 represents the best outcome.

The rank ANCOVA model to estimate the effect of treatment was stratified by study (ASCEND, CAPACITY 1 and CAPACITY 2) and region of the world (US vs. rest of the world) and was adjusted for baseline percent predicted FVC to adjust for baseline variability in lung function.

The stratification assumes an underlying common effect across strata, and combines the result from the strata in a single measure of effect.

The effect of treatment on ranks observed in the three studies separately has been supplied as part of our earlier response to clarification question (sent to NICE 3 March; A28, Tables 8 to 10). The effect on ranks (difference in standardised ranks) was similar across studies.

The p-values reported in Table 18 of our initial submission were intended to refer to the analysis of standardised ranks, as indicated by the footnote to the table. There was, however, a typographical error in the results for the CAPACITY studies: the correct p-values are:

- CAPACITY 2: 0.001;
- CAPACITY 1: **0.501 (instead of 0.440)**, and**;**
- Pooled analysis: **0.005 (instead of 0.003)**, as indicated in Figure 2 of the 2001 Noble publication).

The p-values for the CAPACITY data referred to in Table 18 of the submission refer to the Cochran-Mantel Haenszel test for categorical endpoints, as reported in Table 2 of the Noble publication, and are meant for the analysis of the ordinal categorical endpoint.

Pooling of lung function decline data (ordinal data analysis): Table 18 of the initial submission document also reported analysis of change from baseline in percent predicted FVC, or death, as ordinal categorical data. This analysis was performed and reported for the CAPACITY studies (Noble 2011, Table 2). The change in ordinal endpoint was estimated for each study by a Cochran-Mantel-Haenszel row mean score test to compare pirfenidone 2403 mg/day vs placebo based on five ordinal categories (severe decline, ≥20%; moderate decline, <20% but ≥10%; mild decline; <10% but ≥0; mild improvement, >0 but <10%; and moderate improvement, ≥10%).

Analogously to the other endpoints this, analysis was stratified by region of the world.
The corresponding p-values are as reported by Noble: 0.001 for CAPACITY 2 and 0.440 for CAPACITY 1 [Noble 2011].

A24. ***Priority question:*** *Section 4.5, pp.76-80: Please provide full details of the nature and grade/severity of adverse events that led to discontinuation in each of the trial arms.*

REVISED QUESTION FROM NICE: *Please provide details of the adverse events that led to discontinuation, and whether any would be classified as severe, for ASCEND and CAPACITY 1 & 2. Data in a format similar to that presented in Taniguchi 2010 (table 2) for SP3 would be acceptable.*

Summary tables of adverse events which led to discontinuation in the ASCEND, CAPACITY 1 and CAPACITY 2 trials are provided in Table 1 to Table 6. Tables are provided, by trial, for adverse events which led to discontinuation of study treatment (Table 1 to Table 3), and adverse events which led to patient discontinuation from the study (Table 4 to Table 6). Each table includes detail on the adverse event intensity, including those that were classified as severe.

Table 1: Adverse Event Leading to Discontinuation of Study Drug (as recorded on the Early Discontinuation of Treatment CRF) in ASCEND

Reason for Discontinuation	Intensity	Pirfenidone 2403 mg/day (N=278)	Placebo (N=277)	Total (N=555)
Any Adverse Events	- Any Intensity -	35 (12.6%)	24 (8.7%)	59 (10.6%)
	Mild	9 (3.2%)	2 (0.7%)	11 (2%)
	Moderate	19 (6.8%)	5 (1.8%)	24 (4.3%)
	Severe	5 (1.8%)	8 (2.9%)	13 (2.3%)
	Life Threatening	2 (0.7%)	9 (3.2%)	11 (2%)
Abdominal Discomfort	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Mild	1 (0.4%)	0	1 (0.2%)
Abdominal Pain Upper	- Any Intensity -	0	2 (0.7%)	2 (0.4%)
	Mild	0	1 (0.4%)	1 (0.2%)
	Moderate	0	1 (0.4%)	1 (0.2%)
Acute Respiratory Failure	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Life Threatening	0	1 (0.4%)	1 (0.2%)
Anorexia	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Mild	1 (0.4%)	0	1 (0.2%)
Aspartate Aminotransferase Increased	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Bladder Pain	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Mild	1 (0.4%)	0	1 (0.2%)
Breast Cancer	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Moderate	0	1 (0.4%)	1 (0.2%)
Bronchopneumonia	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Life Threatening	0	1 (0.4%)	1 (0.2%)
Dermatomyositis	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Dysgeusia	- Any Intensity -	2 (0.7%)	0	2 (0.4%)
	Mild	2 (0.7%)	0	2 (0.4%)
Gamma-Glutamyltransferase Increased	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)

Reason for Discontinuation	Intensity	Pirfenidone 2403 mg/day (N=278)	Placebo (N=277)	Total (N=555)
Gastrointestinal Disorder	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Gastrooesophageal Reflux Disease	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
General Physical Health Deterioration	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Haemorrhage Intracranial	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Life Threatening	0	1 (0.4%)	1 (0.2%)
Hepatic Enzyme Increased	- Any Intensity -	3 (1.1%)	0	3 (0.5%)
	Moderate	3 (1.1%)	0	3 (0.5%)
Idiopathic Pulmonary Fibrosis	- Any Intensity -	0	12 (4.3%)	12 (2.2%)
	Moderate	0	1 (0.4%)	1 (0.2%)
	Severe	0	5 (1.8%)	5 (0.9%)
	Life Threatening	0	6 (2.2%)	6 (1.1%)
Malaise	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Severe	1 (0.4%)	0	1 (0.2%)
Nausea	- Any Intensity -	2 (0.7%)	0	2 (0.4%)
	Mild	1 (0.4%)	0	1 (0.2%)
	Severe	1 (0.4%)	0	1 (0.2%)
Oedema Peripheral	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Mild	0	1 (0.4%)	1 (0.2%)
Photosensitivity Reaction	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Pneumonia	- Any Intensity -	3 (1.1%)	1 (0.4%)	4 (0.7%)
	Severe	1 (0.4%)	1 (0.4%)	2 (0.4%)
	Life Threatening	2 (0.7%)	0	2 (0.4%)
Pneumonia Respiratory Syncytial Viral	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Severe	0	1 (0.4%)	1 (0.2%)
Pruritus	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)

Reason for Discontinuation	Intensity	Pirfenidone 2403 mg/day (N=278)	Placebo (N=277)	Total (N=555)
Psychotic Disorder	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Pulmonary Alveolar Haemorrhage	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Severe	0	1 (0.4%)	1 (0.2%)
Pyrexia	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Rash	- Any Intensity -	3 (1.1%)	0	3 (0.5%)
	Mild	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
	Severe	1 (0.4%)	0	1 (0.2%)
Rash Pruritic	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Mild	1 (0.4%)	0	1 (0.2%)
Renal Failure	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Severe	1 (0.4%)	0	1 (0.2%)
Throat Tightness	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Urticaria	- Any Intensity -	1 (0.4%)	1 (0.4%)	2 (0.4%)
	Mild	1 (0.4%)	0	1 (0.2%)
	Moderate	0	1 (0.4%)	1 (0.2%)
Vomiting	- Any Intensity -	1 (0.4%)	1 (0.4%)	2 (0.4%)
	Moderate	1 (0.4%)	1 (0.4%)	2 (0.4%)
Weight Decreased	- Any Intensity -	3 (1.1%)	0	3 (0.5%)
	Moderate	3 (1.1%)	0	3 (0.5%)

Note: Patients with a Discontinuation do not Always have a Corresponding AE.
[a] Patients discontinued the study drug due to AE whereas this action was not recorded on the AE Case Report Form.

Table 2: Adverse Event Leading to Discontinuation of Study Drug (as recorded on the Early Discontinuation of Treatment CRF) in CAPACITY 1

Reason for Discontinuation	Intensity	Pirfenidone 2403 mg/day (N=171)	Placebo (N=173)	Total (N=344)
Any Adverse Events	- Any Intensity -	24 (14.0%)	14 (8.1%)	38 (11.0%)
	Mild	2 (1.2%)	1 (0.6%)	3 (0.9%)
	Moderate	7 (4.1%)	3 (1.7%)	10 (2.9%)
	Severe	8 (4.7%)	6 (3.5%)	14 (4.1%)
	Life Threatening	5 (2.9%)	2 (1.2%)	7 (2%)
	Missing[a]	2 (1.2%)	2 (1.2%)	4 (1.2%)
Alanine Aminotransferase Increased	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Severe	0	1 (0.6%)	1 (0.3%)
Back Pain	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Mild	1 (0.6%)	0	1 (0.3%)
Bladder Cancer	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Severe	1 (0.6%)	0	1 (0.3%)
Blood Creatinine Increased	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Severe	0	1 (0.6%)	1 (0.3%)
Cerebrovascular Accident	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Life Threatening	1 (0.6%)	0	1 (0.3%)
Dermatitis	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Moderate	1 (0.6%)	0	1 (0.3%)
Diarrhoea	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Moderate	0	1 (0.6%)	1 (0.3%)
Femur Fracture	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Severe	1 (0.6%)	0	1 (0.3%)
Flatulence	- Any Intensity -	1 (0.6%)	0	1 (0.3%)

Reason for Discontinuation	Intensity	Pirfenidone 2403 mg/day (N=171)	Placebo (N=173)	Total (N=344)
	Mild	1 (0.6%)	0	1 (0.3%)
Hepatic Enzyme Increased	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Moderate	0	1 (0.6%)	1 (0.3%)
Idiopathic Pulmonary Fibrosis	- Any Intensity -	4 (2.3%)	4 (2.3%)	8 (2.3%)
	Severe	3 (1.8%)	2 (1.2%)	5 (1.5%)
	Life Threatening	1 (0.6%)	2 (1.2%)	3 (0.9%)
Liver Function Test Abnormal	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Severe	1 (0.6%)	0	1 (0.3%)
Metastatic Neoplasm	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Severe	0	1 (0.6%)	1 (0.3%)
Myocardial Infarction	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Life Threatening	1 (0.6%)	0	1 (0.3%)
Nausea	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Moderate	1 (0.6%)	0	1 (0.3%)
Ovarian Cancer	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Severe	0	1 (0.6%)	1 (0.3%)
Photosensitivity Reaction	- Any Intensity -	2 (1.2%)	0	2 (0.6%)
	Moderate	2 (1.2%)	0	2 (0.6%)
Pleural Effusion	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Moderate	0	1 (0.6%)	1 (0.3%)
Rash	- Any Intensity -	2 (1.2%)	0	2 (0.6%)
	Moderate	2 (1.2%)	0	2 (0.6%)
Rectal Haemorrhage	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Mild	0	1 (0.6%)	1 (0.3%)
Renal Failure Acute	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Severe	1 (0.6%)	0	1 (0.3%)
Respiratory Failure	- Any Intensity -	2 (1.2%)	0	2 (0.6%)
	Severe	1 (0.6%)	0	1 (0.3%)
	Life Threatening	1 (0.6%)	0	1 (0.3%)
Septic Shock	- Any Intensity -	1 (0.6%)	0	1 (0.3%)

Reason for Discontinuation	Intensity	Pirfenidone 2403 mg/day (N=171)	Placebo (N=173)	Total (N=344)
	Life Threatening	1 (0.6%)	0	1 (0.3%)
Weight Decreased	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Moderate	1 (0.6%)	0	1 (0.3%)
Note: Patients with a Discontinuation do not Always have a Corresponding AE. [a] Patients discontinued the study drug due to AE whereas this action was not recorded on the AE Case Report Form.				

Table 3: Adverse Event Leading to Discontinuation of Study Drug (as recorded on the Early Discontinuation of Treatment CRF) in CAPACITY 2

Reason for Discontinuation	Intensity	Pirfenidone 1197 mg/day (N=87)	Pirfenidone 2403 mg/day (N=174)	Placebo (N=174)	Total (N=435)
Any Adverse Events	- Any Intensity -	11 (12.6%)	21 (12.1%)	14 (8.0%)	46 (10.6%)
	Mild	1 (1.1%)	0	0	1 (0.2%)
	Moderate	1 (1.1%)	10 (5.7%)	3 (1.7%)	14 (3.2%)
	Severe	3 (3.4%)	9 (5.2%)	6 (3.4%)	18 (4.1%)
	Life Threatening	4 (4.6%)	2 (1.1%)	5 (2.9%)	11 (2.5%)
	Missing[a]	2 (2.3%)	0	0	2 (0.5%)
Acute Respiratory Distress Syndrome	- Any Intensity -	0	0	1 (0.6%)	1 (0.2%)
	Life Threatening	0	0	1 (0.6%)	1 (0.2%)
Acute Respiratory Failure	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Severe	0	1 (0.6%)	0	1 (0.2%)
Angina Pectoris	- Any Intensity -	0	0	1 (0.6%)	1 (0.2%)
	Moderate	0	0	1 (0.6%)	1 (0.2%)
Bladder Cancer	- Any Intensity -	0	2 (1.1%)	0	2 (0.5%)
	Severe	0	1 (0.6%)	0	1 (0.2%)
	Life Threatening	0	1 (0.6%)	0	1 (0.2%)
Bone Marrow Failure	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Moderate	0	1 (0.6%)	0	1 (0.2%)
Clostridium Difficile Colitis	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Life Threatening	1 (1.1%)	0	0	1 (0.2%)

Colon Cancer	- Any Intensity -	0	0	1 (0.6%)	1 (0.2%)
	Severe	0	0	1 (0.6%)	1 (0.2%)
Confusional State	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Moderate	1 (1.1%)	0	0	1 (0.2%)
Diarrhoea	- Any Intensity -	0	0	1 (0.6%)	1 (0.2%)
	Moderate	0	0	1 (0.6%)	1 (0.2%)
Dyspnoea	- Any Intensity -	0	0	1 (0.6%)	1 (0.2%)
	Moderate	0	0	1 (0.6%)	1 (0.2%)
Gamma-Glutamyltransferase Increased	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Severe	1 (1.1%)	0	0	1 (0.2%)
Reason for Discontinuation	**Intensity**	**Pirfenidone 1197 mg/day (N=87)**	**Pirfenidone 2403 mg/day (N=174)**	**Placebo (N=174)**	**Total (N=435)**
Gastritis	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Moderate	0	1 (0.6%)	0	1 (0.2%)
Gastrooesophageal Reflux Disease	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Severe	0	1 (0.6%)	0	1 (0.2%)
Headache	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Moderate	0	1 (0.6%)	0	1 (0.2%)
Idiopathic Pulmonary Fibrosis	- Any Intensity -	1 (1.1%)	2 (1.1%)	3 (1.7%)	6 (1.4%)
	Severe	0	2 (1.1%)	1 (0.6%)	3 (0.7%)
	Life Threatening	1 (1.1%)	0	2 (1.1%)	3 (0.7%)
Liver Function Test Abnormal	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Severe	1 (1.1%)	0	0	1 (0.2%)
Lobar Pneumonia	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Life Threatening	1 (1.1%)	0	0	1 (0.2%)
Lung Cancer Metastatic	- Any Intensity -	0	0	1 (0.6%)	1 (0.2%)
	Life Threatening	0	0	1 (0.6%)	1 (0.2%)
Lymph Node Cancer Metastatic	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Severe	1 (1.1%)	0	0	1 (0.2%)
Malaise	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Moderate	0	1 (0.6%)	0	1 (0.2%)
Muscle Rupture	- Any Intensity -	0	0	1 (0.6%)	1 (0.2%)

	Severe	0	0	1 (0.6%)	1 (0.2%)
Myocardial Infarction	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Life Threatening	1 (1.1%)	0	0	1 (0.2%)
Nausea	- Any Intensity -	0	4 (2.3%)	0	4 (0.9%)
	Moderate	0	3 (1.7%)	0	3 (0.7%)
	Severe	0	1 (0.6%)	0	1 (0.2%)
Photosensitivity Reaction	- Any Intensity -	0	1 (0.6%)	1 (0.6%)	2 (0.5%)
	Severe	0	1 (0.6%)	1 (0.6%)	2 (0.5%)
Pneumonia	- Any Intensity -	0	0	2 (1.1%)	2 (0.5%)
	Severe	0	0	1 (0.6%)	1 (0.2%)
	Life Threatening	0	0	1 (0.6%)	1 (0.2%)
Pulmonary Haemorrhage	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Life Threatening	0	1 (0.6%)	0	1 (0.2%)
Rash	- Any Intensity -	0	3 (1.7%)	0	3 (0.7%)
	Moderate	0	2 (1.1%)	0	2 (0.5%)
	Severe	0	1 (0.6%)	0	1 (0.2%)
Rash Papular	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Mild	1 (1.1%)	0	0	1 (0.2%)
Sunburn	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Severe	0	1 (0.6%)	0	1 (0.2%)
Thrombotic Stroke	- Any Intensity -	0	0	1 (0.6%)	1 (0.2%)
	Severe	0	0	1 (0.6%)	1 (0.2%)
Weight Decreased	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Moderate	0	1 (0.6%)	0	1 (0.2%)

Note: Patients with a discontinuation do not always have a corresponding AE.
[a] Patients discontinued the study drug due to AE whereas this action was not recorded on the AE Case Report Form.

Table 4: Adverse Event Leading to Discontinuation of Study (as recorded on the End of Study CRF) in ASCEND

Reason for Discontinuation	Intensity	Pirfenidone 2403 mg/day (N=278)	Placebo (N=277)	Total (N=555)
Any Adverse Events	- Any Intensity -	6 (2.2%)	7 (2.5%)	13 (2.3%)
	Mild	1 (0.4%)	2 (0.7%)	3 (0.5%)
	Moderate	5 (1.8%)	2 (0.7%)	7 (1.3%)
	Severe	0	1 (0.4%)	1 (0.2%)
	Life Threatening	0	2 (0.7%)	2 (0.4%)
Abdominal Pain Upper	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Mild	0	1 (0.4%)	1 (0.2%)
Breast Cancer	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Moderate	0	1 (0.4%)	1 (0.2%)
Dysgeusia	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Mild	1 (0.4%)	0	1 (0.2%)
Gastrointestinal Disorder	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
General Physical Health Deterioration	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Hepatic Neoplasm Malignant	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Idiopathic Pulmonary Fibrosis	- Any Intensity -	0	3 (1.1%)	3 (0.5%)
	Moderate	0	1 (0.4%)	1 (0.2%)
	Life Threatening	0	2 (0.7%)	2 (0.4%)
Oedema Peripheral	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Mild	0	1 (0.4%)	1 (0.2%)
Pneumonia Respiratory Syncytial Viral	- Any Intensity -	0	1 (0.4%)	1 (0.2%)
	Severe	0	1 (0.4%)	1 (0.2%)
Psychotic Disorder	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)
Weight Decreased	- Any Intensity -	1 (0.4%)	0	1 (0.2%)
	Moderate	1 (0.4%)	0	1 (0.2%)

Reason for Discontinuation	Intensity	Pirfenidone 2403 mg/day (N=278)	Placebo (N=277)	Total (N=555)
Note: Patients with a Discontinuation do not Always have a Corresponding AE. [a] Patients discontinued study due to AE whereas this action was not recorded on the AE Case Report Form.				

Table 5: Adverse Event Leading to Discontinuation of Study (as recorded on the End of Study CRF) in CAPACITY 1

Reason for Discontinuation	Intensity	Pirfenidone 2403 mg/day (N=171)	Placebo (N=173)	Total (N=344)
Any Adverse Events	- Any Intensity -	5 (2.9%)	4 (2.3%)	9 (2.6%)
	Mild	1 (0.6%)	1 (0.6%)	2 (0.6%)
	Moderate	2 (1.2%)	0	2 (0.6%)
	Severe	2 (1.2%)	1 (0.6%)	3 (0.9%)
	Life Threatening	0	2 (1.2%)	2 (0.6%)
Flatulence	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Mild	1 (0.6%)	0	1 (0.3%)
Idiopathic Pulmonary Fibrosis	- Any Intensity -	2 (1.2%)	2 (1.2%)	4 (1.2%)
	Moderate	1 (0.6%)	0	1 (0.3%)
	Severe	1 (0.6%)	1 (0.6%)	2 (0.6%)
	Life Threatening	0	1 (0.6%)	1 (0.3%)
Rash	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Moderate	1 (0.6%)	0	1 (0.3%)
Rectal Haemorrhage	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Mild	0	1 (0.6%)	1 (0.3%)
Renal Failure Acute	- Any Intensity -	1 (0.6%)	0	1 (0.3%)
	Severe	1 (0.6%)	0	1 (0.3%)
Respiratory Arrest	- Any Intensity -	0	1 (0.6%)	1 (0.3%)
	Life Threatening	0	1 (0.6%)	1 (0.3%)
Note: Patients with a Discontinuation do not Always have a Corresponding AE. [a] Patients discontinued study due to AE whereas this action was not recorded on the AE Case Report Form.				

Table 6: Adverse Event Leading to Discontinuation of Study (as recorded on the End of Study CRF) in CAPACITY 2

Reason for Discontinuation	Intensity	Pirfenidone 1197 mg/day (N=87)	Pirfenidone 2403 mg/day (N=174)	Placebo (N=174)	Total (N=435)
Any Adverse Events	- Any Intensity -	3 (3.4%)	8 (4.6%)	3 (1.7%)	14 (3.2%)
	Moderate	1 (1.1%)	3 (1.7%)	2 (1.1%)	6 (1.4%)
	Severe	2 (2.3%)	1 (0.6%)	1 (0.6%)	4 (0.9%)
	Life Threatening	0	2 (1.1%)	0	2 (0.5%)
	Missing[a]	0	2 (1.1%)	0	2 (0.5%)
Acute Respiratory Failure	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Severe	1 (1.1%)	0	0	1 (0.2%)
Bladder Cancer	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Life Threatening	0	1 (0.6%)	0	1 (0.2%)
Dyspnoea	- Any Intensity -	0	0	2 (1.1%)	2 (0.5%)
	Moderate	0	0	2 (1.1%)	2 (0.5%)
Gastritis	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Moderate	0	1 (0.6%)	0	1 (0.2%)
Headache	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Moderate	0	1 (0.6%)	0	1 (0.2%)
Idiopathic Pulmonary Fibrosis	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Moderate	1 (1.1%)	0	0	1 (0.2%)
Lymph Node Cancer Metastatic	- Any Intensity -	1 (1.1%)	0	0	1 (0.2%)
	Severe	1 (1.1%)	0	0	1 (0.2%)
Muscle Rupture	- Any Intensity -	0	0	1 (0.6%)	1 (0.2%)
	Severe	0	0	1 (0.6%)	1 (0.2%)
Myocardial Infarction	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Life Threatening	0	1 (0.6%)	0	1 (0.2%)

Reason for Discontinuation	Intensity	Pirfenidone 1197 mg/day (N=87)	Pirfenidone 2403 mg/day (N=174)	Placebo (N=174)	Total (N=435)
Polymyalgia Rheumatica	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Severe	0	1 (0.6%)	0	1 (0.2%)
Rash	- Any Intensity -	0	1 (0.6%)	0	1 (0.2%)
	Moderate	0	1 (0.6%)	0	1 (0.2%)
Note: Patients with a Discontinuation do not Always have a Corresponding AE. [a] Patients discontinued study due to AE whereas this action was not recorded on the AE Case Report Form.					

A28. Section 4.12, p.171: Please provide data on serious adverse events for SP3.

REVISED QUESTION FROM NICE: *Please provide the numbers of serious adverse events, and p values to indicate whether there was any statistically significant difference between arms. Please see the data presented for CAPACITY 1 and 2 in the webappendix (pp.8-9) of the Noble 2011 publication and present the equivalent data for ASCEND, and SP2 and SP3.*

Summary tables for ASCEND, CAPACITY 1 and 2 are provided in our response to A25.

18 March update: A tabulated summary of all serious adverse events from the SP3 study is provided in Table 7.

Table 7: Serious adverse events reported in SP3

Reason for Discontinuation	Pirfenidone 1800 mg/day (N=109)	Placebo (N=107)	p-value (Fisher Exact Test)
Any Adverse Events	27 (24.8%)	24 (22.4%)	0.7496
Alanine aminotransferase increased	1 (0.9%)	0	1.000
Bronchitis	2 (1.8%)	2 (1.9%)	1.000
Bronchopulmonary aspergillosis	0	0	NA
Cataract	1 (0.9%)	0	1.000
Cerebral thrombosis	0	1 (0.9%)	0.4954
Cholangiolitis	1 (0.9%)	0	1.000
Cholecystitis	0	1 (0.9%)	0.4954
Colonic polyp	0	1 (0.9%)	0.4954
Completed suicide	0	1 (0.9%)	0.4954
Dyspnea	0	1 (0.9%)	0.4954
Foot fracture	0	1 (0.9%)	0.4954
Gastric cancer	1 (0.9%)	0	1.000
Gastrointestinal obstruction	1 (0.9%)	0	1.000
Idiopathic pulmonary fibrosis	6 (5.5%)	5 (4.7%)	1.000
Lung neoplasm malignant	1 (0.9%)	1 (0.9%)	1.000
Lung squamous cell carcinoma Stage 1	0	1 (0.9%)	0.4954
Myalgia	1 (0.9%)	0	1.000
Nasopharyngitis	0	0	NA
Osteonecrosis	1 (0.9%)	0	1.000
Pneumonia	6 (5.5%)	3 (2.8%)	0.4986
Pneumonia bacterial	0	0	NA
Pneumonia hemophilus	0	1 (0.9%)	0.4954
Pneumonia mycoplasmal	0	1 (0.9%)	0.4954
Pneumothorax	2 (1.8%)	3 (2.8%)	0.6817
Pyrexia	1 (0.9%)	0	1.000
Respiratory failure	1 (0.9%)	0	1.000
Respiratory tract infection	0	1 (0.9%)	0.4954
Rheumatoid arthritis	0	0	NA
Traumatic shock	1 (0.9%)	0	1.000

A36. ***Priority question**: Section 4.10, p.125: The individual trials report outcomes at different time points. These are synthesised under the assumption that the treatment effects are constant over time.*

a. Please provide evidence to justify this assumption, by considering the effect of including a covariate for trial duration through meta-regression.

In response to this query, we ran NMA models including a covariate for trial duration on three survival endpoints:

- all-cause mortality (ACM)
- IPF-related mortality (IPFRM)
- progression-free survival (PFS)

Background and Methods

As a reminder, Table 8 summarises the trials included in the base case NMAs for ACM, IPFRM and PFS, and takes note of the study duration for the relevant data inputs. Note that in our submission, the all-cause mortality NMA using data up to 52 weeks was chosen as the principal analysis. A sensitivity analysis is presented using data up to 72 weeks.

We ran the base case network (random and fixed effect models) from our original submission.

The network meta-regression model includes a single interaction term for all treatments as described in the NICE DSU TSD3 [Dias 2013]. The covariate, study duration in weeks, was centred in the analysis.

More complex meta-regression models cannot be fitted to this data set. The number of studies is extremely limited, and effect estimates at various trial durations are available only for the pirfenidone studies. In contrast, we have data at only one time point for each of the other treatments (nintedanib at 52 weeks, NAC at 60 weeks and triple therapy at 32 weeks), such that there are no contrasts on study duration. This means that the interaction term is determined by the CAPACITY, ASCEND, and SP3 trials (Table 8).

These meta-regression results should therefore be interpreted with caution as they rely on only the 4 pirfenidone studies. We note that a recommendation in the methodological literature suggests a minimum of 10 observations for each new variable included in a regression model, a recommendation which this analysis falls short on [Harrell 2015]).

Table 8: Summary of evidence (excerpt from Table 13, Response to ERG questions, March 3rd 2016)

	Study duration (weeks)	All-cause mortality	IPF-related mortality	Progression-Free Survival
CAPACITY 1 and 2	72	√	√	√
ASCEND	52	√	√	√
SP3	52	√[a]		√
INPULSIS1 and 2	52	√	√*	√*
TOMORROW	52	√[a]	√[b]	[f]
PANTHER (NAC)	60	√[b]	√[a]	√[a]
PANTHER (Triple)	32	√	√[a]	√

[a] HRs were unavailable: number of events and number of patients were used as an alternative (via the Wood model)
[b] HR was calculated from other available data using the methods of Parmar
[f] only reported the proportion of patients who progressed, rather than the proportion of patients who eithe progressed or died. Although the number of deaths was reported, it was unclear how many patients progresse before they died and therefore PFS cannot be calculated
* Our review – including the direction provided by NICE on 26 February 2016 over email – did not identify individua HRs for the INPULSIS trials for IPF-related mortality and PFS [1]. Based on this, we used the pooled HR.
[1] Post hoc analysis only supplied as part of the submission to NICE by Boehringer Ingelheim.

Results & Discussion

The slopes of the interaction terms are positive but all posterior densities span a wide range of values around 0, implying there is a significant degree of uncertainty in these models (Table 9). For a visual representation of posterior densities for these interaction terms, please refer to Figure 1, Figure 4, Figure 7, Figure 10, Figure 13 and Figure 16 of Appendix A.

Table 9: Summary of results: Posterior intervals for interaction term for study duration (change in logHR per week)

		Estimate (95% credible interval)
All-cause mortality	**RE**	0.013 (-0.037, 0.063)
	FE	0.013 (-0.030, 0.056)
IPF-related mortality	**RE**	0.019 (-0.082, 0.116)
	FE	0.020 (-0.046, 0.086)
Progression-free survival	**RE**	0.005 (-0.025, 0.035)
	FE	0.005 (-0.012, 0.022)

Adjusting for study duration does not meaningfully improve model fit compared to the original NMA models (Table 10), and the between-trial standard deviations (tau) are increased in all cases (vs. the original models) at 52 weeks and at 72 weeks (Table 11).

Table 10: Summary of results: Comparison of DICs

		Meta-regression models	Original models (without adjustment)*	
			52 wks	72 wks
All-Cause Mortality	**RE**	32.54	31.46	30.91
	FE	31.78	30.58	30.11
IPF-Related Mortality	**RE**	31.62	31.51	30.20
	FE	31.32	31.06	29.77
Progression-Free Survival	**RE**	14.86	13.34	12.89
	FE	14.14	12.40	11.98

*as presented in the Response to ERG questions, March 3rd 2016

Table 11: Summary of results: Comparison of posterior summaries of between-trial standard deviation: estimates and 95% credible intervals

		Meta-regression models	Original models (without adjustment)*	
			52 weeks	72 weeks
All-Cause Mortality	**RE**	0.173 (0.026, 0.647)	0.153 (0.025, 0.542)	0.152 (0.025, 0.532)
	FE	NA	NA	NA
IPF-Related Mortality	**RE**	0.421 (0.029, 2.008)	0.333 (0.028, 1.440)	0.306 (0.028, 1.293)
	FE	NA	NA	NA
Progression-Free Survival	**RE**	0.197 (0.025, 0.827)	0.126 (0.024, 0.449)	0.122 (0.024, 0.426)
	FE	NA	NA	NA

Results of the meta-regression models are in line with those form the original models, i.e. pirfenidone reduced all-cause mortality, IPF-related mortality and progression-free survival over 1 year compared with placebo (noting that meta-regression results are presented at the mean time while our original models were reported at 52 weeks (principal analysis) and at 72 weeks in a sensitivity analysis (Table 12)).

Table 12: Summary of results: Comparison of main NMA results
(Bold and italic: credible intervals do not cross 1)

		Meta-regression models			Original models (without adjustment)*					
		Pirf vs placebo	Nin vs placebo	Pirf vs Nin	Pirf vs placebo	Nin vs placebo	Pirf vs Nin	Pirf vs placebo	Nin vs placebo	Pirf vs Nin
		at 54.5 weeks			at 52 weeks			at 72 weeks		
ACM	RE	0.55 (0.28, 1.06)	0.73 (0.43, 1.24)	0.75 (0.30, 1.86)	***0.52 (0.30, 0.88)***	0.71 (0.43, 1.16)	0.73 (0.35, 1.50)	***0.62 (0.38, 0.99)***	0.71 (0.43, 1.16)	0.87 (0.44, 1.72)
	FE	***0.55 (0.31, 0.98)***	0.73 (0.47, 1.14)	0.75 (0.34, 1.64)	***0.52 (0.32, 0.84)***	0.71 (0.46, 1.09)	0.73 (0.38, 1.40)	***0.62 (0.40, 0.95)***	0.71 (0.46, 1.09)	0.87 (0.48, 1.60)
		at 53.8 weeks			at 52 weeks			at 72 weeks		
IPFRM	RE	0.37 (0.08, 1.69)	0.62 (0.18, 1.57)	0.61 (0.10, 4.92)	***0.36 (0.14, 0.90)***	0.60 (0.22, 1.32)	0.61 (0.18, 2.33)	0.48 (0.22, 1.01)	0.60 (0.23, 1.28)	0.80 (0.27, 2.65)
	FE	0.37 (0.13, 1.05)	0.65 (0.38, 1.13)	0.58 (0.17, 2.01)	***0.37 (0.18, 0.76)***	0.63 (0.37, 1.07)	0.59 (0.24, 1.45)	***0.49 (0.27, 0.87)***	0.63 (0.37, 1.07)	0.78 (0.35, 1.71)
		at 56.5 weeks			at 52 weeks			at 72 weeks		
PFS	RE	***0.62 (0.44, 0.87)***	0.76 (0.43, 1.34)	0.82 (0.41, 1.64)	***0.63 (0.50, 0.80)***	0.74 (0.50, 1.08)	0.85 (0.55, 1.34)	***0.63 (0.50, 0.78)***	0.74 (0.51, 1.07)	0.85 (0.55, 1.31)
	FE	***0.62 (0.52, 0.74)***	***0.76 (0.61, 0.94)***	0.82 (0.61, 1.10)	***0.63 (0.53, 0.74)***	***0.74 (0.61, 0.90)***	0.85 (0.66, 1.10)	***0.63 (0.54, 0.73)***	***0.74 (0.61, 0.90)***	0.85 (0.66, 1.09)

*as presented in the Response to ERG questions, March 3rd 2016

Full detailed results (including the data input table; total residual deviance, DIC, pD; details on posterior density for interaction term including a plot of hazard ratio (on log-scale) as a function of trial duration in weeks; cross-tabulation and forest plot of NMA results; and the posterior summaries of the between-trial standard deviation) are supplied in Appendix A for each of the three survival models.

Comparing the meta-regression approach requested in A36a with the approach used in our manufacturer submission, the meta-regression analysis provides a good understanding of the uncertainty inherent in comparing measurements from trials at different time-points, and provides an alternative to the original presented analysis, accounting for differences in time-points.

Use of this method did not, however, meaningfully improve model fit compared to our original approach. Coupled with the limitations in the number of observations available to power the meta-regression, we still consider the original analysis at 52 weeks to be the most useful analysis for evaluation of the relative benefits of pirfenidone, nintedanib and BSC,

For completeness, the impact of using results from the NMA based on the meta-regression models (from Table 12) on the cost-effectiveness analyses are presented in Table 13 to Table 15.

Table 13: Impact of NMA results from meta-regression models on cost-effectiveness – ITT (list prices)

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	5.61	3.93				
PFN	[redacted]	8.67	5.67	[redacted]	3.07	1.74	[redacted]

Table 14: Impact of NMA results from meta-regression models on cost-effectiveness – Mild (list prices)

TRT	Total			Incremental			ICER (£)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)
BSC	[redacted]	7.40	4.98				
PFN	[redacted]	11.26	6.99	[redacted]	3.86	2.01	[redacted]

Table 15: Impact of NMA results from meta-regression models on cost-effectiveness – Moderate (list prices)

TRT	Total			Incremental			ICER (£)	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (QALYs)	Incremental (QALYs)
BSC	[redacted]	5.00	3.56					
NTB	[redacted]	6.12	4.25	[redacted]	1.12	0.69	[redacted]	[redacted]
PFN	[redacted]	7.67	5.14	[redacted]	1.55	0.89	[redacted]	[redacted]

B13. ***Priority question****: In Table 92 of the company submission, the mean actual pirfenidone dose received in the ASCEND and CAPACITY trials is reported. The company states that pirfenidone needs to be titrated in the first 2 weeks. Please provide the mean (and standard error) for the actual dose received:*

a. estimated for the first 3 months of treatment only (to represent the first cycle of the model)

b. excluding the first 3 months of treatment (to represent subsequent cycles in the model).

Part a: The dose reported in Table 92 is a dose post-titration, ie was calculated upon exclusion of the first 14 days of treatment (titration period).

As requested, the dose estimated as an average of the first 3 months of treatment (including the titration period) is provided in Table 16.

Part b: The dose estimated in the period beyond the first 3 months of treatment is provided in Table 17.

Table 16: Average dose in the first three months of treatment (including the 14 days of titration period)

Trial	RDD (mg)	N	ADR [SE] (mg)	NPD
CAPACITY 2	2,403.00	174	1991.90 [30.79]	7.46
CAPACITY 1		171	2049.62 [25.23]	7.68
ASCEND		278	2025.35 [19.04]	7.59
Pooled		623	2022.70 [13.94]	7.58
Key: ADR, actual dose received; NPD, number of pills per day; RDD, recommended daily dose				

Table 17: Average dose beyond the first three months of treatment (excluding the first 91 days)

Trial	RDD (mg)	N	ADR [SE] (mg)	NPD
CAPACITY 2	2,403.00	165	2017.82 [41.19]	7.56
CAPACITY 1		167	2108.57 [35.51]	7.90
ASCEND		269	2101.31 [28.44]	7.87
Pooled		601	2080.41 [19.71]	7.79
Key: ADR, actual dose received; NPD, number of pills per day; RDD, recommended daily dose				

It is noted that the values presented in Table 16 and Table 17 are less than the weighted value presented in Table 92 of our initial submission (7.88 pills per day). The rationale for this is the estimate of 7.88 pills / day being based only on the post-titration period of the study (i.e. after day 14, if the SmPC recommended dosing schedule is followed). The average dose and pills per day will be lower during this period. As the ERG-requested analysis in B13a covers the entire first 3 month period, the average number of pills per day will be lower. Beyond this time frame (as assessed in B13b), dose adjustments contribute to an average number of pills which is lower than that presented in Table 92 of the initial submission. This is outlined in Table 18, where is can be seen that the average dose over the entire treatment period (including the titration period is 7.71 pills per day.

Table 18: Comparison of alternative assessments of average dose – pooled data

Average dose assessment	RDD (mg)	N	ADR [SE] (mg)	NPD
Average dose post titration (Table 92 of manufacturer submission)	2,403.00	619	2104.63 [17.18]	7.88
Average dose over total treatment period	2,403.00	623	2059.44 [16.96]	7.71
Average dose in the first 3 months of treatment (question B13a)	2,403.00	623	2022.70 [13.94]	7.58
Average dose after the first 3 months of treatment (question B13b)	2,403.00	601	2080.41 [19.71]	7.79
Key: ADR, actual dose received; NPD, number of pills per day; RDD, recommended daily dose				

The impact of accounting for the average dose in the first 3 months and post-month 3 period on results of the cost-effectiveness model are presented in Table 19. Based in the lower number of pills per day vs. that assessed in the initial submission, the ICER is each patient population is reduced when these data are incorporated into the model.

Table 19: Impact of stratification of average dose on cost-effectiveness results (list-prices

Population	ICER vs. BSC	
	Per 3 March revisions	With revision to average doses
ITT	[redacted]	[redacted]
Mild	[redacted]	[redacted]
Moderate	[redacted]	[redacted]

References

Dias 2013. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence Synthesis for Decision Making 3: Heterogeneity—Subgroups, MetaRegression, Bias, and Bias-Adjustment, Medical Decision making, p. 618-640, Jul 2013, DOI: 10.1177/0272989X13485157

Harrell FE. Regression Modeling Strategies, Springer Series in Statistics (Cham: Springer International Publishing, 2015), http://link.springer.com/10.1007/978-3-319-19425-7.

Nathan 2016 (Academic in Confidence). Nathan SD, Albera C, Bradford WZ. Effect of Pirfenidone on Mortality: Pooled and Meta-Analyses of Three Phase 3 Clinical Trials in Idiopathic Pulmonary Fibrosis. Manuscript in development for submission to The Lancet Respiratory Medicine [AiC].

NICE 2016. National Institute for Health and Care Excellence. Nintedanib for treating idiopathic pulmonary fibrosis [TA379]

Noble 2011. Noble PW, Albera C, Bradford WZ et al. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. Lancet 2011; 377: 1760–69

Appendix A: Response to A36a – Meta-Regression for Trial Duration

1. All-cause mortality

New input: the data set contains the trial duration in weeks as covariate.

Table 20: The (full) analysis data set used in the JAGS fits:

Study	Treatment	Comparator	HR	logHR	SE	N	n	weeks	t	b
PIPF-004 CAPACITY 2	Pirfenidone 2403mg/day (*)	Placebo	0.508910	-0.6754841	0.4378553	NA	NA	72	2	1
PIPF-006 CAPACITY 1	Pirfenidone 2403mg/day (*)	Placebo	0.866370	-0.1434432	0.3789628	NA	NA	72	2	1
ASCEND	Pirfenidone 2403mg/day (*)	Placebo	0.550000	-0.5978370	0.3793018	NA	NA	52	2	1
INPULSIS 1	Nintedanib 300mg/day	Placebo	0.630000	-0.4620355	0.3942242	NA	NA	52	3	1
INPULSIS 2	Nintedanib 300mg/day	Placebo	0.740000	-0.3011051	0.3103049	NA	NA	52	3	1
PANTHER (Triple therapy)	Triple therapy	Placebo	9.260000	2.2257040	1.0604775	NA	NA	32	5	1
PANTHER (NAC)	NAC	Placebo	1.995622	0.6909556	0.6666667	NA	NA	60	4	1
SP3	Pirfenidone 2403mg/day (*)	Placebo	NA	NA	NA	110	3	52	2	1
SP3	Placebo	Placebo	NA	NA	NA	109	6	52	1	1
TOMORROW	Nintedanib 300mg/day	Placebo	NA	NA	NA	85	7	52	3	1
TOMORROW	Placebo	Placebo	NA	NA	NA	85	9	52	1	1

Meta-Regression for Trial Duration, Random effects model

Total residual deviance (posterior mean): 8.37

DIC: 32.54

pD: 7.51

Table 21: Posterior density for interaction term: trial duration in weeks (centred): All-cause mortality (base case network with adjustment, RE model)

mean	50%	2.5%	97.5%
0.013	0.013	-0.037	0.063

Figure 1: Plot of posterior density for interaction term: trial duration in weeks: All-cause mortality (base case network with adjustment, RE model)

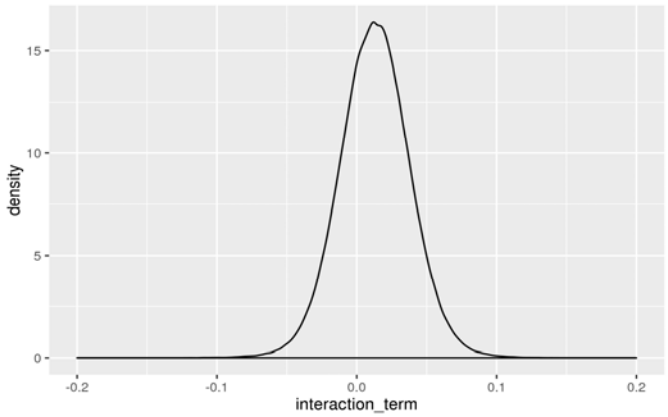


Figure 2: Plot of hazard ratio (on log-scale) as a function of trial duration in weeks: All-cause mortality (base case network with adjustment, RE model)

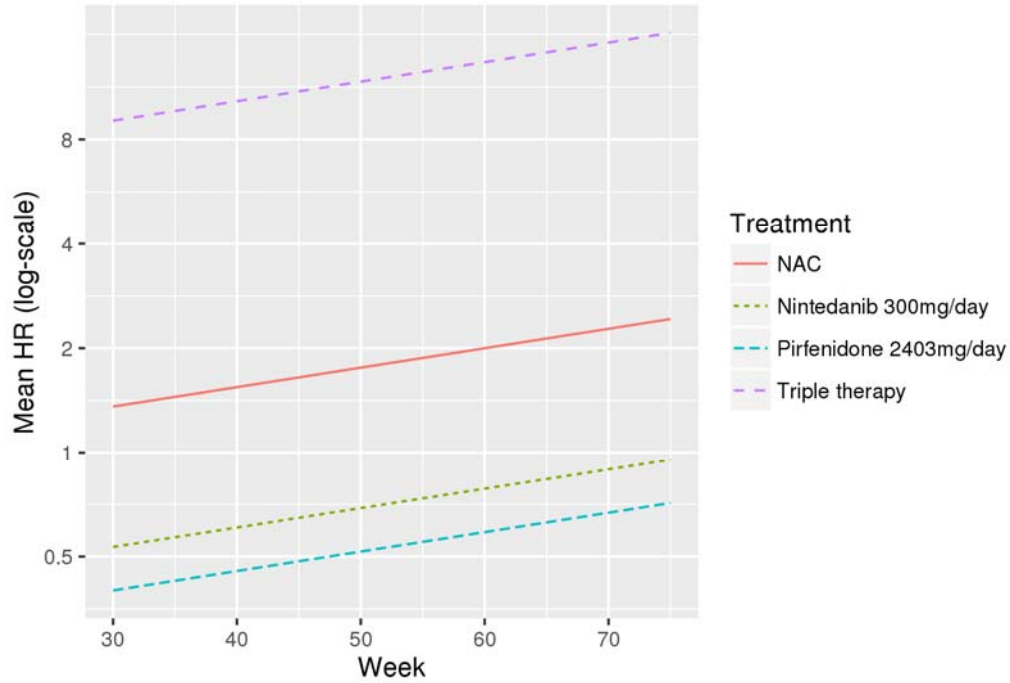


Table 22: Hazard ratios and 95% credible intervals (CrI) at the mean covariate value (54.5 weeks): All-cause mortality (base case network with adjustment, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.55 (0.28, 1.06)	0.73 (0.43, 1.24)	1.87 (0.45, 7.67)	12.42 (1.12, 138.26)
Pirfenidone 2403mg/day	1.82 (0.95, 3.55)		1.33 (0.54, 3.32)	3.41 (0.77, 15.01)	22.61 (1.55, 334.78)
Nintedanib 300mg/day	1.37 (0.81, 2.32)	0.75 (0.3, 1.86)		2.55 (0.55, 11.83)	16.96 (1.52, 188.92)
NAC	0.54 (0.13, 2.21)	0.29 (0.07, 1.3)	0.39 (0.08, 1.82)		6.67 (0.36, 121.84)
Triple therapy	0.08 (0.01, 0.9)	0.04 (0, 0.65)	0.06 (0.01, 0.66)	0.15 (0.01, 2.74)	

Figure 3: Forest plot of hazard ratios and 95% credible intervals (CrI) at the mean covariate value (54.5 weeks): All-cause mortality (base case network with adjustment, RE model)

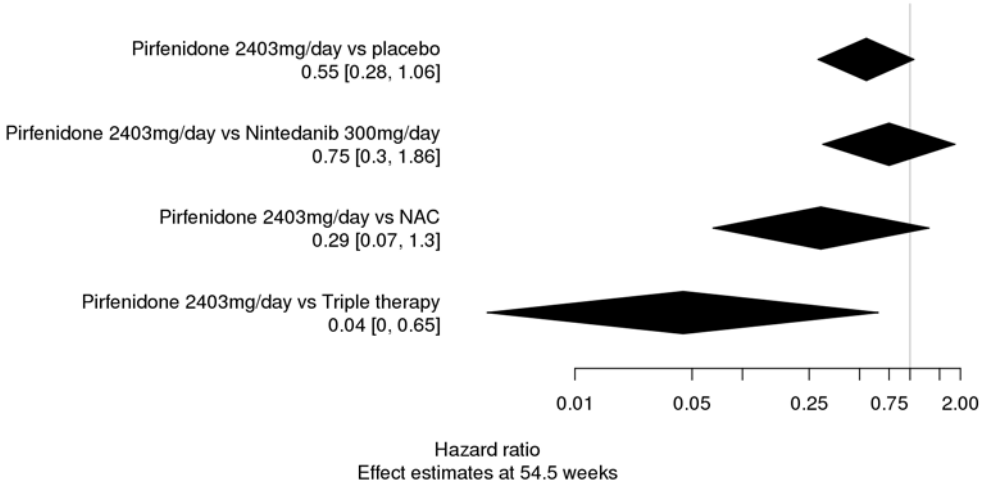


Table 23: Posterior summaries of between-trial standard deviation: All-cause mortality (base case network with adjustment, RE model)

mean	50%	2.5%	97.5%
0.173	0.116	0.026	0.647

Meta-Regression for Trial Duration, Fixed effect model

Total residual deviance (posterior mean): 8.03

DIC: 31.78

pD: 7.09

Table 24: Posterior density for interaction term: trial duration in weeks (centered): All-cause mortality (base case network with adjustment, FE model)

mean	50%	2.5%	97.5%
0.013	0.013	-0.03	0.056

Figure 4: Plot of posterior density for interaction term: trial duration in weeks: All-cause mortality (base case network with adjustment, FE model)

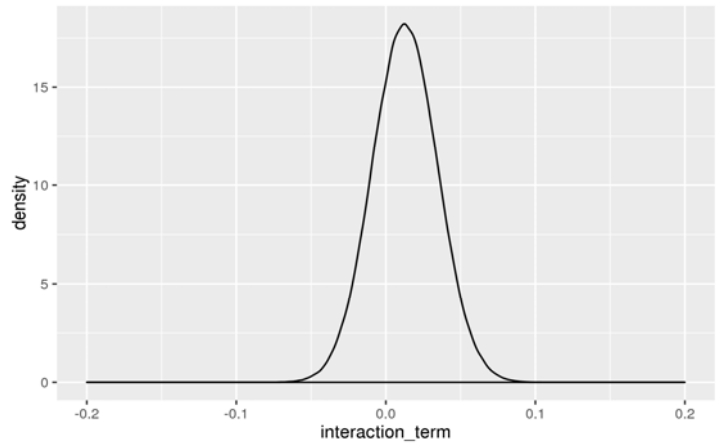


Figure 5: Plot of hazard ratio (on log-scale) as a function of trial duration in weeks: All-cause mortality (base case network with adjustment, FE model)

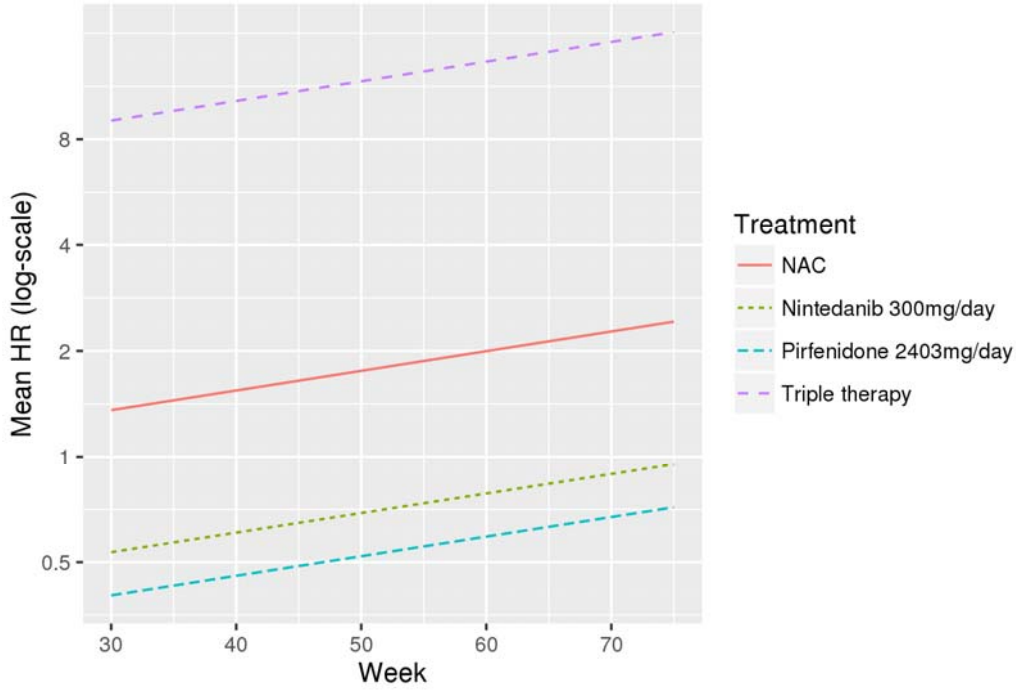
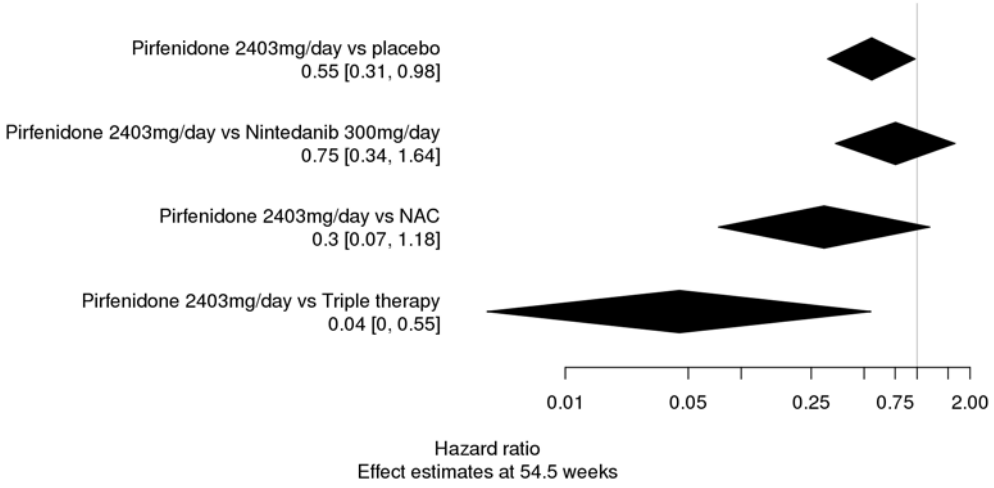


Table 25: Hazard ratios and 95% credible intervals (CrI) at the mean covariate value (54.5 weeks): All-cause mortality (base case network with adjustment, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.55 (0.31, 0.98)	0.73 (0.47, 1.14)	1.86 (0.49, 7.04)	12.39 (1.26, 123)
Pirfenidone 2403mg/day	1.81 (1.03, 3.23)		1.33 (0.61, 2.91)	3.38 (0.85, 13.46)	22.44 (1.83, 278.49)
Nintedanib 300mg/day	1.37 (0.88, 2.14)	0.75 (0.34, 1.64)		2.54 (0.61, 10.51)	16.91 (1.72, 166.6)
NAC	0.54 (0.14, 2.03)	0.3 (0.07, 1.18)	0.39 (0.1, 1.63)		6.64 (0.43, 101.89)
Triple therapy	0.08 (0.01, 0.79)	0.04 (0, 0.55)	0.06 (0.01, 0.58)	0.15 (0.01, 2.31)	

Figure 6: Forest plot of hazard ratios and 95% credible intervals (CrI) at the mean covariate value (54.5 weeks): All-cause mortality (base case network with adjustment, FE model)



2. IPF-related mortality

New input: the data set contains the trial duration in weeks as covariate.

Table 26: The (full) analysis data set used in the JAGS fits:

Study	Treatment	Comparator	HR	logHR	SE	N	n	weeks	t	b
PIPF-004 CAPACITY 2	Pirfenidone 2403mg/day (*)	Placebo	0.3500000	-1.0498221	0.5850752	NA	NA	72	2	1
PIPF-006 CAPACITY 1	Pirfenidone 2403mg/day (*)	Placebo	0.6900000	-0.3710637	0.4302117	NA	NA	72	2	1
ASCEND	Pirfenidone 2403mg/day (*)	Placebo	0.3600000	-1.0216512	0.5743197	NA	NA	52	2	1
Pooled INPULSIS	Nintedanib 300mg/day	Placebo	0.7400000	-0.3011051	0.3021147	NA	NA	52	3	1
TOMORROW	Nintedanib 300mg/day	Placebo	0.3043678	-1.1895183	0.6324555	NA	NA	52	3	1
PANTHER (Triple therapy)	Triple therapy	Placebo	NA	NA	NA	77	7	32	5	1
PANTHER (Triple therapy)	Placebo	Placebo	NA	NA	NA	78	1	32	1	1
PANTHER (NAC)	NAC	Placebo	NA	NA	NA	133	5	60	4	1
PANTHER (NAC)	Placebo	Placebo	NA	NA	NA	131	3	60	1	1

Meta-Regression for Trial Duration, Random effects model

Total residual deviance (posterior mean): 9.31

DIC: 31.62

pD: 8.19

Table 27: Posterior density for interaction term: trial duration in weeks (centered): IPF-related mortality (base case network with adjustment, RE model)

mean	50%	2.5%	97.5%
0.019	0.019	-0.082	0.116

Figure 7: Plot of posterior density for interaction term: trial duration in weeks: IPF-related mortality (base case network with adjustment, RE model)

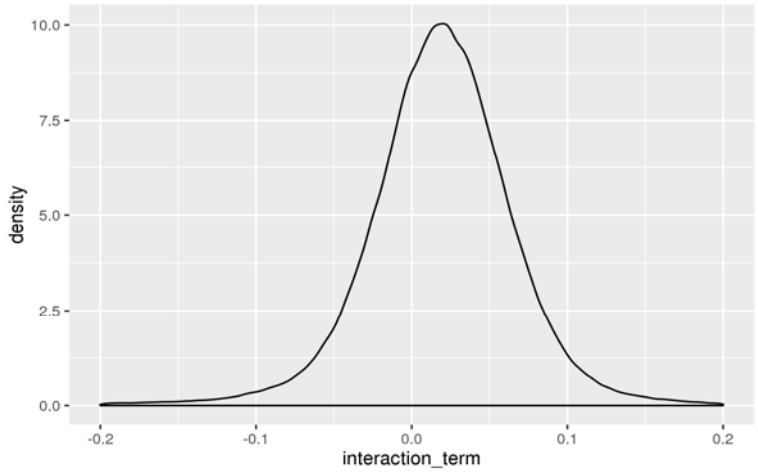


Figure 8: Plot of hazard ratio (on log-scale) as a function of trial duration in weeks: IPF-related mortality (base case network with adjustment, RE model)

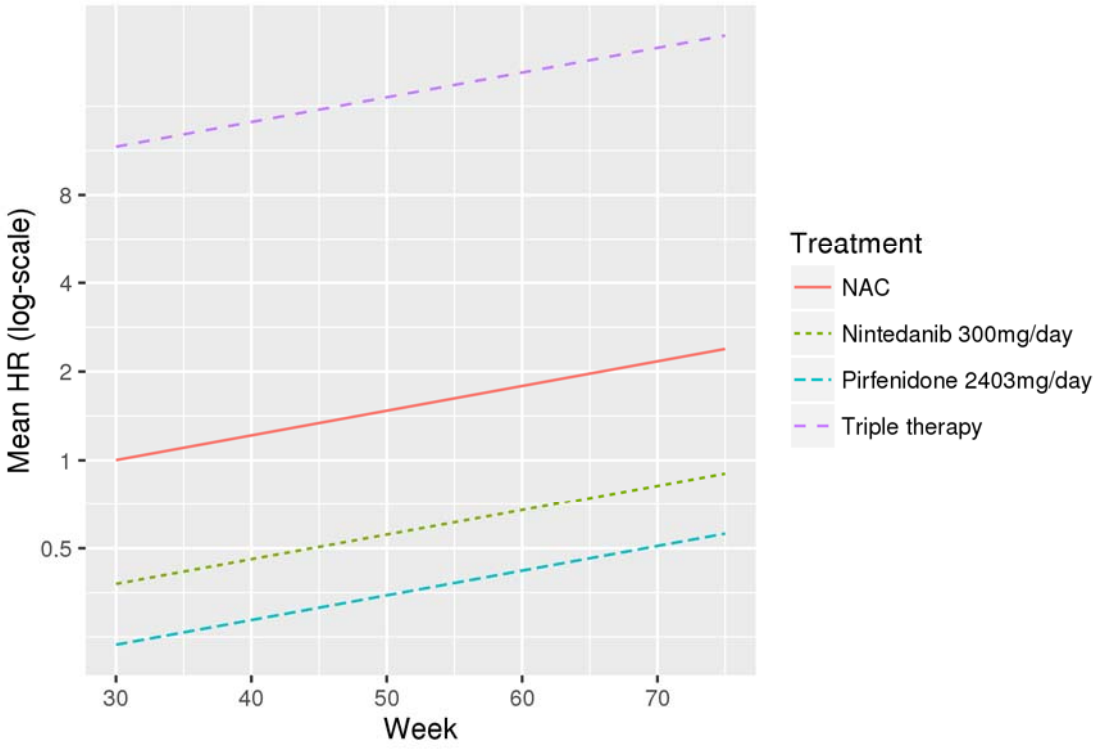


Table 28: Hazard ratios and 95% credible intervals (CrI) at the mean covariate value (53.8 weeks): IPF-related mortality (base case network with adjustment, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.37 (0.08, 1.69)	0.62 (0.18, 1.57)	1.53 (0.21, 13.86)	16.25 (0.67, 1069.07)
Pirfenidone 2403mg/day	2.68 (0.59, 12.04)		1.64 (0.2, 10.35)	4.14 (0.46, 43.75)	45.15 (0.66, 6152.76)
Nintedanib 300mg/day	1.62 (0.64, 5.62)	0.61 (0.1, 4.92)		2.53 (0.27, 33.2)	26.8 (1.14, 1907.38)
NAC	0.65 (0.07, 4.82)	0.24 (0.02, 2.17)	0.4 (0.03, 3.65)		10.81 (0.16, 1354.1)
Triple therapy	0.06 (0, 1.49)	0.02 (0, 1.52)	0.04 (0, 0.88)	0.09 (0, 6.42)	

Figure 9: Forest plot of hazard ratios and 95% credible intervals (CrI) at the mean covariate value (53.8 weeks): IPF-related mortality (base case network with adjustment, RE model)

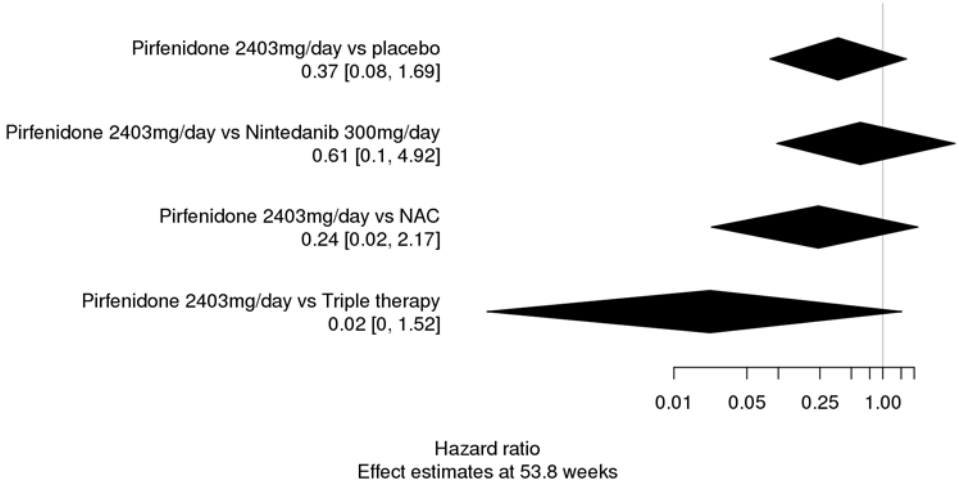


Table 29: Posterior summaries of between-trial standard deviation: IPF-related mortality (base case network with adjustment, RE model)

mean	50%	2.5%	97.5%
0.421	0.215	0.029	2.008

Meta-Regression for Trial Duration, Fixed effect model

Total residual deviance (posterior mean): 9.73

DIC: 31.32

pD: 7.48

Table 30: Posterior density for interaction term: trial duration in weeks (centered): IPF-related mortality (base case network with adjustment, FE model)

mean	50%	2.5%	97.5%
0.02	0.02	-0.046	0.086

Figure 10: Plot of posterior density for interaction term: trial duration in weeks: IPF-related mortality (base case network with adjustment, FE model)

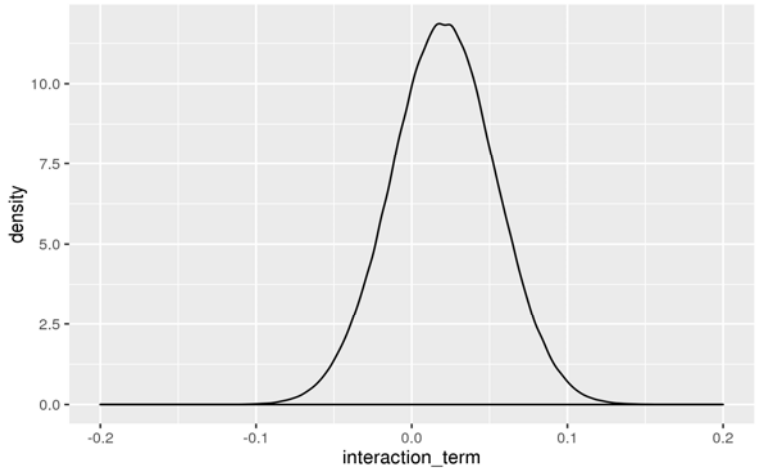


Figure 11: Plot of hazard ratio (on log-scale) as a function of trial duration in weeks: IPF-related mortality (base case network with adjustment, FE model)

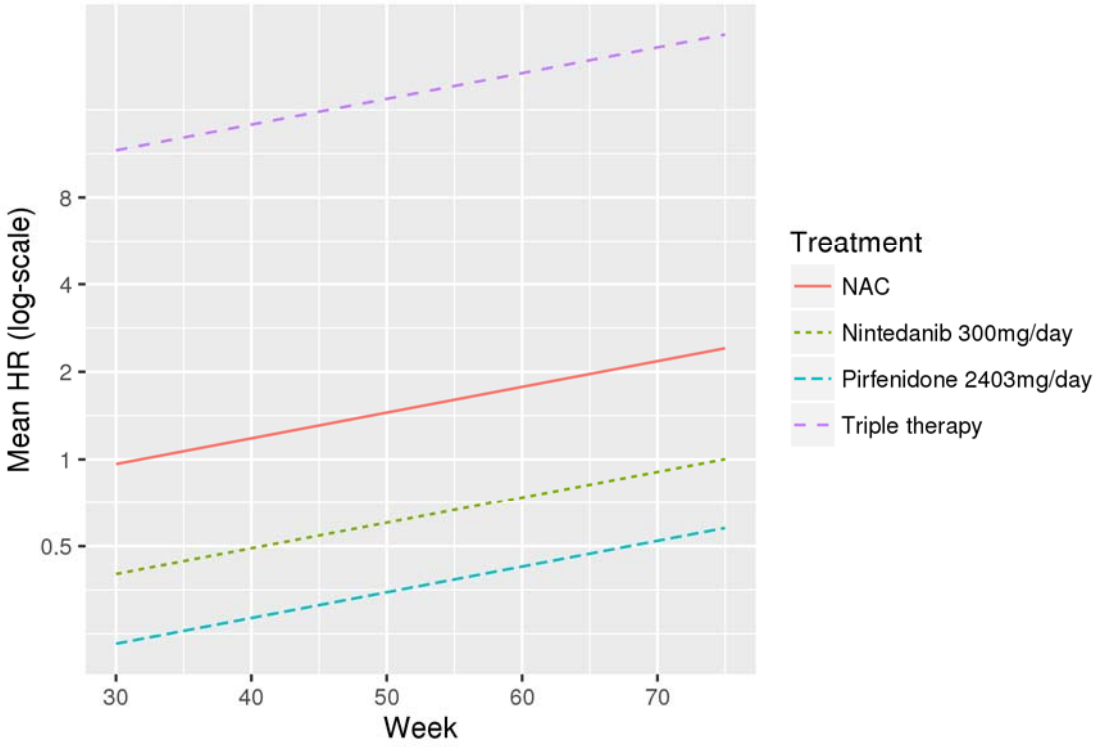
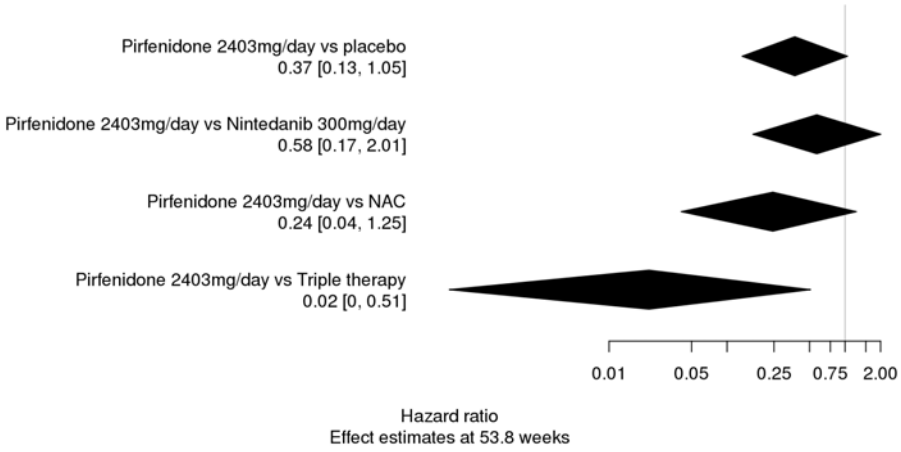


Table 31: Hazard ratios and 95% credible intervals (CrI) at the mean covariate value (53.8 weeks): IPF-related mortality (base case network with adjustment, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.37 (0.13, 1.05)	0.65 (0.38, 1.13)	1.53 (0.34, 8.29)	16.53 (1.4, 548.47)
Pirfenidone 2403mg/day	2.67 (0.95, 7.45)		1.74 (0.5, 6.04)	4.09 (0.8, 24.5)	45.94 (1.96, 2261.26)
Nintedanib 300mg/day	1.54 (0.89, 2.66)	0.58 (0.17, 2.01)		2.36 (0.46, 14.21)	25.42 (2.18, 842.76)
NAC	0.65 (0.12, 2.97)	0.24 (0.04, 1.25)	0.42 (0.07, 2.19)		11.01 (0.45, 539.16)
Triple therapy	0.06 (0, 0.71)	0.02 (0, 0.51)	0.04 (0, 0.46)	0.09 (0, 2.23)	

Figure 12: Forest plot of hazard ratios and 95% credible intervals (CrI) at the mean covariate value (53.8 weeks): IPF-related mortality (base case network with adjustment, FE model)



3. Progression free survival

New input: the data set contains the trial duration in weeks as covariate.

Table 32: The (full) analysis data set used in the JAGS fits:

Study	Treatment	Comparator	HR	logHR	SE	N	n	weeks	t	b
PIPF-004 CAPACITY 2	Pirfenidone 2403mg/day (*)	Placebo	0.58	-0.5447272	0.1862180	NA	NA	72	2	1
PIPF-006 CAPACITY 1	Pirfenidone 2403mg/day (*)	Placebo	0.78	-0.2484614	0.2024753	NA	NA	72	2	1
ASCEND	Pirfenidone 2403mg/day (*)	Placebo	0.57	-0.5621189	0.1486265	NA	NA	52	2	1
SP3	Pirfenidone 2403mg/day (*)	Placebo	0.65	-0.4307829	0.1546293	NA	NA	52	2	1
Pooled INPULSIS	Nintedanib 300mg/day	Placebo	0.74	-0.3011051	0.1020372	NA	NA	52	3	1
PANTHER (Triple therapy)	Triple therapy	Placebo	1.46	0.3784364	0.3754703	NA	NA	32	5	1
PANTHER (NAC)	NAC	Placebo	NA	NA	NA	133	36	60	4	1
PANTHER (NAC)	Placebo	Placebo	NA	NA	NA	131	35	60	1	1

Meta-Regression for Trial Duration, Random effects model

Total residual deviance (posterior mean): 7.52

DIC: 14.86

pD: 6.74

Table 33: Posterior density for interaction term: trial duration in weeks (centered): Progression-free survival (base case network with adjustment, RE model)

mean	50%	2.5%	97.5%
0.005	0.005	-0.025	0.035

Figure 13: Plot of posterior density for interaction term: trial duration in weeks: Progression-free survival (base case network with adjustment, RE model)

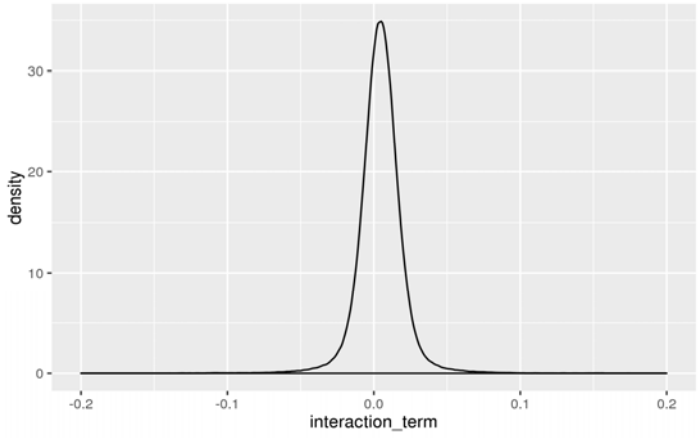


Figure 14: Plot of hazard ratio (on log-scale) as a function of trial duration in weeks: Progression-free survival (base case network with adjustment, RE model)

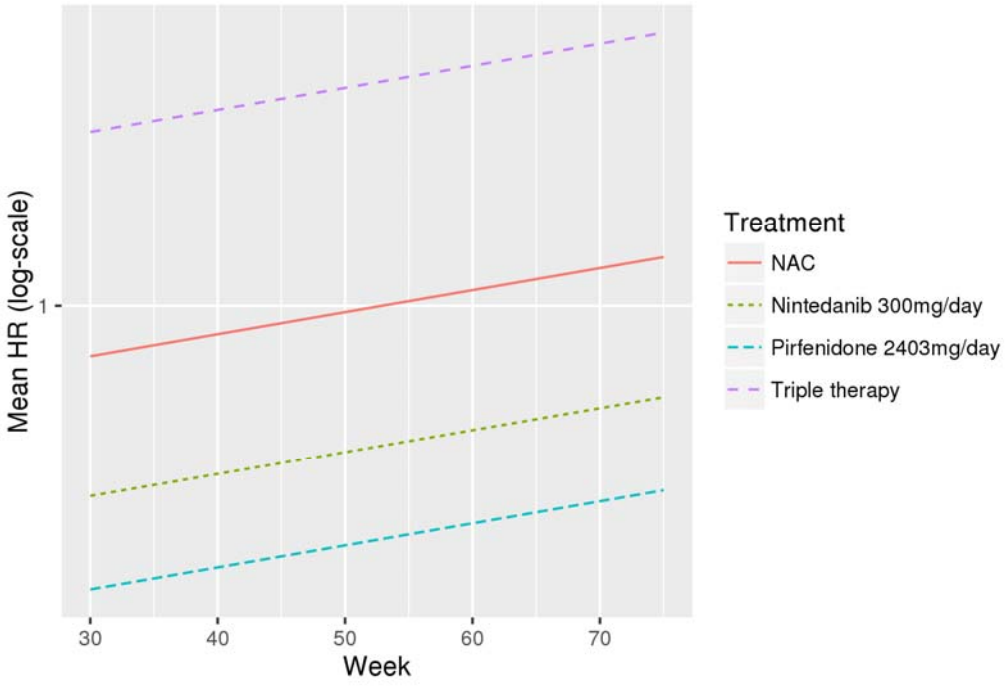


Table 34: Hazard ratios and 95% credible intervals (CrI) at the mean covariate value (56.5 weeks): Progression-free survival (base case network with adjustment, RE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.62 (0.44, 0.87)	0.76 (0.43, 1.34)	1 (0.51, 2.03)	1.63 (0.52, 5.16)
Pirfenidone 2403mg/day	1.61 (1.15, 2.25)		1.22 (0.61, 2.43)	1.61 (0.77, 3.47)	2.63 (0.73, 9.44)
Nintedanib 300mg/day	1.32 (0.75, 2.35)	0.82 (0.41, 1.64)		1.33 (0.54, 3.32)	2.16 (0.65, 7.13)
NAC	1 (0.49, 1.97)	0.62 (0.29, 1.3)	0.75 (0.3, 1.84)		1.63 (0.4, 6.47)
Triple therapy	0.61 (0.19, 1.93)	0.38 (0.11, 1.36)	0.46 (0.14, 1.53)	0.61 (0.15, 2.52)	

Figure 15: Forest plot of hazard ratios and 95% credible intervals (CrI) at the mean covariate value (56.5 weeks): Progression-free survival (base case network with adjustment, RE model)

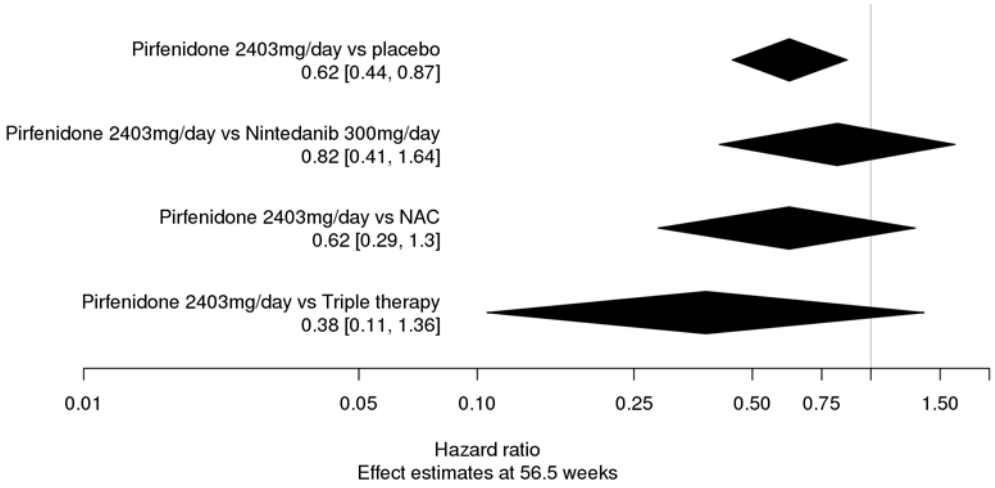


Table 35: Posterior summaries of between-trial standard deviation: Progression-free survival (base case network with adjustment, RE model)

mean	50%	2.5%	97.5%
0.197	0.109	0.025	0.827

Meta-Regression for Trial Duration, Fixed effect model

Total residual deviance (posterior mean): 7.54

DIC: 14.14

pD: 6

Table 36: Posterior density for interaction term: trial duration in weeks (centered): Progression-free survival (base case network with adjustment, FE model)

mean	50%	2.5%	97.5%
0.005	0.005	-0.012	0.022

Figure 16: Plot of posterior density for interaction term: trial duration in weeks: Progression-free survival (base case network with adjustment, FE model)

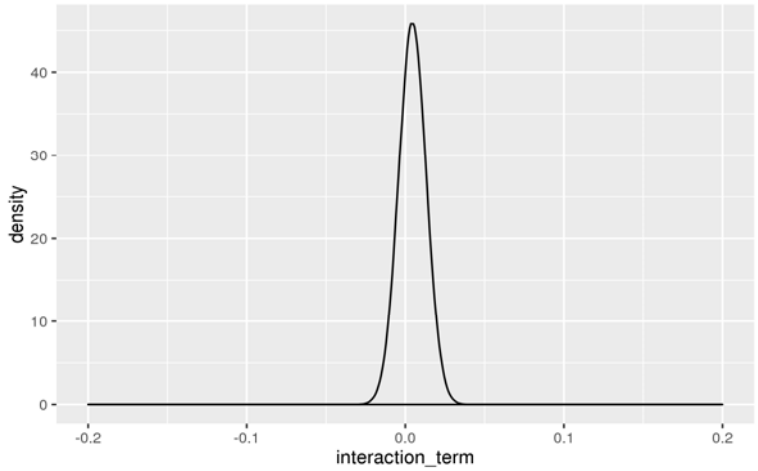


Figure 17: Plot of hazard ratio (on log-scale) as a function of trial duration in weeks: Progression-free survival (base case network with adjustment, FE model)

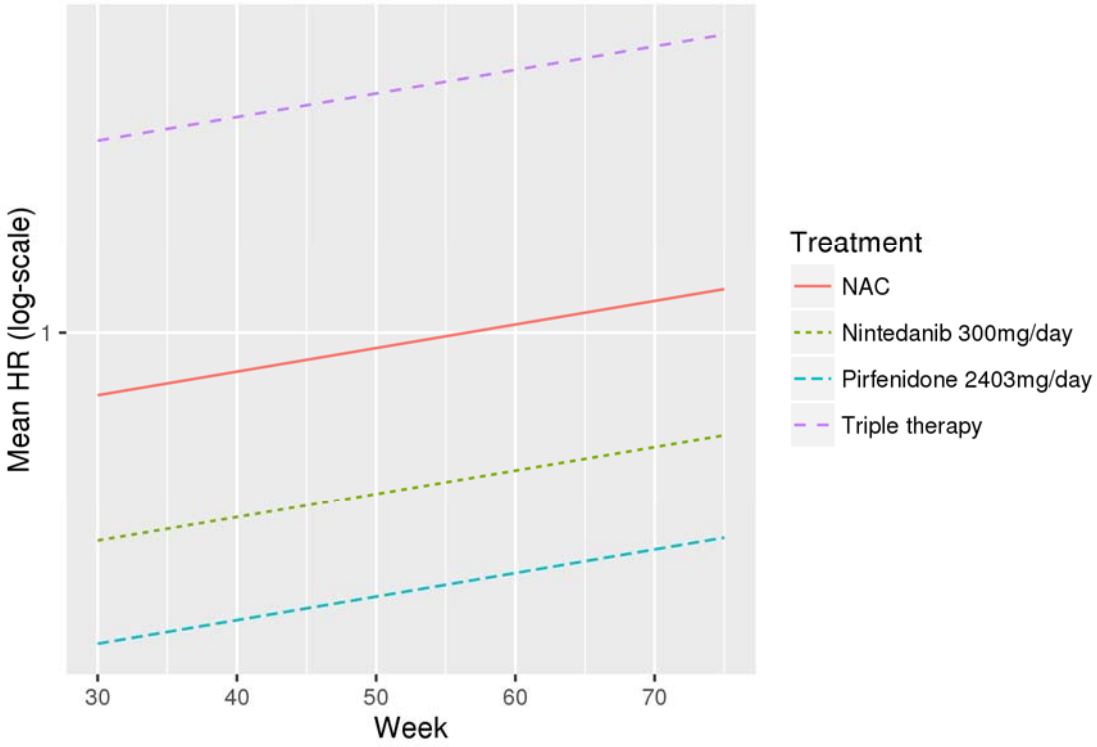
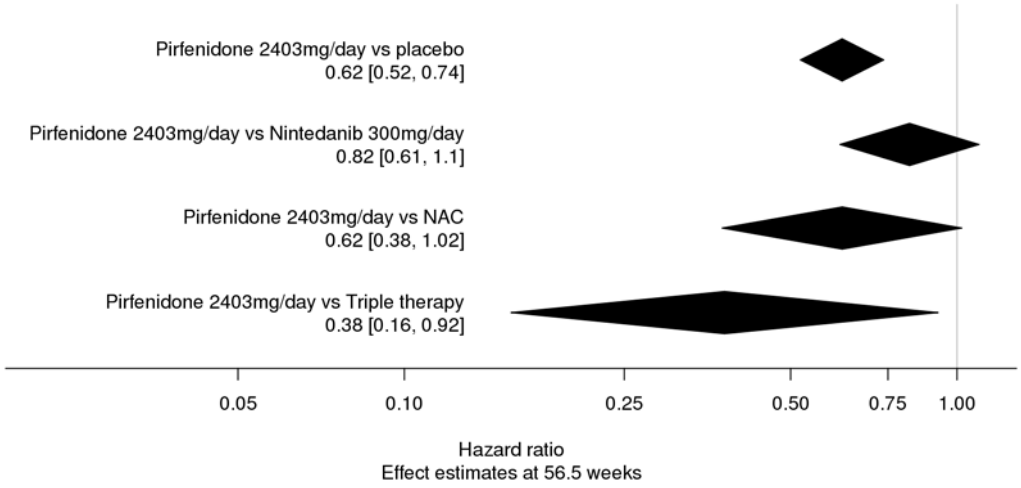


Table 37: Hazard ratios and 95% credible intervals (CrI) at the mean covariate value (56.5 weeks): Progression-free survival (base case network with adjustment, FE model)

Comparator	Placebo	Pirfenidone 2403mg/day	Nintedanib 300mg/day	NAC	Triple therapy
Placebo		0.62 (0.52, 0.74)	0.76 (0.61, 0.94)	1 (0.62, 1.61)	1.63 (0.7, 3.81)
Pirfenidone 2403mg/day	1.61 (1.36, 1.92)		1.22 (0.91, 1.63)	1.61 (0.98, 2.66)	2.63 (1.08, 6.41)
Nintedanib 300mg/day	1.32 (1.07, 1.64)	0.82 (0.61, 1.1)		1.32 (0.78, 2.25)	2.16 (0.94, 4.98)
NAC	1 (0.62, 1.61)	0.62 (0.38, 1.02)	0.76 (0.44, 1.28)		1.63 (0.6, 4.41)
Triple therapy	0.61 (0.26, 1.43)	0.38 (0.16, 0.92)	0.46 (0.2, 1.07)	0.61 (0.23, 1.66)	

Figure 18: Forest plot of hazard ratios and 95% credible intervals (CrI) at the mean covariate value (56.5 weeks): Progression-free survival (base case network with adjustment, FE model)



Appendix G - professional organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Thank you for agreeing to make a submission on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your submission, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: Dr Michael Gibbons

Name of your organisation: British Thoracic Society
Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? Yes
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)? Yes
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? Yes
- If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc)?
- other? (please specify)

Appendix G - professional organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS?
Patients are treated with pirfenidone when their FVC is between 50-80%.

Is there significant geographical variation in current practice? **Possibly.**

Are there differences of opinion between professionals as to what current practice should be? **No**

What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

For patients with an FVC > 80% there are no treatment options, despite good clinical evidence of efficacy.

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? **Clearly as the condition progresses, life expectancy decreases. The logical follow on from this is that preventing disease progression prolongs life, and starting therapy earlier, further prolongs life**.

Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology? **Data from the ASCEND trial demonstrates that pirfenidone is beneficial for patients with FVC up to 90%.**

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? **Specialist clinics.**

Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)? **Specialist nursing to aid side effect management.**

If the technology is already available, is there variation in how it is being used in the NHS? **Unknown.**

Is it always used within its licensed indications? **Yes**. If not, under what circumstances does this occur?

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

NICE Guideline: Idiopathic pulmonary fibrosis in adults: diagnosis and management (CG163)
2015: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline

Appendix G - professional organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use? **There are currently no licensed alternatives. Patients require regular blood monitoring**

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

Currently, we are able to prescribe pirfenidone for patients with an FVC <80%. It should be noted that for patients with an FVC >80%, regular lung function tests are required. These have to be performed frequently in order to identify at which time-point to intervene. This requires significant extra support from pulmonary physiologists to undertake the tests, so as not to miss a significant loss of lung function during the period prior to the patient being seen by the medical team. Pirfenidone is meant to be stopped if patients show a drop in FVC >10% despite treatment but there is no mechanism for untreated patients who are deteriorating being started on treatment, unless the FVC falls. It would be reasonable to allow for use of the drug in patients in whom DLco or six minute walk distance is falling and/or there is progression in fibrosis on HRCT despite a well-preserved FVC.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. **Registry studies have demonstrated efficacy.**

Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? **Overall, the pirfenidone trials reflect current practice, for patients that are truly eligible for commencement of the drugs. This requires an (ILD) MDT approach.**

What, in your view, are the most important outcomes, and were they measured in the trials? **Appropriate outcomes were used .**

If surrogate measures of outcome were used, do they adequately predict long-term outcomes? **We would like to state the following: mortality, progression free survival and exacerbation rate are likely to be very low in this patient cohort. For instance the risk of acute exacerbation is increased in those with FVC<70%. These outcome parameters are likely to be unhelpful in patients with FVC>80%.**

What is the relative significance of any side effects or adverse reactions? **About 1 in 5 patients are intolerant.**

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

In what ways do these affect the management of the condition and the patient's quality of life? If they are intolerant, there are no licensed alternative disease-modifying drugs. Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice? **Patients do have side effects as a result of pirfenidone. It is important both to try and avoid these, and to intervene early when they appear. This requires a strong clinical team with good systems and processes in place.**

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? **Data from (our) BTS registry demonstrate that >40% of IPF patients have FVC > 80% at presentation and have clinically significant disease. The ASCEND study shows that Pirfenidone treatment benefits patients with FVC>80%. CAPACITY studies also show treatment benefit in FVC>80% cohort**

This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

Implementation issues

The NHS is required by the Department of Health to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone. How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? **It would allow access to the drug for patients with an FVC 80-90%. However, there are limitations of lung function measurements. The starting FVC is unknown in any one patient. If a patient has an FVC 85%, this is outside of the current treatment window, but the starting FVC may have been >100% which means the patient has lost >15% in FVC. This is clinically significant but on current criteria, patients with IPF would be excluded from treatment with Pirfenidone (or Nintedanib).**

Would NHS staff need extra education and training? **No.**
Would any additional resources be required (for example, facilities or equipment)? **No**

Equality

Appendix G - professional organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that this appraisal:

- could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed;
- could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology;
- could lead to recommendations that have any adverse impact on people with a particular disability or disabilities.

Please tell us what evidence should be obtained to enable the Committee to identify and consider such impacts.

Comparator is Nintedanib (FVC 50-80%); however the INPULSIS studies also show treatment benefit of Nintedanib in patients with FVC>80%. Furthermore, the effect of treatment is the same independent of the starting FVC (provided FVC>50% as per the study). Therefore this comparator is of limited use.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Thank you for agreeing to give us a statement on your view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name:
Dr Toby Maher

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? x
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)? X
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)?
- other? (please specify)

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry:

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

What is the expected place of the technology in current practice?

Patients with idiopathic pulmonary fibrosis in England are generally diagnosed and their initial treatment managed in Specialist Interstitial Lung Disease Centres.

Diagnosis and management is usually in accordance with the NICE guidelines on IPF.

The only pharmacotherapies currently available are pirfenidone and nintedanib. These are used for patients as per the relevant NICE TAs. Because of the lung function criteria stipulated in these TAs less than half of patients have access to treatment. The remainder have, in some cases, the option of clinical trials, otherwise the only available treatments are symptom based and palliative.

The advantages and disadvantages of the technology

IPF is a progressive disease with a median survival of 3 years. The majority of patients die from respiratory failure and their last year of life is typically limited by severe dyspnoea. Only two treatments have been proven to alter the natural history of IPF; pirfenidone and nintedanib. Both drugs slow the rate of disease decline by approximately 50% and both can therefore be expected to extend both total life expectancy and also the period of life free of respiratory failure. The greatest benefit

in terms of survival gain is likeliest to be seen in those diagnosed at an early stage in the disease process when lung function remains preserved.
The downside of pirfenidone is its potential to cause intrusive side effects including gastrointestinal upset and photosensitive rash. As a consequence of side effects approximately one third of patients discontinue pirfenidone within 2 – 4 months of initiating treatment. The majority of the remainder however, tolerate long term treatment.

Equality and Diversity

Current lung function prediction indices tend to perform poorly in the elderly and in ethnic minority patients. Using predicted averages of Forced Vital Capacity may inappropriately preclude access to therapy in these patient groups.

Any additional sources of evidence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

n/a

Implementation issues

In England, the existence of specialist ILD services should ensure that the necessary infrastructure is in place to deliver this technology to patients with IPF

Appendix G – NHS organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Thank you for agreeing to give us your views on the technology and the way it should be used in the NHS.

Clinical Commissioning Groups (CCGs) provide a unique perspective on the technology, which is not typically available from the published literature. NICE believes it is important to involve NHS organisations that are responsible for commissioning and delivering care in the NHS in the process of making decisions about how technologies should be used in the NHS.

To help you give your views, we have provided a template. The questions are there as prompts to guide you. You do not have to answer every question. Short, focused answers, giving a CCG perspective on the issues you think the committee needs to consider, are what we need.

About you

Your name: Dr Anna Murphy

Name of your organisation: Specialist Respiratory CRG, NHS England

Please indicate your position in the organisation:

- Consultant Respiratory Pharmacist, providing views on behalf of the Specialist Respiratory CRG.
- I am a specialist pharmacist working with and supporting patients with Idiopathic Pulmonary Fibrosis, who are managed with pirfenidone treatment.

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences in opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Diagnosis and initiation of anti-fibrotic therapy for patients with IPF occurs is designated specialist ILD centres in England. Treatment is offered in accordance with the NICE IPF guidelines and in line with TA282. Monitoring of treatment and disease progression is usually performed by local secondary care services. The prescription of pirfenidone remains under the oversight of the specialist centres with drug being delivered to patients via Homecare services.

For patients with an FVC >80% current treatment is best supportive care and monitoring for disease progression. For those with an FVC<50% predicted they are

usually in the final stages of disease and treatment typically consists of a combination of best supportive care, domicillary oxygen and palliative care.

At present no alternatives to pirfenidone are available but following the recent TA, nintedanib will become an option for those patients with an FVC between 50 – 80% predicted

To what extent and in which population(s) is the technology being used in your local health economy?

University Hospitals of Leicester NHS Trust is currently has 47 patients receiving pirfenidone treatment for their IPF. A recent audit has been conducted to ensure compliance with NICE TA282 and the licensing authorisation for pirfenidone:

Data was collected for 60 patients receiving pirfenidone between August 2013 and April 2015. Of these patients 8 were female and 52 male. The age range was from 53 to 90 years old with an average age of 70 years. 22 patients were ex-smokers and the remaining 38 were non-smokers. During the period reviewed 33 (55%) patients discontinued therapy: 7 (12%) patients stopped therapy due to a drop in their FVC of >10% (3 patients subsequently died), 22 (36%) patients stopped due to adverse drug reactions and 4 (7%) patients died whilst taking therapy.

Results related to the standards set in the audit:
1. 97% diagnosis of IPF confirmed at MDT prior to starting pirfenidone (2 patients have lost MDT notes, excluding these patients the 100% standard is met)
2. 98% FVC predicted between 50% and 80% prior to starting treatment. 1 patient was outside this standard with a FVC>80% at baseline.
3. 100% of patient's therapy was discontinued due to FVC drop >10% within any 12 month period. Of the patients with available data: If % drop in predicted FVC is used, 3 patients (6.4%) had a drop > 10% on treatment and all were stopped. However, if % drop in actual FVC is used as an alternative then 18 from 47 patients had a >10% drop in FVC and 7 (14.9%) remain on treatment. With this alternative it suggests that only 61% of patients with >10% drop in FVC have had their therapy discontinued.
4. 98.3% underwent appropriate liver function test monitoring during therapy
5. 100% of patients had a eGFR of >30ml/min prior to starting and during therapy.
6. 100% were a non or ex-smoker
7. 100% of patients remaining on therapy were titrated to full therapeutic dose of 2403mg daily
8. 96% of patients' weight was monitored during treatment for evidence of significant weight loss.

There is little variation in terms of prescribing pirfenidone according to the criteria set by NICE and licensing authorisation. The audit demonstrates that it is being prescribed according to licensed indications.

The prescribing of pirfenidone and other high cost medication requires appropriate resource allocation to enhance the service. Consideration must be given to these resources including; the increased time in healthcare professional (HCP) support to

patients in terms of counselling and on-going support, drug dispensing and organisation of drug delivery services to patient, blood monitoring and clinic follow-up and review.

Potential impact on the NHS if NICE recommends the technology

What impact would the guidance have on the delivery of care for patients with this condition?

The British Thoracic Society prospective IPF Registry collects data on incident cases of IPF from a large number of UK centres. Data from the registry indicates that 42% of patients with IPF have an FVC>80% (n=100 at 21 July 2014). This is in keeping with data from individual IPF Specialist Centres at which the proportion of patients with an FVC >80% is consistently >30%. With this in mind the current TA 282 guidance denies access to pirfenidone for a significant proportion of patients with IPF. Therefore, extending the NICE criteria for inclusion of patients to be eligible for pirfenidone potentially will increase patient numbers by approximately 30% if all patients with a FVC>80% are included or an estimated 15% if only those with FVC between 80 and 90% predicted normal value included.

The recommended review of TA282 would ensure that all patients eligible and those who should receive positive clinical outcomes from treatment will be able to be prescribed pirfenidone and specialist centres will be appropriately reimbursed financially.

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Specialist ILD clinics (as per Specialist Respiratory CRG recommendations).

Would there be any requirements for additional resources (for example, staff, support services, facilities or equipment)?

An increased number of patients requiring treatment will lead to the requirement for extra resources. The main impact will be on staff time in managing the patient case load.

Can you estimate the likely budget impact? If this is not possible, please comment on what factors should be considered (for example, costs, and epidemiological and clinical assumptions).

The likely budget impact is mainly related to drug acquisition costs. There is currently a confidential patient access scheme (PAS) contracted with individual centres, which determines the reimbursed cost. If the assumption was made, that only 30% of IPF patients have an FVC>80%, then of the 5000 patients who are estimated to be diagnosed with IPF each year in the UK (Navaratnam 2011), 1500 patients would have a FVC>80%. Of this population, at least 870 patients will have an FVC between 80-90% and therefore be eligible for pirfenidone under the reviewed NICE recommendation each year. This would equate to an estimated 10% increase in annual drug costs, taking into account a % of the population will discontinue

medication due to adverse effects, FVC decline and death. Additional costs will be related to blood test monitoring and staff time for outpatient clinic organisation.

Would implementing this technology have resource implications for other services (for example, the trade-off between using funds to buy more diabetes nurses versus more insulin pumps, or the loss of funds to other programmes)?

Minimal impact on pathology services due to an increase number of blood tests required for hepatic monitoring.

Would there be any need for education and training of NHS staff?

No further training will be required as HCPs who currently prescribe and monitor pirfenidone within specialist centres will continue. Information will be provided to primary care from specialists as per current practice.

Equality

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that this appraisal:

- could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed;
- could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology;
- could lead to recommendations that have any adverse impact on people with a particular disability or disabilities.

Please tell us what evidence should be obtained to enable the Committee to identify and consider such impacts.

n/a

Other Issues

Please include here any other issues you would like the Appraisal Committee to consider when appraising this technology?

Appendix G – NHS organisation submission template

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Patient/carer expert statement (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Thank you for agreeing to give us your views on this treatment that is being appraised by NICE and how it could be used in the NHS. Patients, carers and patient organisations can provide a unique perspective on conditions and their treatment that is not typically available from other sources. We are interested in hearing about:

- the experience of having the condition or caring for someone with the condition
- the experience of receiving NHS care for the condition
- the experience of having specific treatments for the condition
- the outcomes of treatment that are important to patients or carers (which might differ from those measured in clinical studies, including health-related quality of life)
- preferences for different treatments and how they are given
- expectations about the risks and benefits of the treatment.

We have already asked your nominating organisation to provide an organisation's view. We are asking you to give your views as an individual whether you are:

- a patient
- a carer (who may be voicing views for a patient who is unable to) or
- somebody who works or volunteers for a patient organisation.

To help you give your views, we have provided a questionnaire. You do not have to answer every question — the questions are there as prompts to guide you. The response area will expand as you type. The length of your response should not normally exceed 10 pages.

1. About you

Your name: Ronald Flewett
Name of your nominating organisation: Pulmonary Fibrosis Trust
Do you know if your nominating organisation has submitted a statement?

☐X Yes ☐ No

Do you wish to agree with your nominating organisation's statement?

☐X Yes ☐ No

(We would encourage you to complete this form even if you agree with your nominating organisation's statement.)

Are you:

- a patient with the condition?

☐X Yes ☐ No

- a carer of a patient with the condition?

☐ Yes ☐ No

- a patient organisation employee or volunteer?

☐ Yes ☐X No

Do you have experience of the treatment being appraised?

☐X Yes ☐ No

If you wrote the organisation submission and do not have anything to add, tick here ☐ (If you tick this box, the rest of this form will be deleted after submission.)

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: No

2. Living with the condition

What is your experience of living with the condition as a patient or carer?

I was diagnosed with the condition in April 2014, after undergoing many tests including VATS Biopsy, Breathing Tests, Lung Function tests, Lung wash, Echocardiogram and Bronchoscopy. During these tests I was not aware of why and what was going on where it would lead to. Since diagnosis I have reduced my working hours, and now work part time. I have to plan my week, a busy day like today will be followed by two days of rest. I have to exercise or do daily house chores within my new limits, and have learnt to conserve energy to enable me to do the things are most important to me.

I am no longer in a position where I can just assume I can get up and although my brain is telling me to go for a walk, mow the lawn wash the car my body will no longer entertain doing this in one week let alone one day.

3. Current practice in treating the condition

Which treatment outcomes are important to you? (That is, what would you like treatment to achieve?) Which of these are most important? If possible, please explain why.

Pro Long life, possible cure, and reduction of cough I am aware that the cough is one of the symptoms that most sufferers would like to be reduced at best

What is your experience of currently available NHS care and of specific treatments? How acceptable are these treatments – which did you prefer and why?

Pulmonary Rehabilitation has been vital to me. I lost the incentive to exercise and was scared of damaging other organs and where my boundaries of exercise should be. Having attended several classes and the nurses explaining the benefits, and minimum oxygen levels I can take my body to, it re-ignited my fire to return to exercising.

The extra benefit is weight loss that has occurred since returning to the gym.

Lansoprazole to me has been a huge benefit, since diagnosed GERD attacks have been reduced thus reducing the acid in the lungs.

Pirfenidone has reduced slowed and stabilized the process of scarring of the lungs. Although suffering from some side effects the benefits outweigh the effects.

4. What do you consider to be the advantages of the treatment being appraised?

Benefits of a treatment might include its effect on:

- the course and/or outcome of the condition

In my case Pirfenidone has slowed and stabilized the process of scarring of the lungs. I am aware that others have considerable problems where side effects have been intolerably, no effect on the scarring have either been taken off the medicine or asked to come off

- physical symptoms

I am aware of fatigued, insomnia and tiredness that can occur as a result of the medication which I have suffered. I apply Factor 50 + Sun cream every morning and have had no rashes from sunlight. I am aware from support groups, of people suffering Dizziness, Loss of appetite, weight loss and skin Rashes.

- pain
- level of disability
- mental health

I sometimes have problems with the disease and the route it may take me, and why me? But at the end of the day, it's a card I have been dealt so turn that massive Negative into a Positive and that's why I am here to day.

- quality of life (such as lifestyle and work)

I get very frustrated with my quality of life, my biggest problem is that I look so well, most people (Family and friends) do not believe I have a terminal illness, this is my biggest obstacle

- other people (for example, family, friends and employers)

My full time employees made me redundant after 3 months of diagnosis, but with hindsight I would not have been able to maintain a full time job, Support and understanding from them would have been good at the time. I felt I was on the scrapheap at 53.

- ease of use (for example, tablets rather than injection)

In my opinion tablet is the best form

- where the treatment has to be used (for example, at home rather than in hospital)
- any other issues not listed above

Please list the benefits that you expect to gain from using the treatment being appraised.

My expectations of the treatment would be prolonging my life. Having been told that I may have 3 – 5 years to live my goal is to get to 60 with a reasonable quality of life

Please explain any advantages that you think this treatment has over other NHS treatments in England.

As above it can prolong your life and potentially give a better quality of life and slow the scarring

If you know of any differences in opinion between you and other patients or carers about the benefits of the treatment being appraised, please tell us about them.

I believe that people cannot cope with the side effects mainly Fatigued, weight loss, loss of appetite, dizziness.

5. What do you consider to be the disadvantages of the treatment being appraised?

Disadvantages of a treatment might include:

- aspects of the condition that the treatment cannot help with or might make worse

Severe Fatigue insomnia, and cough

- difficulties in taking or using the treatment (for example, injection rather than tablets)

I am happier with the medication being in tablet form.

- side effects (for example, type or number of problems, how often, for how long, how severe. Please describe which side effects patients might be willing to accept or tolerate and which would be difficult to accept or tolerate)

It is my belief that because the medication is in tablet form, users do not believe there should be side severe side effects. If people are on Chemotherapy they put up with the side effects, (Hair Falling out, Loss of weight, Nausea Tiredness etc)

- where the treatment has to be used (for example, in hospital rather than at home)

Home

- impact on others (for example, family, friends and employers)

My family and friends because I look well, still think I can do the things I used to. I try to explain that I cannot and have to do things a lot slower they tell me to stop moaning and get on with it

- financial impact on the patient and/or their family (for example, the cost of travel to hospital or paying a carer)

Having lost a full time job and now work part time the financial impact at such a young age has been hard to cope with. I am unable to claim benefits such PIP as being moderate IPF, I do not qualify

- any other issues not listed above

Please list any concerns you have about current NHS treatments in England.

No Concerns

Please list any concerns you have about the treatment being appraised.

I have no concerns about the treatment, I suppose because it is working for me it hard to criticise.

If you know of any differences in opinion between you and other patients or carers about the disadvantages of the treatment being appraised, please tell us about them.

I am aware of many differences for the medications in relation to the side effects. I believe the medication is a very strong drug with the potential of slowing the scarring and prolonging the life of the user.

6. Patient population

Do you think some patients might benefit more from the treatment than others? If so, please describe them and explain why.

I believe that patients need to give the medication time to work. I have heard of several instance because side effects early on users have asked to come off, and not given the medication time to work. I was aware from the start the hospital need to have history of whether the medication is working or not, this takes a year of breathing and Lung function tests.

Do you think some patients might benefit less from the treatment than others? If so, please describe them and explain why.

I have heard the FVC levels have fallen by more than 10% in the first year so patients have been taken off.

I think the medication needs to be treated with respect. i.e adhere to the rules of taking the medication in the leaflet

7. Research evidence on patient or carer views of the treatment

Are you familiar with the published research literature for the treatment?

☐ Yes ☐X No

If you answered 'no', please skip the rest of section 7 and move on to section 8.

Please comment on whether your experience of using the treatment as part of routine NHS care reflects the experience of patients in the clinical trials.

Do you think the clinical trials have captured outcomes that are important to patients? Are you aware of any limitations in how the treatment has been assessed in clinical trials?

If the treatment being appraised is already available in the NHS, are there any side effects that were not apparent in the clinical trials but have emerged during routine NHS care?

Are you aware of any relevant research on patient or carer views of the condition or existing treatments?

☐ Yes ☐ No

If yes, please provide references to the relevant studies.

8. Equality

NICE is committed to promoting equality of opportunity and eliminating discrimination. Please let us know if you think that recommendations from this appraisal could have an adverse impact on any particular groups of people, who they are and why.

9. Other issues

Do you consider the treatment to be innovative?

☐X Yes ☐ No

If yes, please explain what makes it significantly different from other treatments for the condition.

Yes I consider the Treatment to be innovative. There is no cure for IPF and this is the medication on the market that has the potential of increasing life expectancy

Is there anything else that you would like the Appraisal Committee to consider?

10. Key messages

In no more than 5 bullet points, please summarise the key messages of your submission.

- Find a Cure
- More research into IPF
- Make IPF a well-known illness
- Having access to all levels of care as a terminal ill patient.
- All patients need access to a specialist hospital

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Patient/carer expert statement (STA)

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

Thank you for agreeing to give us your views on this treatment that is being appraised by NICE and how it could be used in the NHS. Patients, carers and patient organisations can provide a unique perspective on conditions and their treatment that is not typically available from other sources. We are interested in hearing about:

- the experience of having the condition or caring for someone with the condition
- the experience of receiving NHS care for the condition
- the experience of having specific treatments for the condition
- the outcomes of treatment that are important to patients or carers (which might differ from those measured in clinical studies, including health-related quality of life)
- preferences for different treatments and how they are given
- expectations about the risks and benefits of the treatment.

We have already asked your nominating organisation to provide an organisation's view. We are asking you to give your views as an individual whether you are:

- a patient
- a carer (who may be voicing views for a patient who is unable to) or
- somebody who works or volunteers for a patient organisation.

To help you give your views, we have provided a questionnaire. You do not have to answer every question — the questions are there as prompts to guide you. The response area will expand as you type. The length of your response should not normally exceed 10 pages.

1. About you

Your name: Malcolm Weallans
Name of your nominating organisation: Pulmonary Fibrosis Trust
Do you know if your nominating organisation has submitted a statement?

☐ Yes ☐ No

Do you wish to agree with your nominating organisation's statement?

☐ Yes ☐ No

(We would encourage you to complete this form even if you agree with your nominating organisation's statement.)

Are you:

- a patient with the condition?

☐ Yes ☐ No

- a carer of a patient with the condition?

☐ Yes ☐ No

- a patient organisation employee or volunteer?

☐ Yes ☐ No

Do you have experience of the treatment being appraised?

☐ Yes ☐ No

If you wrote the organisation submission and do not have anything to add, tick here ☒ (If you tick this box, the rest of this form will be deleted after submission.)

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: None

2. Living with the condition

What is your experience of living with the condition as a patient or carer?

I was first diagnosed with Idiopathic Pulmonary Fibrosis in November 2006 but I am sure from the way the illness has progressed that I had the condition for many years before that. I have been treated at Windsor Chest Clinic and more recently by Royal Brompton hospital. I was part of The NICE guideline development group for IPF from 2011 to 2013 and took part in the initial assessment of Pirfenidone as a patient expert. I have been very active in the wider patient community through social networking groups and was a founder of Pulmonary Fibrosis Trust, the first UK national charity to focus specifically on helping patients to live with this condition.

3. Current practice in treating the condition

Which treatment outcomes are important to you? (That is, what would you like treatment to achieve?) Which of these are most important? If possible, please explain why.

This condition has no known cure, and very few recommended treatment options. The condition is also degenerative but nobody can tell for an individual which of the few treatment options are best. Many patients, myself included, continue to look well and show very few symptoms. Others have severe cough symptoms and/or shortness of breath on exercise. In general it is the shortness of breath on exercise which declines over the years.

I consider it important with this illness that patients should be kept as comfortable as possible and helped to cope with the symptoms that they experience.

What is your experience of currently available NHS care and of specific treatments? How acceptable are these treatments – which did you prefer and why?

When I was first diagnosed there was limited treatment available. Many patients were treated with long term steroids, azathioprine, and N-acetyl cysteine. However this was deemed unsuitable in 2013.

When Pirfenidone was first made available in August 2013 many patients started taking it, myself included. Some experienced side effects immediately, whilst others saw very little effect. Personally I was one of the more fortunate in that I did not suffer any of the side effects initially. However as time went on I became aware that I was losing appetite and weight. I was also feeling very lethargic. In august 2014 I had a lung function test which showed that my FVC had dropped from 50.5 to 43% predicted. I therefore decided in conjunction with a clinical nurse specialist from Royal Brompton to stop taking Pirfenidone as a trial to see if I could get my appetite to return. I felt much better within a few days so I decided to stop taking Pirfenidone on a permanent basis. My FVC has not returned to original levels, I am still suffering symptoms of little appetite but my weight seems to have stabilised, and I have more energy than I had.

I have discussed these symptoms with many other patients who have felt similarly. Some have stopped treatment, some have continued treatment, and some have changed to taking a newly available drug Nintedanib.

I am unlikely to be offered treatment with Nintedanib as my FVC is now below 50%.

These drugs do give patients a degree of hope that they may delay the inevitable. But personally I prefer to maintain quality of life over the side effects.

4. What do you consider to be the advantages of the treatment being appraised?

Benefits of a treatment might include its effect on: • the course and/or outcome of the condition

- physical symptoms
- pain
- level of disability
- mental health
- quality of life (such as lifestyle and work)
- other people (for example, family, friends and employers)
- ease of use (for example, tablets rather than injection)
- where the treatment has to be used (for example, at home rather than in hospital)
- any other issues not listed above

Please list the benefits that you expect to gain from using the treatment being appraised.

The benefits of taking a drug such as Pirfenidone have been shown to be an increased life expectancy for some patients, but not all. There is the natural effect of taking something being better than taking nothing which is an advantage for many. However we should be careful to measure the quality of life and if this is becoming seriously impaired withdrawal of the treatment may be better for the patient.

Please explain any advantages that you think this treatment has over other NHS treatments in England.

I do not have enough experience of the only other approved treatment.

If you know of any differences in opinion between you and other patients or carers about the benefits of the treatment being appraised, please tell us about them.

There are many differences of opinion about treatment, even among the medical profession. Patients are no exception and many of them are influenced by the advice their personal consultant gives them.

5. What do you consider to be the disadvantages of the treatment being appraised?

Disadvantages of a treatment might include:

- aspects of the condition that the treatment cannot help with or might make worse
- difficulties in taking or using the treatment (for example, injection rather than tablets)

- side effects (for example, type or number of problems, how often, for how long, how severe. Please describe which side effects patients might be willing to accept or tolerate and which would be difficult to accept or tolerate)
- where the treatment has to be used (for example, in hospital rather than at home)
- impact on others (for example, family, friends and employers)
- financial impact on the patient and/or their family (for example, the cost of travel to hospital or paying a carer)
- any other issues not listed above

Please list any concerns you have about current NHS treatments in England.

I am very concerned that consultants are not prepared to withdraw treatment with a drug such as Pirfenidone. I recently visited a patient who was taking Pirfenidone even though he was bed bound. He had lost a great deal of weigh, had no appetite at all, and asked me if I could tell him how much longer he had to endure. He would not stop taking the drug unless the specialist said so and the specialist would not say no until he had seen the patient.

Please list any concerns you have about the treatment being appraised.

I am concerned that many patients are being told that that they have to plaster themselves with Factor 50 sun cream every day. Not all patients suffer the side effect of photosensitivity, But they will never know and I am concerned at the effect of using an unlicensed product in such a thorough way.

If you know of any differences in opinion between you and other patients or carers about the disadvantages of the treatment being appraised, please tell us about them.

There are differences of opinion about the disadvantages if this treatment in the medical profession. Patients are no different.

6. Patient population

Do you think some patients might benefit more from the treatment than others? If so, please describe them and explain why.

I really feel that if this drug is going to be helpful it it more likely to be of assistance in early stages. Currently patients have to be at least moderately ill with the condition before they can get any treatment.

Do you think some patients might benefit less from the treatment than others? If so, please describe them and explain why.

I think we need to see some more research on the two available treatments to see if we can determine which treatment will work with the right group of patients.

7. Research evidence on patient or carer views of the treatment

Are you familiar with the published research literature for the treatment?

☐ Yes ☐ No

If you answered 'no', please skip the rest of section 7 and move on to section 8.

Please comment on whether your experience of using the treatment as part of routine NHS care reflects the experience of patients in the clinical trials.

I am aware that patients on the clinical trials experienced similar symptoms to those I have experienced. I am not sure that the clinical trials repots paid enough attention to the issues of weight and appetite loss.

Do you think the clinical trials have captured outcomes that are important to patients? Are you aware of any limitations in how the treatment has been assessed in clinical trials?

We are still learning about patient reported outcome measures for pulmonary fibrosis and how these could be used to improve clinical trials

If the treatment being appraised is already available in the NHS, are there any side effects that were not apparent in the clinical trials but have emerged during routine NHS care?

I am not aware of any side effects that were not apparent in the clinical trials.

Are you aware of any relevant research on patient or carer views of the condition or existing treatments?

☐ Yes ☐ No

If yes, please provide references to the relevant studies.

Anne-Marie Russell is researching this very important area.

8. Equality

NICE is committed to promoting equality of opportunity and eliminating discrimination. Please let us know if you think that recommendations from this appraisal could have an adverse impact on any particular groups of people, who they are and why.

9. Other issues

Do you consider the treatment to be innovative?

☐ Yes ☐ No

If yes, please explain what makes it significantly different from other treatments for the condition.

There are few other treatments to compare it with.

Is there anything else that you would like the Appraisal Committee to consider?

No

10. Key messages

In no more than 5 bullet points, please summarise the key messages of your submission.

- Like all drugs the use of Pirfenidone needs to be carefully monitored
- Not all patients will benefit.
- Care needs to be taken when avoiding side effects.
- Pirfenidone should be made available earlier.
-

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282): A Single Technology Appraisal

Produced by School of Health and Related Research (ScHARR), The University of Sheffield

Authors Sarah Davis, Senior Lecturer in Health Economics, ScHARR, University of Sheffield, Sheffield, UK

Rachid Rafia, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Christopher Carroll, Reader in Systematic Review and Evidence Synthesis, ScHARR, University of Sheffield, Sheffield, UK

Munira Essat, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Jean Sanderson, Research Associate in Statistics, ScHARR, University of Sheffield, Sheffield, UK

Naila Dracup, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK

Dr Stephen Bianchi, Consultant in Respiratory Medicine, Sheffield Teaching Hospitals NHS Foundation

Professor David Thickett, Consultant in Respiratory Medicine, University Hospitals of Birmingham NHS Foundation Trust and Chair in Respiratory Medicine, University of Birmingham

Correspondence Author Sarah Davis, Senior Lecturer in Health Economics, ScHARR, University of Sheffield, Sheffield, UK

Date completed Date completed (05/03/2016)

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 142/06/02.

Declared competing interests of the authors

None of the authors have any conflicts of interest to declare.

Acknowledgements

We would like to thank Paul Tappenden, ScHARR, for providing comments on the draft report and Andrea Shippam, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Davis S, Rafia R, Carroll C, Essat M, Sanderson J, Dracup N, Bianchi S, Thickett D. Pirfenidone for treating idiopathic pulmonary fibrosis: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2016.

Contributions of authors

Christopher Carroll and Munira Essat summarised and critiqued the clinical effectiveness data reported within the company's submission. Sarah Davis and Rachid Rafia critiqued the health economic analysis submitted by the company. Jean Sanderson critiqued the statistical analyses undertaken by the company. Naila Dracup critiqued the company's search strategy. Dr Stephen Bianchi and Professor David Thickett acted as clinical advisors to the ERG. All authors were involved in drafting and commenting on the final report.

Abbreviations

AEs	Adverse events
BSC	Best supportive care
CADTH	Canadian Agency for Drugs and Technologies in Health
COPD	Chronic obstructive pulmonary disease
CrI	Credible interval
CRD	Centre for Reviews and Dissemination
CS	Company's submission
CSR	Clinical study review
DLco	Diffusing capacity for carbon monoxide
ERG	Evidence Review Group
EQ-5D	EuroQoL five dimensions questionnaire
FVC	Forced vital capacity
HR	Hazard ratios
HRCT	High resolution computed tomography
HRQoL	Health-related quality of life
HTA	Health Technology Appraisal
ICER	Incremental cost-effectiveness ratio
IPD	Individual patient data
IPF	Idiopathic pulmonary fibrosis
ITT	Intention to treat
KM	Kaplan-Meier
LOCF	Last observation carried forward
NAC	N-acetylcysteine
NICE	National Institute for Health and Care Excellence
NHS	National Health Service
NMA	Network meta-analysis
PBO	Placebo
PBS	Pharmaceutical Benefits Scheme
PEY	Person exposure years
PFN	Pirfenidone
PFS	Progression-free survival
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSS	Personal Social Services
QALY	Quality-adjusted life year
RCT	Randomised controlled trial

SD	Standard deviation
6MWD	Six minute walking distance
SGRQ	St George's Respiratory Questionnaire
STA	Single Technology Appraisal
TA	Technology Appraisal
VC	Vital capacity
UCSD SOBQ	University of California San Diego Shortness of Breath Questionnaire

CONTENTS

TABLES

FIGURES

1 SUMMARY

1.1 Critique of the decision problem in the company's submission

The population addressed in the company's submission (CS) is adults with mild to moderate idiopathic pulmonary fibrosis (IPF), as specified in the final scope issued by the National Institute for Health and Care Excellence (NICE).

The Evidence Review Group (ERG) notes that patients included in the main clinical trials for pirfenidone, may not be wholly representative of the population likely to receive pirfenidone in clinical practice as real-life patients often have comorbidities, more severe disease, take concomitant medications and have a higher mortality risk compared with those patients enrolled within the clinical trials. Patients with obstructive airway disease were excluded from the clinical trials. However, clinical advisors to the ERG stated that patients with obstructive airway disease may be offered pirfenidone in current clinical practice, provided that they meet the treatment criteria laid out in technology appraisal (TA) 282.

The final NICE scope specified that if evidence allows, subgroup analysis by disease severity, defined by forced vital capacity (FVC) (such as above and below or 80% FVC) and/or diffusing capacity for carbon monoxide (DLco), should be considered. However, the CS states that available data only allowed subgroups by FVC to be assessed.

In the company's health economic analysis, the CS presents results for three populations: mild to moderate IPF (described as the intention to treat [ITT] population), mild IPF (percent predicted FVC >80%) and moderate IPF (percent predicted FVC of 50 – 80%). No subgroups results are presented by DLco status.

The intervention specified in the final NICE scope is pirfenidone and the comparators specified are best supportive care (BSC) and nintedanib. Nintedanib is only listed in the scope as a comparator for the subgroup of patients with a percent predicted FVC of between 50% and 80% as this is the population recommended for treatment in the NICE appraisal of nintedanib (TA379).

Within the economic analysis nintedanib and BSC have been included as comparators for the subgroup of patients with moderate IPF (percent predicted FVC of 50 – 80%) and BSC has been included as a comparator for the subgroup of patients with mild IPF (percent predicted FVC >80%). The ERG considers the comparators chosen for the mild and moderate subgroups to be appropriate.

For the economic analysis considering the ITT population, which includes patients with both mild and moderate IPF, only BSC is included as a comparator. The ERG does not consider this analysis to be

relevant to the decision problem as nintedanib is a valid comparator for the subgroup of the ITT population with moderate IPF (percent predicted FVC of 50 – 80%). The ERG considers that it is more appropriate to conduct an economic analysis separately within the mild and moderate subgroups as the comparators vary by subgroup.

In general, the CS adequately addresses the range of outcomes specified in the final NICE scope. The majority of the outcomes were reported for both the direct comparison with placebo from the pirfenidone clinical trial programme and for the indirect comparison with nintedanib from the network meta-analysis (NMA).

The definition of progression-free survival (PFS) used across the pirfenidone trials was not consistent; however, where possible, individual patient data (IPD) were re-analysed to provide results based on a consistent definition. However, this could not be done for all of the trials which contributed to the NMA. The ERG considers that the NMAs which combined data from studies using different definitions should be interpreted with caution.

1.2 Summary of clinical effectiveness evidence submitted by the company

The company submitted a systematic review of randomised controlled trial (RCT) evidence comparing pirfenidone with placebo in adults with mild or moderate IPF. The review identified three multi-centre international RCTs: ASCEND and CAPACITY 1 compared pirfenidone at the licensed dose of 2,403mg per day with placebo, whilst CAPACITY 2 compared pirfenidone at doses of 2,403mg per day and 1,197mg per day with placebo. It also identified two multi-centre Japanese RCTs, which compared lower doses of pirfenidone with placebo: SP3 evaluated pirfenidone doses of 1,800mg per day and 1,200mg per day and SP2 1,800mg per day only. The five trials included more than 1,700 patients with IPF. The ASCEND and SP3 trials had 52 weeks follow-up, the CAPACITY trials had 72 weeks follow-up, and the SP2 trial was terminated early at 36 weeks. The company critically appraised all five RCTs and assessed the overall risk of bias in all trials to be low.

The primary efficacy outcome for all of these trials, except SP2, was change from baseline in percent predicted FVC. The magnitude of treatment effect was also measured by mean change from baseline in FVC (ml) and the categorical outcome of a $\geq$10% decline in percent predicted FVC or death. These trials also reported all-cause and IPF-related mortality, PFS (using different definitions), 6-Minute Walking Distance (6MWD), DLco, and patient-reported outcomes, as measured by the University of San Diego Shortness of Breath Questionnaire (UCSD SOBQ) for dyspnoea, and the St George's Respiratory Questionnaire (SGRQ).

The company focused on the categorical outcome of a ≥10% decline in percent predicted FVC or death. For this outcome, the ASCEND trial reported a statistically significant difference in favour of pirfenidone compared with placebo at week 52 (absolute difference: 15.3 [95% Confidence Interval (CI) not reported], $p<0.001$), as did CAPACITY 2 at week 72 (absolute difference: 14.4 [95% CI: 7.4 to 21.3], $p=0.001$). CAPACITY 1 reported that there was no statistically significant difference between pirfenidone and placebo at week 72 (absolute difference: 3.8 [95% CI: -2.7 to 10.2], $p=0.440$). ASCEND also reported a significantly higher proportion of patients with no decline in percent predicted FVC (22.7% for pirfenidone versus 9.7% for placebo, $p<0.000001$), whilst CAPACITY 2 reported a higher proportion of patients with no decline in percent predicted FVC for pirfenidone compared with placebo (24.1% versus 13.8%) but did not report a *p*-value. CAPACITY 1 reported no statistically significant difference between pirfenidone and placebo on this outcome measure (25.8% versus 22%, *p*-value not reported). A meta-analysis of the ASCEND trial (52 weeks) and the CAPACITY trials (48 weeks) suggested that, compared with placebo, pirfenidone lowers the proportion of patients experiencing decline in FVC percent predicted of ≥10% (odds ratio [OR]: 0.50, 95% CI: 0.31 to 0.82, *p*-value not reported).

In terms of change from baseline in FVC, ASCEND (52 weeks) and CAPACITY 2 (72 weeks) found statistically significant benefits for those on pirfenidone compared with those on placebo (mean difference [MD] 4.78%; $p<0.001$ for ASCEND and absolute difference 4.4%; relative difference 35.3%; CI 0.7 to 9.1 $p=0.001$ for CAPACITY 2), whilst CAPACITY 1 found no statistically significant difference for pirfenidone compared to placebo (absolute difference: 0.6%; relative difference: 6.5%; 95% CI -3.5 to 4.7, $p=0.501$). Pooled analyses of the CAPACITY trials found statistically significant benefits for those on pirfenidone compared with placebo (absolute difference: 2.5%; relative difference: 22.8%; $p=0.005$). SP3, which reported Vital Capacity (VC), rather than FVC, also reported statistically significant benefits for those on pirfenidone for change from baseline in percent predicted VC at 52 weeks ($p=0.044$); and change from baseline in VC (ml) ($p=0.042$). Meta-analyses of change in percent predicted FVC for CAPACITY 1 & 2 and ASCEND, and change in percent predicted VC for SP3, suggested that pirfenidone reduces the decline in percent predicted FVC compared with placebo up to 52 weeks (MD: 3.4, 95% CI: 1.87 to 4.94, *p*-value not reported). The meta-analysis also suggested that pirfenidone slows the rate of decline in FVC (MD: 0.12, 95% CI: 0.05 to 0.19, *p*-value not reported) up to 52 weeks.

There were fewer overall deaths or treatment-emergent IPF-related deaths in the pirfenidone than the placebo arms of the ASCEND and CAPACITY trials. These differences were not statistically significant in the ASCEND trial at 52 weeks (for all-cause mortality or treatment-emergent IPF-related deaths, $p=0.105$ and $p=0.226$, respectively), but were significant in the pooled analyses for the CAPACITY trials at 52 weeks (for all-cause mortality and treatment-emergent IPF-related deaths,

p=0.047 and *p*=0.012, respectively). There was a significant difference between groups for treatment-emergent IPF-related mortality in the pooled CAPACITY trials at 72 weeks (*p*=0.03). Meta-analysis of CAPACITY 1 & 2 and ASCEND compared with placebo, at 52 weeks, suggests that pirfenidone reduces all-cause mortality (HR: 0.52, 95% CI: 0.31 to 0.88, *p*-value not reported) and IPF-related mortality (HR: 0.37, 95% CI: 0.18 to 0.76, *p*-value not reported).

Four of the key trials reported data for PFS: ASCEND, CAPACITY 1 & 2 and SP3. The definitions of PFS varied across the trials, albeit with a common element of a confirmed ≥10% decline from baseline in percent predicted FVC or VC. ASCEND at 52 weeks (HR 0.57; 95% CI, 0.43–0.77, *p*=0.0001) and CAPACITY 2 at 72 weeks (HR 0.64; 95% CI, 0.44–0.95, *p*=0.023) found statistically significant benefits in terms of PFS for those on pirfenidone compared with those on placebo, whilst the treatment effect for CAPACITY 1 was not statistically (HR: 0.84; 95% CI, 0.58, 1.22, *p*=0.355). *Post hoc* pooled analyses of the CAPACITY trials found statistically significant benefits for those on pirfenidone compared with those on placebo (HR: 0.62; 95% CI: 0.51 to 0.76; *p*<0.0001). Meta-analysis of the four trials, ASCEND, CAPACITY 1 & 2 and SP3 showed pirfenidone improves PFS at 52 weeks (HR 0.63 95% CI, 0.53 to 0.74, *p*-value not reported).

The CS reported the findings from two sets of analyses for 6MWD. The ASCEND and CAPACITY trials all reported findings on the pre-specified outcome of mean change from baseline in 6MWD for pirfenidone 2,403mg per day compared with placebo. ASCEND at 52 weeks (absolute difference: 26.7m; relative reduction: 44.2%; *p*=0.036) and CAPACITY 1 at 72 weeks (absolute difference: 31.8m; relative difference: not reported; *p*<0.001) both reported a statistically significant and clinically important difference between pirfenidone and placebo on this outcome, but CAPACITY 2 did not (absolute difference: 16.4m; relative difference: not reported; *p*=0.171). A pooled analysis of the CAPACITY trials at 72 weeks also reported a statistically significant and clinically important difference between pirfenidone and placebo on this outcome (absolute difference: 24m; relative difference: 31.2%; *p*=0.0009). Meta-analysis of CAPACITY 1 & 2 (data from week 48) and ASCEND (data from week 52) suggested that pirfenidone reduces the decline in 6MWD (MD: 22.9, 95% CI (10.58 to 35.23, *p*-value not reported). A *post hoc* categorical analysis based on a mean decline ≥50 m in 6MWD from baseline, or death found that there was a statistically significant difference between pirfenidone and placebo in ASCEND (52 weeks: absolute difference: 9.8%; relative reduction: 27.5%; *p*=0.04) and CAPACITY 2 (*p*=0.049), but that there was no statistically significant treatment effect for pirfenidone in CAPACITY 1 (*p*=0.10). A pooled analysis of the CAPACITY trials (72 weeks: absolute difference: 12.2%; relative reduction: 26%; *p*=0.001) also reported a statistically significant effect for pirfenidone compared with placebo for this categorical outcome.

All five included trials reported outcome data on acute exacerbations but used different definitions. The rates of acute exacerbation were higher in the ASCEND trial than in the CAPACITY trials, with higher incidence in the placebo group compared with the pirfenidone arms in the ASCEND and CAPACITY 2 trials: no *p*-values were reported. None of these three trials reported statistically significant treatment effects for this outcome measure. A meta-analysis of ASCEND, CAPACITY 1 & 2 and SP3 indicated a treatment effect in favour of pirfenidone, although the result was not statistically significant (OR 0.64 ,95% CI: 0.38 to 1.06, *p*-value not reported). CAPACITY 1 & 2 and SP2 also reported similarities in rates of hospitalisation (due to respiratory or non-respiratory causes) between the pirfenidone and placebo arms.

Neither ASCEND, CAPACITY 1 or CAPACITY 2 showed a statistically significant treatment effect compared to placebo, as assessed using the UCSD SOBQ or the SGRQ, although results of the meta-analysis suggest that pirfenidone is associated with a statistically significant reduction in USCD SOBQ compared with placebo. Four trials (CAPACITY 1 & 2, SP3, SP2) reported data on the change from baseline in DLco. The CAPACITY trials reported the change in percent predicted DLco, whilst SP2 and SP3 reported the mean decline (mL/min/mmHG). None of the trials showed a statistically significant treatment effect compared to placebo for this outcome measure.

The company submitted evidence from an ongoing, non-controlled, open-label extension (OLE) of the ASCEND and CAPACITY trials (RECAP, PIPF-012). The RECAP study is ongoing. The most recent data-cut was performed in June 2015 and the next data-cut is planned in June 2016. Survival data and time-on-treatment data were reported in the CS and were presented for patients who received pirfenidone 2,403mg per day from baseline onwards in CAPACITY and ASCEND, and through the RECAP extension period, for whom data are available through to 8.8 years. Information on survival of patients with IPF was also presented from six registries to explore the relative survival rates of trial patients receiving pirfenidone compared with these "matched" real-world patients receiving BSC. The company stated that results were similar to the comparisons reported for the trials.

The company submitted a review of evidence on the safety of pirfenidone in patients with mild or moderate IPF. The evidence presented was from the following trials: ASCEND, CAPACITY 1 & 2, SP3, SP2, RECAP and a final, non-controlled safety trial, PIPF-002. Adverse events of any intensity with the highest frequency across all trials were nausea, rash, dizziness, dyspepsia and anorexia, and these were all relatively frequent compared with placebo (no statistically significant *p*-values for between-group differences were reported, except for IPF). SP3 and SP2 also reported a very high frequency of photo-sensitivity (much higher than the CAPACITY trials). Similar, albeit slightly higher, frequencies of these and other adverse events were found in an integrated population from the RECAP extension study. Meta-analyses of treatment-emergent serious adverse events using data from

ASCEND, CAPACITY 1 & 2 and SP3 at week 52 showed no difference between the pirfenidone and placebo group (OR: 0.90, 95% CI: 0.70 to 1.15, *p*-value not reported).

In the absence of head-to-head RCTs evaluating nintedanib against pirfenidone the company conducted a Bayesian NMA to perform an indirect treatment comparison. NMAs were conducted for 11 outcomes relevant to the decision problem and the results of four of these outcomes (overall survival [OS], PFS, time to treatment discontinuation and acute exacerbations) were used to inform the economic model. Based on the NMA, the treatment effects for pirfenidone were broadly similar to those for nintedanib for all outcomes, with the pairwise treatment effects indicating that neither treatment is statistically significantly more effective.

NMA of safety data indicated that pirfenidone is associated with reduced odds of diarrhoea compared to nintedanib, and increased odds of rash as compared to nintedanib. For discontinuation due to adverse events and serious cardiac adverse events, the treatment effects for pirfenidone are broadly similar to those for nintedanib, with the pairwise treatment effects indicating that neither treatment is associated with more adverse events.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The final selection of three trials (ASCEND, CAPACITY 1 and CAPACITY 2) for the main clinical efficacy review was considered to be appropriate by the ERG, as was the inclusion of the trials from Japan, SP3 and SP2, as supporting evidence. An additional relevant trial was also identified by the ERG and included as supporting evidence; this was a multicenter Chinese trial, which compared pirfenidone plus N-acetylcysteine (NAC) with placebo plus NAC in adult patients with mild or moderate IPF (Huang 2015). The ERG noted that there were between-trial differences across some baseline characteristics in the three key trials (ASCEND, CAPACITY 1 & 2), such as mean FVC or 6MWD at baseline, but subgroup analyses suggested that these and other variables did not influence treatment effect.

Overall, the ERG assessed the potential risk of bias in ASCEND and CAPACITY 1 & 2 to be low across most domains, with the exception of reporting bias and "other bias", which were judged to be "moderate", on account of inconsistency between some of the outcomes and analyses specified in the trial protocols and those presented in the CS, and the possible influence of uncontrolled variables such as rate of disease progression. The SP3, SP2 and Huang *et al.* (2015) trials were at a higher or more unclear risk of bias across many domains than the ASCEND and CAPACITY trials. These trials all evaluated lower, unlicensed doses of pirfenidone, applied different eligibility criteria and presented noticeable differences from the other three trials in some baseline characteristics of participants.

The ERG agreed with the findings reported for the FVC outcomes for individual trials and noted that the meta-analyses generated small differences compared with the pooled analyses. The ERG also noted that the findings for CAPACITY 1 differed from those reported for CAPACITY 2 and ASCEND. The additional RCT, Huang *et al.* (2015), reported a statistically significant mean change in FVC from baseline in favour of pirfenidone plus NAC compared with placebo plus NAC at 24 weeks (p=0.02) but not at 48 weeks (p=0.11). In response to an ERG request to explain the differences between the trials on this outcome, the company stated that "*the natural variability in rates of FVC percent predicted decline of this heterogeneous disease*" might explain differences in outcomes both within and across trials.

The ERG accepts that there were fewer overall deaths or treatment-emergent IPF-related deaths in the pirfenidone arms than the placebo arms of the ASCEND and CAPACITY trials and that, in some pooled analyses, these differences were statistically significant at the 5% level. However, the ERG noted that these differences were not statistically significant in the ASCEND trial at 52 weeks and most differences that were significant in pooled analyses of the CAPACITY 1 & 2 data at 52 weeks were no longer significant at 72 weeks. However, meta-analysis of CAPACITY 1 & 2 and ASCEND at 52 weeks did suggest that pirfenidone reduces all-cause mortality (HR: 0.52, 95% CI: 0.31 to 0.88, p-value not reported) and IPF-related mortality (HR: 0.37, 95% CI: 0.18 to 0.76, p-value not reported) compared with placebo. Sensitivity analysis of the three trials at 72 weeks gave similar outcomes in favour of pirfenidone for both all-cause mortality (HR:0.64, 95% CI: 0.41 to 0.99, p-value not reported) and IPF-related mortality (HR: 0.49, CI: 0.27 to 0.87, p-value not reported), but the reduction in mortality was lower at 72 weeks compared with 52 weeks. The ERG noted that there appears to be a markedly increased rate of mortality in the CAPACITY trials between the data reported for 52 weeks and for 72 weeks, the reasons for which are unclear. SP3, SP2 and Huang *et al.* (2015) all reported all-cause mortality and found no differences between the pirfenidone and placebo arms.

The results for PFS were consistent across trials and analyses demonstrated a beneficial effect on this outcome for pirfenidone compared with placebo. The exception, again, was the CAPACITY 1 trial, which reported that the difference between pirfenidone 2,403mg per day and placebo was not significant (p=0.355).

The results for 6MWD were consistent in terms of direction of effect (favouring pirfenidone) but statistical significance varied between trials and between 6MWD outcome measures. The ASCEND and CAPACITY trials all reported findings on the pre-specified outcome of mean change from baseline in 6MWD for pirfenidone 2,403mg per day compared with placebo. ASCEND (absolute difference 26.7 p=0.036) and CAPACITY 1 (absolute difference 31.8 p<0.001) reported a statistically

significant and clinically important difference between pirfenidone and placebo on this outcome, but CAPACITY 2 did not (p=0.171). A *post hoc* categorical analysis based on a mean decline ≥50 m in 6MWD from baseline, or death, found that there was still a statistically significant difference between pirfenidone and placebo in ASCEND (p=0.04), but treatment effect for CAPACITY 1 was not statistically significant (p=0.10) and the treatment effect for CAPACITY 2 was statistically significant (p=0.049). An additional small RCT of pirfenidone in combination with NAC in adults with mild and moderate IPF identified by the ERG (Huang *et al.* 2015) also reported no statistically significant effect for pirfenidone (plus NAC) compared with placebo (plus NAC) on 6MWD outcomes (p=0.43).

The ERG noted that pirfenidone does not have a significant treatment effect compared to placebo, as assessed by a number of other outcomes: rates of acute exacerbations; patient-reported outcomes as measured by the SGRQ; or DLco. For the UCSD SOBQ the treatment effects were not statistically significant for any of the individual trials, but results of the meta-analysis suggest that pirfenidone is associated with a statistically significant reduction in in USCD SOBQ compared with placebo.

The ERG noted how the effect of the, "*intrinsic variability in rates of FVC decline*" (Noble 2011) might explain differences in some outcomes across trials. Participants in the trials included in the CS were not stratified by rate of progression, so it is possible, for example, that the placebo arm might have had more participants with more rapidly progressing disease than the intervention arm. As a result, the true treatment effect of the intervention relative to placebo is uncertain. This could work either for or against the intervention.

A *post hoc* pooled analysis of ASCEND and CAPACITY 1 & 2 found no evidence for differential treatment effects according to disease severity, as assessed using three key efficacy outcomes; absolute ≥10% FVC decline, ≥50m 6MWD decline, and ≥20-point worsening of dyspnoea as measured by UCSD SOBQ. For these analyses disease severity was categorised according to baseline percent predicted FVC of ≤80% (moderate IPF) and >80% (mild IPF). In response to a clarification request from the ERG, the company also provided subgroup analyses according to disease severity for OS and PFS from the ASCEND and CAPACITY trials, although exact numbers within each subgroup in each trial arm were not reported. The findings did not suggest differential treatment effects according to disease severity for either outcome (as judged by the reported HR and 95% CI), however a treatment-by-subgroup interaction test was not reported so it is unclear if the difference between these subgroups was statistically significant.

The ERG noted that, overall, some adverse events (AEs) were frequent, especially nausea, rash, dizziness, dyspepsia, anorexia and photosensitivity, but that these were generally mild or moderate in severity. The ERG requested from the company more detailed data on serious adverse events and

adverse events leading to discontinuation. The most frequently-reported serious adverse events in the pirfenidone arms of the ASCEND and CAPACITY trials, other than worsening of IPF, were pneumonia, prostate cancer, angina pectoris, coronary artery disease, congestive cardiac failure, atrial fibrillation and pneumothorax. The AEs leading to discontinuation of treatment in ≥1% of patients in pirfenidone groups were pneumonia, rash, raised hepatic enzyme levels and decreased weight (in ASCEND), photosensitivity, rash and respiratory failure (in CAPACITY 1) and bladder cancer, nausea and rash (in CAPACITY 2). The majority of safety data were from trials with a follow-up of no more than 72 weeks, but the CS did present analyses that included more than 300 patients who had received pirfenidone for more than four years. However, the results for these patients were not presented separately. The ERG noted that the two ongoing studies to evaluate safety would address some outstanding issues: the non-randomised, non-controlled, OLE study that included a set of patients who completed either ASCEND, CAPACITY 1 or 2 (RECAP) and PIPF-002, an ongoing open-label compassionate-use study in US patients with either IPF or secondary pulmonary fibrosis.

The ERG considers that the NMA appears to be of good methodological quality, and the choice of random effects model was appropriate given the stated concerns in terms of heterogeneity between the studies. The ERG's key concerns were in the use of the earlier 52 week follow up data for key time-to-event outcomes (all-cause mortality and PFS), rather than the full 72 week data available, and the difference in the treatment effects observed at these two time points despite the claim of proportional hazards over both the observed and unobserved time period.

1.4 Summary of cost effectiveness evidence submitted by the company

The company submitted a fully executable economic model as part of their submission to NICE. The analysis was undertaken from the perspective of the UK National Health Service (NHS) and Personal Social Services (PSS) over a lifetime horizon. The company's analysis is presented for three populations: (1) the ITT trial population, which is comprised of adults with mild to moderate IPF; (2) people with a percent predicted FVC above 80% at baseline (considered to be mild IPF), and; (3) people with a percent predicted FVC of 50 to 80% at baseline (considered to be moderate IPF). Within all three analyses, comparators include BSC (defined in the trial as symptom relief, pulmonary rehabilitation, management of comorbidities, and end of life care, including oxygen therapy). Nintedanib is included only in the analysis of people with a percent predicted FVC of 50 - 80% at baseline; it is excluded from the analyses of ITT-trial population and people with a percent predicted FVC above 80% at baseline. In the company's base-case, people initiating pirfenidone are assumed to discontinue treatment at the rate observed in the trials; and therefore no stopping rule is applied in the base-case. The stopping rule defined by NICE which formed the basis for the positive recommendation for pirfenidone and nintedanib is however applied to nintedanib in the company's

base-case. A scenario analysis is also presented where the stopping rule is applied to both nintedanib and pirfenidone.

Within the ITT-trial population (adults with mild to moderate IPF), the company's model estimates the ICER for pirfenidone against BSC to be [redacted] per QALY gained (probabilistic ICER = [redacted] per QALY gained) using the list price for pirfenidone.

Within the subgroup of people with a percent predicted above 80% at baseline (considered to be mild IPF), the company's model estimates the ICER for pirfenidone against BSC to be [redacted] per QALY gained (probabilistic ICER = [redacted] per QALY gained) using the list price for pirfenidone.

Within the subgroup of people with a percent predicted FVC of 50 - 80% at baseline (considered to be moderate IPF), the CS estimates that BSC provided the least number of QALYs, followed by nintedanib and pirfenidone. Using the company's model estimates, based on a fully incremental analysis, nintedanib is ruled out due to extended dominance. The company's model estimates the ICER for pirfenidone against BSC to be [redacted] per QALY gained (probabilistic ICER = [redacted] per QALY gained).

Based on the company model when incorporating the PAS for pirfenidone, the ICER for pirfenidone versus BSC was £21,387 per QALY in the ITT population and £24,187 per QALY in the mild subgroup (percent predicted FVC >80% at baseline) and £21,318 per QALY in the moderate subgroup (percent predicted FVC of 50 - 80% at baseline). The results for pirfenidone versus nintedanib when incorporating the nintedanib and pirfenidone PAS (moderate subgroup) are reported in the confidential appendix.

The company presented a series of scenario analyses. The ICERs were mostly sensitive to the assumption regarding the time horizon, the duration over which the treatment effect is assumed to remain constant, the parametric distributions for OS in people initiating pirfenidone, the treatment effects taken from the NMAs for OS, and the inclusion of stopping rules for pirfenidone and nintedanib.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG critically appraised the company's health economic analysis and the model upon which this analysis is based. The ERG has a number of concerns regarding the structure and parameterisation of the company's model. These include:

- the inability of the model to capture the progressive nature of IPF
- the absence of a stopping rule for pirfenidone in the company's base-case

- the inadequacy of the partitioned survival approach when implementing the stopping rule
- the assumption that treatment effect is constant over the entire model duration
- the estimation of the treatment effect

The ERG further observes that under the company's base-case assumptions, there are discrepancies between the model's prediction of OS for people initiating BSC and the observed trial data for OS in patients who were randomised to placebo. The CS does not comment on these discrepancies and instead focuses on a comparison of the model prediction with registry data for patients receiving BSC, even though the registry data does not match the trial data for people randomised to placebo.

1.6 ERG commentary on the robustness of evidence submitted by the company

The ERG notes the following strengths and weaknesses in the evidence submitted by the company.

1.6.1 Strengths

- The CS reports a generally good quality systematic review of the RCT evidence.
- The three principal RCTs are generally at a low risk of bias.
- Generally, there are no major safety concerns, and some long-term safety evidence is available.
- Evidence in the model for pirfenidone is based upon long-term data for people included in RECAP.
- Results from NMAs are used to inform the relative treatment effects for the comparators.
- Whilst EQ-5D data were not directly available in the trials, SGRQ data from the trials were mapped onto the EQ-5D using a mapping algorithm developed in people with IPF.

1.6.2 Weaknesses and areas of uncertainty

- There is a moderate risk of reporting bias in the three key RCTs and unclear, moderate or high risk of bias across some domains in the three supporting RCTs.
- There are difficulties in controlling for the rate of disease progression among IPF trial participants, which might moderate outcomes, however the extent of this is unclear.
- The efficacy findings are not consistent across individual trials; one of the key trials reports no statistically significant treatment effect for pirfenidone compared with placebo on the primary outcomes measures relating to FVC or the secondary outcome of PFS.
- Individual trials do not report any statistically significant treatment effect compared to placebo for mortality outcomes; a statistically significant treatment effect is only observed when pooling or meta-analysing studies.

- The treatment effects for a number of clinically important and patient-reported outcomes were either not statistically significant (DLco and SGRQ) or did not meet the threshold for a clinically important difference (UCD SOBQ).
- It is unclear how long the treatment effect might be sustained.
- Simplification of a progressive disease into two discrete health states (pre- and post-progression) fails to capture the ongoing progressive nature of IPF and the impact of different levels of disease severity on quality of life and costs.
- The implementation of the stopping rule in the company's model lacks validity.
- There is uncertainty around the treatment effects due to the heterogeneity between trials included in the NMAs in terms of study duration, outcome definition and handling of missing data.
- The duration of extrapolation of the treatment effect is associated with considerable uncertainty.
- There are discrepancies between the modelled OS in people initiating BSC and the OS observed in the clinical trials.
- The treatment effects for the subgroup of people with a percent predicted FVC above 80% are uncertain.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

A number of analyses were undertaken by the ERG which informed the ERG's preferred base-case. The main changes within the ERG's preferred base-case were:

- use of treatment effects estimated from the NMA from the CODA samples of the predictive distributions, using data up to 72 weeks, excluding SP3
- exploration of different durations for the extrapolation of the treatment effect (2 years and entire model duration)
- use of the Gompertz distribution for OS (rather than the Weibull)
- capping utility estimates for individuals at a maximum of 1.0
- adjustment of utility by age
- inclusion of the costs associated with end of life care for all people irrespective of the cause of death
- amendments to dose reductions/interruptions assumed in the company's model for pirfenidone and nintedanib
- amendment of minor programming errors in the economic model

The ERG's preferred scenario led to a higher ICER for pirfenidone against BSC (approximately two-fold compared with ICERs reported by the company) for all three populations (ITT, FVC of 50 - 80%

at baseline, FVC >80% at baseline). For the ITT population the ICERs incorporating the PAS ranged from £27,124 to £115,751. For the mild population (percent predicted FVC >80%) the ICERs incorporating the PAS ranged from £31,722 to £186,260. Fort the moderate population (percent predicted FVC 50 – 80%), the ICERs for pirfenidone versus BSC ranged from £31,722 - £186,260 when incorporating the PAS. Results incorporating the PAS for pirfenidone versus nintedanib in the moderate population are presented in the confidential appendix.

A key uncertainty in the company's model concerns the duration of the extrapolation of the treatment effect. As reported in the company's scenario analyses and the ERG's exploratory analyses, truncating the duration over which the treatment effect applies increases the ICERs for pirfenidone versus BSC. A further important limitation in the company's model relates to the implementation of stopping rules for pirfenidone and nintedanib. The inclusion of the stopping rule in the economic model lacks validity in that the modelled stopping rule impacts on costs but not health outcomes. The ERG considers that the analysis incorporating the stopping rule as implemented in the economic model provides a lower bound of the plausible ICER (i.e. most optimistic scenario).

2 BACKGROUND

Pirfenidone is licensed in the EU for the treatment of mild to moderate IPF in adults.[1] Pirfenidone was previously appraised as part of the NICE Single Technology (STA) process (TA282), with guidance issued in April 2013.[2] Pirfenidone was recommended as an option for treating idiopathic pulmonary fibrosis only if the person has a forced vital capacity (FVC) > 50% and ≤ 80% predicted and the company provides pirfenidone with the discount agreed in the Patient Access Scheme (PAS). The review of TA282 was prompted by publication of the ASCEND study.[3] This report provides a review of the company's submission (CS)[4] provided by the company for pirfenidone (including any additional material submitted by the company in response to clarification requests) during NICE's review of TA282.

2.1 Critique of company's description of underlying health problem

The ERG considers that in general the company's description of the underlying health problem is appropriate and relevant to the decision problem. The ERG notes that whilst the CS states that median 5-year survival is 20%, the source paper by Kim *et al.*, estimates median 5-year survival to be between 20% and 40%.[5] Kim *et al.* also state that survival estimates are dependent on whether survival is estimated from diagnosis, symptom onset or first radiographic abnormality.[5]

The CS states that current guidelines do not propose a formal staging system for classification of disease severity. Clinical advisors to the ERG agreed with the statement in the CS that using percent predicted FVC alone to define mild and moderate disease has the potential to misclassify patients for two reasons. Firstly, FVC can be elevated in patients with emphysema, which masks the impact of fibrosis on lung capacity. Secondly, the normal range for percent predicted FVC is 90% to 120%, so some patients who have an FVC of 80% may have lost a third of their baseline lung capacity and others may have only lost a tenth. Therefore, the same percent predicted FVC may result in a different severity of IPF symptoms being experienced in different individual patients. Clinical advisors to the ERG considered that whilst percent predicted FVC had been used to define severity in clinical trials, this measure was not widely used in clinical practice, except to implement the recommendations in TA282. They commented that carbon monoxide diffusing capacity of the lungs (DLco) is clinically more meaningful and that DLco is the primary measure used to determine eligibility for lung transplantation, as some patients can have very low DLco values that suggest lung transplantation would be beneficial whilst maintaining a percent predicted FVC value that, in isolation form other measures, would indicate mild disease.

Clinical advisors to the ERG agreed that the course of IPF is unpredictable and heterogeneous. They also agreed with the statement in the CS that a prior decline in lung function does not predict a future

decline and they noted that this statement is also supported by an analysis by Schmidt *et al.* based on a retrospective analysis of pulmonary function tests from 734 patients recruited across 3 centres.[6]

2.2 Critique of company's overview of current service provision

The ERG considers that in general the company's overview of current service provision is appropriate and relevant to the decision problem. However, some additional clarification on the treatment pathway described in the CS is provided below.

Whilst the ERG agrees that N-acetylcysteine (NAC) is not an appropriate comparator for pirfenidone, NAC is currently used in some patients. Clinical Guideline 163 (CG163) recommends that patients should be advised that *"oral N-acetylcysteine is used for managing idiopathic pulmonary fibrosis, but its benefits are uncertain."* Clinical advisors to the ERG confirmed that it is used in some patients for symptom relief as part of BSC but NAC is not expected to be disease-modifying. There is also a recent clinical trial of NAC versus placebo on top of a background therapy of pirfenidone in both arms, which is yet to report in full, but conference abstracts reporting preliminary results suggest that the combination is generally well tolerated but does not provide additional benefit compared to pirfenidone alone.[7, 8]

Clinical advisors to the ERG also reported that a few patients are currently managed with prednisolone or azathioprine. Whilst these drugs are not recommended in CG163 to modify disease progression, their ongoing use in some patients is a possibility under recommendation 1.5.14 of CG163 which states, *"if people with idiopathic pulmonary fibrosis are already using prednisolone or azathioprine, discuss the potential risks and benefits of discontinuing, continuing or altering therapy."* However, the ERG recognises that the use of prednisolone and azathioprine is likely to be limited to a minority of patients and is not expected to be disease-modifying.

Whilst the ERG agrees that pirfenidone has been the standard of care for patients with moderate IPF since TA282 was published in 2013, the ERG notes that following the publication of TA379 in January 2016, nintedanib is likely to become part of the standard of care in the coming months. In Section 3.6 of the CS, which describes other (non-NICE) guidelines, it is stated that pirfenidone is recommended by the ATS/ERS/JRS/ALAT Clinical Practice Guideline. The ERG notes that nintedanib is also recommended in the same document with both treatments being recommended on the basis of the panel considering that both have 'moderate confidence in effect estimates'.[9]

Clinical advisors to the ERG also noted that now that there are two disease-modifying therapies available for patients with moderate IPF, it is possible that a second therapy may be used in patients who have failed to tolerate one therapy or who have progressed on one therapy but who still meet the

starting criteria for the other therapy. Treatment sequences were not addressed in the original CS. Following a clarification request, the company acknowledged that it is possible that clinicians may sequence pirfenidone and nintedanib within the moderate population (see clarification response,[10] question B7). However, the company went on to state that no sequencing studies exist or are anticipated to become available and it is unclear whether the efficacy would be different when used second-line, particularly given that there remains uncertainty regarding the exact mechanism of both pirfenidone and nintedanib. They conclude that any analysis of treatment sequences would be purely speculative in design.

Clinical stopping rules are applied for pirfenidone in TA282 and for nintedanib in TA379. Both sets of guidance recommend that treatment is discontinued if there is evidence of disease progression which is defined as a decline in percent predicted FVC of 10% or more within any 12 month period. The CS claims that the application of this stopping rule is complicated since progression with treatment does not always constitute treatment failure. This statement is supported by a *post hoc* analysis of data from the ASCEND and CAPACITY 1 and 2 trials which showed that patients who continued with pirfenidone following a ≥10% decline in percent predicted FVC, had a significantly reduced risk of the composite outcome of death or a further 10% decline in percent predicted FVC (p=0.032), compared to those who continued with placebo following a ≥10% decline in percent predicted FVC.[11] However, it should be noted that this *post hoc* analysis may be subject to potential bias as it was based on a small proportion of the trial population (3.9% [=24/623] of patients randomised to pirfenidone and 9.6% [=60/624] of those randomised to placebo) who had experienced a 10% decline in the first 3 or 6 months of the study and who had remained on treatment,[11] and therefore patient characteristics may not be balanced between the two groups being compared. Clinical advisors to the ERG reported that to their knowledge the stopping rule is being rigorously applied in clinical practice, but they agreed that the stopping rule is clinically problematic as a prior decline in lung function does not predict a future decline, and periods of stability can sometimes only be identified retrospectively. They also noted that in clinical practice the stopping rule is only applied to patients with a >10% FVC or >15% DLco decline over any 12 month period when the lung function decline has been confirmed as not being due to a temporary and reversible infection. There will therefore be patients who will either already be defined as having severe disease on DLco criteria who would have been offered therapy due to an eligible FVC measurement or will have developed a DLco <35% but if FVC remains between 50 and 80% will have treatment continued. Similarly there is no necessity to stop a patient's therapy if the FVC declines below 50% if the decline is less than 10% per year.

3 CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 1 summarises the population, intervention, comparators and outcomes specified within the company's decision problem. These are discussed and critiqued in the following sections.

Table 1: Summary of the decision problem (adapted from Table 1 of the CS)

	Final scope issued by NICE[3]	Decision problem addressed in the CS[4]	ERG comments
Population	Adults with mild to moderate IPF	Same as final scope issued by NICE	The population addressed in the CS is consistent with the population specified in the final scope.
Intervention	Pirfenidone	Same as final scope issued by NICE	The intervention in the CS is consistent with the population specified in the final scope.
Comparators	• Best supportive care • Nintedanib (only for people with a percent predicted FVC of 50 - 80%, subject to ongoing NICE appraisal)	Same as final scope issued by NICE	The ERG notes that guidance on the use of nintedanib is now published (TA379) and nintedanib is recommended for people with a percent predicted FVC of 50 - 80%.[12] Therefore its inclusion as a comparator in this subgroup is appropriate. NAC, prednisolone and azathioprine were not considered to be relevant comparators for pirfenidone. The ERG notes that these are used as part of BSC in some patients but they are not expected to be disease-modifying.
Outcomes	Outcome measures to be considered include: • Pulmonary function parameters • Physical function • Exacerbation rate • PFS	Same as final scope issued by NICE	In addition to the outcomes listed in the scope, data are also presented for hospitalisations and all-cause discontinuations.

	• Mortality • Adverse effects of treatment • Health-related quality of life		
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies should be taken into account.	Same as final scope issued by NICE	For those analyses which incorporated the nintedanib PAS, results are provided in a confidential appendix.
Subgroups to be considered	If evidence allows, subgroup analysis by disease severity, defined by FVC (such as above and below or 80% FVC) and/or diffusing capacity for carbon monoxide, will be considered	Same as final scope issued by NICE. Subgroup analysis by FVC and DLco status at baseline was investigated, but the available data only allowed FVC to be assessed and reported in this submission.	In the economic analysis, the CS presents results for three populations: • mild to moderate IPF (described as the ITT population)

			• mild IPF (percent predicted FVC >80%) • moderate IPF (percent predicted FVC > 50% and ≤ 80%).

3.1 Population

The population addressed in the CS is adults with mild to moderate IPF; this is in line with the final NICE scope. Harari and Caminati[13] describe how populations and outcomes compare between clinical trials and observational studies that describe real-life treatment. The studies described by Harari and Caminati include single centre studies, such as the UK named patient programme which existed prior to TA282,[14] and international collaborative registries, such as PASSPORT which included UK sites.[15] They conclude that although the profile of patients treated with pirfenidone seems to be quite similar all over the world, patients treated in real-life scenarios differ from those treated in RCTs as real-life patients often have comorbidities, more severe disease, take concomitant medications and have a higher mortality.[13]

In terms of the patients excluded from the three main trials, the ASCEND trial appears to have been more restrictive as it excluded a larger proportion of patients following screening, with only 36% of those screened undergoing randomisation. The proportions of screened patients included in CAPACITY 1 and CAPACTIY 2 were 61% and 56%, respectively. The ASCEND and CAPACITY 1 and 2 trials all excluded patients with obstructive pulmonary disease (asthma or COPD) and patients with significant comorbidities such as a history of unstable or deteriorating cardiac or pulmonary disease (other than IPF). However, clinical advisors to the ERG stated that they would still treat with pirfenidone if there was evidence of asthma or COPD, provided the patient met the treatment criteria specified in TA282 (i.e. a predicted FVC between 50% and 80%). They also stated that many of the patients treated in routine clinical practice had comorbidities. This suggests that the key clinical trials for pirfenidone excluded some patients who would be treated in clinical practice.

In terms of disease severity, the proportion of patients with mild IPF, (i.e. a percent predicted FVC above 80%) was around 25% according to the figures presented in the CS (see CS, page 114 and Table 67). Clinical advisors to the ERG commented that the proportion of patients with an FVC above 80%, in the absence of emphysema (which elevates FVC), varied somewhat across different areas of the UK but was more likely to be between 30% and 50%. It is therefore possible that the subgroup who present with mild IPF are under-represented within the trial populations. It was also noted that only one of the pirfenidone trials, CAPACITY 2, recruited patients from UK centres (3 of 110 centres were UK).

The final NICE scope also specifies that if evidence allows, subgroup analysis by disease severity, defined by FVC (such as above and below or 80% FVC) and/or diffusing capacity for carbon monoxide, should be considered. The statement of the decision problem (see CS, Table 1, page 18, reproduced in Table 1) states that subgroup analysis by FVC and DLco status at baseline was investigated, but the available data only allowed FVC to be assessed and reported in the CS. In the original CS, some subgroup analyses by percent predicted FVC were presented for a limited number

of outcomes, but subgroup analyses were not presented for all the outcomes specified in the final NICE scope. In response to a request for clarification from the ERG, the company provided additional subgroup analyses which examined subgroups defined by percent predicted FVC (> 80% versus ≤80%) for the outcomes of change in percent predicted FVC, overall survival, and PFS, as requested, and for two additional supportive outcomes (see clarification response,[10] questions A29 and A31).

In the economic analysis, the CS presents results for three populations: (1) mild to moderate IPF (described as the ITT population); (2) mild IPF (percent predicted FVC >80%), and; (3) moderate IPF (percent predicted FVC > 50% and ≤ 80%). No subgroups results are presented by DLco status.

The ERG considers that it is more appropriate to conduct an economic analysis separately within the mild and moderate subgroups, as the comparators vary by subgroup, than to consider the ITT population with nintedanib excluded as comparator.

3.2 Intervention

The intervention is pirfenidone, as per the final NICE scope. Pirfenidone is indicated in adults for the treatment of mild to moderate IPF.[1] The mechanism of action of pirfenidone has not been fully established, however, existing data suggest that pirfenidone exerts both antifibrotic and anti-inflammatory properties.[1]

The previous appraisal of pirfenidone for treating idiopathic pulmonary fibrosis (TA282) recommended pirfenidone as an option only in patients with a percent predicted FVC of between 50% and 80%, which is a subgroup of the population covered by its marketing authorisation.[2] It also recommended that treatment *"should be discontinued if there is evidence of disease progression (a decline in per cent predicted FVC of 10% or more within any 12 month period)."*[2]

The recommended daily dose of pirfenidone for patients with IPF is three 267mg capsules three times a day (a total of 2,403 mg per day).[1] The capsules are taken orally with food.[1] Does adjustments and treatment interruptions are allowed to manage adverse events with re-escalation to the recommended daily dose as tolerated.[1] Treatment with pirfenidone should be initiated and supervised by specialist physicians experienced in the diagnosis and management of IPF.[1] Pirfenidone is linearly priced, with pack size costs for 267mg capsules of £501.92 for 63, £2,007.70 for 252 and £2,151.10 for 270.[4] The cost per day for the licensed dose of 2,403mg per day is £71.70 at the list price.[16]

3.3 Comparators

The comparators listed in Table 1 of the CS are BSC and nintedanib. These comparators are consistent with those defined in the final NICE scope. Nintedanib is only a comparator for people with a percent predicted FVC of between 50% and 80%. This is appropriate as this is the population

covered by the recommendation for nintedanib in TA379.[12] However, it should be noted that nintedanib is indicated in "adults for the treatment of IPF",[17] and the restriction of nintedanib as a comparator to patients with a percent predicted FVC of between 50% and 80% is due to the treatment criteria defend in the TA379, which match those defined for pirfenidone in TA282. The stopping criteria for nintedanib in TA379 also match those for pirfenidone in TA282.

BSC is defined in the CS as information and support, symptom relief, management of comorbidities, withdrawal of therapies suspected to be ineffective or causing harm, end of life care, oxygen therapy and/or pulmonary rehabilitation. The ERG and its clinical advisors considered this to be an appropriate description of BSC in current UK practice. The clinical advisors to the ERG also noted that BSC may vary internationally, particularly in countries without universal access to healthcare, and therefore the BSC received by non-UK trial participants may not reflect UK current practice.

Clinical advisors to the ERG were also asked whether any other therapies are currently used in the UK. As discussed in Section 2, the clinical advisors to the ERG noted that NAC is used off-license in some patients for symptom relief as part of BSC, but that it is not expected to be disease-modifying. They also reported that a few patients are currently managed with prednisolone or azathioprine, but again these treatments are not expected to be disease-modifying. NAC, prednisolone and azathioprine were not considered to be relevant comparators for pirfenidone by the clinical advisors to the ERG and were not included as comparators in the final scope.[3] The ERG therefore agrees with the exclusion of NAC, prednisolone and azathioprine from the list of relevant comparators.

Within the economic analysis nintedanib and BSC have been included as comparators for the subgroup of patients with moderate IPF (percent predicted FVC of 50 – 80%) and BSC has been included as a comparator for the subgroup of patients with mild IPF (percent predicted FVC >80%). The ERG considers the comparators chosen for the mild and moderate subgroups to be appropriate.

For the economic analysis considering the ITT population, which includes patients with both mild and moderate IPF, only BSC is included as a comparator, even though nintedanib is a valid comparator for the subgroup with moderate IPF. The ERG considers that it is more appropriate to conduct an economic analysis separately within the mild and moderate subgroups, as the comparators vary by subgroup, than to consider the ITT population with nintedanib excluded as comparator.

3.4 Outcomes

The outcomes reported in the CS match those described in the final NICE scope.[3] The outcomes presented it the CS are discussed in turn.

3.4.1 Pulmonary function

A number of pulmonary function measures are reported in the CS including;

- mean change in percent predicted FVC/VC from baseline,
- mean change in FVC/ VC (mL)
- decline of ≥10 % in percent predicted FVC
- Mean change in percent predicted DLco
- Mean change in DLco (mL)

Mean change from baseline in predicted FVC/VC and mean change in FVC/VC (L) were included as continuous outcomes in the NMA. A decline of ≥10% in percent predicted FVC was included as a binominal outcome in the NMA. Outcomes relating to DLco were only reported for the direct comparison of pirfenidone against placebo.

The pulmonary function outcome which forms the main focus of the submission is FVC. FVC is an accepted trial endpoint for IPF, and one that has been widely used in trials to date.[18,19] It is widely recognised that the change in FVC over time, rather than the absolute FVC, is the outcome of interest, and a change of ≥10% appears to be accepted as being sufficient to define a true change.[19]

The CS cites evidence to support the claim that FVC is a good surrogate for survival, with a ≥10% decline in percent predicted FVC having been shown to be predictive of higher mortality in a number of studies and smaller changes (5-10%) in percent predicted FVC having been shown to be predictive of mortality in a smaller number of more recent studies (see CS, page 201).

The CS cites one study showing that there is a moderate correlation between changes in percent predicted FVC and changes in a disease specific health-related quality of life (HRQoL) measure, (Spearman correlation coefficient of -0.32), but the correlation between absolute values for percent predicted FVC and HRQoL is weaker (Spearman correlation of -0.16).[20] The ERG notes that whilst some evidence on the validity of FVC as a surrogate for mortality and HRQoL is presented, a systematic search does not appear to have been conducted as other relevant papers presenting data on the correlation between FVC and HRQoL have not been summarised.[21,22] However, the ERG notes that in the appraisal of ninetadnib (TA282), the Appraisal Committee concluded that, *"although it had some limitations, percent predicted FVC is the most reliable and widely used measure of lung function in clinical practice."*[12]

Within the CS, data from trials which reported VC but not FVC have been combined with data from trials that reported FVC. This is justified in the CS by the statement that: *"…there is little difference between VC and FVC in subjects without obstructive pathology."* Clinical advisors to the ERG considered this statement to be reasonable. However, whilst the ASCEND and CAPACITY 1 and 2

trials and SP2 trials excluded patients with obstructive airway disease, the exclusion criteria for SP3 are not as clear regarding the exclusion of patients with COPD or emphysema (CS Appendices, Tables A5.1. to A5.4). Therefore, the ERG considers that the combination of VC data from SP3 with FVC data from the ASCEND and CAPACITY trials is questionable.

Although there are some data to suggest that DLco is a good prognostic indicator for mortality in IPF,[23, 24] it is not as well accepted as a clinical trial endpoint.[18, 19] Clinical advisors to the ERG agreed that DLco is harder to measure and is more variable than FVC. The variability of DLco has commonly been recognised to be as high as 15%,[19] whereas the minimal clinically important difference for FVC is reported to be between 2% and 6%.[20] The ERG therefore concludes that whilst DLco may provide important relevant information in clinical practice, it is reasonable for the CS to focus on FVC as the main measure of pulmonary function as it is more accepted as a reliable outcome in a clinical trial setting.

The clinical advisors to the ERG also noted that there is up to 10% variation in FVC testing in real-life clinical settings and therefore when using a >10% decline in FVC to define disease progression, this should not be based on a single FVC reading and any decline should be confirmed as not being due to a temporary and reversible infection.

3.4.2 Physical functioning

The measure of physical functioning reported is the 6 minute walking distance (6MWD). Results are reported both for the mean change in 6MWD from baseline and for a categorical analysis of change from baseline using a threshold of a decrement of ≥50m. Mean change in 6MWD from baseline was included as an outcome in the NMA but loss of ≥50m in 6MWD was only reported for the direct comparison of pirfenidone against placebo.

In the appraisal of nintedanib (TA282), the Appraisal Committee heard from clinical experts that the 6MWD was an unreliable measure.[12] However, in the previous appraisal of pirfenidone, the Committee accepted the use of 6MWD as a covariate to predict survival in the microsimulation model.[3] This opinion is supported by an analysis by du Bois 2011, which showed that a decrement in 6MWD of greater than 50 metres over 24 weeks was associated with a HR for overall mortality at 1 year of 4.27 (p=0.001) when compared with a decrement of less than 25 metres.[25] However, the statistical significance of a decrement of greater than 50 metres when compared to a decrement of between 25 and 50 metres was not demonstrated.[25] Therefore, a decrement of more than 50m in 6MWD may not result in a statistically significantly higher risk of mortality compared with a decrement of less that 50m in 6MWD. The same study also found moderate correlations between changes in 6MWD and changes in disease-specific HRQoL measures which were statistically

significant.[25] The ERG notes that the CS states that the minimal clinically important difference (MCID) for the 6MWD was estimated to be 24-45 metres and therefore differences in the proportions experiencing a decrement of ≥50m and mean differences in 6MWD of ≥50m are likely to be clinically significant.

3.4.3 Exacerbation rate

Acute exacerbations are reported, however, the CS states that the outcome was defined differently across the trials and was not collected systematically in all trials. Acute exacerbation rate was included in the NMA.

In the nintedanib appraisal (TA282), the Committee concluded that exacerbations are an important clinical event, but can be difficult to define, particularly in trials.[12] In the company's clarification response (see clarification response,[10] question A15), the company states that acute exacerbations are notoriously difficult to diagnose, there is no universally agreed definition, and exacerbations meeting the strict definitions employed in trials are rare (<1% in the nintedanib trials). Clinical advisors to the ERG believed that this is because the definitions of acute exacerbations used in trials generally require other causes of respiratory symptoms, such as infection, to be ruled out, but this is a very restrictive definition as it is very hard in practice to rule out infection as a cause. However, in clinical practice, patients experience periods of acute worsening of symptoms with breathlessness that needs treatment and these are recognised by clinicians as acute exacerbation even though they may not meet the strict criteria applied in the trials.

3.4.4 Progression-free survival

Progression-free survival (PFS) is reported as per the NICE scope, however as noted in the CS, the definition of PFS varied between studies. PFS was included in the NMA, but this involved combining data from trials which used different definitions. Where possible, the data available were re-analysed to provide estimates using a consistent definition (that used in the ASCEND trial), but this was not possible for all of the trials included in the NMA. The various definitions for which data are presented are summarised in.Table 2.

Table 2 Summary of definitions used for progression-free survival

Trial	Definition specified in the final trial protocol	Other definitions for which results are provided [a]
ASCEND	confirmed ≥10% decline from baseline in %FVC, or confirmed ≥50 m decline from baseline in 6MWT distance, or death	Definition used in CAPACITY trials Definition(s) [b] used in SP3 / PANTHER
CAPACITY 1 and 2 [c]	confirmed ≥10% decline in percent predicted FVC, or ≥15% decline in percent predicted DLco or death	Definition used in ASCEND Definition(s) [b] used in SP3 / PANTHER
SP3	decline of 10% or more in VC or death	
PANTHER	decline of 10% or more in FVC or death	
INPULSIS trials	None pre-specified	Definition from CAPACITY [d]
TOMORROW	None	None

[a] in the CS or in the company response to the clarification request (clarification response,[10] question A33)
[b] SP3 used VC and PANTHER used FVC but the description of the re-analysis using this definition in Table 14 of the response to the clarification request simply states FVC/VC
[c] The definition in the protocol for the CAPACITY trials was updated by a protocol amendment so the definition in the final protocol is recorded here
[d] taken from the nintedanib company's submission for TA282[26]

3.4.5 Mortality

A number of measures are reported for mortality including all-cause mortality, IPF-related mortality and treatment-emergent IPF-related mortality. All-cause mortality and IPF-related mortality were included in the NMA but treatment-emergent IPF-related mortality was only included in the direct comparison of pirfenidone against placebo.

In the CS, treatment-emergent mortality was defined as occurring between randomisation and 28 days after the last dose of study drug. Treatment emergent IPF-related mortality is defined as a secondary efficacy outcome in the ASCEND protocol. In the protocols for the CAPACITY trials, deaths are described as a safety outcome.[27] Definitions are provided for the terms 'treatment-emergent' and 'IPF-related', but treatment-emergent IPF-related mortality is not specifically defined as an outcome.[28, 29] The clinical advisors to the ERG considered that all-cause mortality was the most important outcome for patients with IPF.

3.4.6 Adverse events

Adverse events (AE) of treatment are reported from the pirfenidone clinical trial programme in Section 4.1.12 of the CS and the AE data for nintedanib applied in the model are described in Section 5.3 of the CS but AEs are not reported systematically for the nintedanib studies. For ASCEND and CAPCATIY 1 and 2, the AEs summarised in Tables 60 and 61 of the CS, were treatment-emergent AEs with 'treatment-emergent' being defined as occurring after first dose and within 28 days after the last dose of study treatment. Additional data on AEs that led to discontinuation were provided in response to a clarification request (see clarification response,[10] question A24). Additional data on treatment-emergent serious adverse events reported in ≥2 patients were also provided in response to a clarification request (see clarification response,[10] question A25 and clarification response addendum,[30] question A28). For SP2, published AE data were presented, however additional summaries on serious AEs and AEs that led to discontinuation could not be provided by the company due to restrictions on access to data from this study.

The NMAs reported in the CS did not include AEs; however, additional NMAs were presented in the clarification response (see clarification response,[10] question A39) for the AEs of diarrhoea, rash, discontinuation of treatment due to AEs and serious cardiac AEs. The ERG considers it reasonable for additional NMAs to be presented for diarrhoea, rash and serious cardiac AEs as these data are useful for informing the indirect comparison with nintedanib within the company's model.

3.4.7 Health-related quality of life

EQ-5D data were not collected in the CAPACITY 1 and 2 or ASCEND trials (see CS, page 224).

CAPACITY 1 and 2 measured HRQoL using the St George's Respiratory Questionnaire (SGRQ); these data are reported in the CS. The change from baseline in the total SGRQ score was also included as an outcome in the NMA. A recent article examining the psychometric properties of the SGRQ in patients with IPF concluded that whilst it was not developed specifically for use in patients with IPF, and further research is needed to confirm the SGRQ's utility in IPF, at present, *"the balance of data suggests that the SGRQ may be a suitable secondary endpoint for measuring HRQoL in therapeutic trials of IPF."*[21] The MCID for the SGRQ in patients with IPF is reported to be 7 for the total SGRQ score.[22]

CAPACITY 1 and 2 and ASCEND measured HRQoL using the University of California San Diego Shortness of Breath Questionnaire (UCSD SOBQ); these data are reported in the CS. The mean change in the UCSD SOBQ was also included in the NMA. The MCID for UCSD SOBQ in patients with IPF is reported to be in the range of 5 to 11 points.[31]

WHO QOL data were also collected in the CAPACITY studies (see CS, page 224), but the results are not presented in the section reporting HRQoL outcomes (CS, pages 109 to 112).

3.4.8 Additional outcomes not specified in the scope

All-cause discontinuations and hospitalisations are reported in addition to the outcomes specified in the final NICE scope. All-cause discontinuations were included as an outcome in the NMA but hospitalisations were not.

3.4.9 Inclusion of outcomes in the indirect comparison

The majority of the outcomes were reported for both the direct comparison with placebo from the pirfenidone clinical trial programme and for the indirect comparison with nintedanib from the NMA. Outcomes addressed in the submission but not included in the NMA were DLco, treatment emergent IPF-related mortality, categorical change in 6MWD (decline of more or less than 50m), and hospitalisations.

3.5 Other relevant factors

Patient Access Schemes were agreed for pirfenidone at the time of TA282 and for nintedanib at the time of TA379. In both cases, the technologies were recommended only when the technology is provided with the discount agreed in the PAS. The company submitted a revised PAS which was accepted by the Department of Health. Further details on the PAS can be found in the confidential appendix.

No equality issues were raised in the CS.

4 CLINICAL EFFECTIVENESS

This chapter presents a summary and critique of the reviews submitted by the company on the efficacy and safety of pirfenidone in adults with mild to moderate IPF. The critique was performed following the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and checklist.[32]

4.1 Critique of the methods of review(s)

The CS[4] reports the methods and results of three separate reviews:

(i) A review of the efficacy evidence from RCTs (see CS, [4] Sections 4.1-4.10);

(ii) A review of the efficacy and safety evidence from non-randomised and non-controlled studies (see CS,[4] Section 4.11), and;

(iii) A review of safety evidence from RCTs and a non-randomised study (see CS,[4] Section 4.12).

Each review applied slightly different inclusion criteria depending on the intended analysis and the included study designs.

The main review of efficacy evidence from RCTs was a generally well-reported systematic review. Following a request for clarification from the ERG regarding certain process elements adopted by the company, the ERG considered the review to be generally sound (see clarification response,[10] questions A1-A7). The key trials were listed as ASCEND (Phase III),[33, 34] CAPACITY 1 & 2 (Phase III),[35-37] SP3 (Phase III),[38] and SP2 (Phase II).[39] All studies compared pirfenidone with placebo. The NMA included five additional relevant RCTs (further details are provided in Section 4.6).

The review of the efficacy evidence from non-randomised and non-controlled studies consisted of a single open-label, non-controlled extension study (RECAP),[40] which was designed to assess long-term safety with some efficacy outcomes listed as secondary outcomes, plus data from six registries. This review was not considered to be a systematic review because it was unclear how the evidence was identified, selected and relevant data extracted; no inclusion or exclusion criteria were provided; and a list of excluded studies or registries was not provided. Quality assessment of the RECAP study[40] was not performed by the company.

The review of the safety evidence was also not considered by the ERG to be a systematic review because it was unclear from the original submission how the included non-RCT evidence, RECAP[40], plus the addition of a new study, PIPF-002[41], were identified and selected, no detailed inclusion or exclusion criteria or details of data extraction were provided, and a list of potentially relevant excluded studies was not provided.

4.1.1 Searches

The company conducted a systematic literature review search for evidence on the comparative efficacy and safety of interventions in IPF in April 2015.

The ERG notes that the search strategy was developed using the PICOS (patient – intervention – comparator – outcome – study types) elements of the systematic review. The strategy was structured to search for the concepts:

1. Idiopathic pulmonary fibrosis AND randomised controlled trials

   OR

2. Pirfenidone

The following sources were searched:

- Cochrane Central Register of Controlled Trials (CENTRAL)
- Cochrane Database of Systematic Reviews (CDSR) Embase
- Database of Abstracts of Reviews of Effectiveness (DARE)
- Embase 1974 to 2015 November 16
- Health Technology Assessment (HTA) database
- Ovid MEDLINE® in-process and other non-indexed citations
- Ovid MEDLINE® 1946 to present
- European Respiratory Society congress abstracts
- British Thoracic Society congress abstracts
- American Thoracic Society conference abstracts
- World Association for Sarcoidosis and Other Granulomatous Disorders conference abstracts

Reference lists of identified relevant studies, papers and review articles were also hand-searched for potentially relevant additional studies that may have been missed in the database searches. Update searching or forward citation searching was not reported to have taken place.

The CS[4] reports that the following databases were omitted from the search, despite having been included in the searches for the original NICE Technology Appraisal guidance on Pirfenidone for treating idiopathic pulmonary fibrosis (2013)[42]:

- World Association for Sarcoidosis and Other Granulomatous Disorders conference abstracts
- Allied and Complementary Medicine Database (AMED)
- British Nursing Index (BNI)
- Health Management Information Consortium (HMIC)
- PsycINFO
- Journals@OVID Full text
- Cochrane Methodology Register
- NHS Economic Evaluation Database (NHS EED)
- About the Cochrane Collaboration

The CS[4] states that the reason for excluding these databases was because their focus was not considered appropriate for the objectives of this specific systematic review. The ERG agrees with the decision to omit these databases based on the specific focus of these databases and they are not amongst the minimum databases suggested by the NICE Guide to the Methods of Technology Appraisal 2013[43] or the Centre for Reviews and Dissemination (CRD) guidance.[44]

When attempting to reproduce and verify the company's searches, the ERG identified a number of potentially relevant studies that met the inclusion criteria via searching the Web of Science.[45-47]

The searches were limited to information published, added to the databases, updated or indexed from January 2011 onwards. This date limit was applied because the original InterMune NICE STA submission searches were conducted in October 2011.[27] However, the ERG noted that the approach to searching differed from the approach that was undertaken for the original submission, which searched for: IPF AND pirfenidone AND RCTs.

The searches were comprehensive and the reporting of the search strategies is clearly reported, reproducible and transparent. The ERG obtained a similar number of records when re-running the searches. The ERG did not identify any errors in the execution of the searches in relation to Boolean or database specific syntax operators and the translation of the strategy across all of the databases from Medline is consistent.

The ERG also re-ran an amended version of the search for pirfenidone on Medline and Embase changing the fields from .ti,ab,kf,rn (Medline) and .ti,ab,kw,rn (Embase) to the more sensitive .af search field. This did not however affect the results.

Additional studies that were published after the systematic searches had been conducted were included in the meta-analysis and the PRISMA chart states that 23 studies were identified via 'other sources' but does not specify the methods of retrieval of these studies. The company stated in their clarification response that the included studies that were published after the search date were obtained via 'internal analyses' (see clarification response,[10] question A2). However, the ERG would recommend update searching or forward citation searching in order to maximise the transparency of reporting and reduce the risk of confirmation bias.

The ERG found that, despite these omissions, the numbers of results retrieved by the company were in accordance with the results obtained when all terms were entered correctly and the searches were re-run by the ERG.

The CS[4] does not report whether a published search filter was utilised in order to identify RCTs of IPF. The search filter appears to be a slightly modified version of the Cochrane highly sensitive search filter. The company stated in their clarification response (see clarification response,[10] question A3) that this filter had been amended to increase the sensitivity of the search.[10, 44]

The reporting of the conference abstract searches in the European Respiratory Society (ERS) Annual Conference Abstracts contains information about how many retrieval hits were obtained and how many records were retrieved for further consideration. However, the searches conducted in the ERS Annual Congress and Conference advanced search feature only contains information about how many results were retrieved and not how many were considered.

The ERG queried the lack of searches for ongoing and unpublished clinical trials in research registers including the metaRegister of Controlled Trials, the EU Clinical Trials Register and the World Health Organization. The company agreed that this was an oversight and searched ClinicalTrials.gov, ICTRP AND PharmNet Bund and provided the results in Appendix A of their clarification response.[10]

The searches conducted by the ERG in research registers identified an additional, potentially relevant RCT published in 2015, which compared pirfenidone plus N-acetylcysteine (NAC) with placebo and NAC in adult patients with mild or moderate IPF (percent predicted FVC at baseline was 75.55±14.72 in the pirfenidone and NAC group and 79.07±18.25 in the NAC and placebo group) (Huang *et al.*

2015[48]). Details and results from this trial have therefore been reported as supporting evidence by the ERG. This was also identified in the additional references provided by the company after conducting searches of research registers as part of the company's response to the ERG's clarification questions.

The ERG's view is that it is likely that all relevant RCTs will have been identified from the searches described in the CS[4] and the company response to the clarification request.[10] The ERG obtained a similar number of records when re-running the searches. No search strategies were reported for AEs; however, the ERG believes that searching for pirfenidone as a standalone concept maximises the sensitivity of this search and would be likely to capture any potentially relevant information in relation to AEs.

4.1.2 Inclusion criteria

The inclusion criteria for the review of pirfenidone RCTs are described in Section 4.1 of the CS[4] (Table 8, page 53) and reproduced in Table 3. These criteria describe RCTs measuring the efficacy and safety of pirfenidone compared with nintedanib or BSC (placebo) in adult patients with mild or moderate IPF. The five RCTs satisfying these criteria are: ASCEND (Phase III),[34] CAPACITY 1 & 2 (Phase III),[49] SP3 (Phase III),[38] and SP2 (Phase II).[39] All of these trials compared pirfenidone with placebo. These RCTs included four different doses of pirfendione: 2,403mg per day, 1,197mg per day, 1,800mg per day and 1,200mg per day. The NMA to evaluate efficacy applied different criteria (see CS,[4] Table 38, page 123) and is covered in detail in Section 4.6 of this report.

The review of the efficacy evidence from non-randomised and non-controlled studies did not specify any inclusion criteria (see CS,[4] Section 4.1). This review reported a single open-label, non-controlled extension study, RECAP,[40] whose participants were recruited from the ASCEND[34] and CAPACITY trials.[49] Further evidence was reported from the Edinburgh registry, INOVA registry and the EuroIPF registry, as well as three additional, "supportive" registries: CPRD, Strand *et al*[41] and Kondoh *et al*.[50] According to the inclusion criteria outlined in Section 4.1 of the CS,[4] non-randomised studies were explicitly excluded. The search conducted for the clinical efficacy review would have enabled the identification of the RECAP[40] non-RCT, but it is unclear whether additional, relevant evidence might have been excluded.

The inclusion criteria for the review of safety evidence from RCTs and non-randomised studies were not specified. The safety review included the five pirfenidone RCTs from the main clinical efficacy review, as well as the non-randomised studies RECAP[40] and an additional non-randomised, non-controlled study, PIPF-02.[41] However, as noted above, the methods by which these non-randomised studies were identified and the criteria by which they were selected, and others were excluded, are not reported.

Table 3: Inclusion and exclusion criteria for the pirfenidone RCT direct comparison clinical efficacy systematic review (reproduced from CS, Section 4.1, Table 8, page 53,)[4]

	Inclusion criteria	Exclusion criteria
Population	Adults (aged 18 or older) with suspected or diagnosed IPF	Studies of children and young people <18 years Studies of people with a diagnosis of pulmonary fibrosis as a complication of either of the following: • Connective tissue disorders • A known exogenous agent (for example, drug induced disease or asbestosis)
Intervention	Pirfenidone	Any studies not containing pirfenidone
Comparators	Any comparator: • Best supportive care* (placebo) • Nintedanib	N/A
Outcomes	**Pulmonary function parameters** • Lung capacity (VC/FVC) • Categorical declines in FVC • Gas transfer (carbon monoxide diffusing capacity [DLco]) **Physical function** • Physical functioning (6MWD) **Exacerbation rate** • Hospitalisations • Acute exacerbations **Progression-free survival** **Mortality** • All-cause mortality • IPF-related mortality **AEs of treatment** **HRQoL** • St George's Respiratory Questionnaire (SGRQ) • University of California, San Diego (UCSD) Shortness of Breath Questionnaire (SOBQ) • EuroQoL five dimensions questionnaire (EQ-5D)	• Anticoagulation for the treatment of pulmonary hypertension • Treatment of lung cancer • Lung transplantation other than timing and referral
Study design	• Studies in humans • Phase II or III RCTs • Studies published as abstracts, conference presentations or press releases were eligible if adequate data were provided • Systematic reviews of RCTs**	• Cross-over RCTs
Language	No language limits	No language limits

*Best supportive care is defined as information and support, symptom relief, management of comorbidities, withdrawal of therapies suspected to be ineffective or causing harm, end of life care, oxygen therapy and/or pulmonary rehabilitation

**Systematic reviews were eligible for inclusion as a source of references to primary studies

4.1.3 Critique of study selection and data extraction

Following an ERG request for the company to clarify the processes undertaken, the ERG was satisfied that standard systematic review good practice was followed in study selection: relevant papers were independently selected for inclusion at title, abstract and full text stage by two reviewers, with any discrepancies between reviewers resolved through discussion or the intervention of a third reviewer (see clarification response,[10] question A4). In a first screen, "obviously irrelevant" studies were excluded by a single information specialist (see clarification response,[10] question A4).

No information was given in any of the reviews regarding the data extraction process (for example, the number of reviewers involved, or actions taken to minimise error). This was addressed in response to clarification requests from the ERG, in which the company detailed standard processes for data extraction in systematic review (see clarification response,[10] question A5). Data extraction was performed by one reviewer and independently checked for errors against the original trial report by a second reviewer. Any discrepancies were resolved through discussion or through the intervention of a third reviewer.

During the clarification stage, discrepancies and inadequacies in some of the numbers reported in the PRISMA flowchart were acknowledged and addressed by the company, and an updated PRISMA flowchart was provided (see clarification response,[10] question A6).

4.1.4 Quality assessment

For the review of clinical efficacy evidence, the company conducted a critical appraisal of the five pirfenidone trials using a version of the Cochrane risk of bias assessment tool (see CS,[4] Section 4.6 and Appendix 6). The process was conducted according to standard systematic review practice, by two reviewers working independently, with any discrepancies resolved by discussion or reference to a third reviewer (see CS,[4] Appendix 6). The CS concluded that all five trials were at "low risk of bias" across the domains assessed, although the adequacy of randomisation and blinding was assessed as "unclear" for the SP3 trial.[38]

The ERG accepts these assessments for the ASCEND[33, 34] and CAPACITY trials,[35, 37, 51] for the domains of selection bias (randomisation, allocation concealment); performance and detection bias (blinding); and attrition bias (drop-out, ITT analysis and management of missing data). However, the ERG disagrees with assessments regarding reporting bias and other types of bias, especially given the absence of adequate information concerning some analyses and some secondary outcomes in both the publicly-available protocols for each trial from the clinical trials register (https://clinicaltrials.gov/ct2/home) and those protocols made available alongside the final publications or provided by the company in response to requests by the ERG (see clarification

responses[10], questions A8-A10). For example, the SGRQ outcome measure that is reported in the CS[4] is absent from all forms of protocol, as well as the actual CAPACITY trials publication[35, 37, 51] (although this outcome is listed in the CSR).

The effect of the "intrinsic variability in rates of FVC decline" acknowledged in the CAPACITY trials' publication,[49] and the company's clarification response (question A26:[10] "the natural variability in rates of FVC percent predicted decline of this heterogeneous disease"), which might explain differences in outcomes both within and across trials, must also be taken into account as a potential moderator influencing results.

Overall, however, the ERG assessed the potential risk of bias in ASCEND[33, 34] and CAPACITY 1 & 2[35, 37, 51] to be low or low-to-moderate. The details of the ERG assessment are provided in Table 4.

The SP3[38] and SP2[39] trials, by contrast, are at a higher or more unclear risk of bias across many domains compared with the ASCEND[33, 34] and CAPACITY trials,[35, 37, 51] principally because of the inadequacy of the information contained within the published manuscripts and the protocols provided by the company in response to a request by the ERG (see clarification response,[10] question A12). These issues particularly affect selection, detection and attrition bias; the last named on account of the smaller sample sizes, the rates of attrition and the application of the last observation carried forward (LOCF) method to impute missing data, which might potentially overestimate treatment effect in a progressive disease such as IPF[51] (see Table 4).

Finally, the supporting trial reported by Huang *et al.,*[48] was generally found to be at moderate risk of bias across most domains as a result of the lack of detail within the available protocol and the publication.

Table 4: ERG risk of bias assessment (Cochrane tool): ASCEND, CAPACITY 1 & 2, SP3 and SP2

Risk of bias	ASCEND[33, 34]	CAPACITY 1[36, 49]	CAPACITY 2[33, 49]	SP3[38]	SP2[39]	Huang 2015[48]
Selection bias	LOW Randomisation codes were generated by computer with the use of a permuted-block design, and the study drug was assigned by means of an interactive voice-response system. Protocol: 3, page 38: Patients will be randomised at the Day 1 Visit (see Section 4.3.4.1) in a 1:1 ratio to receive either pirfenidone 2403 mg per day or placebo equivalent using an automated system. All randomisation codes will be generated by a statistician independent of the trial conduct.	LOW The randomisation code (permuted block design with five patients per block in study 004 and four per block in study 006) was computer generated, stratified by region, by an independent statistician. Study centres, using an interactive voice response system, assigned study drug bottles to patients. The independent statistician had no role other than assignation of the randomisation code and study drug bottle numbers. All personnel involved in the study were masked to treatment group assignment until after final database lock.		UNCLEAR A multicentre, double-blind, placebo-controlled, randomised phase III clinical trial, page.821 etc.; Eligible patients were allocated to three groups: high dose (1,800 mg/day), low dose (1,200 mg/day) and placebo, in a ratio of 2:1:2, respectively, with a modified minimisation method, including some random allocation based on biased coin design to balance baseline SpO_2, page 822): insufficient information and he protocol does not provide any specific information[52]	MODERATE Patients were randomly assigned into pirfenidone or placebo (2:1) groups using a modified permuted-block randomisation method with block sizes of six.(page 1042), but it is not stated who does this. Investigators? Independent body? The protocol does not provide any specific information [53]	MODERATE "Patients were randomly assigned into pirfenidone or placebo (1:1) groups using a modified permuted-block randomisation method with block sizes of 4", but it is not stated who does this. Investigators? Independent body? The protocol does not provide any specific information (https://clinicaltrials.gov/ct2/show/record/NCT01504334)
Performance bias	LOW Publication main text provides no information; page 31 NEJM protocol: Patients will receive blinded study treatment from the time of randomization until the week 52 Visit. Page 37: There will be 270 capsules per bottle, which will be labeled for investigational use only. Pirfenidone 267-mg and placebo will be supplied in opaque, hard, white gelatin capsules that are visually indistinguishable. 3, page 38: Pirfenidone and placebo will both be supplied in capsules that are visually indistinguishable. Pirfenidone and placebo packaging and labeling will be identical. There was no evaluation of blinding.	LOW All personnel involved in the study were masked to treatment group assignment until after final database lock. Available information from publication and protocols is too limited to give this a "low risk of bias" assessment, but sufficient information was given in the CSR. There was no evaluation of blinding.		LOW Allocation and blinding covered in detail in the protocol, sections 14.3.1, 14.3.2, 14.3.3 and 14.3.5 [52]	LOW Allocation and blinding covered in detail in the protocol, sections 6.3.1, 6.3.2 and 6.3.4 [53]	MODERATE Matching placebo tablets, but no other details of blinding and no evaluation of blinding
Detection bias	LOW Central reviewers at Biomedical Systems, who were unaware of study-group assignments, evaluated all FVC results for adequacy and repeatability, according to the criteria of the American	LOW Mortality was pre-specified as an exploratory endpoint, and death related to idiopathic pulmonary fibrosis was assigned by investigators masked to assignment. ASCEND publication: "The primary cause of death and its relation to idiopathic pulmonary		LOW / MODERATE Protocol states that: "14.3.5 Blindedness will be maintained with respect to all study personnel except the study drug allocation manager" [52], but publication acknowledges limitation of ,"The lack of a central pathology review" (page 824); plus	UNCLEAR Protocol indicates that outcome assessors were unblinded: "6.3.4 Blindedness will be maintained with respect to all study personnel except the study drug allocation manager and the efficacy and safety	MODERATE Protocol states that outcome assessors were blinded, but there are no details

Risk of bias	ASCEND[33, 34]	CAPACITY 1[36, 49]	CAPACITY 2[33, 49]	SP3[38]	SP2[39]	Huang 2015[48]
	Thoracic Society The primary cause of death and its relation to idiopathic pulmonary fibrosis were assessed in a blinded fashion by an independent mortality assessment committee in the ASCEND trial and by the site investigators in the CAPACITY trials[33, 36, 49]	fibrosis were assessed in a blinded fashion … by the site investigators in the CAPACITY trials" [33, 36, 49]		problems with un-validated measure of lowest SpO2 during the 6MET	evaluation committee."[53]	
Attrition bias	LOW / MODERATE 522 patients (94.1%) completed the study: 261 patients (93.9%) in the pirfenidone group and 261 patients (94.2%) in the placebo group. Study treatment was discontinued prematurely in 55 patients (19.8%) in the pirfenidone group and in 39 patients (14.1%) in the placebo group. Adherence to the study treatment was high; 237 patients (85.3%) and 256 (92.4%) patients in the pirfenidone and placebo groups, respectively, received at least 80% of the prescribed doses of the assigned study drug.	LOW / MODERATE 409 (94%) of 435 patients in CAPACITY 2 and 322 (94%) of 344 in CAPACITY 1 completed the study. 109 patients (14%) discontinued treatment prematurely: 13 (15%), 30 (17%), and 18 (10%) in the pirfenidone 1197 mg/day, pirfenidone 2403 mg/day, and placebo groups, respectively in CAPACITY 2; and 31 (18%) and 17 (10%) in the pirfenidone and placebo groups, respectively, in CAPACITY 1		MODERATE 30%+ rate of attrition and LOCF used to impute missing data (for a progression disease, this might overestimate treatment effect) if patient data were available for 4 weeks after the baseline (page 823)	MODERATE 20%+ rate of attrition and LOCF used to impute missing data (for a progression disease, this might overestimate treatment effect) For missing values, the principle of last observation carry forward was adopted (page 1042)	UNCLEAR Up to 16% attrition, but it is not clear from the publication or protocol how missing data were managed
Reporting bias	MODERATE Two primary, five secondary outcomes – only the basic primary outcome listed in NCT protocol; others in NEJM protocol, but SGRQ not in any protocol; plus acute exacerbations / hospitalisations – are recorded at Follow-Ups, but not specified as outcomes. Only pre-specified analyses listed in protocols relate to mortality.	MODERATE All protocol outcomes listed in primary publication ,[49] but SGRQ was not in any protocol and was not reported in the primary publication, but was only mentioned in CSR; plus acute exacerbations / hospitalisations are only reported as part of the "Worsening of IPF" composite outcome		LOW All of the outcomes reported in the protocol (Shinogi 2006[52]) were reported in the publication	LOW All of the outcomes reported in the protocol [53] were reported in the publication	LOW All of the outcomes reported in the protocol (https://clinicaltrials.gov/ct2/show/record/NCT01504334) were reported in the publication
Other bias	UNCLEAR "Intrinsic variability in rates of FVC decline" acknowledged as potential moderator of results, and possible explanation for differences across trials in certain outcomes (Noble 2011, pages 1767 and 1768) [49] . This represents a potential uncontrolled moderator of outcomes. Claim that this is controlled for by FVC and DLco eligibility criteria (more severe and progressive population) is questionable.	UNCLEAR "Intrinsic variability in rates of FVC decline" acknowledged as potential moderator of results, and possible explanation for differences across trials in certain outcomes (Noble 2011, pages 1767 and 1768) [49] . This represents a potential uncontrolled moderator of outcomes		MODERATE Acknowledged issue: (page 824 Taniguchi):[38] A selection bias, as patients enrolled in this study needed to be able to perform the 6MET at baseline in accordance with the protocol; the results in this selected group of patients with mild functional impairment may not therefore be applicable to all patients with IPF with varying degrees of pulmonary	UNCLEAR Trial discontinued early due to excessive rates of exacerbations in the placebo arm, so outcomes etc. were not measured at all planned time-points, only at 6 and 9 months Substantial links to study sponsor.	MODERATE Difficulty controlling for natural variability in IPF disease and speed of progression: *post hoc* analyses excluding patients with most substantial decline, produced different findings Some links to industry

Risk of bias	ASCEND[33, 34]	CAPACITY 1[36, 49]	CAPACITY 2[33, 49]	SP3[38]	SP2[39]	Huang 2015[48]
	Composite outcomes do not appear under Outcomes in protocol – first appearance is under efficacy analyses, page 60, 5.4.2 Efficacy Analyses: 5.4.2.1 Primary Efficacy Outcome Variable and Analysis in protocol analysis plan and represents a modification from the CAPACITY trials[33, 36, 49] – it does not appear as an outcome in the protocol or publication			symptoms and functional impairment.	Per protocol drop-outs based on the outcome measure (page 1042)	

4.1.5 Evidence synthesis

The synthesis for the review of clinical efficacy was a basic descriptive summary of the evidence from the five pirfenidone trials for the following outcomes: change from baseline in percent predicted FVC; all-cause and IPF-related mortality; PFS; acute exacerbation; hospitalisation; changes from baseline in 6MWD, the UCSD SOBQ and the SGRQ. This approach to evidence synthesis was neither described nor justified in the CS.[4]

Meta-analyses using both fixed and random effects models comparing pirfenidone with placebo were performed for selected outcomes and time-points, based on available trial data, and the methods used were described in the CS[4] (Section 4.9 and Appendix 9). Data were combined from CAPACITY 1 & 2[33, 36, 49] and ASCEND[33, 34] using the UK licence dosage (2,403mg/day) and from SP3[38] which uses an unlicensed dosage (1,800 mg/day). The company considered this to be appropriate as the dose by weight would be similar for all studies given the lower body weight of the Japanese population compared with the North American and European population. An NMA comparing effects across all treatments was also performed by the company. This is critiqued in Sections 4.6 and 4.7 of this report.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Review of clinical efficacy (relevant pirfenidone RCT evidence)

The CS[4] provides a detailed description of trials identified by the company as satisfying the requirements of the final NICE scope,[3] i.e. pirfenidone compared with placebo or nintedanib (see Table 5). No trial compared pirfenidone with nintedanib. Five RCTs compared pirfenidone at various doses with placebo: ASCEND (Phase III),[34] CAPACITY 1 & CAPACITY 2 (Phase III),[49] SP3 (Phase III),[38] and SP2 (Phase II).[39] Three trials were international and multicentre (ASCEND[34] and CAPACITY 1 & 2[33, 36, 49]), although only CAPACITY 2 included any UK centres[35] (three of 110 centres across both CAPACITY trials).[49] The inclusion criteria in all three trials were adult patients with mild or moderate IPF based on percentage predicted FVC of ≥50% (in ASCEND[34] this had an upper limit of ≤90%). Two trials were conducted exclusively in Japan (SP3[38] and SP2[39]) and did not report baseline levels of FVC or VC. One trial was conducted in China and evaluated pirfenidone in combination with N-acetylcysteine (NAC). The trials varied in criteria relating to lung function, concomitant medications permitted for IPF, and the investigated doses of pirfenidone (for the purposes of this appraisal, ASCEND,[34] and CAPACITY 1 & 2[49] all evaluated the efficacy of the licensed dose of 2,403mg/d; the SP2,[39] SP3[38] and the Huang *et al.* [48] trial evaluated lower doses; the applicability of these lower doses to clinical practice in England and Wales is unclear.

Table 5: Characteristics of included pirfenidone RCTs (reproduced in part from CS,[4] Tables 10 and 15, pages 59 and 82)

Trial No. of patients	Location	Inclusion criteria		Exclusion criteria	Intervention and co-interventions (No. of patients)	Comparator (No. of patients)	Follow-up
		IPF diagnosis	Lung function parameters	Patient factors			
ASCEND (PIPF-016)[33, 34] n=555	International multi-centre	– Confident clinical and radiographic diagnosis of IPF, confirmed centrally with diagnosis of IPF >6 months but <48 months. – No improvement of IPF in preceding year.	– FVC (% predicted value) 50-90% – DLco 30-90% – 6MWT ≥150 m	– Abnormal lab parameters – Obstructive airway disease – History of unstable /deteriorating cardiac or pulmonary disease – History of severe hepatic impairment/ end-stage liver disease/end-stage renal disease requiring dialysis	Pirfenidone 2,403mg/day (n=277) Concomitant treatment with any investigational drug for the treatment of IPF was prohibited. However, concomitant medications used in the treatment of IPF were permitted if given for a non-IPF indication and there was no clinically acceptable alternative.	Placebo (n=278)	52 weeks
CAPACITY 1 (PIPF-006)[36, 49] n=344	International multi-centre		– FVC (% predicted value) ≥ 50% – DLco ≥35% – FVC or DLco ≤90% – 6MWT ≥150 m		Pirfenidone 2,403mg/day (n=171) Concomitant treatments for IPF were prohibited, with exceptions of short courses of azathioprine, cyclophosphamide, corticosteroids, or acetylcysteine for protocol-defined acute exacerbation of IPF, acute respiratory decompensation, or progression of disease.	Placebo (n=173)	72 weeks
CAPACITY 2 (PIPF-004)[35, 49] n=435	International multi-centre	– Confident clinical and radiographic diagnosis of IPF, confirmed locally (diagnosis previous 48 months) – No improvement of IPF in preceding year	– FVC (% predicted value) ≥50% – DLco ≥35% – FVC or DLco ≤90%		Pirfenidone 2,403mg/day (n=174) Pirfenidone 1,197mg/day (n=87) As CAPACITY 1	Placebo (n=174)	72 weeks

IPF: Idiopathic Pulmonary Fibrosis; FVC: Forced Vital Capacity; DLco: Diffusing capacity of the lungs for carbon monoxide; 6MWT: 6-minute walking test

SP2[39] n=107	Japan, multi-centre	Confident clinical and radiographic diagnosis of IPF (as per guideline consensus)	– Adequate oxygenation at rest (PaO2 70 mm Hg) and SpO2 ≤ 90% during exertion	– Coexisting pulmonary hypertension, asthma, tuberculosis, sarcoid, bronchiectasis or respiratory infection; – Comorbid conditions including malignancy, severe hepatic, renal, Diabetes Mellitus or cardiac disease	Pirfenidone 1800mg/day (n=72) Concomitant prednisone ≤10mg/day was allowed. The following immunosuppressants or other anti-inflammatory/ antifibrotic drugs were not allowed: cyclophosphamide, azathioprine, methotrexate, d-penicallimine, cochicine, erythromycin, IFNs, N-acetylcysteine, cyclosporine, tacrolimus and other investigational drugs for IPF.	Placebo (n=35)	36 weeks (trial terminated early due to adverse events)
SP3[38] n=275*	Japan, multi-centre	Confident clinical and radiographic diagnosis of IPF (as per ATS/ERS guideline consensus) No decrease in symptoms during the preceding 6 months	– O2 desaturation of 5% between resting SpO2 and min SpO2 during 6 min exercise test (6MET) – SpO2 >85% during 6MET (air).	– Coexisting pulmonary hypertension, asthma, tuberculosis, sarcoid, bronchiectasis or respiratory infection; – Comorbid conditions including malignancy, severe hepatic, renal, Diabetes Mellitus or cardiac disease	Pirfenidone 1800mg/day (n=108) Pirfenidone 1200mg/day (n=55) Concomitant corticosteroid ≤10mg/day (as the prednisone equivalent) was allowed. However, concomitant immunosuppressants or other investigational drugs for IPF were not allowed.	Placebo (n=104)	52 weeks

Huang 2015[48] n=76	China, multi-centre	The diagnosis of IPF was in accordance with evidence-based guidelines for the diagnosis and management of IPF published in 2011.	– percentage of predicted forced vital capacity (FVC) of at least 45%, – percentage of predicted carbon monoxide diffusing capacity – (DLCO) of at least 30%, and PaO2 of at least 50mmHg when – the patient is at rest and breathing room air	– aggravated dyspnea during the preceding 6 months; – currently in a period of acute exacerbation of IPF (AEIPF); – fasting blood glucose level of more than 11.1 mmol/L – comorbid conditions including malignancy, bleeding tendency, severe hepatic dysfunction or renal or cardiac disease; – use of immune-suppressants, antifibrotic drugs	Pirfenidone 1800mg/day (n=38) All patients were treated with 600 mg of N-acetylcysteine (NAC) 3 times daily as a baseline treatment.	Placebo (n=38)	48 weeks

**8 patients were excluded after randomisation for being ineligible; IPF: Idiopathic Pulmonary Fibrosis; FVC: Forced Vital Capacity; DLco: Diffusing capacity of the lungs for carbon monoxide; 6MWT: 6-minute walking test; ATS/ERS: American Thoracic Society/European Respiratory Society; PaO2: Partial pressure arterial oxygen; SpO_2: Blood oxygen saturation level; 6MET: 6-minute walking test*

The exclusion of certain patients otherwise eligible for pirfenidone, based on co-morbidities, such as obstructive airways disease, must also be taken into account when judging the generalisability of the trials' findings.

The outcomes reported in the CS[4] are generally consistent with those that are listed in the final NICE scope.[3] The ASCEND,[34] CAPACITY[49] and Huang *et al*[48] trials use change from baseline in percent predicted FVC as an endpoint, while SP3[38] and SP2[39] use VC. The CS states that the decision to use VC in the SP3[38] and SP2[39] trials was dictated by the ATS international consensus statement published in 2000, which recommended measurement of VC.[54] The CS[4] did not state when the recommended measurement changed to FVC or provide any reference to substantiate the change. The CS[4] states that VC and FVC should be treated as comparable endpoints as there is little difference between VC and FVC in subjects without obstructive pathology. Whilst the clinical advisors to the ERG agreed with this statement, the ERG noted that the exclusion criteria for SP3[38] were not as explicit regarding the exclusion of patients with emphysema as the exclusion criteria for the other pirfenidone trials. Therefore, the ERG considers that the synthesis of VC data from SP3[38] with FVC data from the ASCEND[34] and CAPACITY trials[49] is questionable.

The outcomes listed in the trial protocols publicly-available from the clinical trials register (https://clinicaltrials.gov/ct2/home) are not entirely consistent with those reported in the CS.[4] For example, the principal efficacy outcome of "percent predicted FVC and death" does not appear in any protocol but appears to be a *post hoc* composite efficacy outcome in the CS[4] (see Table 6), which according to the company was used in order to impute a FVC measurement for patients who have died (see clarification response[10], questions, A11 and A13). Furthermore, neither of the secondary outcomes of "treatment-emergent IPF-related mortality" nor the SGRQ was listed in the protocols, but both appear *post hoc* as outcomes in the CS[4] (as well as in the ASCEND[34] publication, but not in the CAPACITY trials' publication,[49] see Table 7).

The following outcome was listed in protocols but was not reported in the results for the CAPACITY 1 & 2[49] and SP3[38] trials: Change in Worst Oxygen Saturation by Pulse Oximetry (SpO_2) measurement observed during the 6-Minute Walk Test. The CAPACITY trial protocols[35, 37, 51] also listed lung transplantation as a secondary outcome, but this is not included as an outcome in the CS[4] (pages 53 and 66). The CS[4] lists fibrosis by use of high resolution computed tomography (HRCT) (see CS, Table 12, page 68) as an outcome, but this only appears to be used as a diagnostic criterion for IPF or as part of the definitions of acute exacerbations (see CS,[4] pages 104-105).

Definitions of outcomes are first provided under the trial results section of the CS[4] (Section 4.7, pages 90-113). The outcomes, and the definitions applied in each of the trials, taken from the CS and the

original protocols and publications, are summarised in Table 6 and Table 7. The Huang *et al.,* trial[48] has been omitted from these tables because it is being used as supporting evidence only.

Table 6: Primary efficacy outcomes and measures in ASCEND, CAPACITY 1, CAPACITY 2, SP3 and SP2

Outcome	ASCEND[33, 34]	CAPACITY 1[36, 49]	CAPACITY 2[35, 49]	SP3[38]	SP2[39]
Protocol-listed outcome	Change in percent predicted FVC from baseline to week 52†	Mean and absolute change in percent predicted FVC from baseline to week 72		No protocols available	
Reported outcomes	Change in percent predicted FVC **and death** from baseline to week 52	Change in percent predicted FVC **and death** from baseline to week 52		Change in VC from baseline to week 52	Change in the lowest SpO_2 during 6MWT. Full definition given in Azuma, page 1041 Change in VC from baseline was listed as a secondary outcome
	Categorical decline of ≥10% in percent predicted FVC	Categorical decline of ≥10% in percent predicted FVC. This was listed as a secondary outcome in the protocols and publication, defined as "Categorical Assessment of Absolute Change in Percent Predicted Forced Vital Capacity (FVC) based on the change in baseline percent predicted FVC at week 72, patients were assigned to 1 of 5 categories: mild decline (<10% but >=0% decline), moderate decline (<20% but >=10% decline), severe decline (>=20% decline), mild improvement (>0% but <10% improvement), or moderate improvement (>=10% improvement). Those who died or had a lung transplant before week 72 were included in the severe decline category. The results indicate the number of patients who experienced a Categorical Change in Percent Predicted Forced Vital Capacity" [35, 36, 49]			
Magnitude of treatment effect	The magnitude of the treatment effect was estimated by comparing the distribution of patients in the pirfenidone group with those in the placebo group across two thresholds of change at week 52: an absolute decline of 10 percentage points in the percentage of the predicted **FVC or death**, or no decline in the percentage of the predicted FVC (King 2014, page 2085)[34]	Estimated by use of differences in treatment group means and categorical change in FVC (page 1763, Noble 2011)[49]			

† This outcome was not reported in the ASCEND publication; the data were only made available by Roche in the CS,[4] Table 20 and pages 93-94.

Table 7: Secondary efficacy outcomes and measures in ASCEND, CAPACITY 1, CAPACITY 2, SP3 and SP2

Outcome	ASCEND[33, 34]	CAPACITY 1[36, 49]	CAPACITY 2[35, 49]	SP3[38]	SP2[39]
All-cause mortality	Yes				
IPF-related death	Yes	Yes*		No	
Treatment-emergent IPF mortality	Yes. Defined as death occurring after randomisation and within 28 days of the last dose of the study drug (CS, page 96). Listed only in the ASCEND NEJM protocol but reported for all mortality outcomes in ASCEND publication and separately, applied and not-applied, to all-cause and IPF-related mortality in the CAPACITY publication: appears to be a *post hoc* outcome measure.				
Progression-free Survival (PFS)	Defined in the CS (page 99)[4] as a confirmed ≥10% decline from baseline in %FVC, confirmed ≥50 m decline from baseline in 6MWD, or death	PFS is defined as the first occurrence of a 10% absolute decline from baseline in percent predicted Forced Vital Capacity, a 15% absolute decline from baseline in percent predicted hemoglobin(Hgb)-corrected carbon monoxide diffusing capacity (DLco), or death		Defined as VC decline of ≥10% or death. When the VC data could not be obtained due to worsening of respiratory symptoms, including acute exacerbation, the case was also classified as disease progression. (Taniguchi, page 822)[38]	No
Acute Exacerbations	Identified via a *post hoc* analysis of adverse events based on the MedDRA lower level term "acute exacerbation of IPF".(CS, page 104)	Definition not provided in protocols or publication (where it is reported only as part of a composite measure*). CS (page 104)[4] defines this outcome as requiring all of the following within a 4-week interval: Worsening of PaO_2 (≥8 mm Hg drop from the most recent value); clinically significant worsening of dyspnoea; new, superimposed ground-glass opacities on HRCT in one or more lobes; all other cardiac, thromboembolic, aspiration, infectious processes ruled out		†Definition not provided in protocols. CS (page 104)[4] defines this outcome as requiring all of the following within a month: increase in dyspnoea; new, ground-glass opacities on HRCT in addition to previous honeycomb lesion; all oxygen partial pressure in resting arterial blood (PaO2) is lower by more than 10 Torr than previous one; exclusion of obvious causes, such as infection, pneumothorax, cancer, pulmonary embolism or congestive heart failure; the serum levels of CRP, LDH are usually elevated as well as serum markers of interstitial pneumonias, such as KL-6, Sp-A or Sp-D	†Definition not provided in protocols. CS (page 104)[4] defines this outcome as requiring all of the following: worsening, otherwise unexplained clinical features within 1 month; progression of dyspnoea over a few days to less than 5 weeks; new radiographic/HRCT parenchymal abnormalities without pneumothorax or pleural effusion (e.g., new, superimposed ground-glass opacities); a decrease in the PaO_2 by 10 mm Hg or more; exclusion of apparent infection based on absence of Aspergillus and pneumococcus antibodies in blood, urine for Legionella pneumophila, and sputum cultures

Outcome	**ASCEND**[33, 34]	**CAPACITY 1**[36, 49]	**CAPACITY 2**[35, 49]	**SP3**[38]	**SP2**[39]
Hospitalisations	No	Non-respiratory and *respiratory hospitalisations. Only the latter was listed in the protocols.		No	Respiratory hospitalisations
6MWD (6-Minute Walking Distance Test)	Defined as the change from Baseline to week 52 in distance walked during the 6-Minute Walk Test as measured in metres (m).	Defined as the change from baseline to week 72 in distance walked during the 6-Minute Walk Test as measured in meters (m).		The change in the lowest SpO_2 during the 6MET (the original primary endpoint, which was altered after the study started but before un-blinding, Taniguchi, page 822)	No
FVC/VC	No	No		No	Yes
SGRQ (St. George's Respiratory Questionnaire)	No	Yes. Not listed in protocols and not reported in the primary publication: a *post hoc* outcome measure.		No	
Dyspnoea using UCSD SOBQ)	The SOBQ is used to assess shortness of breath with various activities of daily living (for example, brushing ones teeth or mowing the lawn). Patients rated the severity of their shortness of breath experienced on an average day during the past week on a 6 point scale (0 to 5),with 0= not at all breathless, 4= severely breathless and 5= Maximally or unable to do because of breathlessness			No	
Gas transfer (DLco)	Excluded from this trial, see Clarification response,[10] question A9	The change from baseline in Percent Predicted Hemoglobin (Hb)-Corrected Carbon Monoxide Diffusing Capacity (DLco) of the Lungs.			

**Listed under the Worsening of IPF outcome in the CAPACITY 1 and 2 protocols; †Tertiary outcomes: PFS and change in the lowest SpO_2 during the 6MET were the designated secondary outcomes*

4.2.2 Results

Participants' baseline characteristics

More than 620 participants received the licensed 2,403mg/day dose during the three international RCTs compared with more than 620 control patients who received placebo in these trials. Another 322 participants received lower doses of pirfenidone in the CAPACITY 2,[49] SP2[39] and SP3[38] trials.

The final selection of three trials (ASCEND,[34] CAPACITY 1[36, 49] and CAPACITY 2[35, 49]) for the main clinical efficacy review was considered to be appropriate by the ERG. However, there are some between-trial differences across some baseline characteristics (see Table 8). The ASCEND trial[34] participants had a lower mean percentage predicted FVC (range across arms of 67.8-68.6) than the CAPACITY trials[49] (range across arms of 73.1-76.4) and lower pre-enrollment corticosteroid use (range across arms of 0.7%-2.2%) than the CAPACITY trials[49] (range across arms of 5.2%-12.9%). CAPACITY 1[49] participants had a lower mean 6MWD (range across arms of 378.0-399.1) than in ASCEND[34] and CAPACITY 2[35] (range across arms of 410.0-420.7), and there was a relatively lower proportion of patients in CAPACITY 2[35] requiring supplemental oxygen use (range across arms 14.0%-17.0%) than in ASCEND[34] and CAPACITY 1[49] (range across arms of 27.4%-28.1%). All of these variables, with the exception of corticosteroid use, are accepted potential treatment effect modifiers and therefore were the subject of subgroup analyses in the CS,[4] (Section 4.8, pages 114-117).

The ERG considers the relevance of the smaller SP3[38] and SP2[39] trials, which were conducted exclusively in Japan, to be more questionable. These trials evaluate lower, unlicensed doses of pirfenidone, apply different eligibility criteria and present noticeable differences from the other three trials in some baseline characteristics of participants (see Table 9), for example, higher proportions of male participants (range across arms of 78%-94% for SP2[39] and SP3[38] compared with 68%-80% for ASCEND[33, 34] and CAPACITY 1 and 2[49]) and smokers (60%-86% compared with 58%-66%); higher mean percentages of predicted DLco compared with ASCEND[34] and the CAPACITY trials[49] (52.1-57.7 compared with 43.7-47.8), lower trial corticosteroid use (SP3[38] only, 4.8-10.9 compared with 21.0-36.5 in the ASCEND[34] and CAPACITY trials[49]), and smaller proportions having received surgical lung biopsies (21.0%-29.1% compared with 28.5%-55%, see Table 8).

Baseline data from participants on patient-reported outcome measures, such as the SGRQ and UCSD SOBQ, were not reported in the CS.[4]

The Huang *et al.* trial[48] comparing pirfenidone plus NAC with placebo plus NAC reported comparability between arms across all baseline characteristics except for smoking status.[48]

Table 8: Characteristics of participants in ASCEND and CAPACITY 1 & 2 (reproduced from CS,[4] Table 16, pages 84-85)

Baseline characteristic	ASCEND[33, 34]		CAPACITY 2[35, 49]			CAPACITY 1[36, 49]	
	PFN (n=278)	PBO (n=277)	PFN (n=174)	PFN (1,197mg/d) (n=87)	PBO (n=174)	PFN 2,403mg/day (n=171)	PBO (n=173)
Age, mean years ± SD	68.4 ± 6.7	67.8 ± 7.3	65.7 ± 8.2	68.0 ± 7.6	66.3 ± 7.5	66.8 ± 7.9	67.0 ± 7.8
Male, n (%)	222 (79.9)	213 (76.9)	118 (68)	65 (75)	128 (74)	123 (72)	124 (72)
Percentage of predicted FVC, mean % ± SD	67.8 ± 11.2	68.6 ± 10.9	74.5 ± 14.5	76.4 ± 14.4	76.2 ± 15.5	74.9 ± 13.2	73.1 ± 14.2
Percentage of predicted DLco, mean % ± SD	43.7 ± 10.5	44.2 ± 12.5	46.4 ± 9.5	47.2 ± 8.2	46.1 ± 10.2	47.8 ± 9.8	47.4 ± 9.2
Dyspnoea score, mean ± SD	34.0 ± 21.9	36.6 ± 21.7	NR	NR	NR	NR	NR
Mean 6MWD, m ± SD	415.0 ± 98.5	420.7 ± 98.1	411.1 ± 91.8	417.5 ± 112.8	410.0 ± 90.0	378.0 ± 82.2	399.1 ± 89.7
Supplemental O_2 use, n (%)	78 (28.1)	76 (27.4)	29 (16.7)	15 (17)	25 (14)	48 (28)	49 (28)
HRCT definite IPF, n (%)	266 (95.7)	262 (94.6)	159 (91)	83 (95)	164 (94)	149 (87)	158 (91)
Surgical lung biopsy, n (%)	86 (30.9)	79 (28.5)	86 (49)	32 (37)	85 (49)	94 (55)	94 (54)
Time since IPF diagnosis, years ± SD	1.7 ± 1.1	1.7 ± 1.1	1.3 ± 0.96	1.4 ± 1.16	1.4 ± 1.12	1.2 ± 1.09	1.1 ± 1.04
Former smoker, n (%)	184 (66.2)	169 (61.0)	110 (63)	57 (66)	114 (66)	112 (66)	101 (58)
Pre-enrolment corticosteroid use, n (%)	6 (2.2)	2 (0.7)	14 (8.0)	10 (11.5)	9 (5.2)	22 (12.9)	17 (10.0)
Concomitant corticosteroid use, n (%)	82 (29.5)	101 (36.5)	38 (21.8)	24 (27.6)	52 (29.9)	42 (24.6)	50 (29.0)

PFN: pirfenidone 2,403mg/day; PBO: placebo; mg/d: milligrams per day

Table 9: Characteristics of participants in SP2 and SP3 (reproduced from CS,[4] Table 16, pages 84-85)

Baseline characteristic	SP3[38]			SP2[39]	
	PFN (1,800mg/d) (n=108)	PFN (1,200mg/d) (n=55)	PBO (n=104)	PFN (1,800mg/d) (n=72)	PBO (n=35)
Age, mean years ± SD	65.4 ± 6.2	63.9 ± 7.5	64.7 ± 7.3	64.0 ± 7.1	64.3 ± 7.6
Male, n (%)	85 (78.7)	47 (85.5)	81 (77.9)	62 (86.0)	33 (94.0)
Percentage of predicted VC, mean % ± SD	77.3 ± 16.8	76.2 ± 18.7	79.1 ± 17.4	81.6 ± 20.3	78.4 ± 17.2
Percentage of predicted TLC, mean % ± SD	73.2 ± 16.5	72.4 ± 15.6	75.2 ± 15.7	78.5 ± 17.9	73.9 ± 16.4
Percentage of predicted DLco, mean % ± SD	52.1 ± 16.8	53.6 ± 19.1	55.2 ± 18.2	57.6 ± 17.2	57.7 ± 13.8
Lowest SpO_2 during 6MWT, mean % ± SD	89.0 ± 2.3	88.8 ± 2.4	89.0 ± 2.0	87.1 ± 3.9	87.1 ± 4.2
Desaturation <88% during 6MWT, n (%)	34 (31.5)	19 (34.5)	24 (23.1)	NR	NR
Mean P(A-a)O_2 ± SD	18.4 ± 11.3	16.9 ± 9.6	17.4 ± 9.7	NR	NR
Percentage of predicted SpO_2, mean % ± SD	89.0 ± 2.3	88.8 ± 2.4	89.0 ± 2.0	NR	NR
Mean PaO_2 at rest, mmHg ± SD	79.8 ± 10.2	81.6 ± 8.4	81.0 ± 9.5	80.3 ± 7.7	82.0 ± 17.6
Mean VC, mL ± SD	2400.8 ± 638.4	2437 ± 684.8	2472.3 ± 698.9	NR	NR
Surgical lung biopsy, n (%)	26 (24.1)	16 (29.1)	28 (26.9)	15 (21.0)	8 (23.0)
IPF diagnosis, n (%)					
≤1 year	38 (35.2)	20 (36.4)	41 (39.4)	20 (28.0)	6 (17.0)
1-3 years	29 (26.9)	13 (23.6)	25 (24.0)	17 (24.0)	10 (29.0)
>3 years	41 (38.0)	22 (40.0)	38 (36.5)	35 (49.0)	19 (54.0)
Former smoker, n (%)	81 (75.0)	33 (60.0)	70 (67.3)	57 (79.0)	30 (86.0)
Pre-enrolment corticosteroid use, n (%)	9 (8.3)	6 (10.9)	6 (5.8)	10 (14.0)	5 (14.0)
Concomitant corticosteroid use, n (%)	8 (7.4)	6 (10.9)	5 (4.8)	NR	NR

PFN: pirfenidone; PBO: placebo; mg/d: milligrams per day

Participant flow and numbers

The loss to follow-up in the three trials was reported in the participant flow figures in the CS (Section 4.5, pages 77-81),[4] which were reproduced from the original publications. The ASCEND[33, 34] and CAPACITY trials all reported two types of patient trial discontinuation. Some patients discontinued the trial due to AEs, being lost to follow-up, withdrawing themselves or being withdrawn by the clinician. These were designated as the "discontinued study" group and did not include patients who had died or underwent lung transplantation. A second group of patients discontinued study treatment, principally on account of AEs, but also due to reasons such as death and lung transplantation. These were designated as the "discontinued treatment" group. However, they were deemed to have completed the study and were included in the analysis. The ASCEND[34] and CAPACITY trials[49] therefore experienced only a small loss of patients to follow-up in terms of those who "discontinued the study": approximately 5%-8% in any arm (see Table 10), compared with between 22% and 37% for any arm in the SP3[38] and SP2[39] trials. However, the rate of attrition was substantially higher (up to 22%) in the ASCEND[34] and CAPACITY trials[49] for the "discontinued treatment" groups (see Table 10). The overall rate of attrition for participants who either "discontinued study" or "discontinued treatment" was between 23% and 29% in any arm of the ASCEND[34] and CAPACITY trials[49] (see Table 10). However, the rates of attrition were essentially similar across intervention and placebo arms.

The primary approach for managing missing values in the efficacy analysis in ASCEND and the CAPACITY trials was to use the ITT population (which consisted of all patients who signed the informed consent form and were randomised). Last observation carried forward (LOCF) was used in SP2 and SP3. The safety analysis population included all patients who signed informed consent and received any amount of study drug (see CS,[4] Table 13). In the analyses of mean change, missing values owing to death were assigned the worst possible outcome (e.g. FVC=0%). Missing values with reasons other than death were imputed as the average value for the three patients with the smallest sum of squared differences at each visit. For the ranked ANCOVA analyses, missing values owing to death were assigned the worst ranks, with early deaths ranked worse than later deaths.

Table 10: Patient loss to follow-up in trials

Trial	Follow-up	Arms	Baseline n	Completed study n (%)	Completed treatment n (%)	Completed study and treatment n (%)
ASCEND[33, 34]	52 weeks	PFN 2,043mg/d	278	261 (94)	223 (80)	206 (74)
		PBO	277	261 (94)	238 (86)	222 (76)
CAPACITY 1[36, 49]	72 weeks	PFN 2,043mg/d	171	158 (92)	137 (80)	124 (72)
		PBO	173	164 (95)	142 (82)	133 (77)
CAPACITY 2[35, 49]	72 weeks	PFN 2,043mg/d	174	161 (93)	136 (78)	123 (71)
		PFN 1,197mg/d	87	82 (95)	70 (80)	65 (75)
		PBO	174	166 (95)	143 (82)	135 (77)
SP3[38]	52 weeks	PFN 1,800mg/d	108	68 (63)	Not reported	Not reported
		PFN 1,200mg/d	55	40 (73)		
		PBO	104	73 (70)		
SP2[39]	9 months	PFN 1,800mg/d	72	56 (78)		
		PBO	35	27 (78)		

PFN: pirfenidone; PBO: placebo; mg/d: milligrams per day

There was only general consistency across trials in terms of the primary and secondary outcomes designated in protocols and reported in publications, so for this reason the efficacy results are structured by clinical area or outcome measure, reflecting the structure of the CS.[4]

4.2.2.1 Lung function

Change from baseline in percent predicted FVC/VC

This outcome was reported by four of the five trials: for FVC by ASCEND[33, 34] and CAPACITY 1 &2[33, 36, 49] and for VC by SP3.[38]

The protocol made publicly available in the clinical trials register reported the primary efficacy outcome in the ASCEND trial[33] as change in percent predicted FVC from baseline to week 52 (see Table 6). The protocol[55] that accompanied the publication stated (Section 13.2, page 29): *"The clinical study protocol (dated 16 March 2011, section 5.4.2.1) describes a supportive analysis of FVC as the change from Baseline to Week 52 in FVC volume (in mL). Based on new findings from external sources, the analysis of FVC volume will be based on relative change (%) rather than actual volume (mL). A categorical analysis of relative change from Baseline has been added."* [55] The primary efficacy outcome in the protocols and publication for CAPACITY 1 & 2[35, 37, 51] was the change in percent predicted FVC from baseline to week 72.[49] In SP3,[38] the primary efficacy outcome reported was the change from baseline in VC in the pirfenidone 1,800mg per day group compared with the placebo group at 52 weeks.[38]

The ASCEND manuscript[34] did not report the change in percent predicted FVC, but this was reported in the CS,[4] principally to inform the NMA (see CS,[4] Table 20). At week 52, the mean difference in change from baseline in percent predicted FVC for pirfenidone 2,403mg per day compared with placebo was statistically significant in ASCEND[34] (mean difference 4.78%; p<0.001, see Table 11).

Table 11: Change from baseline in percent predicted FVC in ASCEND (reproduced from CS,[4] Table 20, page 94)

Study (source)	Treatment	Time point	Mean change from baseline	SE	Mean difference from PBO	*p*-value
ASCEND*[56] (Data on file[1])	PFN 2,403mg/day (n=278)	52 weeks	-6.17	0.875	4.781	<0.001
	PBO (n=277)		-10.95	0.877		
* The ASCEND manuscript did not report the change in percent predicted FVC but this was analysed to inform the NMA.						
[1] Roche 2016a[56]						

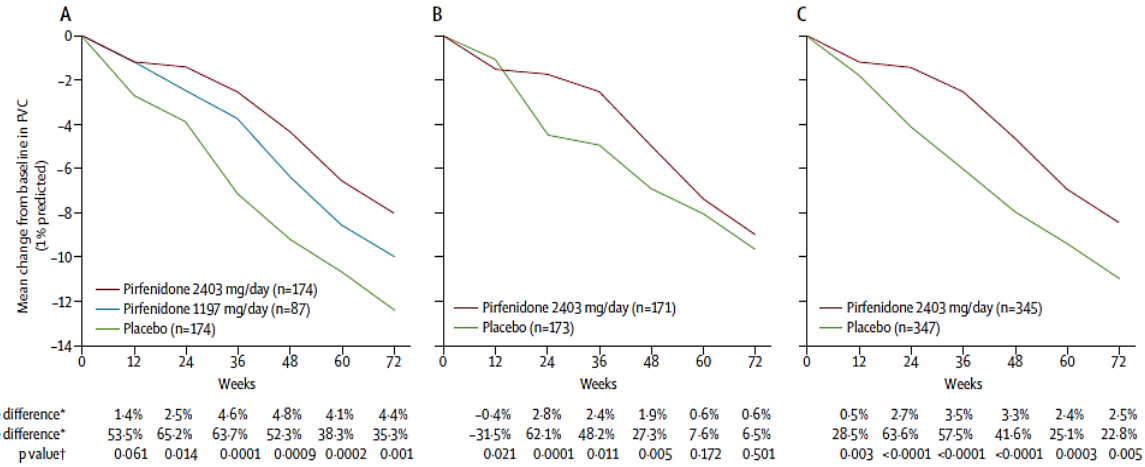
At week 72, the absolute difference in change in percent predicted FVC for pirfenidone 2,403mg per day compared with placebo in CAPACITY 1[49] was not statistically significant (absolute difference:

0.6%; relative difference: 6.5%; 95% CI -3·5 to 4·7, *p*=0.501), see Figure 1 reproduced from Noble 2011[49]).

At week 72, the absolute difference in change in percent predicted FVC for pirfenidone 2,403mg per day compared with placebo in CAPACITY 2[49] was statistically significant (absolute difference 4.4%; relative difference 35.3%; 95% CI 0·7 to 9·1, *p*=0.001). Outcomes in the pirfenidone 1,197mg/day group were intermediate to the pirfenidone 2,403mg/day and placebo groups.

At week 72, the absolute difference in change in percent predicted FVC for pirfenidone 2,403mg per day compared with placebo in a reported pooled analysis of the CAPACITY 1 & 2 trials[49] was statistically significant (absolute difference: 2.5%; relative difference: 22.8%; *p*=0.005, rank ANCOVA, see Figure 1 reproduced from Noble 2011[49]).

Figure 1: Change from baseline in percent predicted FVC in the CAPACITY 2 (A), CAPACITY 1 (B), and in the pooled population (C) (reproduced from Noble *et al.* 2011[49] and CS,[4] page 93)



Absolute difference*	1·4%	2·5%	4·6%	4·8%	4·1%	4·4%	-0·4%	2·8%	2·4%	1·9%	0·6%	0·6%	0·5%	2·7%	3·5%	3·3%	2·4%	2·5%
Relative difference*	53·5%	65·2%	63·7%	52·3%	38·3%	35·3%	-31·5%	62·1%	48·2%	27·3%	7·6%	6·5%	28·5%	63·6%	57·5%	41·6%	25·1%	22·8%
p value†	0·061	0·014	0·0001	0·0009	0·0002	0·001	0·021	0·0001	0·011	0·005	0·172	0·501	0·003	<0·0001	<0·0001	<0·0001	0·0003	0·005

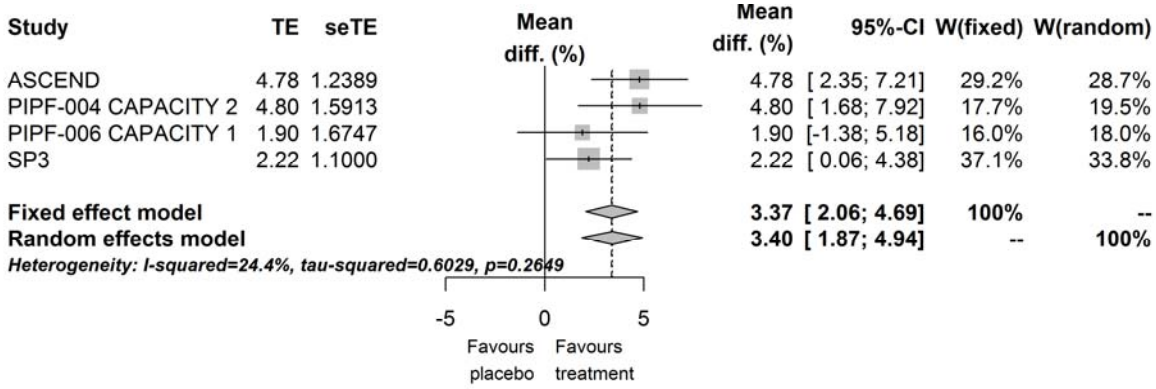
**Pirfenidone 2,403 mg/day versus placebo †Rank ANCOVA (pirfenidone 2,403mg/day vs placebo). 95% CIs were only calculated for absolute differences for the Week 72 time point in CAPACITY 2 (95% CI: 0.7-9.1) and CAPACITY 1 (95% CI: -3.5-4.7)*

At week 52, in SP3,[38] an analysis of the mean decline from baseline in percent predicted VC showed a significant treatment effect of pirfenidone 1,800mg/day compared with placebo, respectively: -2.91% ± 0.77 compared with -5.13% ± 0.78 (*p*=0.044, ANCOVA, see CS,[4] page 93)

The company conducted a meta-analyses using change in percent predicted FVC for ASCEND[33, 34] and CAPACITY 1 & 2[49] and change in percent predicted VC for SP3.[38] Both ASCEND[33, 34] and SP3[38] reported data at week 52, whilst the primary analysis in the CAPACITY trials[49] was at week 72. However, data at week 48 were used for the CAPACITY trials[49] to facilitate a like-for-like comparison between all four studies. The results of the meta-analysis are presented in Figure 2. The results suggest that the decline in percent predicted FVC in patients receiving pirfenidone

(2,403mg/day) was 3.4% less (95% CI: 1.87 to 4.94, *p*-value not reported) than in patients receiving placebo.

Figure 2: Forest plot of the mean difference in change from baseline in percent predicted FVC/VC (%) up to week 52 (reproduced from CS,[4] Appendix 9)



TE, Treatment effect; SE, Standard error

The ERG notes that both CAPACITY 1 and 2[49] report smaller treatment effects at week 72 (MD: 0.6 % in CAPACITY 1 and MD: 4.4% in CAPACITY 2) than at week 48. Selecting the 48 week data for inclusion in the meta-analysis therefore provides a larger estimate of overall treatment effect than would have estimated had the longer-term follow up data been used.

Mean change from baseline in FVC/VC (ml)

This outcome was reported by all five trials: FVC by CAPACITY 1,[49] CAPACITY 2[49] and ASCEND,[34] and VC by SP2[39] and SP3.[38]

Data from 48 weeks from the CAPACITY trials[49] were used in the NMA to allow comparison of studies across a similar time point (see Section 4.6), but the 72-week data are reported here.

All trials showed a statistically significant difference at the 5% level in favour of pirfenidone compared with placebo for change in FVC/VC, except CAPACITY 1[49] (absolute difference -5%; relative difference -1.4%; *p*=0.508). Detailed results are presented in Table 12.

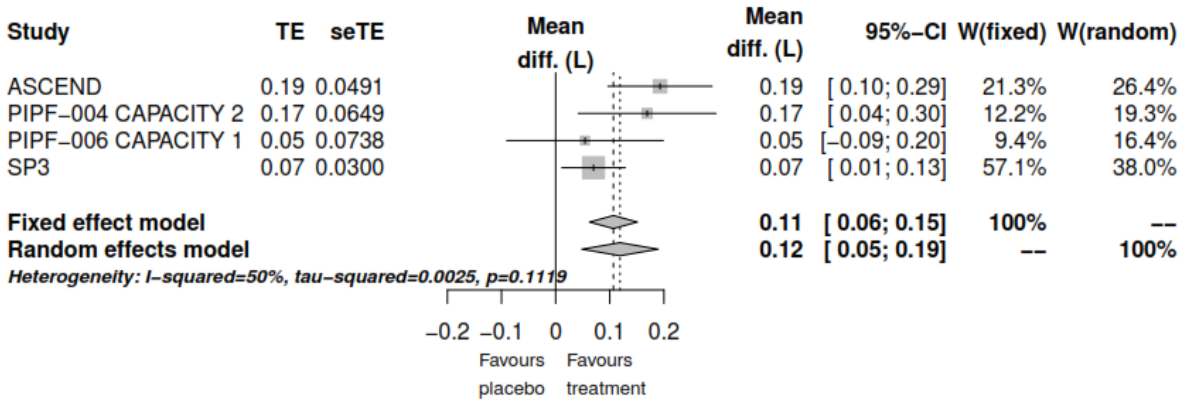
Table 12: Mean change from baseline in FVC/VC (ml) (reproduced from CS,[4] Table 21)

Study	Time point	Treatment group	Mean decline in FVC/VC	Difference, *p*-value†
ASCEND[34]	52 weeks	PFN 2,403mg/day (N=278)	FVC: 235 ml	Absolute difference: 193ml Relative difference: 45.1% *p*<0.001
		PBO (N=277)	FVC: 428 ml	
CAPACITY 1[49]	72 weeks	PFN 2,403mg/day (N=171)	FVC: 379 ml	Absolute difference: -5ml Relative difference: -1.4% *p*-value=0.508
		PBO (N=173)	FVC: 373 ml	
CAPACITY 2[49]	72 weeks	PFN 2,403mg/day (N=174)	FVC: 318 ml	Absolute difference:157ml Relative difference: 33% *p*-value=0.004
		PBO (N=174)	FVC: 475 ml	
SP3[38]	52 weeks	PFN 1,800mg/day (N=108)	VC: 90 ml	PFN 1,800 mg/day vs. PBO: Absolute difference: 70ml Relative difference: NR *p*=0.042
		PFN 1,200mg/day (N=55)	VC: 80 ml	
		PBO (N=104)	VC: 160 ml	
SP2[39]	9 months	PFN 1,800mg/day (N=72)	VC: 30 ml	Absolute difference: 100ml Relative difference: NR *p*=0.037
		PBO (N=35)	VC: 130 ml	

PFN: pirfenidone; PBO: placebo; NR: not reported
†Rank ANCOVA: ASCEND, CAPACITY 1 & 2 (pirfenidone 2,403mg/day vs placebo); SP2 and SP3 (pirfenidone 1,800mg/day vs. placebo)

A meta-analysis for change in FVC/VC (L) was conducted using data form ASCEND[34] and CAPACITY 1 & 2[49] (FVC (L)) and SP3[38] (VC (L)). Both ASCEND[34] and SP3[38] reported data for this outcome at week 52 and data at week 48 were used for the CAPACITY trials.[49] The meta-analysis suggests that on average, over 52 weeks, FVC in patients receiving pirfenidone (2,403mg/day) decline by 0.12L less than patients receiving placebo (95% CI: 0.05 to 0.19, *p*-value not reported), suggesting that pirfenidone slows the decline in lung function (see Figure 3). However, there was moderate heterogeneity between the trials (I^2=50%). In addition, the ERG noted that as with mean difference in change from baseline in percent predicted FVC (%), both CAPACITY 1[49] (MD: 0.005L) and CAPACITY 2[49] (MD: 0.16L) report smaller treatment effects at week 72 than that at week 48.

Figure 3: Forest plot of the mean difference in change from baseline in FVC/VC (L) up to week 52 (reproduced from CS,[4] Appendix 9)



TE, Treatment effect; SE, Standard error

FVC categorical decline of ≥10% percent predicted or death

This outcome was only reported for the ASCEND[34] and CAPACITY 1 & 2 trials.[49] The CS[4] states that a decline in percentage predicted FVC of ≥10% is a decrement that is recognised as clinically significant (see CS,[4] Section 4.7, page 90).

ASCEND[34] reported a statistically significant difference in favour of pirfenidone compared with placebo in terms of those who had experienced a decline in FVC by ≥10% or had died at week 52 (absolute difference: 15.3 [95% CI not reported], $p<0.001$). ASCEND also reported a significantly higher proportion of patients with no decline in percent predicted FVC for pirfenidone compared with placebo (22.7% versus 9.7%, $p<0.000001$).[34]

CAPACITY 1[49] reported that there was no statistically significant difference between pirfenidone and placebo in terms of those who had experienced a decline in FVC by ≥10% at week 72 (absolute difference: 3.8 [95% CI: -2.7 to 10.2], $p=0.440$). CAPACITY 1[49] also reported no statistically significant difference in the proportion of patients with no decline in percent predicted FVC for pirfenidone compared with placebo (25.8% versus 22%, *p*-value not reported).

CAPACITY 2[49] reported a statistically significant difference in favour of pirfenidone compared with placebo in terms of those who had experienced a decline in FVC by ≥10% at week 72 (absolute difference: 14.4 [95% CI: 7.4 to 21.3], $p=0.001$). CAPACITY 2[49] also reported a higher proportion of patients with no decline in percent predicted FVC for pirfenidone compared with placebo (24.1% versus 13.8%, *p*-value not reported).

Table 13: Categorical analysis of change from baseline in percent predicted FVC or death (reproduced from CS,[4] Table 18)

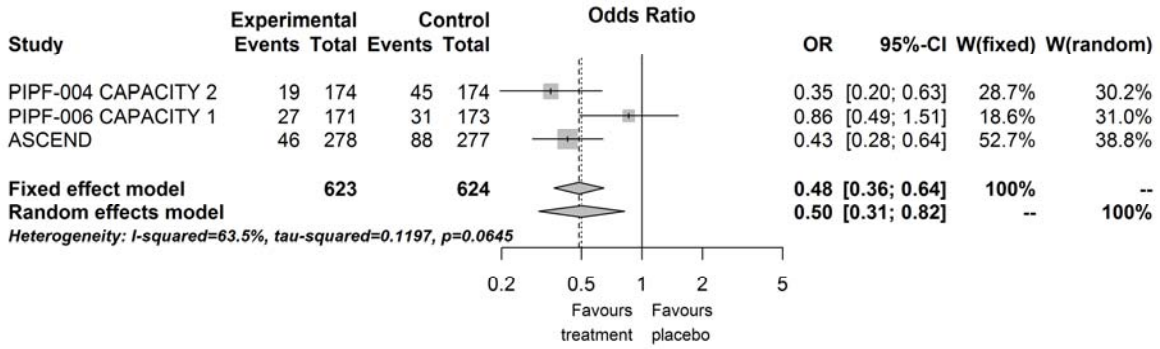
Study	Time point	Treatment group	Decline ≥10% FVC or death, n (%)	No decline* in FVC, n (%)	*p*-value†
ASCEND[34]	52 weeks	PFN 2,403mg/day (N=278)	46 (16.5)	63 (22.7)	*p*<0.000001
		PBO (N=277)	88 (31.8)	27 (9.7)	
CAPACITY 1[49] §	72 weeks	PFN 2,403mg/day (N=171)	39 (22.8)	44 (25.8)	*p*=0.440
		PBO (N=173)	46 (26.6)	38 (22.0)	
CAPACITY 2[49] §	72 weeks	PFN 2,403mg/day (N=174)	35 (20.1)	42 (24.1)	*p*=0.001
		PBO (N=174)	60 (34.5)	24 (13.8)	
Pooled CAPACITY 1 & 2[49]	72 weeks	PFN 2,403mg/day (N=345)	74 (21)	86 (24.9)	*p*=0.003
		PBO (N=347)	106 (31)	62 (17.9)	

PFN: pirfenidone; PBO: placebo
**Change in predicted FVC ≥10%; CAPACITY trials data not reported in original publication (Noble 2011[49])*
†Rank ANCOVA (pirfenidone 2,403mg/day vs placebo). It is unclear if this p value relates to the "Decline or death" or the "No decline" comparison: the numbers in the CS, Table 18 refer to the "No decline" comparison in ASCEND (King 2014[34]), but the "Decline or death" comparison for the CAPACITY trials (Noble 2011[49])
§ Note: these data are from the original publication (Noble 2011[49]), which only reports decline of >10% FVC and not decline of >10% or death

A pooled analysis of ASCEND[34] (week 52) and CAPACITY 1 & 2[49] (week 48) reported a statistically significant difference in favour of pirfenidone compared with placebo in terms of those who had experienced a decline in FVC by ≥10% or had died (absolute difference: 10.0 [95% CI not reported], *p*<0.003), and reported a higher proportion of patients with no decline in percent predicted FVC for pirfenidone compared with placebo (24.9% versus 17.9%, *p*-value not reported). This analysis is described as "pre-specified" in the CS[4] (page 91), but this is inaccurate: there is no reference to this analysis for this outcome in the any of the ASCEND protocols,[33, 55] which only refer to pooling of these trials for mortality (see Section 5.4.2.3.2 in the protocols). The protocol that accompanied the ASCEND publication (Section 13.2, page 29) stated that, *"The clinical study protocol (dated 16 March 2011, section 5.4.2.1) describes a supportive analysis of FVC as the change from baseline to Week 52 in FVC volume (in mL) … A categorical analysis of relative change from baseline has been added"*.[55]

A meta-analysis was conducted using data from 52 weeks for the ASCEND trial[34] and 48 weeks from the CAPACITY trial.[49] The results suggested that compared with placebo, pirfenidone lowers the proportion of patients experiencing decline in FVC percent predicted of ≥10% (OR: 0.50, 95% CI: 0.31 to 0.82, *p*-value not reported, see Figure 4). However, heterogeneity between the trials (I^2=63.5%) was moderately high.

Figure 4: Forest plot of odds ratios for FVC categorical decline of ≥10% percent predicted up to week 52 (reproduced from CS,[4] Appendix 9)



The pooled analysis of the two CAPACITY trials[49] at week 72, showed a lower proportion of patients experienced a decline of ≥10% in percent predicted FVC in the pirfenidone 2,403mg/day group (21% compared with 31%, respectively *p*=0.003).[49]

4.2.2.2 Mortality

All-cause and IPF-related mortality

All five trials provided data on mortality, although none of the studies was powered to assess the effect of pirfenidone on this outcome. No definition of IPF-related mortality was provided in the CS[4] or in the relevant publications.

ASCEND[34] reported all-cause mortality and so-called treatment-emergent IPF-related mortality (i.e. defined as the time after randomisation until 28 days after the final dose of the study drug) at 52 weeks; and CAPACITY 1 & 2[49] reported all-cause mortality, treatment-emergent all-cause mortality, IPF-related mortality and so-called treatment-emergent IPF-related mortality for 52 and 72 weeks.

Details of the all-cause mortality and TE IPF-related mortality at the common time point of 52 weeks, as well as the evidence from CAPACITY 1 & 2[49] at 72 weeks, are presented in Table 14.

The ASCEND trial[34] reported that at 52 weeks there were fewer overall deaths and TE IPF-related deaths in the pirfenidone group than the placebo group, but these differences were not statistically significant (p=0.105 and p=0.226 respectively).

In the pooled analysis of CAPACITY 1 & 2[49] at 52 weeks, there were fewer overall deaths and TE IPF-related deaths in the pirfenidone groups compared with the placebo groups and this difference was statistically significant in both groups (p=0.047 and p=0.012 respectively).

In the pooled analysis of CAPACITY 1 & 2[49] at 72 weeks, there were fewer overall deaths and TE IPF-related deaths in the pirfenidone groups compared with the placebo groups. Overall, there was a 23% reduction in all-cause mortality versus placebo among patients treated with pirfenidone 2,403mg/day (HR=0.77; 95% CI: 0.47 to 1.28; p=0.315), a 38% reduction in IPF-related mortality (HR=0.62; 95% CI: 0.35 to 1.13; p=0.117) and a 35% reduction in TE all-cause mortality (HR=0.65; 95% CI: 0.36 to 1.16; p=0.141). However, none of these differences were statistically significant.

For TE IPF-related mortality, the HR between the pirfenidone and placebo groups at week 72 also favoured pirfenidone and was statistically significant (HR=0.48; 95% CI: 0.24 to 0.95; p=0.03, see Table 14).

There appears to be a markedly increased rate of mortality for the CAPACITY trials[49] between the data reported in the CS[4] for 52 weeks (Table 23, page 97) and the data reported in the publication for 72 weeks.[49] If one assumes that the reported "all-cause mortality" is actually "treatment emergent all-cause mortality" (these distinctions exist in the CAPACITY trial publication,[49]), then there is a substantial increase in death rates in the pirfenidone group, from 11 at 52 weeks to 19 at 72 weeks, compared with a much smaller increase in the placebo group from 22 at 52 weeks to 29 at 72 weeks (the p-values for the differences between groups are 0.047 and 0.315 for 52 weeks and 72 weeks, respectively). The numbers for non-treatment emergent all-cause mortality are higher (see Table 14). In the same way, TE IPF-related mortality in the pirfenidone group increases from 4 deaths at 52 weeks to 12 deaths at 72 weeks in the pirfenidone group, and from 15 at 52 weeks to 25 at 72 weeks in the placebo group (p-values for the differences between groups are 0.012 and 0.030 for 52 and 72 weeks, respectively). No explanation is provided in the CS[4] for these relative increases in rates of mortality, particularly for the pirfenidone groups, between weeks 52 and 72 in the CAPACITY trials.[49]

In the pooled analysis of the data from 52 weeks for ASCEND[34] and CAPACITY 1 & 2[49] (required by the Food and Drug Administration (FDA)[57] and finalised as an analysis in the Statistical Analysis Plan only on 1st January 2014, according to the company's clarification response[10] (question A22), there

were significantly fewer overall deaths (p=0.047) and TE IPF-related deaths (p=0.012) in the pirfenidone groups compared with the placebo groups.

SP3[38] and SP2[39] reported all-cause mortality; there was no significant difference between groups. SP3[38] reported three deaths, four deaths and four deaths in the high-dose (1,800mg/d), low-dose (1,200mg/d) and placebo groups respectively, at 52 weeks,[38] and SP2[39] reported one death in the placebo group only, at 9 months.[39]

Table 14: Mortality rates in the CAPACITY 1 & 2 studies at week 52 and week 72 and the ASCEND and pooled populations at week 52 (reproduced from CS,[4] Table 23, page 97 and Noble 2011[49])

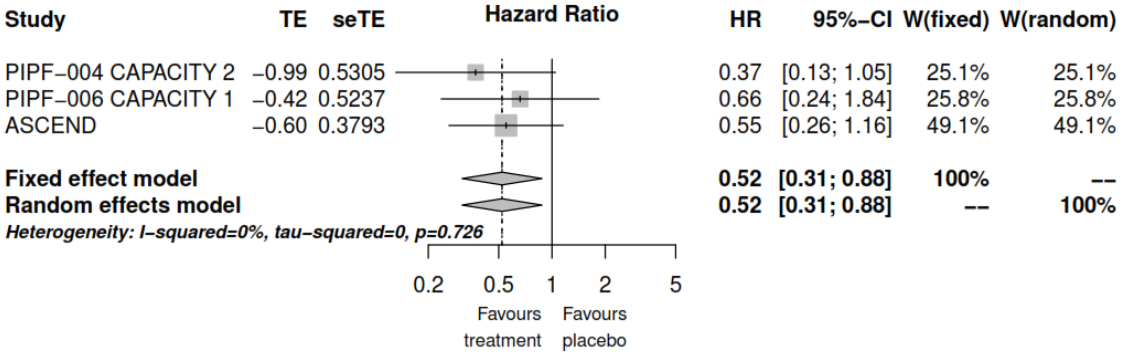
Patients	Time-point	PFN n (%)	PBO n (%)	HR (95% CI)*	*p*-value**
ASCEND[33, 34]	52 weeks	n=278	n=277		
All-cause mortality		11 (4.0)	20 (7.2)	0.55 (0.26, 1.15)	0.105
TE IPF-related mortality		3 (1.1)	7 (2.5)	0.44 (0.11, 1.72)	0.226
CAPACITY 1 & 2[49†]	52 weeks	n=345	n=347		
All-cause mortality		11 (3.2)	22 (6.3)	0.49 (0.24-1.01)	0.047
TE IPF-related mortality		4 (1.2)	15 (4.3)	0.27 (0.09–0.81)	0.012
All-cause mortality	72 weeks	27 (8)	34 (10)	0.77 (0.47-1.28)	0.315
IPF-related mortality		18 (5)	28 (8)	0.62 (0.35-1.13)	0.117
TE all-cause mortality		19 (6)	29 (8)	0.65 (0.36-1.16)	0.141
TE IPF-related mortality		12 (3)	25 (7)	0.48 (0.24-0.95)	0.030
Pooled data for ASCEND, [34] CAPACITY 1 & 2[49]	52 weeks	n=623	n=624		
All-cause mortality		22 (3.5)	42 (6.7)	0.52 (0.31–0.87)	0.011
TE IPF-related mortality		7 (1.1)	22 (3.5)	0.32 (0.14–0.76)	0.006

†Data in the CAPACITY 1 & 2 studies were censored at one year, but the 72-week data were published in Noble 2011[49]
**Cox proportional hazards model*
***Log-rank test (pirfenidone 2,403mg per day vs placebo)*
Abbreviations: PFN: pirfendione; PBO: placebo; TE- treatment-emergent

Meta-analysis was conducted using CAPACITY 1 & 2[49] and ASCEND[34] to assess the effect of pirfenidone on all-cause mortality; the trials reported HRs and the proportion of deaths. The company

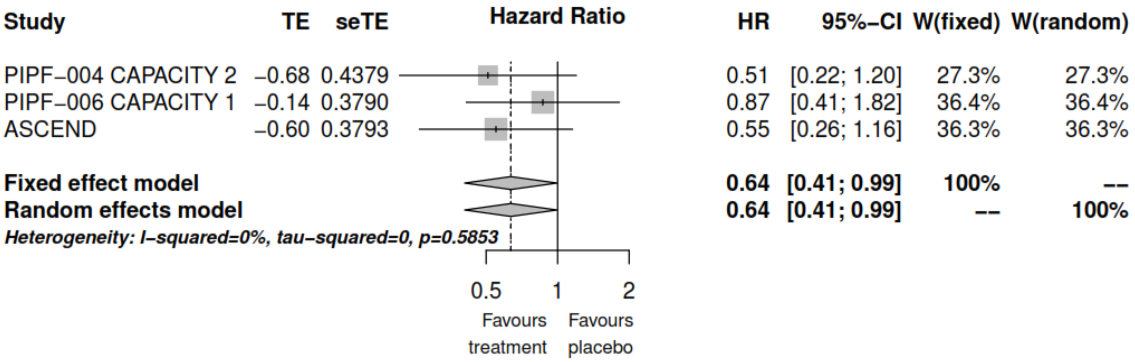
excluded SP3[38] from the analyses as it only reported the proportion of deaths. However, the ERG noted that SP3[38] was included in the company's NMA, where they used the method of Woods *et al.*[58] to combine the proportions reported in SP3[38] with HR. The results of the meta-analysis suggest that pirfenidone (2,403mg/day) compared with placebo reduces all-cause mortality (HR: 0.52, 95% CI: 0.31 to 0.88, *p*-value not reported) at 52 weeks (see Figure 5). A sensitivity analyses of the 3 trials based on data at 72 weeks for the CAPACITY trials[49] also favours pirfenidone (HR: 0.64, 95% CI: 0.41 to 0.99, *p*-value not reported, see Figure 6), however the reduction in mortality is lower than that observed using the 52 week data. Under the assumption of proportional hazards, we would expect the treatment effect to be constant over time.

Figure 5: Forest plot of hazard ratios for all-cause mortality (CAPACITY data at week 52) (reproduced from CS,[4] Appendix 9)



TE, Treatment effect (log hazard ratio); SE, Standard error

Figure 6: Forest plot of hazard ratios for all-cause mortality (CAPACITY data at week 72) (reproduced from CS,[4] Appendix 9)



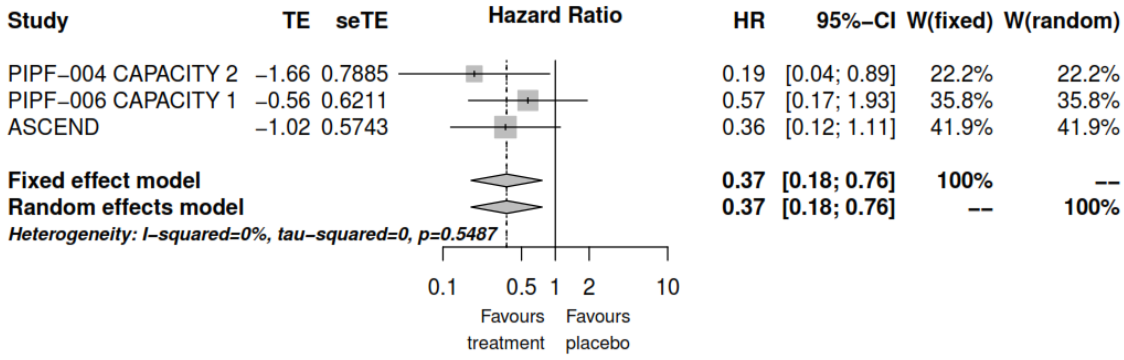
TE, Treatment effect (log hazard ratio); SE, Standard error

Meta-analysis of IPF-related mortality was also conducted using data from CAPACITY 1 & 2[49] and ASCEND.[34] All three trials reported data for 'IPF-related mortality' and 'IPF-related treatment emergent deaths', where treatment-emergent was defined as "the period from baseline to 28 days after

the last dose of the study drug." 'IPF-related mortality' is used in this analysis in line with an ITT approach for analysis. Meta-analysis of the 3 trials[34, 49] at 52 weeks suggests that pirfenidone compared with placebo reduces IPF-related mortality (HR: 0.37, 95%CI: 0.18 to 0.76, *p*-value not reported, see Figure 7).

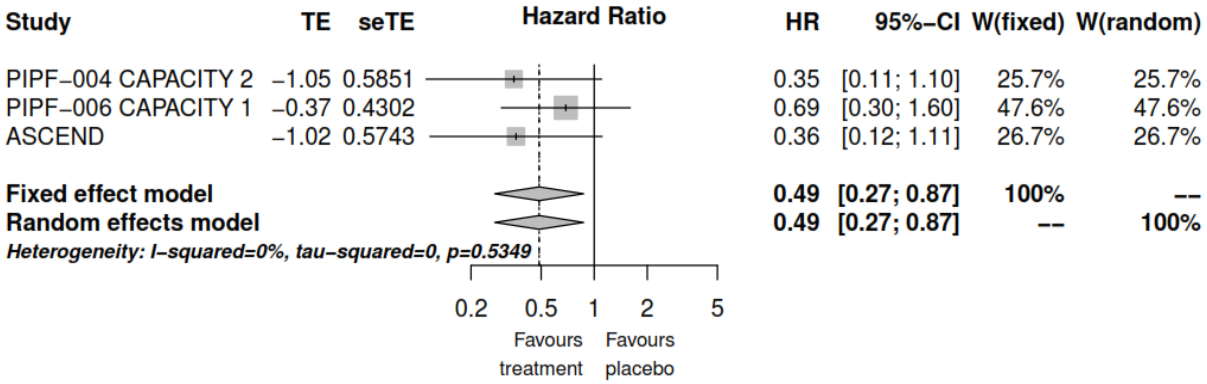
A sensitivity analyses of the three trials[34, 49] based on data at 72 weeks for the CAPACITY trials[49] also favours pirfenidone (HR: 0.49, 95% CI: 0.27 to 0.87, *p*-value not reported, see Figure 8), however, as with the all-cause mortality outcome, the reduction in mortality is lower than that observed using the 52 week data. Under the assumption of proportional hazards, we would expect the treatment effect to be constant over time.

Figure 7: Forest plot of hazard ratios for IPF-related mortality (CAPACITY data at week 52) (reproduced from CS,[4] Appendix 9)



TE, Treatment effect (log hazard ratio); SE, Standard error

Figure 8: Forest plot of hazard ratios for IPF-related mortality (CAPACITY data at week 72) (reproduced from CS,[4] Appendix 9)



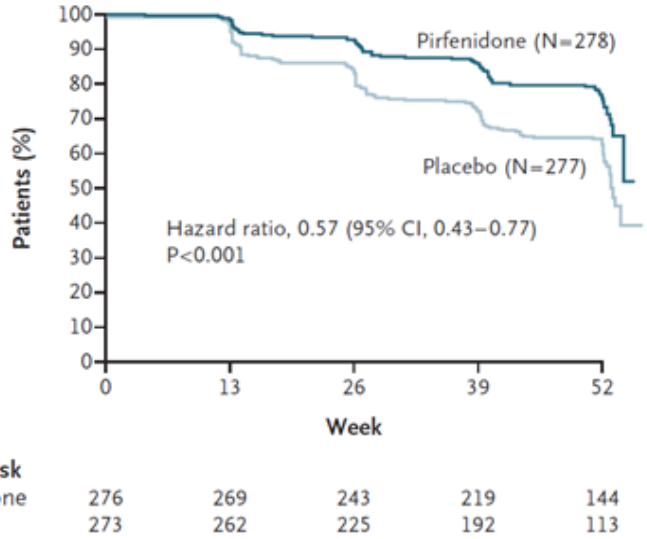
TE, Treatment effect (log hazard ratio); SE, Standard error

4.2.2.3 Progression-free survival (PFS)

The CS makes a case for the inclusion of this outcome based on similarities between IPF and *"the fundamental hallmarks of cancer biology"* (CS,[4] page 99). Four trials reported data for PFS: ASCEND,[34] CAPACITY 1 & 2[49] and SP3.[38] The definitions of PFS varied across the trials. ASCEND[34] defined PFS as the time to the first occurrence of any of the following: a confirmed ≥10% decline from baseline in percent predicted FVC, confirmed ≥50 m decline from baseline in 6MWD, or death.[34] The CAPACITY 1 & 2[49] defined PFS as the time to the first occurrence of any of the following: a confirmed ≥10% decline in percent predicted FVC, ≥15% decline in % predicted DLco or death.[49] In a *post hoc* analysis, the ASCEND[34] definition of PFS was applied to the CAPACITY trials[49] at week 52 and week 72 (see Figure 13). The SP3 trial[38] defined PFS as VC decline of ≥10% or death.

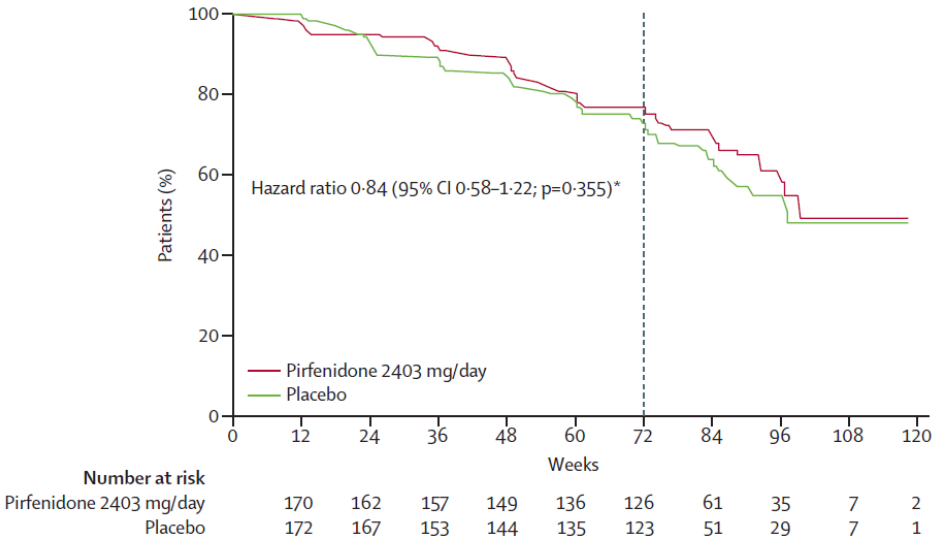
In ASCEND,[34] at 52 weeks, across all randomised patients, there was a statistically significant reduction in the risk of disease progression or death for patients receiving pirfenidone compared with those receiving placebo (HR 0.57; 95% CI, 0.43–0.77, p=0.0001, log-rank test, see Figure 9:).[34] That is, for each component of the composite endpoint, fewer patients in the pirfenidone group than in the placebo group had a qualifying event: death (3.6% versus 5.1%); a confirmed absolute decrease of ≥10% in percent predicted FVC (6.5% versus 17.7%); or a confirmed decrease of 50 m or more in the 6MWD (16.5% versus 19.5%).[34]

Figure 9: Kaplan–Meier estimates for PFS in all randomised patients from ASCEND (reproduced from CS,[4] Figure 10 and King 2014[34])



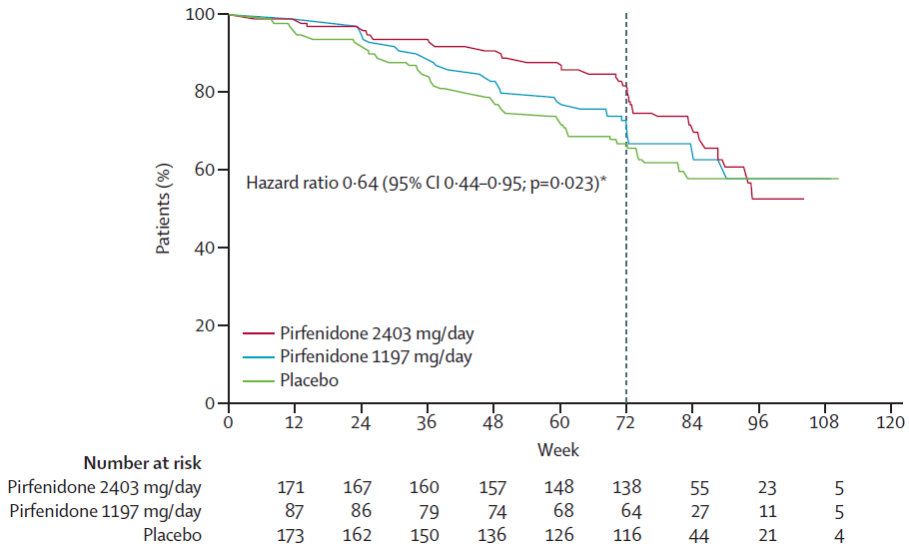
In CAPACITY 1, at 72 weeks, across all randomised patients, there was no statistically significant reduction in the risk of disease progression or death for pirfenidone compared with placebo (HR: 0.84; 95% CI, 0.58 to 1.22, p=0.355, see Figure 10).[49]

Figure 10: Kaplan-Meier estimates for PFS in CAPACITY 1 (reproduced from CS,[4] Figure 12)



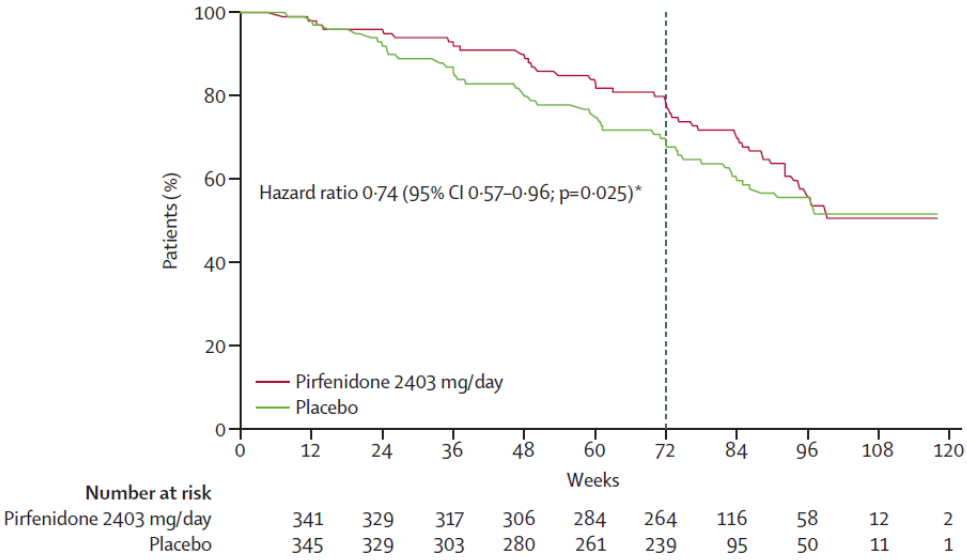
In CAPACITY 2,[49] at 72 weeks, across all randomised patients, there was a statistically significant reduction in the risk of disease progression or death for pirfenidone compared with placebo (HR 0.64; 95% CI, 0.44 to 0.95, *p*=0.023, log-rank test, see Figure 11).

Figure 11: Kaplan-Meier estimates for PFS in CAPACITY 2 (reproduced from CS,[4] Figure 11)



In the pooled population from CAPACITY 1 & 2,[49] at 72 weeks, there was a statistically significant reduction in the risk of disease progression or death for pirfenidone compared with placebo (HR=0.74; 95% CI: 0.57 to 0.96; *p*=0.025, see Figure 12).

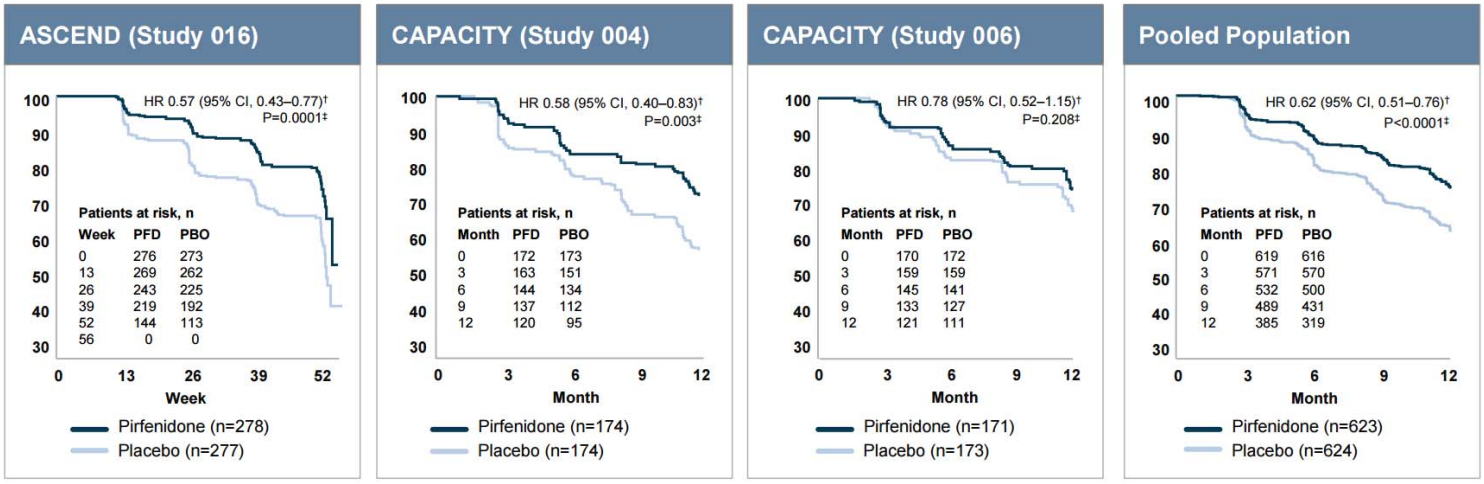
Figure 12: Kaplan-Meier estimates for PFS in the CAPACITY 1 & 2 pooled population (reproduced from CS,[4] Figure 13)



As noted above, an exploratory *post hoc* analysis of PFS was conducted on data from the 52-week CAPACITY 1 & 2[49] populations using the ASCEND[34] definition for disease progression (time to the first occurrence of death, confirmed ≥10% decline in percent predicted FVC, or confirmed ≥50 m decrement in 6MWD). The company justified replacing the DLco criteria with the 6MWD criteria with reference to the relationship between 6MWD and survival.[59] The use of 52-week data and the application of this definition of PFS, which included criteria relating to 6MWD rather than DLco, resulted in reduced HRs and *p*-values in the CAPACITY trials.[49] For CAPACITY 1[49] from HR 0.84 (95% CI, 0.58 to 1.22, *p*=0.355) (original definition using DLco criteria) to HR 0.78 (95% CI, 0.52 to 1.15, *p*=0.208) (using 6MWD criteria), and for CAPACITY 2 from HR 0.64 (95% CI, 0.44 to 0.95, *p*=0.023) to HR 0.58 (95% CI, 0.40 to 0.83, *p*=0.003). See the CS[4] (and Figure 13).

A *post hoc* pooled analysis of these data on PFS from ASCEND,[34] CAPACITY 1 and CAPACITY 2[49] at week 52 was also undertaken: there was a statistically significant reduction in the risk of disease progression or death for pirfenidone compared with placebo (HR: 0.62; 95% CI: 0.51 to 0.76; *p*<0.0001, see Figure 13).

Figure 13: Post hoc analysis of progression-free survival at week 52 in ASCEND, CAPACITY trials, and in the pooled population (reproduced from CS,[4] Figure 14)

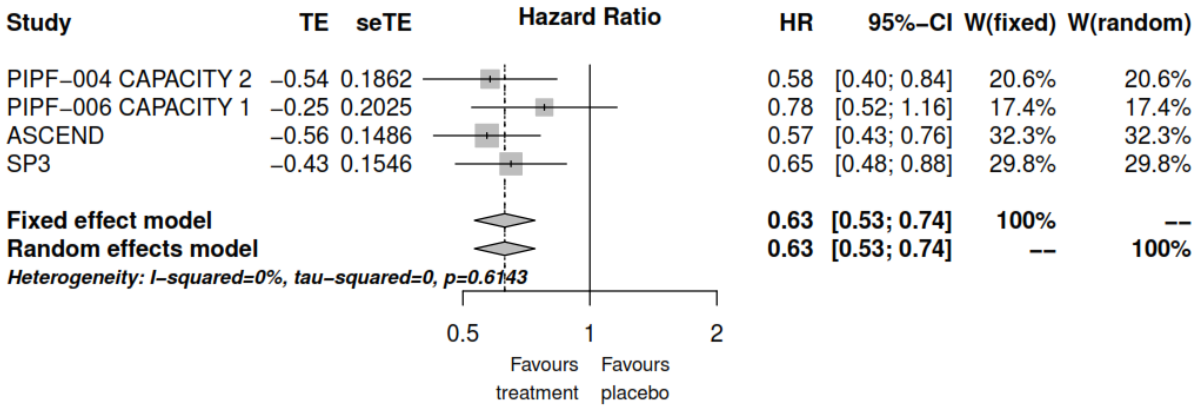


† Cox proportional hazards model
‡ Log-rank test

In SP3,[38] pirfenidone 1,800 mg per day significantly reduced the risk of disease progression or death (defined as VC decline of ≥10% or death) by 55% compared with placebo (HR 0.45; 95% CI 0.11 to 0.79; *p*=0.028, log-rank test).

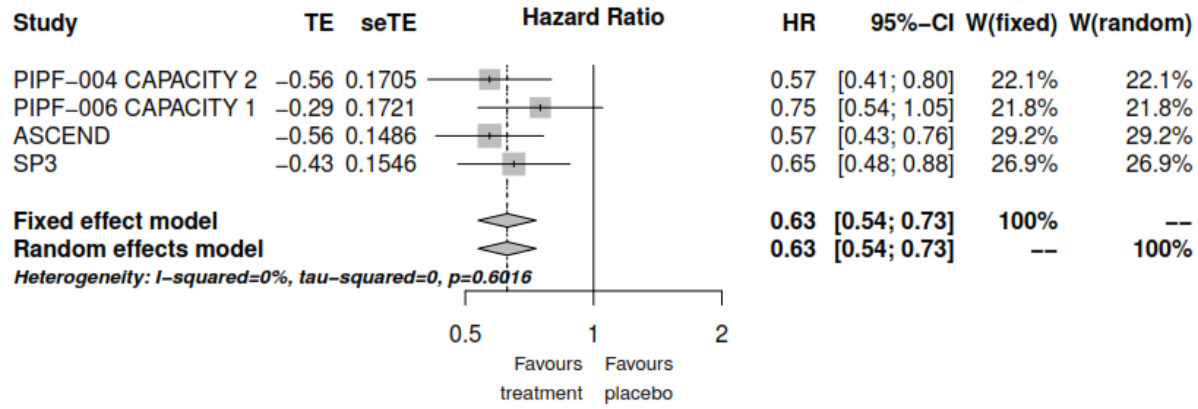
A meta-analysis based on data at 52 weeks was conducted using all four trials (ASCEND,[34] CAPACITY 1&2[49] and SP3[38]). The results of the meta-analysis suggest that pirfenidone compared with placebo reduces the risk of disease progression or death (HR: 0.63, 95% CI: 0.53 to 0.74, *p*-value not reported, see Figure 14). These results are in line with the *post hoc* pooled analysis of ASCEND,[34] CAPACITY 1 and CAPACITY 2[49] at week 52. A sensitivity analysis based on 72 week results for the CAPACITY trials[49] and 52 week results for ASCEND[34] with the assumption that the proportional hazards assumption holds up to 72 weeks gave the same results (see Figure 15). However, as noted above, the definition of PFS varied across the trials and the CS applied the ASCEND[34] definition of PFS to the CAPACITY trials[49] at both week 52 and week 72. The SP3 trial[38] defined PFS as VC decline of ≥10% or death. Hence, the ERG believes caution should be applied when interpreting these results.

Figure 14: Forest plot of hazard ratios for progression-free survival (CAPACITY data at week 52) (reproduced from CS,[4] Appendix 9)



TE, Treatment effect (log hazard ratio); SE, Standard Error

Figure 15: Forest plot of hazard ratios for progression-free survival (CAPACITY data at week 72) (reproduced from CS,[4] Appendix 9)



TE, Treatment effect (log hazard ratio); SE, Standard Error

4.2.2.4 Acute exacerbations

All five trials provided data on acute exacerbations, although the criteria for this outcome varied across the trials. The definitions are provided in secondary outcomes Table 7. For ASCEND,[34] acute exacerbations were identified *"via a post hoc analysis of adverse events based on the MedDRA lower level term 'acute exacerbation of IPF'"* (CS,[4] page 104). The publications for ASCEND[34] and the CAPACITY studies[49] did not report the incidence of acute exacerbations, and the latter trials recorded this outcome only as part of the protocols' composite outcome "Worsening of IPF" (see Table 7). These data were therefore extracted from the CSRs and presented in the CS[4] (Table 27, page 106), for use in the pairwise meta-analysis and NMA.

The rates of acute exacerbation were much higher in the ASCEND trial than in the CAPACITY trials,[49] with a higher incidence in the placebo than the pirfenidone arms in the ASCEND[34] and CAPACITY 2 trials[49]: no *p*-values were reported (see Table 15).

Table 15: CSR data for acute exacerbations for ASCEND, CAPACITY 1 & 2 (reproduced from CS,[4] Table 27)

Trial	Intervention	Time point	n
ASCEND (Data on file)	PFN n=278	52 weeks	24
	PBO n=277		40
CAPACITY 1 (Data on file)	PFN n=171	52 weeks	1
	PBO n=173		0
CAPACITY 2 (Data on file)	PFN n=174	52 weeks	0
	PBO n=174		3

- For ASCEND, acute exacerbations were not reported in the primary manuscript King 2014[34]. Acute exacerbations at 52 weeks were available as data on file.
- For CAPACITY 1 & 2, acute exacerbations were not reported in the primary manuscript Noble 2011[49]. Data at 52 weeks were available as data on file and were handled as separate studies.

PFN: pirfenidone 2,403mg/d; PBO: placebo

In SP3,[38] according to the CS[4] and Taniguchi *et al*[38], the incidence of acute exacerbation during the study or within 28 days after the termination of the study was 5.6% (n=6), 5.5% (n=3) and 4.8% (n=5) in the pirfenidone 1,800mg/day, pirfenidone 1,200mg/day and placebo groups, respectively. No statistically significant differences were seen between the three groups. According to a published abstract,[60]stepwise multivariate analysis revealed that decline in VC ≥10% within 6 months was a significant risk factor for acute exacerbations (HR, 3.951, *p*=0.012).

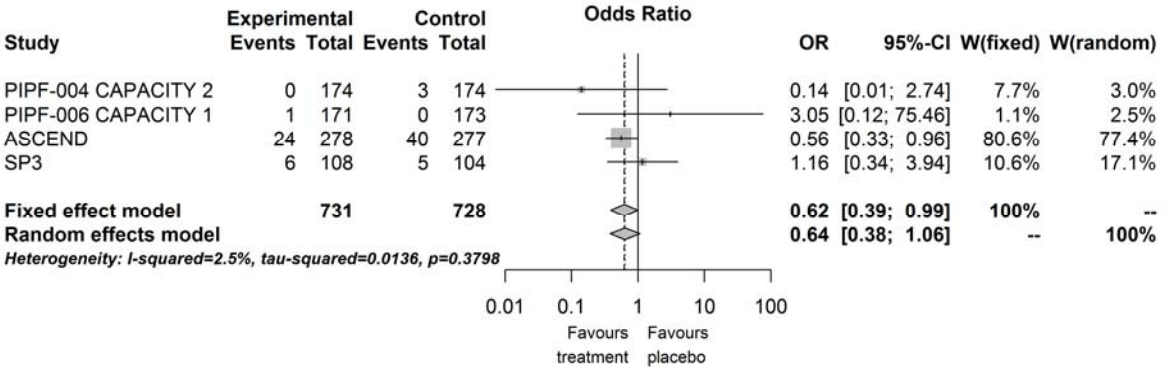
In SP2,[39] according to the CS[4] and Azuma *et al*[39], the incidence of acute exacerbation of IPF was 14% (n=5) in the placebo group and was zero in the pirfenidone group during the 9 months (*p*=0.0031).

There was no consistency in the frequency of acute exacerbation reported across trials. This might be explained by the different definitions used and the difficulty in diagnosis (see clarification response,[10] question A15).

A meta-analysis was conducted based at 52 weeks, using data from ASCEND,[34] CAPACITY 1 & 2[49] and SP3[38] (see Table 15). However, according to the CS,[4] Appendix 9, page 73), data used for the CAPACITY trials[49] were from 48 weeks. The results show that pirfenidone is associated with a reduced risk of acute exacerbations of IPF compared with placebo, with a OR of 0.64 (95% CI: 0.38 to 1.06, *p*-value not reported, see Figure 16). Analyses using a fixed effects model suggest a

statistically significant treatment effect in favour of pirfenidone (OR: 0.62: 95% CI: 0.39 to 0.99, *p*-value not reported). Caution should be applied when interpreting these results as the definition of 'acute exacerbation' varied across trials and there were very few events in the CAPACITY 1 and 2[49] trials whilst in ASCEND[34] the event rate was high.

Figure 16: Forest plot of odds ratios for acute exacerbations up to week 52 (reproduced from CS,[4] Appendix 9)



4.2.2.5 Hospitalisations

This outcome was only reported for CAPACITY 1 & 2[49] and SP2.[39] The protocols for CAPACITY 1 & 2[49] included respiratory hospitalisations as part of the "Worsening of IPF" outcome and SP2[39] reported respiratory hospitalisations, but the CS[4] also reported non-respiratory hospitalisations for the CAPACITY trials[49] (see Table 7).

The CS[4] reported *post hoc* analyses for this outcome (pages 106 and 107), including number of patients hospitalised; number of hospitalisations; mean length of stay in hospital and total number of days in hospital. The data for respiratory and non-respiratory hospitalisations are reported in Table 16. In the pooled CAPACITY 1 & 2[49] population, the number of patients with at least one hospitalisation for respiratory causes (14.8% for pirfenidone versus 15% for placebo) or non-respiratory causes (20.9% versus 16.1% respectively) was similar across treatment arms. However, the duration of hospital stay was consistently numerically greater in the placebo arms.

Table 16: Post hoc analysis of data on hospitalisations in CAPACITY 1 & 2 (reproduced from CS, Table 28)[4]

Study arm	CAPACITY 1[49]		CAPACITY 2[49]		Pooled	
	PFN n=171	PBO n=173	PFN n=174	PBO n=174	PFN n=345	PBO n=347
Respiratory hospitalisations (RH)						
Number of patients with at least 1 RH	22 (12.9%)	23 (16.7%)	29 (16.7%)	29 (16.7%)	51 (14.8%)	52 (15.0%)
Number of RH	31	37	34	40	65	77
Mean length of RH (days)	8.5	17.3	7.6	12.1	8.0	14.6
Total number of days in hospital	264	640	259	484	522	1124
Non-respiratory hospitalisations (NRH)						
Average number of NRH days per patient	1.5	3.7	1.5	2.8	1.5	3.2
Number of patients with at least 1 NRH	37 (21.6%)	25 (14.5%)	35 (20.1%)	31 (17.8%)	72 (20.9%)	56 (16.1%)
Number of NRH	48	31	38	42	86	73
Mean length of NRH (days)	10.1	20.8	7.2	16.0	8.8	8.0
Total number of days in hospital	485	645	274	672	758	1317
Average number of NRH days per patient	2.8	3.7	1.6	3.9	2.2	3.8

PFN: pirfenidone 2,403mg/d; PBO: placebo

In SP2,[39] five patients in the placebo arm and none in the pirfenidone treatment were hospitalised due to exacerbations (Azuma 2005[39]). The company did not conduct a meta-analysis as data were only available for the CAPACITY trials.

4.2.2.6 Patient-Reported Outcomes (Quality of Life)

University of San Diego (UCSD) Shortness of Breath Questionnaire (SOBQ)

The ASCEND[34] and CAPACITY trials[39] reported this outcome. The CS[4] states (pages 111 and 112) that the SOBQ can be used to formulate clinically relevant inferences about IPF patients; that the total score in this questionnaire increases with increased dyspnoea, and an increment of 20 points is considered a clinically relevant threshold based on estimates of the minimal important difference for the USCD SOBQ that range from 5-11.[31] In ASCEND,[34] the proportion of patients with ≥20 point increase in shortness of breath as measured by SOBQ at week 52 was smaller in patients receiving pirfenidone than in those receiving placebo, but this difference was not statistically significant (p=0.1577, see Table 17).

Table 17: Categorical outcomes for UCSD SOBQ in ASCEND at week 52† (reproduced from CS,[4] Table 34)

Outcomes, n (%)	PFN (n=278)	PBO (n=277)	*p*-value*
Worsening score ≥20 points or death	81 (29.1)	100 (36.1)	0.1577
Worsening score <20 to 0 points	124 (44.6)	115 (41.5)	
No worsening (score change <0 points)	73 (26.3)	62 (22.4)	

†Missing data due to reasons other than death were imputed using the sum of squared differences (SSD) method and included in the ≥20 points category
*p-value by rank ANCOVA

PFN: pirfenidone 2,403mg/d; PBO: placebo

In CAPACITY 1 & 2,[49] there were no significant differences between the pirfenidone and placebo groups for the change from baseline to week 72 (see Table 18). There was therefore no evidence of a treatment effect in any of the three key trials.

Table 18: Mean change in UCSD SOBQ score from baseline for the relevant RCTs (ITT population, reproduced from CS,[4] Table 35)

Study	Time point	Treatment group	Mean change in dyspnoea score	*p*-value*
CAPACITY 1[49]	72 weeks	PFN (n=171)	11.9	p=0.604
		PBO (n=173)	13.9	
CAPACITY 2[49]	72 weeks	PFN (n=174)	12.1	p=0.509
		PBO (n=174)	15.2	

*Rank ANCOVA (PFN vs placebo)

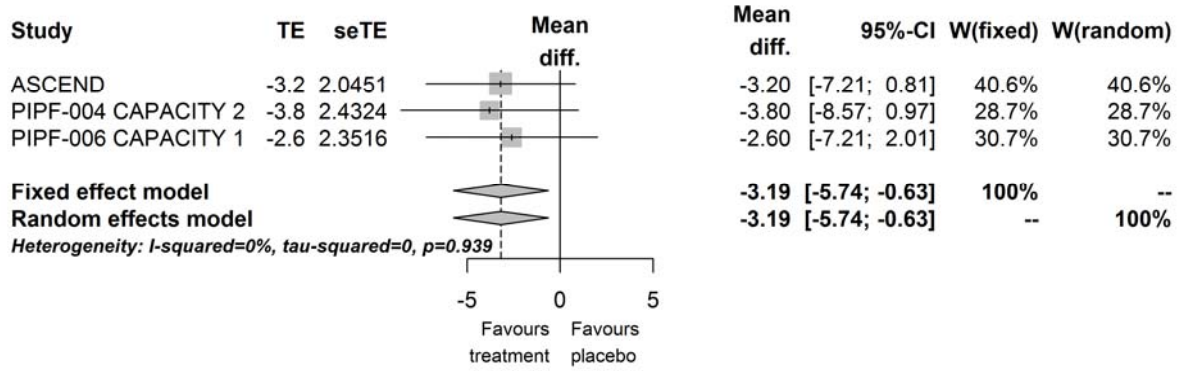
PFN: pirfenidone 2,403mg/d; PBO: placebo

The CS,[4] (page 112) reported that pooled data from the three studies showed pirfenidone treatment reduced the proportion of patients who experienced a ≥20 point increase or death (*p*=0.0471).[37]

The meta-analysis included data from the ASCEND trial[34] at 52 weeks and the CAPACITY trials[49] at 48 weeks. The results suggest that, at this time point, pirfenidone reduces the decline in USCD SOBQ compared with placebo (Mean difference: -3.19 (95% CI: -5.74, to -0.63, *p*-value not reported, see Figure 17), although the mean difference in the individual studies was not statistically significant. The ERG notes that both CAPACITY 1 & 2[49] report smaller treatment effects at week 72 (MD: 2.0% % in CAPACITY 1 and MD: 3.1% in CAPACITY 2) than at week 48 and so the observed statistically

significant difference does not necessarily hold for time points beyond 48/52 weeks, The results of the meta-analysis are consistent with the pooled analysis of CAPACITY 1 & 2[49] and ASCEND.[34]

Figure 17: Forest plot of the mean difference in change from baseline in UCSD SOBQ up to week 52 (reproduced from CS,[4] Appendix 9)



TE, Treatment effect; SE, Standard error

St. George's Respiratory Questionnaire

Only CAPACITY 1 & 2[49] reported data for this outcome. However, it was not listed in any protocols and was not reported in the original publication. It therefore appears to be a *post hoc* analysis. The CS[4] (page 111) states that this measure has demonstrated a strong correlation between physical impairment and disease severity, clinical symptoms, and functional disability in patients with IPF. At week 72, the difference in change in SGRQ between pirfenidone and placebo was not statistically significant in either trial (see Table 19).

Table 19: SGRQ measure of change in health status from baseline to week 72 in CAPACITY 1 & 2 (reproduced from CS,[4] Table 32, page 111)

	Change from baseline to week 72 (mean ± SD)		*p*-value*
	PFN	**PBO**	
CAPACITY 1[49]	**(n=166)**	**(n=169)**	
SGRQ	7.2 ± 16.85	7.3 ± 20.37	0.766
CAPACITY 2[49]	**(n=163)**	**(n=165)**	
SGRQ	7.6 ± 18.89	9.0 ± 18.86	0.495
*Rank ANCOVA stratified by geographic region (USA and rest of world). Missing data due to a patient's death were ranked as worse than any non-death and according to time until death			

PFN: pirfenidone 2,403 mg/d; PBO: placebo

As only the CAPACITY trials[49] reported data for this outcome, the company did not conduct a meta-analysis.

4.2.2.7 6-Minute Walking Distance (6MWD) or 6-Minute Walking Test (6MWT)

Three trials reported data on this outcome: ASCEND[34] and CAPACITY 1 & 2.[49] Data were analysed according to categories of decrements of ≥50 metres or <50 metres, and mean change from baseline.

Categorical analysis of change from baseline in 6MWD

The CS[4] states that a decrement of ≥50 metres in 6MWD is considered an appropriate and clinically relevant threshold for a categorical assessment of response to therapy because it has been associated with an increased risk of mortality.[25] Categorical analysis of 6MWD data was carried out *post hoc* in the CAPACITY 1 & 2 studies,[49] but was pre-specified as a secondary endpoint in ASCEND in the protocol accompanying the publication,[55] but not in the clinical trials register protocol.[33] However, the CS[4] (Table 29, page 108) reported findings for these trials for a *post hoc* composite outcome of mean decline ≥50 m from baseline in 6MWD or death.

At week 52, the absolute difference in the proportion of patients with a mean decline ≥50 m from baseline, or death, for pirfenidone 2,403mg per day compared with placebo in ASCEND[34] was statistically significant (absolute difference: 9.8%; relative reduction: 27.5%; *p*=0.04, see Table 20).

Table 20: Proportion of patients with a mean decline of ≥50 m in 6MWD from baseline or death in ASCEND and CAPACITY 1 &2 (ITT population) (reproduced from CS,[4] Table 29)

Trial	Time point	Treatment group	Mean decline of ≥50 m in 6MWD or death, n (%)	Difference, *p*-value
ASCEND[34]	52 weeks	PFN (n=278)	72 (25.9)	Absolute difference: 9.8% Relative reduction: 27.5% *p*=0.04*
		PBO (n=277)	99 (35.7)	
CAPACITY 1[49]	72 weeks	PFN (n=169)	56 (33.1)	*p*=0.10**
		PBO (n=168)	79 (47.0)	
CAPACITY 2[49]	72 weeks	PFN (n=170)	62 (36.5)	*p*=0.049**
		PBO (n=170)	80 (47.1)	
Pooled CAPACITY 1 & 2[49]	72 weeks	PFN (n=339)	118 (34.8)	Absolute difference: 12.2% Relative risk: 26% *p*=0.001**
		PBO (n=338)	159 (47.0)	

PFN: pirfenidone; PBO: placebo
**Rank ANCOVA (pirfenidone 2,403mg/day vs placebo)*
***Cochran-Mantel-Haenszel test*

At week 72, the difference in the proportion of patients with a mean decline ≥50 m from baseline, or death, for pirfenidone 2,403mg per day compared with placebo in CAPACITY 1[49] was not

statistically significant (p=0.10, see Table 20). At week 72, the difference in the proportion of patients with a mean decline ≥50 m from baseline, or death, for pirfenidone 2,403mg per day compared with placebo in CAPACITY 2[49] was statistically significance (p=0.049, see Table 20).

At week 72, the absolute difference in the proportion of patients with a mean decline ≥50 m from baseline, or death, for pirfenidone 2,403mg per day compared with placebo across CAPACITY 1 & 2[49] was statistically significant (absolute difference: 12.2%; relative reduction: 26%; p=0.001, see Table 20).

The CS[4] (page 108) reported that a *post hoc* pooled analysis of ASCEND[34] and CAPACITY 1 & 2 (data from weeks 52 and 48 respectively)[49] reported a statistically significant improvement in 6MWD for patients receiving pirfenidone 2,403mg per day compared with placebo (p=0.0004). The CS[4] cites Nathan 2014[19] as the supporting study, but the reference provided does not contain this analysis; the source of this analysis and its data is therefore unclear.

Mean change in 6MWD from baseline

The CS[4] states that the reliability and validity of 6MWD as a responsive measure of disease status and a valid endpoint for clinical trials has been demonstrated in a recent study, where the minimally clinical important difference (MCID) was estimated at 24-45 meters.[25]

At week 52, in ASCEND,[34] the absolute difference in the mean change from baseline in 6MWD for pirfenidone 2,403mg per day compared with placebo was statistically significant (absolute difference: 26.7m; relative reduction: 44.2%; p=0.036) and satisfied the lower end of the MCID (see Table 21).

At week 72, in CAPACITY 1,[49] the absolute difference in the mean change from baseline in 6MWD for pirfenidone 2,403mg per day compared with placebo was statistically significant (absolute difference: 31.8m; relative difference: not reported; p<0.001) and satisfied the MCID (see Table 21). Therefore, the CAPACITY 1[49] results for the categorical analysis of 6MWD (not statistically significant) and the mean change in 6MWD (statistically significant) were different in terms of statistical significance.

Table 21: Mean change from baseline in 6MWD in ASCEND and CAPACITY 1 & 2 (reproduced from CS,[4] Table 30)

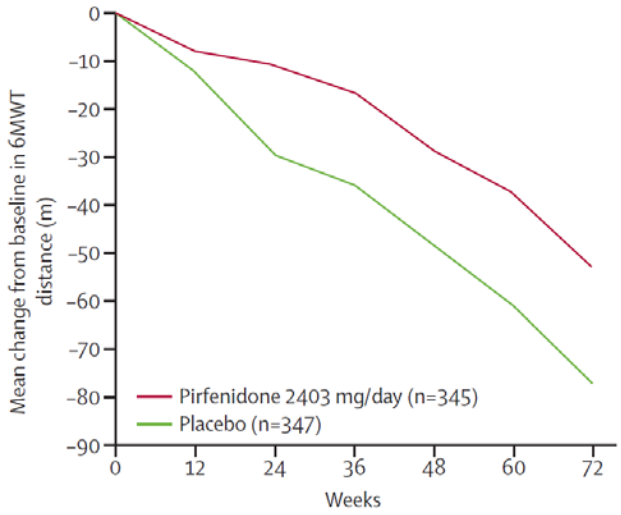
Study	Time point	Treatment group	Mean decline, metres	Difference, *p*-value†
ASCEND[34]	52 weeks	PFN (n=278)	33.5 m	Absolute difference: 26.7 m Relative reduction: 44.2% *p*=0.036
		PBO (n=277)	60.2 m	
CAPACITY 1[49]	72 weeks	PFN (n=174)	45.1 m	Absolute difference: 31.8 Relative difference: NR *p*<0.001
		PBO (n=174)	76.9 m	
CAPACITY 2[49]	72 weeks	PFN (n=171)	60.4 m	Absolute difference: 16.4 m Relative difference: NR *p*=0.171
		PBO (n=173)	76.8 m	
Pooled CAPACITY 1 & 2	72 weeks	PFN (n=345)	52.8 m	Absolute difference: 24 m Relative difference: 31.2% *p*=0.0009
		PBO (n=347)	76.8 m	

PFN: pirfenidone; PBO: placebo; m: metres
†Rank ANCOVA (pirfenidone 2,403 mg/day vs placebo)

However, at week 72, in CAPACITY 2,[49] the absolute difference in the mean change from baseline in 6MWD for pirfenidone 2,403mg per day compared with placebo was not statistically significant (absolute difference: 16.4m; relative difference: not reported; *p*=0.171) and did not satisfy the lower end of the MCID (see Table 21). Therefore, the CAPACITY 2[49] results for the categorical analysis of 6MWD (statistically significant) and the mean change in 6MWD (not statistically significant) were different in terms of statistical significance.

At week 72, in the pooled analysis of CAPACITY 1 & 2,[49] the absolute difference in the mean change from baseline in 6MWD for pirfenidone 2,403mg per day compared with placebo was statistically significant (absolute difference: 24m; relative difference: 31.2%; *p*=0.0009) and satisfied only the lowest threshold of the MCID (see Table 21 and Figure 18).

Figure 18: Mean change from baseline in 6MWD in CAPACITY 1 & 2 pooled population (reproduced from CS,[4] Figure 15)

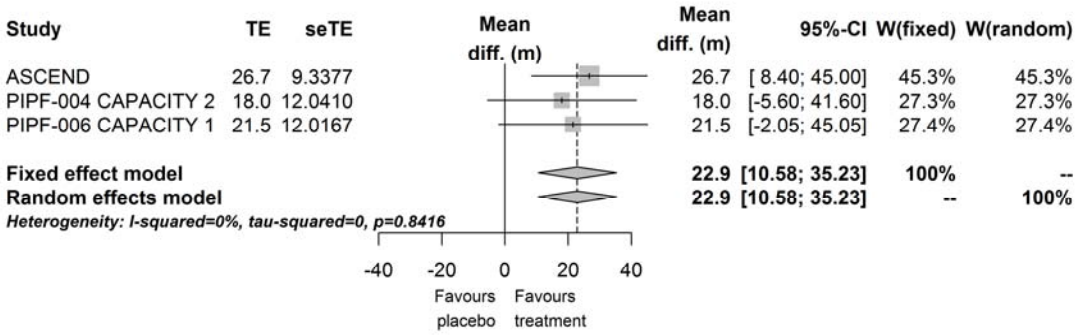


Absolute difference* (m)	3·9	18·6	18·7	19·8	23·3	24·0
Relative difference*	32·2%	62·8%	52·5%	40·6%	38·2%	31·2%
p value†	0·760	0·042	0·053	0·004	0·002	0·0009

**Pirfenidone 2,403 mg/day vs placebo*
†Rank ANCOVA (pirfenidone 2,403 mg/day vs placebo)

Three trials (ASCEND,[34] CAPACITY 1 & 2[49]) were included in the meta-analysis to assess change in distance walked from baseline in the 6MWT. Data at week 48 from the CAPACITY trials[49] were combined with data from week 52 in the ASCEND trial.[34] The meta-analysis suggests that, on average, patients receiving pirfenidone declined by 22.9m less than patients receiving placebo with a 95% CI of (10.58m to 35.23m, *p*-value not reported, see Figure 19).

Figure 19: Forest plot of the mean difference in change from baseline in 6MWD up to week 52 (reproduced from CS,[4] Appendix 9)



TE, Treatment effect; SE, Standard error

4.2.2.8 Measurement of the carbon monoxide diffusing capacity of the lungs (DLco)

Four trials (CAPACITY 1 & 2,[49] SP3,[38] SP2[39]) reported data on the change from baseline in DLco. The CAPACITY trials reported the change in % predicted DLco, while SP2[39] and SP3[38] reported the mean decline (mL/min/mmHG). None of the trials showed a statistically significant treatment effect compared to placebo for this outcome measure.

CAPACITY 1[49] reported a mean change of -9.8% for pirfenidone and -9.2% for placebo, respectively (p=0.996); and CAPACITY 2[49] reported a mean change of -7.9% for pirfenidone and -9.9% for placebo (p=0.145). A published, pooled analysis also indicated that there was no evidence of a statistically significant treatment effect for this outcome (p=0.301).[49] In both the SP2[39] and SP3 trials,[38] there was no statistically significant difference in mean decline of DLco between pirfenidone 1,800mg/day and placebo. The company did not conduct a meta-analysis for DLco as the measurements were not considered comparable.

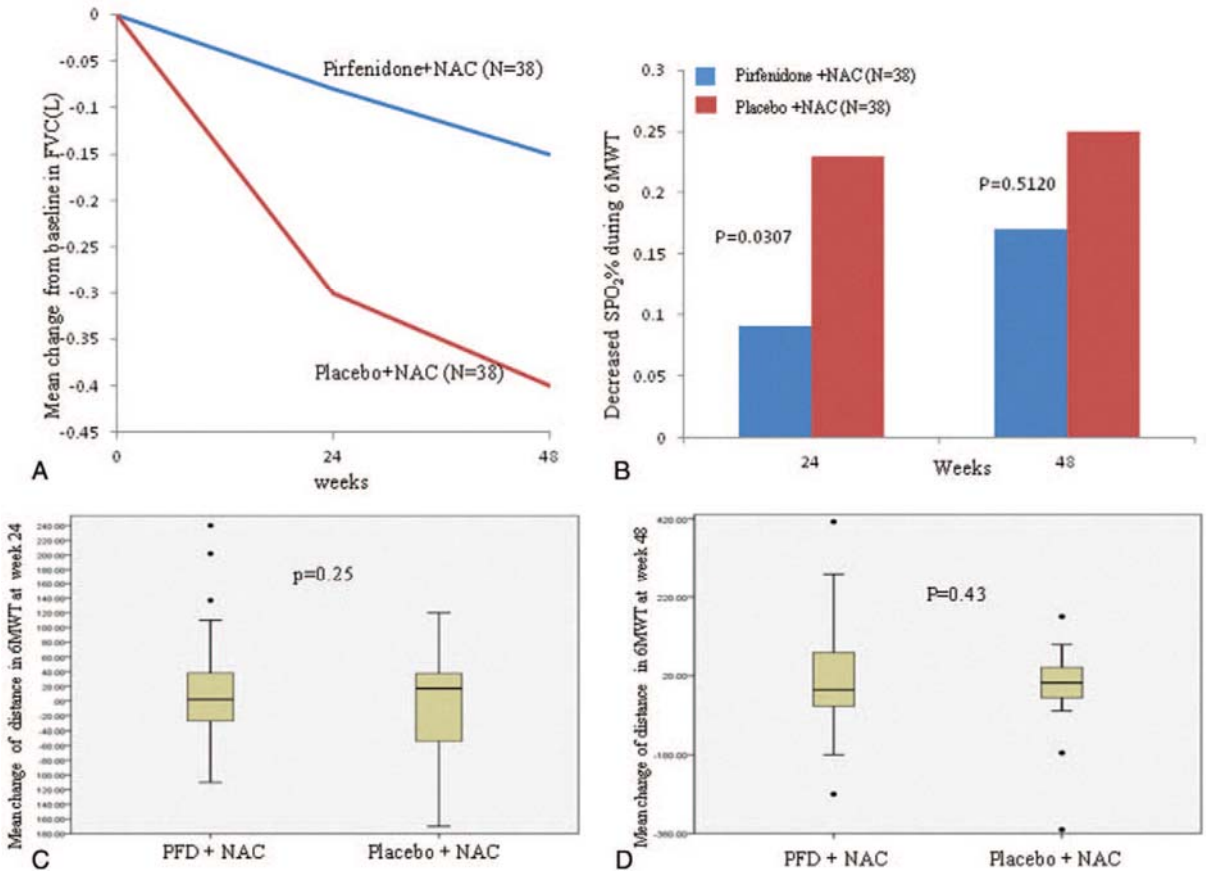
4.2.2.9 Supporting evidence from the Huang et al. trial [48] of pirfenidone plus NAC versus placebo plus NAC

For the purposes of this appraisal, as supporting evidence, only details of the Huang *et al.* efficacy results for FVC, 6-Minute Walking Test (6MWT) and PFS, are presented here[48] (see Figure 20).

The Huang *et al.* trial[48] reported a statistically significant mean change in FVC from baseline in favour of pirfenidone plus NAC compared with placebo plus NAC at 24 weeks (p=0.02) but not at 48 weeks (p=0.11). The authors performed *post hoc* analyses to explore possible reasons behind the change from week 24 to week 48. In doing so, they identified four patients (three in the pirfenidone group and one in the placebo group) who experienced a substantial decline in pulmonary function test parameters (including FVC and DLco) due to AEs after 24 weeks but before 48 weeks. When these patients, were excluded from the analyses, the authors reported that they found a significant treatment effect at both 24 weeks (p=0.018) and 48 weeks (p=0.048).

This trial[48] also reported that there was no statistically significant mean change in 6MWT from baseline for pirfenidone plus NAC compared with placebo plus NAC at either 24 (p=0.25) or 48 weeks (p=0.43, see Figure 20).

Figure 20: Mean change from baseline in FVC at 24 and 48 weeks and in 6MWT at 24 and 48 weeks (reproduced from Huang 2015,[48] Figure 2A-D)



FVC: Forced vital capacity; 6MWT: 6-minute walking test

PFS was also evaluated, defined as the time until the first occurrence of any one of the following: a confirmed ≥10% decline in the percentage predicted FVC, a confirmed ≥15% decline in the percentage predicted DLco (corrected based on the patient's actual haemoglobin levels), a confirmed progression of fibrosis defined by the HRCT fibrosis score, AE-IPF, or death. For PFS, pirfenidone plus NAC had a significant treatment benefit compared with placebo plus NAC (HR=1.88, 95% CI: 1.092–3.242, *p*=0.02). No significant differences were observed in the percent change in the secondary outcomes of arterial blood gas (ABG) ($PaCO_2$, PaO_2, and SaO_2) levels, the dyspnoea score, the HRCT findings, the SGRQ score, or the number of IPF-related adverse events between the pirfenidone and placebo groups.

4.3 Subgroup analyses

4.3.1 Pre-specified analyses

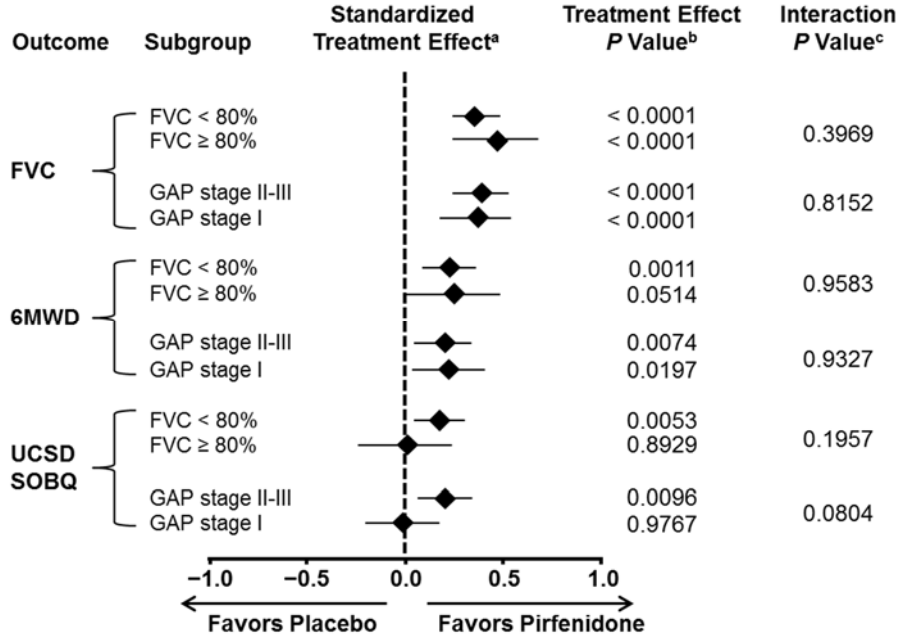
No subgroup analyses were pre-specified for the ASCEND,[34] SP3[38] or SP2[39] trials. Subgroup analyses based on pooled CAPACITY 1 & 2 data were reported in the CS,[4] Figure 16 (page 114) for the primary efficacy outcome variable (difference between pirfenidone and placebo in mean change from

baseline to week 72 in percent predicted FVC). There was no evidence for differential treatment effects according to: sex (p=0.263), age (p=0.864), race (p=0.807), geographic region (p=0.359), and baseline IPF severity (p=0.352). However there was evidence of an interaction between treatment and time from IPF diagnosis (p=0.021), with patients diagnosed >1 year before randomisation experiencing greater effect). In response to a clarification request by the ERG concerning some of the subgroups in this analysis, the company stated: *"As results reported in Figure 16 deviated from [the] more robust approach for the primary outcome, we believe they should not be further used for assessment of robustness and consistency of results in subpopulations"* (see clarification response,[10] question A29).

4.3.2 Post hoc analyses

A *post hoc* analysis of pooled data from ASCEND[34] and CAPACITY 1 & 2[49] was conducted to examine the effects of pirfenidone on patients stratified by earlier versus more advanced disease severity, i.e. "earlier" being "mild" IPF: baseline FVC ≥80% (pirfenidone, n=146; placebo, n=170); and "more advanced" being "moderate" IPF: baseline FVC <80% (pirfenidone, n=477; placebo n=454). According to the CS,[4] (page 115), baseline characteristics and demographics were similar across groups. Efficacy outcomes of interest included absolute ≥10% FVC decline, ≥50m 6MWD decline, and ≥20-point worsening of dyspnoea as measured by UCSD SOBQ. Treatment-by-subgroup interactions were tested based on rank ANCOVA models. Missing values were imputed by using the sum of squared differences method. Factors in the model include study, geographic region, treatment group, subgroups, and treatment-by-subgroup interaction. A proportional hazards model estimated the HR between subgroups. The analysis indicated that there was no significant difference (treatment-by-subgroup interaction) between those patients with baseline FVC ≥ 80% predicted and those with FVC < 80% predicted (see Figure 21).

Figure 21: Treatment effect of pirfenidone by baseline disease severity from pooled data of ASCEND, CAPACITY 1 & 2 (reproduced from CS,[4] Figure 17)



A separate *post hoc* analysis (unpublished) was conducted to evaluate the outcomes for patients who experienced a ≥10% decline in percent predicted FVC during the first 6 months of treatment across the three ASCEND and CAPACITY 1 & 2 trials.[61, 62] Eight-four out of 1,247 patients experienced a ≥10% decline in % FVC during the first 6 months of treatment across these trials: 24 had received pirfenidone (it is unclear if any of these had received the 1,197mg per day dose) and 60 had received placebo. Of these, one (4.2%) had experienced >10% decline in FVC in the pirfenidone group, and 15 (25%) in the placebo group (p=0.032) (see Table 22). The CS[4] states that these findings suggest a potential benefit to continued treatment with pirfenidone despite an initial decline in FVC; this is not consistent with the stopping rule currently recommended in NICE TA282.[2]

Table 22: Outcomes following previous ≥10% decline in FVC at 6 months in ASCEND and CAPACITY 1 & 2 (reproduced from CS,[4] Table 37)

Outcome, n (%)	PFN (n=24)		PBO (n=60)		Relative Difference*	*p*-value
≥10% decline in FVC or death	1	(4.2)	15	(25.0)	−83.3%	0.032
Death	0	(0)	10	(16.7)	−100%	0.056
>0% to <10% decline in FVC	9	(37.5)	23	(38.3)	-2.2%	ND
No further decline in FVC	14	(58.3)	22	(36.7)	59.1	0.089
**Relative difference calculated using the following formula: 100 × [pirfenidone − placebo]/[placebo]*						

These results were supported by an additional *post hoc* analysis,[37] which evaluated the effect of pirfenidone on subgroups based on age, smoker status, and baseline disease status; this analysis found no evidence for differential effects between subgroups.

Exploratory subgroup analyses were conducted in SP3[38] and SP2[39] also. Both analyses found that, in terms of percent predicted VC, IPF patients with baseline percent predicted VC ≥ 70% had better outcomes in terms of VC and PFS at week 52 than those patients with a baseline percent predicted VC <70% although for SP2 the actual data were not reported.[39]

In response to a clarification request from the ERG (see clarification response,[10] question A31), the company also provided results on OS (see Table 23) and PFS (see Table 24) from the ASCEND[34] and CAPACITY trials[49] on groups with a baseline percent predicted FVC of ≤80% (moderate IPF) and >80% (mild IPF). However, numbers within each trial and trial arm were not reported.

The findings did not suggest differential treatment effects according to disease severity for either outcome (as judged by the reported HR and 95% CI), however a treatment-by-subgroup interaction test was not reported so it is unclear if the difference between these subgroups was statistically significant.

Table 23: OS results by baseline FVC percent predicted subgroup at 52 and 72 weeks (reproduced from Clarification response,[10] question A31, Table 12)

Study & time point	Baseline FVC ≤80% predicted			Baseline FVC >80% predicted		
	Adjusted HR	95% CI	*p*-value	Adjusted HR	95% CI	*p*-value
CAPACITY 1[49]						
52 weeks	0.6	0.17-2.04	0.4051	0.77*	0.11-5.59	0.7932
72 weeks	0.89	0.40-1.99	0.7763	0.77	0.11-5.59	0.7932
CAPACITY 2[49]						
52 weeks	0.25	0.08-0.76	0.0080	NE**	**	**
72 weeks	0.29	0.10-0.79	0.0102	4.04***	0.42-38.87***	0.1900***
ASCEND[34]						
52 weeks	0.63	0.29-1.34	0.2215	<0.01	0.00-NE	0.1231
72 weeks	N/A	N/A	N/A	N/A	N/A	N/A
Pooled trials						
52 weeks	0.48	0.27-0.83	0.0071	0.59	0.14-2.51	0.4682
72 weeks	0.58	0.36-0.94	0.0240	0.90	0.27-2.99	0.8610

NE: not evaluable
** Only two deaths occurred in CAPACITY 1 before 52 weeks*
*** There were no additional deaths observed in either arm of CAPACITY 2 between 52 and 72 weeks in patients with FVC >80% predicted*
**** Low number of events*

Table 24: PFS results by baseline FVC percent predicted subgroup at 52 and 72 weeks (reproduced from clarification response,[10] question A31, Table 11)

Study & time point	Baseline FVC ≤80% predicted			Baseline FVC >80% predicted		
	Adjusted HR	95% CI	*p*-value	Adjusted HR	95% CI	*p*-value
CAPACITY 1[49]						
52 weeks	0.84	0.53-1.32	0.4438	0.63	0.29-1.41	0.2571
72 weeks	0.85	0.58-1.26	0.4128	0.56	0.28-1.11	0.0919
CAPACITY 2[49]						
52 weeks	0.60	0.40-0.92	0.0159	0.40	0.18-0.89	0.0193
72 weeks	0.58	0.39-0.86	0.0590	0.48	0.25-0.92	0.0233
ASCEND[34]						
52 weeks	0.56	0.41-0.76	0.0002	0.64	0.30-1.40	0.2584
72 weeks	N/A	N/A	N/A	N/A	N/A	N/A
Pooled trials						
52 weeks	0.62	0.52-0.78	<0.0001	0.54	0.35-0.75	0.0069
72 weeks	0.64	0.52-0.79	<0.0001	0.53	0.35-0.79	0.0017

4.4 Non-randomised and non-controlled evidence

The CS[4] reported findings from RECAP (PIPF-012),[40] a non-randomised, non-controlled, OLE of the ASCEND[34] and CAPACITY trials.[49] The study was designed to assess the long-term safety of pirfenidone 2,403mg per day in patients with IPF. To be included in the extension study, patients must have "completed treatment", that is, they must have received ≥80% of scheduled doses (of either

active treatment or placebo) and completed the week 72 final study visit in CAPACITY 1 or 2[49] (CS,[4] page 159, see Table 25).

Table 25: Summary of RECAP study design

	RECAP (PIFP-012) (Costabel, 2014[40]; Kreuter 2014[63])
Study design	Open-label, uncontrolled, Phase III extension study in which eligible patients receive treatment with pirfenidone 2,403mg/day
Intervention	Eligible patients received pirfenidone 2,403mg/day Concomitant therapy with corticosteroids, azathioprine, cyclophosphamide, and/or NAC were permitted if judged appropriate by investigator
Population	IPF patients that completed the ASCEND[34] or CAPACITY 1 & 2
Objectives	**Primary objective:** To examine the long-term safety and tolerability of pirfenidone in patients with IPF who were previously randomised to the placebo group in either CAPACITY 1 or 2 studies (later adjusted to allow enrolment from the ASCEND trial, Kreuter 2014[63]) **Secondary objective:** To obtain additional efficacy data for pirfenidone 2,403mg/day in patients with IPF
Inclusion/Exclusion criteria	**Inclusion criteria:** • Completes the ASCEND or CAPACITY studies final visit • In the opinion of the principal investigator has been generally compliant (received ≥80% of scheduled doses) with study requirements during the qualifying study, or must be considered eligible to enrol in RECAP by the InterMune medical monitor • Is able to provide informed consent and comply with the requirements of the study **Exclusion criteria:** • Pregnant or lactating women • In the opinion of the PI, is not a suitable candidate for study participation • Known hypersensitivity to any of the components of the study drug • Participates in another interventional clinical trial between the end of participation in ASCEND or either CAPACITY studies and time of enrolment in RECAP • Receives concomitant medications defined in the protocol • Permanently discontinues study drug during the ASCEND or CAPACITY studies for any reason

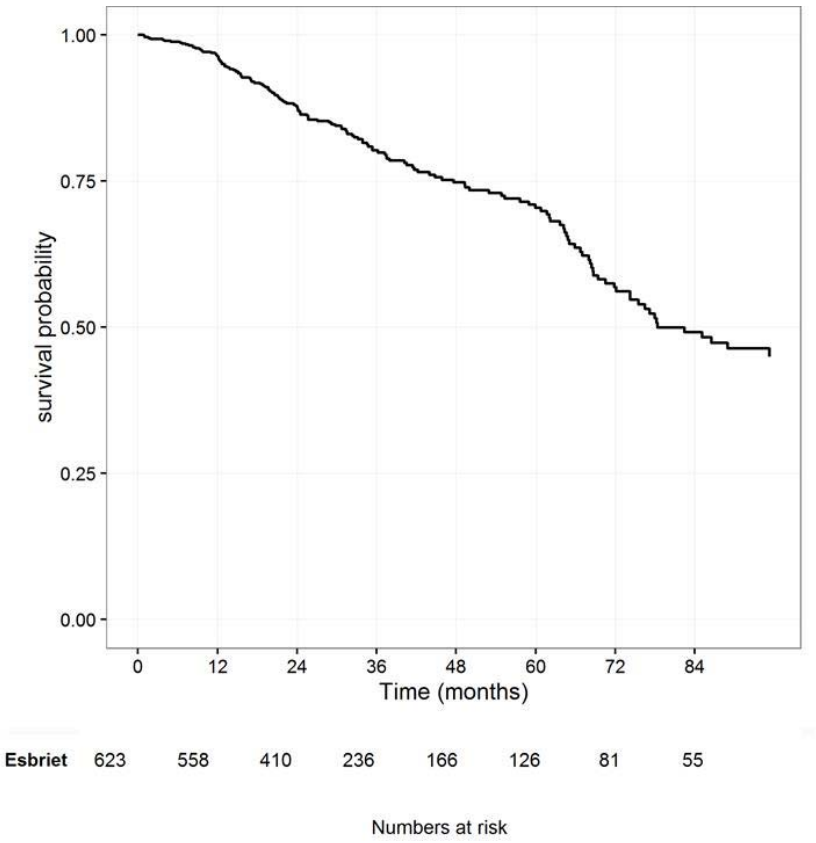
To facilitate comparison with outcomes from the 72-week CAPACITY trials,[49] subgroup analyses were conducted for those who had received placebo in the original trials and who either had baseline FVC and DLco values that met ASCEND[34] or CAPACITY[49] entry criteria (n=178) or did not (n=96)[40,63], although no results were reported in the CS.[4] The publication by Kreuter et al[63] found that discontinuation rates were highest in those patients who had originally received placebo and especially those who did not meet the ASCEND[34] or CAPACITY[49] entry criteria.

In total 603 patients were enrolled in RECAP from the CAPACITY trials.[63] Participants from the ASCEND trial[34] have also been eligible since 2014. The CS[4] (page 158) states that no published data analysis including ASCEND is available to date, but the text refers to CAPACITY/ASCEND data. No results were reported for this specific population in the CS.[4]

The RECAP study[40, 56] is ongoing. The most recent data-cut was performed in June 2015 and the next data-cut is planned in June 2016; some analyses based on summary data from this data-cut were provided by the company as an unpublished conference presentation.[64] As noted in Table 23, the primary objective was to evaluate the safety of pirfenidone 2,403mg per day: data on AEs are included in the integrated analyses set reported under Section 4.5, Table 29.

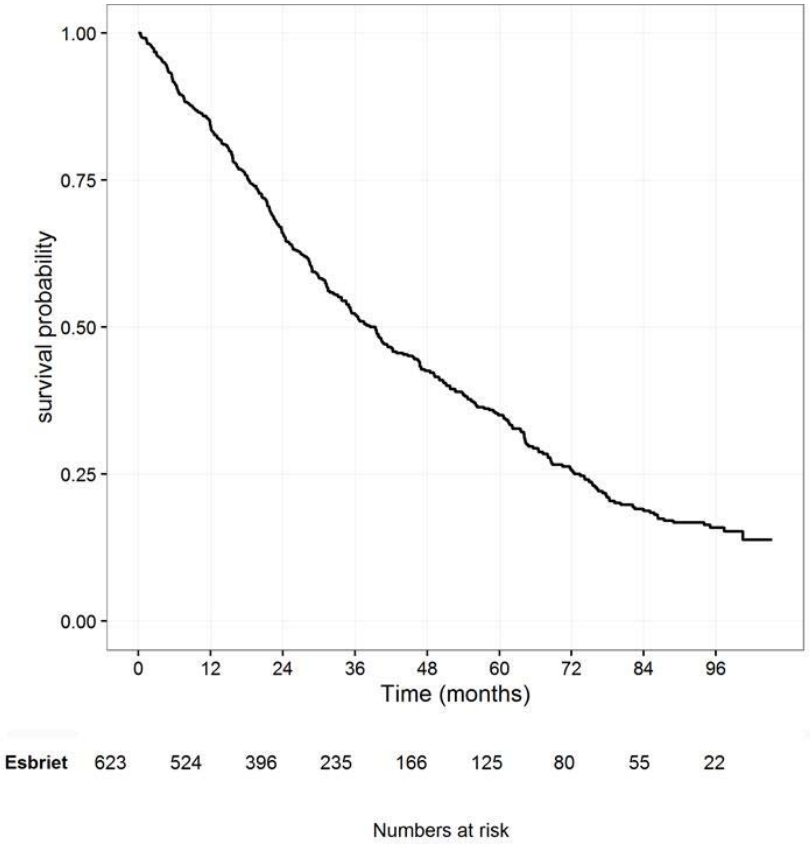
Survival data and time-on-treatment data were reported in the CS[4] (pages 159-161) and are presented here for patients who received pirfenidone 2,403mg per day from baseline onwards in CAPACITY and ASCEND, and through the RECAP extension period,[40, 56] for whom data are available through to 8.8 years (see Figure 22).

Figure 22: RECAP Kaplan Meier estimates for Overall Survival: patients continuing on pirfenidone 2,403mg per day (data cut: June 2015, reproduced from CS,[4] Figure 35)



Time on treatment data for these patients from the latest data-cut of RECAP are presented in Figure 23.

Figure 23: RECAP Kaplan Meier estimates for time on treatment: patients continuing on pirfenidone 2,403mg per day (data cut: June 2015) (reproduced from CS,[4] Figure 36)



The CS,[4] (pages 116-117)[4] also reported the following data [redacted]

[redacted]No details were provided in the CS[4] about how this study was identified or whether any other potentially relevant studies were excluded (for example, a second non-randomised open-label study, PIPF-002,[65] is only mentioned in the safety section of the CS[4] (Section 4.12); it was not reported how data were extracted or analysed, and no critical appraisal was conducted by the company or reported in the CS,[4] so the risk of bias affecting the RECAP study[40] is unknown.

However, it has been stated by Kreuter et al[63] that, *"the RECAP data must be interpreted with caution due to possible selection bias with regard to both pirfenidone (patients selected for tolerability and treatment response) and placebo (selection for mild progression because death or significant worsening led to informed drop out)."*

Registry data

The CS[4] then used IPD from selected registries with the aim of providing potential long-term comparative data for RECAP[40] based on "best supportive care." The CS[4] stated that the company contacted the holders of various registries reporting outcomes for patients with IPF in real-world practice were contacted, resulting in the availability of patient-level information from three registries: the Edinburgh registry, INOVA registry and the EuroIPF registry (see Table 26).

Table 26: Summary of available registries for best supportive care, registries with patient level demographic data (reproduced from CS,[4] Table 58, pages 164-165)

	Edinburgh	**INOVA**	**EuroIPF**
Geographic Region	UK	USA	Europe
Dates of registry information	1 January 2001 – 30 May 2014	November 1996 - June 2015	2008 - 2011
Patient population	• Incident IPF cases with a definite or possible UIP pattern on HRCT based on the 2011 ATS/ERS diagnostic guidelines for IPF • Event time available • Patients diagnosed up to 48 months prior to data collection date	Confirmed as incident IPF cases based on the 2011 ATS/ERS/JRS/ALAT diagnostic guidelines for IPF.	Verified diagnosis of IPF
n	323	815	409
Follow-up	Patients were followed from index date (date of IPF diagnosis) to date of death or May 30, 2014. Vital status was ascertained on May 30, 2014. Patients were censored on May 30, 2014, if their death could not be confirmed. None of the patients seen at this center underwent lung transplantation during the follow-up period, so this was not included as a censoring criterion for this cohort.	Patients were followed from index date (date of IPF diagnosis) to date of death or date of last visit. Date of last vital status is provided in the dataset. Patients were censored on their date of last visit, if their death could not be confirmed. If patients had a transplant, it was indicated in the dataset, but no dates were provided for treatment or transplant.	Patients were followed from index date (date of inclusion in registry) to date of death or date of last visit. Date of last visit and vital status check was provided. Patients were censored on date of last visit, if their death could not be confirmed
Treatments received during follow-up	BSC only	BSC only	BSC only
Inclusion/exclusion criteria applied to match ASCEND/CAPACITY	• FVC/VC<90% and/or DLco<90% • FVC/VC>50% • DLco >30% • FEV1/FVC>0.7 • Age 40 - 80 • Gender known	• FVC/VC<90% and/or DLco <90% • FVC/VC>50% • DLco >30% • FEV1/FVC>0.7 • Age 40 - 80 • Gender known	• FVC/VC<90% and/or DLco <90% • FVC/VC>50% • DLco >30% • FEV1/FVC>0.7 • Age 40 - 80 • Gender known

		• Event time available	• Event time available
Number of patients following application of ASCEND/ CAPACITY inclusion/ exclusion criteria	182	286	115
Parameters included in the propensity score model	• Age • Sex • Baseline %predicted FVC • Baseline %predicted DLco • First order interaction terms	• Age • Sex • Baseline %predicted FVC • Baseline %predicted DLco • Baseline FEV/FVC • First order interaction terms	• Age • Sex • Baseline %predicted FVC • Baseline %predicted DLco • Baseline FEV/FVC • Baseline smoking status • First order interaction terms
Number of patients remaining after trimming	125	254	89
Age, mean years ± SD	69.4 ± 7.6	66.2 ± 7.9	66.3 ± 8.4
Male (%)	72%	80%	85%
FVC ± SD	81.2 ± 12.4	70.9 ± 12.8	75.4 ± 14.3
DLco **± SD**	51.6 ± 11.8	46.5 ± 11.1	46.0 ± 10.6
FEV1/FVC ± SD	0.83 ± 0.07	0.83 ± 0.06	0.83 ± 0.07
Propensity score model	logOdds(Trial=1) = Age + Sex + DLco + FVC + Age* DLco + Age*FVC	logOdds(Trial=1) = Age + Sex + DLco + FVC +FEV/FVC + Age* DLco + Sex*FEV/FVC	logOdds(Trial=1) = Age + Sex + DLco + FVC +FEV/FVC + Smoke + Age*FVC + Age*Sex + Age*FEV/FVC + Sex*FVC + Sex*Smoke + DLco *Smoke
Key: *DLco, carbon monoxide diffusing capacity of the lungs; FEV, forced expiratory volume; FVC, forced vital capacity; HRCT, high-resolution computed tomography; IPF, idiopathic pulmonary fibrosis; SD, standard deviation; UIP, usual interstitial pneumonia.*			

To improve the comparability between the data from the ASCEND[34] and CAPACITY trials[49] and these registry data, a two-stage process was conducted. The company selected from these registries: (1) those patients that were considered most likely to satisfy the eligibility for the RECAP trial by applying the ASCEND and CAPACITY trials' inclusion and exclusion criteria, and; (2) applied a propensity score model that calculated the probability of being included in a clinical trial based on baseline characteristics, and excluding patients with unusual profiles based upon propensity-score based trimming (see CS,[4] page 162).

The CS[4] (page 163) argued that, based on the kernel density distributions for each of the logistic models post trimming for each of three registries, the INOVA and EuroIPF registries provided the most comparable patient sample to the patients in the pirfenidone Phase III RCTs. The comparative effectiveness estimated across the three registries was comparable with or better than the comparative effectiveness observed in the pooled ASCEND/CAPACITY data.[34, 49] Results were similar comparing the pooled hazard ratio versus BSC from ASCEND/CAPACITY[34, 49] and INOVA which represented the study with the largest sample size and most similar patient characteristics post trimming (HR 0.52 versus [redacted] see Table 27).[4] The CS[4] accepts that there are limitations in comparing data from a Phase III trial with real-world evidence.

Table 27: Overall survival comparison: pirfenidone versus BSC (from registry data) (reproduced from CS,[4] Table 57, page 163)

Outcome	Edinburgh registry	INOVA registry	EuroIPF registry	Pooled CAPACITY and ASCEND data
Hazard ratio for pirfenidone vs BSC (post trimming unadjusted data)	[redacted]	[redacted]	[redacted]	0.64 (0.41;0.99) at 72 weeks
Hazard ratio for pirfenidone vs BSC (post trimming data using propensity score model to adjust for remaining imbalances)	[redacted]	[redacted]	[redacted]	0.52 (0.31; 0.88) at 52 weeks
Key: BSC, best supportive care.				

In addition to the registries where IPD were available, three additional sources of supportive information were provided on probability of survival:

1. CPRD data (n=4,527) were obtained from 2000 to 2012 (inclusive), before pirfenidone was available in the UK .[66] Patients were selected based on the following criteria:

- A clinical or referral event record for IPF as defined by Read (general practices coding system in the UK) as specified in Navaratnam 2011.[67]
- No clinical or referral codes for connective tissue disease, extrinsic allergic alveolitis, sarcoidosis, pneumoconiosis, or asbestosis at any time in the patient record
- IPF events whilst alive and registered at an up-to-standard general practice
- At least 1 year of registration prior to the index date (date of IPF record)

To improve the similarity between the CAPACITY[49] and CPRD cohorts, the following restrictions were applied to the CPRD data:

- Survival times were adjusted using random-sampling of diagnosis to randomisation collected in the CAPACITY studies (n=2,888)[49]
- Patients with an FVC<50% were excluded, this was determined based on data within 1 month of the patient's index date (n=193)

Full propensity scoring was not possible as only FVC data were available for patients within the CPRD dataset. Standard care patients were followed up to 9.53 years; a median survival of 3.41 years was observed (95% CI: 2.67, 4.93).

2. Strand *et al.*[41] report overall survival for patients prospectively enrolled from the National Jewish Health Institutional Review Board-approved ILD database for patients between January 1, 1985 and January 1, 2011 diagnosed with IPF according to consensus guidelines. Median survival was 4.4 years (95% CI: 4.1-5.2) for IPF.

3. Kondoh *et al.*[50] retrospectively studied patients diagnosed with IPF based on ATS/ERS criteria.[50] Median survival was 3.7 years. A stepwise multivariate Cox regression model demonstrated the prognostic significance of FVC progression (10% decline in FVC at 6 months), acute exacerbations, BMI and disease severity measured via the modified MRC scale.

A summary of the characteristics of the patients contained within the three additional registries, and the patients in the CAPACITY/ASCEND trials,[34, 49] is provided in Table 28. The CS[4] states that the patients within the Strand registry[41] appear to be most similar to those in the CAPACITY[49] and ASCEND[34] trials.

Table 28: Summary of available registries for best supportive care, registries without patient level demographic data (reproduced from CS,[4] Table 59, page 168)

	CPRD	Strand[41]	Kondoh[50]	CAPACITY / ASCEND[34, 49]
Geographic region	UK	USA	Japan	Global
Data collection dates	2000 - 2012	Jan 1985 – Jan 2011	Jan 2000 - Dec 2005	
Patient population	ICD10 codes: H563.00 H563.11 H563.12 H563100 H563z00	Subgroup diagnosed with IPF according to consensus guidelines including ATS/ERS	Patients diagnosed with IPF based on ATS/ERS criteria	Diagnosis of IPF in accordance with the ATS international consensus statement
n	193 in FVC reported and ≥50 subgroup	321	74	623 on high dose PFN arms
Age, mean years ± SD	73.5 ± 9.2	66.1 ± 9.1	64.1 ± 7.4	67.2 ± 7.6
Male (%)	68%	75%	82%	74%
FVC ± SD	79.3 ± 15.7	71.4 ± 17.4	77.0 ± 19.2	67.8 ± 11.2
DLco ± SD	NR	52.3 ± 18.7	59.3 ± 18.7	47.1 ± 9.7

Key: DLco, carbon monoxide diffusing capacity of the lungs; FEV, forced expiratory volume; FVC, forced vital capacity; HRCT, high-resolution computed tomography; IPF, idiopathic pulmonary fibrosis; SD, standard deviation; UIP, usual interstitial pneumonia; PFN, pirfenidone; NR, not reported

4.5 Safety evidence

Serious adverse events

In the ASCEND trial[34] there were 55 patients (19.8%) and 69 patients (24.9%) in the pirfenidone and placebo groups, respectively, who experienced a serious AE (see Table 29). The most common serious AE was "worsening of IPF", which was reported in 7 patients (2.5%) in the pirfenidone group and 27 patients (9.7%) in the placebo group. According to trial protocols, "worsening of IPF" is defined as, "acute IPF exacerbation, IPF-related death, lung transplant or respiratory hospitalization, whichever comes first."[36] "Worsening of IPF" was not specifically categorised as either an efficacy outcome (see CS, [4] pages 96-99 and 104-107) or a safety outcome (unless it could be designated as certainly due to the drug), but its presence was simply reported by the investigator (see clarification response,[10] question A18). The other most frequently-reported serious AEs in the pirfenidone arm were pneumonia, prostate cancer, angina pectoris, nausea, congestive cardiac failure and rib fracture (see Table 29). Other than the more frequent occurrence of "worsening of IPF" in the placebo arm, none of the differences in serious AEs between arms was statistically significant.

Table 29: Serious treatment-emergent adverse events reported by ≥2 patients in ASCEND at 52 weeks (reproduced from clarification response,[10] question A25)

Adverse event	Number of patients, n (%)		Rate ratio (95% CI)	Pr>chi2
	PNF 2,403mg/d (n=278)	Placebo (n=277)		
Worsening of Idiopathic Pulmonary Fibrosis	7 (2.5)	27 (9.7)	0.26 (0.11, 0.58)	<0.001
Pneumonia	11 (4.0)	14 (5.1)	0.78 (0.36, 1.69)	0.533
Prostate Cancer (*M)	2 (0.7)	4 (1.4)	0.50 (0.09, 2.70)	0.409
Angina Pectoris	3 (1.1)	0 (0.0)		0.083
Nausea	3 (1.1)	0 (0.0)		0.083
Atrial Fibrillation	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Bronchitis	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Dyspnoea	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Pulmonary Embolism	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Septic Shock	1 (0.4)	2 (0.7)	0.50 (0.05, 5.46)	0.561
Cardiac Failure Congestive	2 (0.7)	0 (0.0)		0.157
Rib Fracture	2 (0.7)	0 (0.0)		0.157
Aortic Aneurysm	0 (0.0)	2 (0.7)	0.00	0.156
Gastroenteritis Viral	0 (0.0)	2 (0.7)	0.00	0.156

*Each patient is counted only once for each preferred term. For terms followed by (*M), percentages are based on the number of males within each treatment group. Preferred terms are listed in order of decreasing frequency in the total study population.*
TE SAE = treatment-emergent serious adverse events, defined as occurring after the first dose and within 28 days after the last dose of study treatment.

The CS[4] did not report any serious AEs for CAPACITY 1 & 2, but these were reported in the publication[49] and the clarification response,[10] (question A25). The principal serious AEs for pirfendione, excepting IPF, occurring in >2 patients in any pirfenidone group are reported in Table 30.

Table 30: Serious treatment-emergent adverse events reported by ≥2 patients in CAPACITY 1 & 2 at 72 weeks[49]

	CAPACITY 1[49]		CAPACITY 2[49]		
Adverse event, n (%)	**PFN 2,403mg/d (n=171)**	**PBO (n=173)**	**PFN 2,403mg/d (n=174)**	**PFN 1,197mg/d (n=87)**	**PBO (n=174)**
Pneumonia	7 (4.1)	7 (4.0)	4 (2.3)	3 (3.4)	6 (3.4)
Respiratory failure	4 (2.3)	6 (3.5)	2 (1.1)	3 (3.4)	2 (1.1)
Angina pectoris			2 (1.1)	2 (2.3)	1 (0.6)
Atrial fibrillation	2 (1.1)	1 (0.6)	1 (0.6)	3 (3.4)	1 (0.6)
Coronary artery disease	6 (3.5)	0 (0)	0	3 (3.4)	2 (1.1)
Acute renal failure	2 (1.2)	2 (1.2)	1 (0.6)	2 (2.3)	0 (0)
Fall	2 (1.2)	1 (0.6)			
Hypotension	2 (1.2)	1 (0.6)			
Colitis	2 (1.2)	0 (0)			
Hip fracture	2 (1.2)	0 (0)			
Prostate cancer	2 (1.6)*	0 (0)			
Interverterbral disc profusion	2 (1.2)	0 (0)			
Liver test function abnormal	2 (1.2)	0 (0)			
Nephrolithiasis	2 (1.2)	0 (0)			
Sick sinus syndrome	2 (1.2)	0 (0)			
Pneumothorax			3 (1.7)	2 (2.3)	0
Pulmonary embolism			1 (0.6)	3 (3.4)	1 (0.6)
Syncope			3 (1.7)	1 (1.1)	1 (0.6)
Chest pain			3 (1.7)	0	0
Bladder cancer			2 (1.1)	0	0
Gastroesophageal reflux disease			2 (1.1)	0	0

* Male patients only

None of the differences in serious AEs between arms, including IPF, were statistically significant within the CAPACITY trials[49] (see clarification response,[10] question A25).

The publications for SP3[38] did not report AE data, but these were provided by the company in response to a request for clarification from the ERG (see clarification response addendum,[30] question A28)[10]: serious AEs occurring in ≥1% of participants in the pirfenidone arm were pneumonia (5.5% for pirfenidone versus 2.8% for placebo), bronchitis (1.8% versus 1.9%), worsening of IPF (5.5% versus 4.7%) and pneumothorax (1.8% versus 2.8%). The SP2[39] publications did not report any serious AEs and there was no additional information available for this trial (see clarification response,[10] question A25).

The company conducted a meta-analyses for treatment-emergent serious AEs using data from ASCEND,[34] CAPACITY 1 & 2[49] and SP3[38] at week 52 (see Figure 24) and at 72 weeks (see Figure 25) using data from CAPACITY 1 & 2[49] only (see CS,[4] Appendix 9, page 76). Both analyses showed no difference between the pirfenidone and placebo groups (OR: 0.90, 95% CI: 0.70 to 1.15, *p*-value not reported) and (OR: 1.06, 95% CI: 0.77 to 1.46, *p*-value not reported).

Figure 24: Forest plot of odds ratios for treatment emergent serious adverse events at week 52 (reproduced from CS,[4] Appendix 9)

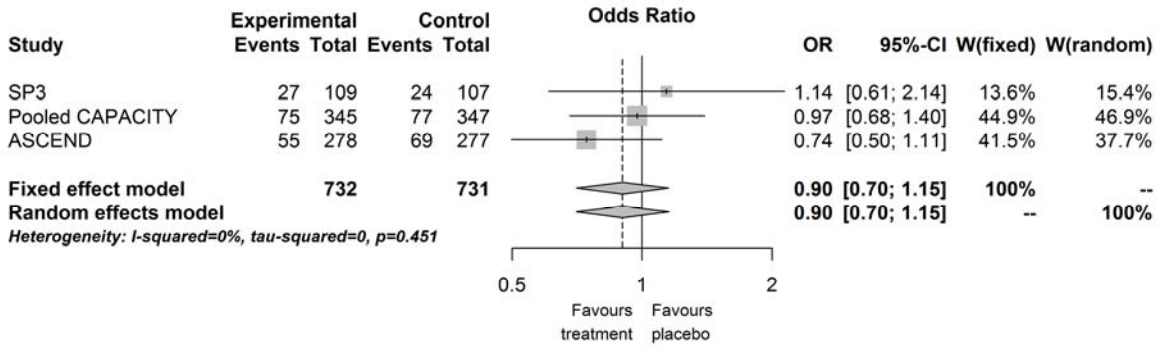
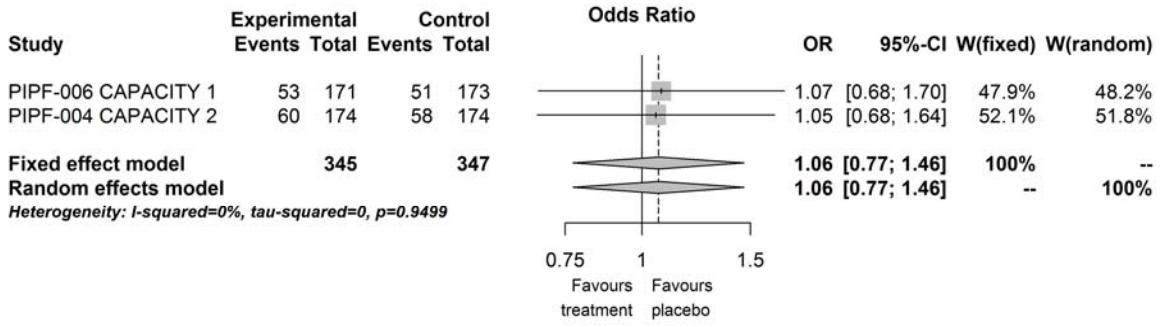


Figure 25: Forest plot of odds ratios for treatment emergent serious adverse events at week 72 (reproduced from CS,[4] Appendix 9)



Adverse events leading to discontinuation of treatment

In the ASCEND trial,[34] the proportion of patients discontinuing treatment due to an AE was 14.4% (n=40) in the pirfenidone group and 10.8% (n=30) in the placebo group. The most common AE leading to treatment discontinuation was worsening IPF (1.1% [n=3] in the pirfenidone group versus 5.4% [n=15] in the placebo group), but again the caveats should be noted regarding the categorisation of this event as a safety outcome. The only other AEs leading to discontinuation of treatment in ≥1% of patients in the pirfenidone group were elevated hepatic enzymes levels, pneumonia, rash and decreased weight, which occurred in 3 patients (1.1%) in each trial arm.

In the CAPACITY trials,[49] treatment was discontinued due to AEs in 15% (n=51) of 345 patients in the pooled pirfenidone 2,403mg/day group compared with 9% (n=30) of 347 patients in the placebo group. The most common AE leading to discontinuation was worsening of IPF (3% in both groups). The other AEs leading to discontinuation of treatment in ≥1% of patients in the pirfenidone group were provided by the company in response to a request for clarification from the ERG (see clarification response addendum,[30] question A24). In CAPACITY 1,[49] these were elevated IPF (2.3% in each arm), photosensitivity, rash and respiratory failure, which each occurred in 2 patients (1.2%) in the pirfenidone trial arm but not at all in the placebo arm. In CAPACITY 2,[49] for the 2,403mg per day dose, these were elevated IPF (1.1% for pirfenidone versus 1.7% for placebo), bladder cancer (1.1% vs 0%), nausea (2.3% versus 0%) and rash (1.7% versus 0%). The following substantial laboratory abnormalities (Grade 4 or a shift of 3 grades e.g. from 0 to 3) occurred more frequently in the CAPACITY 1 and 2 pooled pirfenidone 2,403mg/day group compared with placebo: hyperglycaemia (1% [n=4] versus <1% [n=3], respectively); hyponatraemia (1% [n=5] versus 0%); hypophosphatemia (2% [n=6] versus <1% [n=3]); and lymphopenia (1% [n=5] versus 0). However, none were associated with clinically significant consequences. More patients in the pooled pirfenidone-treated group than in the pooled placebo group had elevations in alanine aminotransferase and aspartate aminotransferase of more than 3x the upper limit of normal (4% [n=14] versus <1% [n=2]). However, all reports were reversible and without clinical sequelae.

SP2[39] reported that 11 patients discontinued pirfenidone treatment, compared with 2 patients in the placebo arm, due to AEs.[39] The CS (page 172) stated that skin photosensitivity was the AE that was principally responsible for discontinuing or reducing pirfenidone dose; full data on AE discontinuations were provided in the publication:[39] the principal AEs affecting discontinuation from pirfenidone treatment were: photosensitivity (n=5); vomiting (n=1); fever (n=1); abnormality of hepatic function (n=1); dizziness (n=1); facial paralysis (n=1) and hepatoma (n=1). There were no instances of any of these events in the placebo arm.

SP3[38] reported that 15 patients in the high dose group (1,800mg/d) and 9 patients in the low dose group (1,200mg/d), compared with 7 patients in the placebo group, discontinued the study due to AEs. The CS did not report details of these adverse events, but the publication did so:[38] the principal adverse events affecting discontinuation from pirfenidone treatment were: photosensitivity (n=5); lung carcinoma (n=3); fever (n=2); respiratory failure (n=2); rash (n=2) and an increase in aspartate aminotransferase (AST) and/or alanine amino-transferase (ALT) (n=2).

All adverse events

The most common "treatment-emergent" AEs with higher incidence in the pirfenidone group were primarily gastrointestinal and skin-related events. The CS[4] reported data for any AE with a frequency of at least 15% in any arm (in ASCEND)[34] or a frequency of at least 10% and 1.5 times in the pirfenidone arm compared with the placebo arm (in the CAPACITY trials).[49] Nausea was the most frequent AE: 36% in the pirfenidone arm compared with 13.4% in the placebo arm in ASCEND, and 36% in the pirfenidone arm compared with 17% in the placebo arm CAPACITY trials (*p*-values not reported, see Table 31).[49] The second most frequent event was rash: 28.1% in the pirfenidone arm compared with 8.7% in the placebo arm in ASCEND,[34] and 32% in the pirfenidone arm compared with 12% in the placebo arm in the CAPACITY trials.[49] Dyspepsia was also much more frequent in the pirfenidone arms than the placebo arm: 17.6% in the pirfenidone arm compared with 6.1% in the placebo arm in ASCEND,[34] and 19% in the pirfenidone arm compared with 7% in the placebo arm in the CAPACITY trials;[49] as was anorexia: 15.8% in the pirfenidone arm compared with 6.5% in the placebo arm in ASCEND,[34] and 11% in the pirfenidone arm compared with 4% in the placebo arm in the CAPACITY trials;[49] and dizziness: 17.6% in the pirfenidone arm compared with 13% in the placebo arm in ASCEND, and 18% in the pirfenidone arm compared with 10% in the placebo arm in the CAPACITY trials.[49]

Headache, cough, diarrhoea, fatigue and upper respiratory tract infection were all frequent (between 20% and 26%), but were similar across pirfenidone and placebo arms (see Table 31) According to the CS[4] (page 171), no instances of Stevens-Johnson syndrome or toxic epidermal necrosis were reported in the CAPACITY trials.[49]

Table 31 Adverse events in ASCEND at 52 weeks and CAPACITY 1 & 2 at 72 weeks (adapted from CS,[4] Tables 60 and 61)

	ASCEND*[34]		CAPACITY 1 & 2†[49]	
Adverse event, n (%)	**PFN (n=278)**	**PBO (n=277)**	**PFN (n=345)**	**Placebo (n= 347)**
Nausea	100 (36)	37 (13.4)	125 (36)	60 (17)
Rash	78 (28.1)	24 (8.7)	111 (32)	40 (12)
Headache	72 (25.9)	64 (23.1)		
Cough	70 (25.2)	82 (29.6)		
Diarrhoea	62 (22.3)	60 (21.7)		
Upper respiratory tract infection	61 (21.9)	56 (20.2)		
Fatigue	58 (20.9)	48 (17.3)		
Dizziness	49 (17.6)	36 (13)	63 (18)	35 (10)
Dyspepsia	49 (17.6)	17 (6.1)	66 (19)	26 (7)
Anorexia	44 (15.8)	18 (6.5)	37 (11)	13 (4)
Dyspnoea	41 (14.7)	49 (17.7)		
Vomiting			47 (14)	15 (4)
Photosensitivity reaction			42 (12)	6 (2)
Anorexia			37 (11)	13 (4)
Arthralgia			36 (10)	24 (7)
Insomnia			34 (10)	23 (7)
Abdominal distension			33 (10)	20 (6)

** Occurring in ≥15% of patients in either treatment group; † Occurring in ≥10% of patients on pirfenidone and with an incidence of 1.5 x greater than that in patients receiving placebo; PFN: pirfenidone 2,403mg/day; PBO: placebo*

The SP3[38] and SP2[39] trials also reported a relatively high incidence of the following AEs for pirfenidone compared with placebo: photosensitivity; anorexia; dizziness; nausea; heartburn; fatigue and elevated gamma-GTP (see Table 32). *P*-values were reported for the SP2 trial[39] and the incidence of many of the AEs was significantly higher in the pirfenidone group than the placebo group (see Table 32). Respiratory infections were reported to be more common in patients treated with placebo.

The CS[4] (page 173) stated that most of the AEs reported for SP2[39] disappeared with decrease of the dose or temporarily holding the medication.

It is unclear why there is some inconsistency between trials in the frequency of some AEs, such as photosensitivity, nausea and anorexia.

Findings on AEs from Huang *et al*[48] were consistent with the other published trials, including, for example, the significantly higher incidence of rash in patients receiving pirfenidone.

Table 32: Adverse events reported from the SP3 at 52 weeks and SP2 at 26 weeks (adapted from CS,[4] Tables 62 and 63)

	SP3*[38]			SP2†[39]	
Adverse event n (%)	**PFN 1,800mg/d (n=109)**	**PFN 1,200mg/d (n=55)**	**Placebo (n=107)**	**PFN 1,800mg/d (n=72)**	**PBO (n=32)**
Any adverse event§	109 (100.0)	54 (98.2)	106 (99.1)	72 (98.6)	32 (88.9)
Photosensitivity§	56 (51.4)	29 (52.7)	24 (22.4)	32 (43.8)	0 (0)
Anorexia§	18 (16.5)	6 (10.9)	3 (2.8)	23 (31.5)	2 (5.6)
Abdominal discomfort	3 (2.8)	4 (3.7)	0 (0.0)	22 (30.1)	3 (8.3)
Nausea§				16 (21.9)	2 (5.6)
Heartburn				12 (16.4)	1 (2.8)
Fatigue§				16 (21.9)	1 (2.8)
Dizziness	8 (7.3)	0 (0.0)	1 (0.9)		
Nasopharyngitis	54 (49.5)	30 (54.5)	70 (65.4)		
Upper respiratory tract infection	1 (0.9)	3 (5.5)	9 (8.4)	12 (16.4)	3 (8.3)
γ-GTP elevation§	25 (22.9)	12 (21.8)	10 (9.3)	20 (27.4)	3 (8.3)

* With an incidence of ≥5%; † With an incidence of ≥10% at six months; § Difference between pirfenidone 1800mg per day and placebo is significant at level of $p<0.05$ or better in trial SP2; PFN: pirfenidone 2,403mg/day; PBO: placebo

Integrated analysis of safety data from ASCEND, CAPACITY 1 & 2, and two ongoing open-label studies

Data from the three principal Phase III trials (ASCEND,[34] CAPACITY 1 & 2[49]) were analysed together with data from the non-randomised, non-controlled, OLE study that included a set of patients who completed either ASCEND, CAPACITY 1 or 2 (RECAP)[40] (see Section 4.4) and PIPF-002, an ongoing open-label compassionate-use study in US patients with either IPF or secondary pulmonary fibrosis.[68] No critical appraisal was reported for either the RECAP[40] or the PIPF-002 study.[65] Safety outcomes were assessed from baseline until 28 days after study drug discontinuation.

The latest interim analyses of the integrated population were conducted using a data cut-off date of 17 January 2014.[69] A total of 1,299 patients were included in the integrated population and the reported data only concern AEs occurring in at least 15% of patients in the cumulative clinical database. The cumulative total exposure to pirfenidone was 3,160 person exposure years (PEY). The median duration of exposure was 1.7 years (range, 1 week–9.9 years); 545 (42%) patients received pirfenidone for ≥2 years and 325 (25%) patients received pirfenidone for ≥4 years. The majority of patients (n=964, 74.2%) received a mean daily dose between 1,800mg and 2,600mg. Cumulative safety outcomes in the pooled pirfenidone 2,403mg/day and placebo treatment groups in the Phase III studies were presented for comparison (see Table 33).

Table 33: AEs in the integrated population compared with the pooled pirfenidone 2,403mg/day and placebo groups from the ASCEND and CAPACITY 1 & 2 trials* (reproduced from CS,[4] Table 64)

	Integrated population† (n=1,299)	Pooled ASCEND, CAPACITY 1 & 2 population	
		PFN (n=623)	PBO (n=624)
Median duration of exposure, years (range)	1.7 (>0, 9.9)	1.0 (>0, 2.3)	1.0 (>0, 2.3)
Treatment-emergent adverse event, %			
Nausea	37.6	36.1	15.5
Cough	35.1	27.8	29.2
Dyspnoea	30.9	16.9	20.2
Upper respiratory tract infection	30.6	26.8	25.3
Idiopathic pulmonary fibrosis	29.3	13.0	19.9
Fatigue	28.2	26.0	19.1
Diarrhoea	28.1	25.8	20.4
Rash	25.0	30.3	10.3
Bronchitis	23.8	14.1	15.4
Headache	21.6	22.0	19.2
Nasopharyngitis	21.3	16.7	17.9
Dizziness	21.2	18.0	11.4
Dyspepsia	18.4	18.5	6.9
Vomiting	15.9	13.3	6.3
Weight decreased	15.6	10.1	5.4
Back pain	15.4	10.4	10.4
Anorexia	15.2	13.0	5.0

**Occurring in ≥15% of patients in the cumulative clinical database*
†Includes two patients from PIPF-002 with a diagnosis of "pulmonary fibrosis"
PFN: pirfenidone 2,403mg/d; PBO: placebo

The findings for the integrated population are consistent with the findings of the ASCEND[34] and CAPACITY trials[49] (though not always with the SP3[38] and SP2[39] trials), i.e. gastrointestinal and skin-related events were among the most common AEs. The CS[4] (page 174) states that these were mainly mild to moderate in severity, reversible, and rarely led to treatment discontinuation. Elevations in liver enzymes (ALT or AST >3 x Upper Limit of Normal [ULN]) occurred in 40/1,299 (3.1%) patients in the integrated population, compared with 23/623 (3.7%) and 5/624 (0.8%) in the pooled pirfenidone and placebo groups in the Phase III trials. All elevations were reversible without clinical sequelae. Respiratory AEs were more common in the integrated population than the placebo and pirfenidone-treated patients from the pooled Phase III trials. The CS[4] (page 175) states that this finding is expected from a chronic progressive respiratory disease followed over a long period of observation.

The CS[4] stated that the safety and AE profile of pirfenidone is different from that of nintedanib, for which most frequently reported adverse reactions are diarrhoea, nausea and vomiting, abdominal pain, decreased appetite, weight loss and elevation of hepatic enzymes.[1] However, the pirfenidone trials did

report nausea, diarrhoea, vomiting, weight loss and anorexia as frequent events (see Table 30 and Table 31).

4.5.1 Ongoing studies

As noted above, there are two ongoing studies to evaluate safety: the non-randomised, non-controlled, OLE study that included a set of patients who completed either ASCEND, CAPACITY 1 or CAPACITY 2 (RECAP) (see Section 4.5, the final data collection date is listed as December 2015 https://clinicaltrials.gov/ct2/show/record/NCT00662038) and PIPF-002, an ongoing open-label compassionate-use study in US patients with either IPF or secondary pulmonary fibrosis (https://clinicaltrials.gov/ct2/show/NCT00080223), which has a listed completion date of April 2015.

4.6 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

In the absence of any direct head-to-head RCTs comparing pirfenidone and nintedanib, for the treatment of IPF, the company conducted an NMA. This is an extension of the conventional pairwise meta-analysis that can be used to combine direct and indirect evidence about treatment effects across RCTs that share at least one treatment in common with at least one other study.

The company conducted a systematic review to collate the published RCTs which assess the efficacy and safety of therapies prescribed for the treatment of IPF. The inclusion criteria for the NMA systematic review were as follows (see CS,[4] pages 122-123): the population of interest was adults (aged 18 or older) with suspected or diagnosed IPF; the interventions of interest were pirfenidone, double therapy (prednisone and azathioprine), N-acetylcysteine (NAC), nintedanib, and triple therapy (prednisone and azathioprine and NAC); the relevant study designs were Phase II or Phase III RCTs and the outcomes of interest included lung capacity, gas transfer, physical functioning (6MWD), PFS, adverse effects of treatment, HRQoL measured using SGRQ, SOBQ, dyspnoea score or EQ-5D, hospitalisations, acute exacerbations, mortality (all cause or IPF-related), categorical declines in FVC (0%, 5% and 10%), discontinuation and compliance of study treatments.

The systematic review methods undertaken for the NMA (e.g. literature searching, study selection, data extraction and quality assessment) were the same as those undertaken for the pirfenidone systematic review. As noted in Section 4.1.1, adequate systematic searches were undertaken to identify all relevant RCT studies assessing the efficacy and safety of NAC, nintedanib and triple therapy for the treatment of IPF. NAC, triple therapy and double therapy were not included in the NICE scope as comparators of interest, however, the company's literature search was developed to support submissions of pirfenidone to all national agencies and as such some comparators of interest included in the searches were beyond the scope of this appraisal.

Studies included in NMA

The company's systematic review identified 10 RCTs of reasonable methodological quality that compared pirfenidone, nintedanib, NAC, or triple therapy with placebo in patients with IPF. However, the company excluded two of the trials; SP2[39] (pirfenidone) and IFIGENIA[70, 71] (double and triple therapy) from the NMA. INFIGENIA[70, 71] was excluded from the NMA as the trial compares double and triple therapy, which are not comparators of interest for this appraisal. SP2[39] was excluded from the NMA as it was considered as an outlier by the NICE Appraisal Committee for the review of nintedanib (TA379)[12] and there was no useable data at one year as the trial was stopped early at 36 weeks. In addition, a non-valid primary end point, SpO_2, was used.

A total of eight studies were included in the company's NMA: ASCEND[24] (pirfenidone), CAPACITY 1[49] (pirfenidone), CAPACITY 2[49] (pirfenidone), SP3[38] (pirfenidone), IMPULSIS 1[72] (nintedanib), IMPULSIS 2[72] (nintedanib), TOMORROW[73] (nintedanib) and PANTHER[74, 75] (NAC and triple therapy). However, not all trials presented outcome data that could contribute to each NMA for all outcomes.

The ERG notes that although not in the final NICE scope,[3] the evidence network includes NAC and triple therapy. The trials of comparators contributing data to the NMA were all placebo-controlled RCTs and therefore all comparisons were made with placebo (see Figure 26). The ERG therefore believes that PANTHER[74, 75] has little influence on the NMA results for nintedanib and pirfenidone, and therefore data from PANTHER[74, 75] have been excluded from the additional analyses performed by the ERG in Section 4.8. In this section, only data from the trials of relevance to the decision problem are summarised.

A summary of the design and study characteristics of the studies included in the NMA is provided in Table 34.

Table 34: Summary of trials included in the company's NMA: (adapted from CS,[4] Table 12, page 66-67 and Appendix 10)

Study	Design, Location	Population	Treatment, dose and sample size (used in NMA)	Study durations (week)	Key outcomes measured in NMA
Pirfenidone					
ASCEND [34]	Phase III, randomised, double-blind, placebo-controlled trial. **Location** 127 sites (no sites in UK)	• Patients aged 40–80 years with confident clinical and radiographic diagnosis of IPF, in accordance with the International consensus statement [ATS, 2000] of >6 months but <48 months before randomisation, confirmed by central review. • FVC (% predicted value) 50-90% • DLco 30-90% • 6MWD ≥150 m • No improvement of IPF in preceding year.	Pirfenidone 2,403mg/day (n=278) Placebo (n=277)	52 weeks	**Primary outcomes** Change in percent predicted FVC or death at week 52. **Secondary outcomes** Change from baseline to Week 52 in 6MWD and PFS, change in dyspnoea (UCSD SOBQ); rate of death from any cause and the rate of death from IPF.
CAPACITY 2[49]	Phase III, randomised, double-blinded, placebo-controlled trial **Location** 110 centres (including 3 sites in the UK)	• Patients aged 40–80 years with confident clinical and radiographic diagnosis of IPF, in accordance with the International consensus statement [ATS, 2000] in the previous 48 months. • FVC (% predicted value) ≥50% at Screening and Day 1 (before randomisation) • DLco ≥35% • FVC or DLco ≤90% • No improvement of IPF in preceding year	Pirfenidone 2,403mg/day, (n=174) Placebo (n=174)	72 weeks	**Primary outcomes:** Change in percent predicted FVC from baseline to week 72. **Secondary outcomes:** Categorical FVC, PFS, worsening IPF, dyspnoea (SOBQ), 6MWD, worst SpO2 during the 6MWT, % predicted DLco, and fibrosis by use of HRCT.
CAPACITY	Phase III,	• Patients aged 40–80 years with	Pirfenidone	52 weeks	**Primary outcomes:**

1[49]	randomised, double-blinded, placebo-controlled trial **Location** 110 centres(no UK sites)	confident clinical and radiographic diagnosis of IPF, in accordance with the International consensus statement [ATS, 2000] in the previous 48 months. • FVC (% predicted value) ≥50% at Screening and Day 1 (before randomisation) • DLco ≥35% • FVC or DLco ≤90% • No improvement of IPF in preceding year	2,403mg/day (n=171) Placebo (n=173)		Change in percent predicted FVC from baseline to week 72. **Secondary outcomes:** Categorical FVC, PFS, worsening IPF, dyspnoea (SOBQ), 6MWD, worst SpO2 during the 6MWT, % predicted DLco, and fibrosis by use of HRCT.
SP3[38]	Phase III, randomised, double-blind, placebo-controlled trial. **Location** 73 centres in Japan.	• Patients aged 20 -75 years, with confident clinical and radiographic diagnosis of IPF in accordance with the International consensus statement [ATS/ERS, 2000]. • O2 desaturation of 5% between resting SpO2 and min SpO2 during 6MET • SpO2 >85% during 6MET (air). • No decrease in symptoms during the preceding 6 months	Pirfenidone 1,800mg/day (n=108) Placebo (n=104)	52 weeks	**Primary outcomes:** Change in VC from baseline to week 52 (originally was the change in lowest SpO2 during the 6MWT). **Secondary outcomes:** PFS time, change in the lowest SpO2 during the 6MWT
Nintedanib					
TOMORROW[73]	Phase II, randomised, double-blind, placebo-controlled trial **Location** 92 sites in 25	• Patients >40 years of age with diagnosis of IPF in accordance with ATS and ERS criteria and who had received the diagnosis of IPF <5 years before screening • Patients had to have undergone HRCT <1 year before randomisation	Nintedanib 300mg/day (n=86) Placebo (n=85)	52 weeks	**Primary outcome** The annual rate of decline in FVC. **Secondary outcome** Changes from baseline in percent predicted FVC and DLco; changes in SpO_2 and TLC (as measured by body plethysmography); 6MWD; SGRQ; a decrease from baseline in FVC of > 10%

	countries including UK	• FVC (% predicted value) ≥50% • DLco (% predicted value) 30 to 79% • PaO2 when breathing ambient air that was 55 mm Hg or greater at altitudes up to 1500m or a PaO2 of 50mm Hg or greater at altitudes above 1500 m.			or > 200 ml; SpO_2 decrease of more than 4%; incidence of acute exacerbations; survival at 52 weeks; and death from a respiratory cause
INPULSIS 1[72]	Phase III, randomised, double-blind, placebo-controlled trial **Location** 98 study sites including UK	• Age > 40 years; • IPF diagnosed, according to most recent ATS, ERS, JRS, ALAT IPF guideline for diagnosis and management, within 5 years; • Combination of HRCT pattern, and if available surgical lung biopsy pattern, as assessed by central reviewers, are consistent with diagnosis of IPF • Dlco (corrected for Hb): 30%-79% predicted of normal; • FVC> 50% predicted of normal	Nintedanib 150mg/bid (n=309) Placebo (n=204)	52 weeks	**Primary outcome** Annual rate of decline in FVC (mL) from baseline to week 52. **Secondary outcome** Time to the first acute exacerbation, change from baseline in SGRQ total score, acute exacerbations, absolute change from baseline in FVC (mL) and as a % predicted value over the 52-week treatment period, proportion of patients with an FVC response, risk of acute exacerbation, change from baseline in SGRQ domain scores over the 52-week treatment period, death from any cause, death from a respiratory cause, and death that occurred between randomisation and 28 days after the last dose of the study drug.
INPULSIS 2[72]	Phase III, randomised, double-blind, placebo-controlled trial **Location**	• Age > 40 years; • IPF diagnosed, according to most recent ATS, ERS, JRS, ALAT IPF guideline for diagnosis and management, within 5 years; • Combination of HRCT pattern, and if available surgical lung biopsy pattern, as assessed by	Nintedanib 150mg/bid (n=329) Placebo (n=219)	52 weeks	**Primary outcome** Annual rate of decline in FVC (mL) from baseline to week 52. **Secondary outcome** Time to the first acute exacerbation, change from baseline in SGRQ total score, acute exacerbations, absolute

	108 study sites, no sites in UK	central reviewers, are consistent with diagnosis of IPF • Dlco (corrected for Hb): 30%-79% predicted of normal; • FVC> 50% predicted of normal			change from baseline in FVC (mL) and as a % predicted value over the 52 week treatment period, proportion of patients with an FVC response, risk of acute exacerbation, change from baseline in SGRQ domain scores over the 52 week treatment period, death from any cause, death from a respiratory cause, and death that occurred between randomisation and 28 days after the last dose of the study drug.

Note: only trials relevant to the decision problem are reported

ALAT, Latin American Thoracic Association ATS; American Thoracic Society; bid, twice a day; DLco, Diffusing capacity of the lungs for carbon monoxide; EQ-5D, The EuroQoL Group 5-Dimension Self-Report Questionnaire; ERS, European Respiratory Society; 6MWD, 6-Minute walking distance; 6MWT, 6-Minute walking distance; FVC, Forced vital capacity; HRCT, High-resolution computed tomography; ICECAP, Investigating Choice Experiments for the Preferences of Older People Capability Instrument; IPF, Idiopathic pulmonary fibrosis; JRS, Japanese Respiratory Society; MedDRA, Medical Dictionary for Regulatory Activities; 6MET, 6 min exercise test; mL, millilitres; PaO2, Partial pressure arterial oxygen; PFS, progression-free survival;SF-36; Medical Outcomes Study 36-Item Short-Form Health Survey; SGRQ, St. George's Respiratory Questionnaire; SpO2, Peripheral oxygen saturation; tid, three times a day; TLC, total lung capacity; UCSD SOBQ, University of California San Diego Shortness of Breath Questionnaire and VC, Vital capacity

The main differences noted between the studies relate to definition of the outcomes, patient characteristics, methods used for handling missing data, and the time period of outcome assessment. The CS states that "*due to the limited number of studies contributing to each network, a pragmatic approach was adopted, whereby trials were included regardless of minor differences in outcome definitions, timing of assessment and analysis methods. It was assumed that the differences in definitions and methods did not influence the relative treatment effects*" (CS,[4] page 153). The main sources of heterogeneity are discussed in turn below.

Handling of missing data

In ASCEND[34] and CAPACITY 1 & 2[49], missing values as a result of death were assigned the worst rank in the ANCOVA analyses, and worst possible outcome in mean change analyses (e.g., FVC=0) and categorical analyses. Other missing data were imputed with the average value from three patients with the smallest sum of squared differences at each visit with data that were not missing. For the SP3[38] study and the analysis of secondary endpoints in the TOMORROW trial,[73] LOCF imputation was used when data for the entire 52 week period were not available. In the INPULSIS trials,[72] the statistical model used for the primary analysis allowed for missing data, assuming that they were missing at random; missing data were not imputed for the primary analysis. The company[4] acknowledged that the inclusion of all the trials in the NMA regardless of how missing data were handled may produce bias in the results but strict exclusion criteria on the handling of missing data, could lead to the exclusion of most trials from the network.

Study duration

The time of outcome assessment for data included in the NMA varied (see Table 35 and Table 39). The primary endpoint in the CAPACITY trials[49] was evaluated at 72 weeks with assessments for certain endpoints conducted every 12 weeks. The company considered that data at 48 weeks was the most appropriate data cut-off to use in the NMA so that it could be compared with the 52 week data from the other trials. The CS[4] (page 125) assumes that the treatment effect will be similar across these time points. The ERG asked the company to provide additional analyses to explore the sensitivity of the results to this assumption (discussed in Section 4.7). For a highly progressive disease such as IPF, if trials enrol participants at the same point in their disease course then those with a shorter follow-up might be expected to observe fewer negative outcomes (e.g. exacerbations, decline in lung function, deaths) whilst trials with a longer follow-up would be expected to observe worse outcomes.

Outcome definition

The definitions of the outcomes included also varied. In the SP3[38] study, lung function was reported as VC whilst the remaining trials used FVC. The CS[4] (page 93) stated that *"given that there is little difference between VC and FVC in subjects without obstructive pathology,[76] and IPF patients have a*

restrictive pathology, it is appropriate that VC and FVC are treated as comparable endpoints." The ERG noted that the exclusion criteria for SP3[38] were not as explicit regarding the exclusion of patients with obstructive airway disease as the exclusion criteria for the ASCEND[34] and CAPACITY trials.[49] Therefore the ERG considers that the combination of VC data from SP3[38] with FVC data from the ASCEND[34] and CAPACITY trials[49] is questionable.

Definition of PFS and mortality also differed across the studies. PFS was assessed as composite endpoint and in response to clarification question A14,[10] the company provided the definition of PFS used in each of the trials (see Table 35) and stated that *"To maintain similarity as far as possible, for CAPACITY 1 and 2, the PFS estimate based on the definition used in the ASCEND trial was included in the analysis. For the definitions of SP3, PANTHER and INPULSIS, it is assumed that they will lead to similar hazard ratios and odds ratios between a given pair of treatments, and thus that it is appropriate to combine them in an NMA. We believe this to be a reasonable assumption because in a comparison between the CAPACITY and ASCEND trials, the replacement of DLco by 6MWD led to an increase in qualifying events without changing the HR estimate*."[10] In response to clarification question A33,[10] the company demonstrated that the HRs using the ASCEND[34] definition of PFS provide more conservative estimates of treatment effect (as compared to placebo) than those using the definition utilised in SP3[38] and PANTHER.[74, 75]

In response to clarification question A17,[10] the company confirmed that the definition of OS was the same across all the trials in the NMA; this was defined as patients who died due to any cause (all-cause mortality) in the ITT populations.

Table 35: Reported outcomes and definitions adapted from CS,[4] (including response from clarification question A14, and A17 and A32)[10]

Outcome	ASCEND[34]	*CAPACITY 1 & 2[49]	SP3[38]	IMPULSIS 1&2[72]	TOMORROW[73]	PANTHER[74]
Study duration**	52 weeks	72 weeks	52 weeks	52 weeks	52 weeks	60 weeks (NAC), 32 weeks (Triple therapy)
Lung function						
Change in percent predicted FVC	Yes	Yes	Reported change in % predicted VC	Yes	Yes	Yes (NAC only)
Change from baseline in FVC (L)	Yes	Yes	Reported change from baseline in VC (L)	Yes	Yes	Yes
Categorical decline of ≥ 10% in percent predicted FVC	Yes	Yes	No	Yes	Not clearly defined, therefore excluded	Yes (NAC only)
Survival						
All-cause mortality	Defined as rate of death from any cause	Defined as OS	Number of deaths	Defined as OS	Deaths from any cause	
IPF-related death	Reported as treatment-emergent -IPF-related mortality and defined as deaths occurring between randomisation and within 28 days of last dose of study drug	Reported as IPF-related mortality and defined as deaths occurring between randomisation and within 28 days of last dose of study drug	No	Defined as death from respiratory cause		
PFS	Defined as a confirmed ≥10%	Defined as confirmed ≥10% decline in percent predicted	Defined as VC decline of ≥10% or death.)	No	Excluded as only reported the	Defined as decline of

Outcome	ASCEND[34]	*CAPACITY 1 & 2[49]	SP3[38]	IMPULSIS 1&2[72]	TOMORROW[73]	PANTHER[74]
	decline from baseline in percent predicted FVC, confirmed ≥50 m decline from baseline in 6MWD, or death	FVC, ≥15% decline in % predicted DLco or death. In a *post hoc* analysis, the ASCEND definition of PFS was applied to the CAPACITY trials at 52 weeks and at 72 weeks, and used within the NMA			proportion of patients who progressed, rather than the proportion of patients who either progressed or died. It was unclear how many patients progressed before they died and therefore PFS cannot be calculated	≥10% in FVC or death.
Acute Exacerbations	Identified via a *post hoc* analysis of AEs based on the MedDRA lower level term "acute exacerbation of IPF" (CS,[4] page 104)	Defined as requiring all of the following within a 4 week interval: Worsening of PaO_2 (≥8 mm Hg drop from the most recent value) Clinically significant worsening of dyspnoea New, superimposed ground-glass opacities on HRCT in one or more lobes All other cardiac, thromboembolic, aspiration,	Defined as requiring all of the following within a month: increase in dyspnoea; new, ground-glass opacities on HRCT in addition to previous honeycomb lesion; all oxygen partial pressure in resting arterial blood (PaO_2) is lower by more than 10 Torr than previous	Yes	Yes	Yes (NAC only)

Outcome	ASCEND[34]	*CAPACITY 1 & 2[49]	SP3[38]	IMPULSIS 1&2[72]	TOMORROW[73]	PANTHER[74]
		infectious processes ruled out	one; exclusion of obvious causes, such as infection, pneumothorax, cancer, pulmonary embolism or congestive heart failure; the serum levels of CRP, LDH are usually elevated as well as serum markers of interstitial pneumonias, such as KL-6, Sp-A or Sp-D			
Physical function						
6MWD	Defined as the change from Baseline to week 52 in distance walked during the 6MWD test as measured in metres (m).	Defined as the change from baseline to week 48 in distance walked during the 6WMD test as measured in meters (m).	No	No	Yes	Yes
Health Related Quality of Life						
SGRQ	No	Yes	No	Yes	Yes	No
UCSD SOBQ	The SOBQ is used to assess shortness of breath with various activities of daily living (for example, brushing ones teeth or mowing the lawn). Patients rated the severity of their shortness of breath experienced on an average day during the past week on a 6 point scale (0 to		No	No	No	No

Outcome	ASCEND[34]	*CAPACITY 1 & 2[49]	SP3[38]	IMPULSIS 1&2[72]	TOMORROW[73]	PANTHER[74]
	5),with 0= not at all breathless, 4= severely breathless and 5= Maximally or unable to do because of breathlessness					
All cause discontinuation of treatment	Defined as the count of patients who "did not complete the planned observation time"					

**In CAPACITY 1 & 2[49] assessments were conducted every 12 weeks therefore data at 48 weeks was considered most appropriate to use for comparing with 52 week data from other trials*

*** Note that duration of follow up varies by outcomes.*

CRP, C reactive Protein; DLco, Diffusing capacity of the lungs for carbon monoxide; 6MWD, 6-Minute walking distance; FVC, Forced vital capacity; HRCT High-resolution computed tomography; IPF, Idiopathic pulmonary fibrosis; KL-6, Krebs von den Lungen 6; LDH, Lactate dehydrogenase; MedDRA, Medical Dictionary for Regulatory Activities; mmHg, millimetres of mercury; PaO2, Partial pressure arterial oxygen; PFS, progression-free survival; Sp-A; Surfactant protein A; Sp-D, Surfactant protein D and VC, Vital capacity;

Baseline characteristics

The CS,[4] page 128 notes that there were some differences between the baseline populations in the included trials (see Table 36), but there were no major concerns regarding the inclusion of any of these trials in the network. The CS,[4] page 127 notes that that the populations included in the trials are in line with the licensed indications[1] and the scope[3] and all patients had mild to moderate impairment in pulmonary function at baseline. However, the ERG notes that the SP3[38] study was conducted in a Japanese population and used a lower dose of pirfenidone (1,800mg/day) than that licensed in the UK (2,403mg/day). The CS,[4] page 127 notes that the difference in dosage reflects the difference in mean weights in the North American and European population compared to the Japanese population, hence the trials are comparable. However, the ERG notes that the INPULSIS trials,[72] which compared nintedanib with placebo, also had a high Japanese contingent compared with the other trials assessed, but no reported dose adjustments were made in these studies. The ERG is unsure how this would impact on the evaluation of effectiveness and safety of the therapy.

Despite stating that patients had mild to moderate impairment in pulmonary function at baseline, the measure of function was reported inconsistently across trials at baseline (see Table 36). The ASCEND[34] and the CAPACITY trials[49] used percentage predicted FVC and percentage predicted DLco; SP3 trial[38] used percentage predicted total lung capacity and vital capacity; TOMORROW[73] and the INPULSIS[72] trials used percentage predicted FVC and DLco **(**ml/min/mm Hg). As highlighted in CS,[4] page 82 patients recruited in the ASCEND trial[34] were at higher risk of disease progression with a reported percentage predicted FVC approximately 7-8% lower than the CAPACITY trials.[49]

The time since patients were diagnosed with IPF varied between the trials. Approximately half of the patients in the CAPACITY trials had a diagnosis for less than 1 year,[49] whilst the majority of patients in the remaining trials had been diagnosed for just over 1 year and 38% of patients in SP3[38] had disease duration of greater than 3 years. The ERG notes that due to the progressive and unpredicted clinical course of IPF, difference in disease duration will have an impact on outcomes as reported in the company's subgroup analysis: *"There was evidence of an interaction between treatment and time from IPF diagnosis to randomisation, with those patients diagnosed more than a year before randomisation experiencing a significantly greater treatment effect"* (CS,[4] page 113).

Table 36: Summary of baseline characteristic of trials included in the company's NMA: (CS,[4] Table 16, page 83 and Appendix 10)

	CAPACITY 2 [49]		ASCEND [34]		CAPACITY 1 [49]		SP3[38]		PANTHER (NAC) [74]		TOMORROW [73]		INPLUSIS 1[72]		INPLUSIS 2[72]	
	Pirfenidone 2403 mg/day	Placebo	Pirfenidone 2403 mg/day	Placebo	Pirfenidone 2403 mg/day	Placebo	Pirfenidone 1800 mg/day	Placebo	NAC 600mg/tid	Placebo	Nintedanib 300mg	Placebo	Nintedanib 150mg/bid	Placebo	Nintedanib 150mg/bid	Placebo
N	174	174	278	277	171	173	108	104	133	131	86	85	309	204	329	219
Mean Age, years (SD)	65.7 (8.2)	66.3 (7.5)	68.4 (6.7)	67.8 (7.3)	66.8 (7.9)	67.0 (7.8)	65.4 (6.2)	64.7 (7.3)	68.3 (8.4)	67.2 (8.2)	65.4 (7.8)	64.8 (8.6)	66.9 (8.2)	66.9 (8.4)	67.1 (7.5)	66.4 (7.9)
Males (%)	118 (68)	128 (74)	222 (79.9)	213 (76.9)	123 (72)	124 (72)	85 (78.7)	81 (77.9)	107 (80.5)	98 (74.8)	65 (76.5)	63 (74.1)	163 (79.9)	251 (81.2)	171 (78.1)	256 (77.8)
White (%)									94.7	95.5	71.8	76.5	66.2	64.1	51.6	49.2
Previously smoked (%)	63.0	66.0	66.2	61.0	66.0	58.0	75.0	67.3	70.5	71.0			70.6	70.2	63.5	66.3
Never smoked (%)									27.3	25.2			25	23.0	32.4	31.3
Currently smokes (%)									2.3	3.8			4.4	6.8	4.1	2.4
Definite IPF (HRCT)	159 (91)	164 (94)	266 (95.7)	262 (94.6)	149 (87)	158 (91)			103 (77.4)	99 (75.6)	33 (38.8)	24 (28.2)				
Mean time since IPF diagnosis, years (SD)	*1.3 (0.96)	*1.4 (1.12)	*1.7 (1.1)	*1.7 (1.1)	*1.2 (1.09)	*1.1 (1.04)	38 (35.2) ≤ 1y 29 (26.9) 1-3y 41 (38.0) >3y	20 (28.0) ≤ 1y 17 (24.0) 1-3y 35 (49.0) >3y	1.0 (1.0)	1.1 (1.0)	1.0 (1.2)	1.4 (1.5)	1.6 (1.4)	1.7 (1.4)	1.6 (1.3)	1.6 (1.3)
Desaturation <80% during 6MWT							34 (31.5)	24 (23.1)								
Mean (SD) 6MWD (m)	411.1 (91.8)	410.0 (90.0)	415.0 (98.5)	420.7 (98.1)	378.0 (82.2)	399.1 (89.7)			371.4 (115.5)	375.4 (104.7)						

	CAPACITY 2 [49]		ASCEND [34]		CAPACITY 1 [49]		SP3[38]		PANTHER (NAC) [74]		TOMORROW [73]		INPLUSIS 1[72]		INPLUSIS 2[72]	
	Pirfenidone 2403 mg/day	Placebo	Pirfenidone 2403 mg/day	Placebo	Pirfenidone 2403 mg/day	Placebo	Pirfenidone 1800 mg/day	Placebo	NAC 600mg/tid	Placebo	Nintedanib 300mg	Placebo	Nintedanib 150mg/bid	Placebo	Nintedanib 150mg/bid	Placebo
Mean (SD) SpO_2 % predicted							89.0 (2.3)	89.0 (2.0)	95.75 (2.45)	96.12 (2.3)	95.6 (1.7)	95.3 (2.2)	95.9 (1.9)	95.9 (2.0)	95.7 (2.1)	95.8 (2.6)
Mean (SD) FVC % predicted	74.5 (14.5)	76.2 (15.5)	67.8 (11.2)	68.6 (10.9)	74.9 (13.2)	73.1 (14.2)			72.2 (15.9)	73.4 (14.3)	79.1 (18.5)	81.7 (17.6)	80.5 (17.3)	79.5 (17.0)	78.1 (19.0)	80.0 (18.1)
Mean (SD) FVC (L)									2.9 (0.8)	2.9 (0.8)	2.7 (0.8)	2.8 (0.8)	2.76 (0.74)	2.85 (0.82)	2.67 (0.78)	2.62 (0.79)
Mean (SD) DLco % predicted	46.4 (9.5)	46.1 (10.2)	43.7 (10.5)	44.2 (12.5)	47.8 (9.8)	47.4 (9.2)	52.1(16.8)	55.2 (18.2)	44.7 (10.8)	46.0 (12.2)						
Mean (SD) DLco (ml/min/mm Hg)									13.2 (3.7)	13.5 (3.8)	3.7 (1.0)	3.8 (1.1)	4.0 (1.2)	4.0 (1.1)	2.7 (1.3)	3.8 (1.2)
Mean (SD) PaO_2							79.8 (10.2)	17.4 (9.7)	80.7 (10.5)	81.5 (11.8)	79.6 (13.3)	76.5 (14.1)				
Mean (SD) $P(A\text{-}a)O_2$							18.4 (11.3)	17.4 (9.7)	17.81 (9.95)	17.34 (10.96)						
Mean (SD) VC % predicted							77.3 (16.8)	79.1 (17.4)					79.5 (17)	80.5 (17.3)	80 (18.1)	78.1 (19)
Mean (SD) VC (L)							2.40 (0.64)	2.47 (0.70)								

	CAPACITY 2 [49]		ASCEND [34]		CAPACITY 1 [49]		SP3 [38]		PANTHER (NAC) [74]		TOMORROW [73]		INPLUSIS 1 [72]		INPLUSIS 2 [72]	
	Pirfenidone 2403 mg/day	Placebo	Pirfenidone 2403 mg/day	Placebo	Pirfenidone 2403 mg/day	Placebo	Pirfenidone 1800 mg/day	Placebo	NAC 600mg/tid	Placebo	Nintedanib 300mg	Placebo	Nintedanib 150mg/bid	Placebo	Nintedanib 150mg/bid	Placebo
Surgical Lung Biopsy (%)	86 (49)	85 (49)	86 (30.9)	79 (28.5)	94 (55)	94 (54)	26 (24.1)	28 (26.9)			29 (34.1)	19 (22.4)				

Note: Data only reported from arms included in NMA
** Data reported as time since IPF diagnosis, years ± SD (ERG assumes this is mean time)*
DLco, Diffusing capacity of the lungs for carbon monoxide; 6MWD, 6-Minute walking distance; 6MWT, 6-Minute walking distance; FVC, Forced vital capacity; HRCT, High-resolution computed tomography; IPF, Idiopathic pulmonary fibrosis; m, meter; PaO2, Partial pressure arterial oxygen; SD, Standard deviation; SpO2, Peripheral oxygen saturation; TLC, total lung capacity and VC, Vital capacity

Risk of bias

The methodological quality of the studies included in the NMA was assessed in the CS,[4] (page 128-129 and Appendix 10) using standard criteria adapted from the CRD guidance for undertaking systematic reviews.[44] A summary of the quality assessment results, as reported by the company, is provided in Table 37.

The CS[4] noted that a potential risk of bias arises from the different methods used for handling missing data across the studies and the process undertaken for randomisation was unclear in the TOMORROW[73] and the SP3 trial.[38] In the TOMORROW trial,[73] an interactive voice-response system (IVRS) was used to perform randomisation; however, no information was provided on how randomisation was generated. In the SP3[38] study, patients were allocated to treatment groups using a modified minimisation method, including some random allocation based on biased coin design to balance baseline SpO_2. However, for the purpose of these analyses it was assumed that randomisation process was adequate for all (CS,[4] page 128). The ERG agrees that the majority of the studies were of good quality, with low risk of bias, however, the ERG disagrees with categorising SP3[38] as a study with low risk of bias, principally because of the absence of any published protocols and the inadequacy of the information contained within the published manuscripts. Further details are provided in Section 4.2.

Table 37: Quality assessment summary of RCTs included for NMA (reproduced from CS,[4] page 129)

	Was randomisation carried out appropriately?	Was the concealment of treatment allocation adequate?	Groups similar at baseline in terms of prognostic factors?	Were the care providers, participants, and outcome assessors blind to treatment allocation?	Unexpected imbalances in drop-outs between groups?	Authors measured more outcomes than they reported?	Did the analysis include an ITT basis?	Risk of bias of the study
CAPACITY 1 & CAPACITY 2[49]	Yes	Yes	Yes	Yes	No	No	Yes	Low risk
SP3[38]	Unclear	Yes	Yes	Unclear	No	No	Yes	Low risk
ASCEND[34]	Yes	Yes	Yes	Yes	No	No	Yes	Low risk
PANTHER[74, 75]	Yes	Yes	Yes	Unclear	Yes (at interim analysis)	No	No	Some risk of bias
TOMORROW[73]	Unclear	Yes	Yes	Yes	No	No	No	Low risk
INPLUSIS 1 & INPULSIS 2[72]	Yes	Yes	Yes	Yes	No	No	No	Low risk

Scenarios considered

For the statistical analysis (see CS,[4] pages 126-129), the company used a base-case network which included all Phase II and III trials. Sensitivity analyses were performed using a restricted network which was limited to Phase III trials and excluded the triple therapy arm of the PANTHER trial.[77]

The ERG also asked the company to perform a sensitivity analysis without the SP3[38] and PANTHER studies[74, 75] (see clarification response,[10] question A38).[10] The company did not agree on the relevance of excluding SP3[38] from the network, stating in their response that *"SP3 has been recognised as providing valuable evidence in several reviews: the initial NICE technology appraisal of pirfenidone[2]; the nintedanib appraisal[78], and; as part of the EMA's review of the marketing authorisation application for pirfenidone."* The company did, however, provide results excluding the PANTHER study[74, 75] for one outcome (all-cause mortality up to 52 weeks) as proof of concept that excluding PANTHER[74, 75] does not change the comparative efficacy of pirfenidone, nintedanib and placebo. The ERG considers that the stated concerns relating to population difference, statistical methods for handling missing data, and risk of bias provide reason to consider excluding SP3[38] from the analyses and have consequently not included SP3[38] in the ERG base-case network. Table 38 summarises the studies included in the company's base-case network, and how this differs to the company's restricted network and ERG base-case network. Note that the inclusion of studies in the NMA analyses varies by outcome.

Table 38: Summary of the trials used in the network meta-analysis (reproduced from CS, page 124)

Trial (reference) included in CS base-case	CS restricted network?	ERG base-case network?	Treatments		
			Placebo	PFN	NTB
ASCEND (King 2014[34])	Yes	Yes	Yes	Yes	
CAPACITY 1 (Noble 2011[49])	Yes	Yes	Yes	Yes	
CAPACITY 2 (Noble 2011[49])	Yes	Yes	Yes	Yes	
SP3 (Taniguchi 2010[38])	Yes		Yes	Yes	
INPULSIS-1 (Richeldi 2014[72])	Yes	Yes	Yes		Yes
INPULSIS-2 (Richeldi 2014[72])	Yes	Yes	Yes		Yes
TOMORROW (Richeldi 2011[73])		Yes	Yes		Yes
PANTHER NAC (Martinez 2014[74])			Yes		
PANTHER Triple therapy (Raghu 2012[77])	Yes		Yes		

4.7 Critique of the NMA

4.7.1 Efficacy

Summary of analyses undertaken

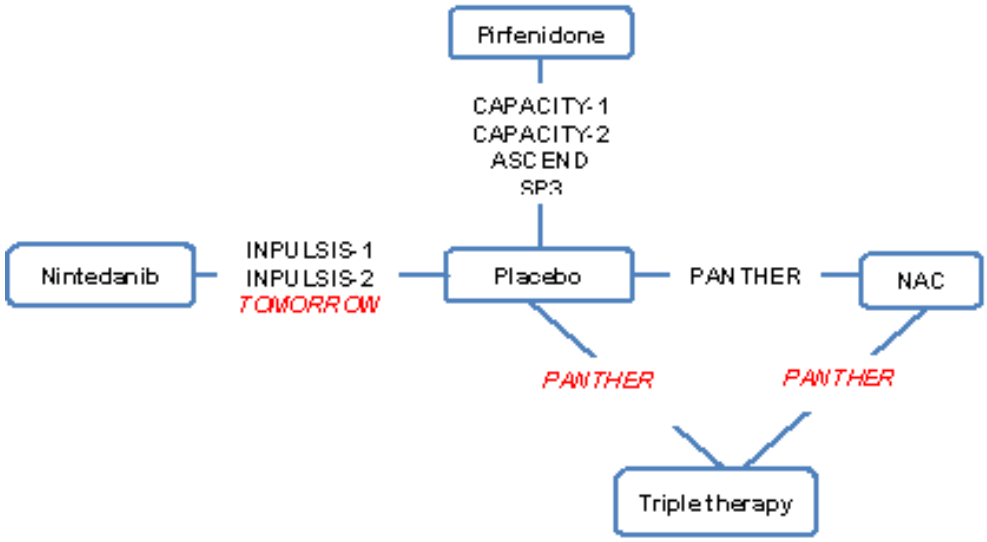
NMA were performed by the company to compare the treatment effects of pirfenidone, nintedanib, NAC, triple therapy and placebo for 11 outcomes relevant to the decision problem, as listed in Table 35. The results of four of these outcomes (OS, PFS, time to treatment discontinuation, acute exacerbations) are used to inform the economic model. Separate NMAs were undertaken for each outcome.

The base-case NMAs included all Phase II and III trials (eight trials in total). The network diagram for these studies is presented in Figure 26, however not all trials reported data that could contribute to all NMA outcomes. Table 39 summarises data available in each trial, for each outcome. A full summary of the NMA results and the number of studies included by scenario is provided in Table 41.

The company also performed sensitivity analyses using a restricted network which was limited to Phase III trials and excluded the triple therapy arm of the PANTHER trial. A full summary of the NMA results for the restricted network is provided in CS appendix 14.

Additional analyses performed by the ERG are summarised in Section 4.8.

Figure 26: Network diagram including all trials for NMA (reproduced from CS, Figure 19 page 125)



All trials are included in the base-case network

Red italicised text indicates trials excluded from the restricted network

Table 39: Summary of evidence for the company's base-case NMAs (adapted from clarification response,[10] question A32, Table 13)

	Study duration (weeks)	All-cause mortality	IPF-related mortality	Progression-Free Survival	Exacerbations	10% categorical decline FVC% pred [d]	FVC% pred	FVC Litres	6MWD	SGRQ	UCSD SOBQ	All cause discontinuation
CAPACITY1 and 2	72	52 and 72	52 and 72	52 and 72	52	48	48	48	48	48	48	48
CAPACITY1 and 2	72	52 and 72	52 and 72	52 and 72	52	48	48	48	48	48	48	48
ASCEND	52	52	52	52	52	52	52	52	52	-	52	52
SP3	52	52[a]	-	52	52	[e]	52	52	-	-	-	-
INPULSIS1	52	52	52*	52*	52	52	52	52	-	52	-	52
INPULSIS 2	52	52			52	52	52	52	-	52	-	52
TOMORROW	52	52[a]	52[b]	[f]	52	[g]	52	52	52	52	-	52
PANTHER (NAC)	60	60[b]	60[a]	60[a]	60	60	60	60	60	60	60	60
PANTHER (Triple)	32	32 (mean)	32 (mean)[a]	60[a]	[h]	[h]	-	60[c]	60[c]	60[c]	60[c]	-

[a] *HRs were unavailable: number of events and number of patients were used as an alternative (via the Woods model)*

[b] *HR was calculated from other available data using the methods of Parmar*

[c] *For FVC (L), 6MWD, SGRQ and UCSD SOBQ, publication presented estimated changes over 60 weeks (based on a repeated measures model)*

[d] *The NMA for this outcome assumes that all patients with missing values are non-responders (i.e. have a decline of more than 10%)*

[e] *Taniguchi 2010 reported some results for FVC 10% (Table E2 in the supplementary appendix) however there was insufficient information to calculate FVC 10% in line with the above definition*

[f] *only reported the proportion of patients who progressed, rather than the proportion of patients who either progressed or died. Although the number of deaths was reported, it was unclear how many patients progressed before they died and therefore PFS cannot be calculated.*

[g] *The outcome is not clearly defined in the nintedanib company submission to NICE.[26] Based on the company's response to clarification questions, the submission may be measuring any decline up to 52 weeks, whereas the other studies are measuring declines at exactly 48/52 weeks.*

[h] *Results were reported but the time point was not comparable.*

***** *pooled HR used for INPULSIS trials. Post hoc analysis only supplied as part of the submission to NICE by Boehringer Ingelheim.[26]*

Methods for the NMA

The CS specified the use of a random effects model for the principal analysis and also performed sensitivity analyses using fixed effects models (results provided in CS Appendix 14). Model fit statistics (total residual deviance and deviance information criterion [DIC]) were not provided in the original CS, but were provided for key outcomes upon clarification. The DIC provides a relative measure of goodness-of-fit that penalises complexity and was used to compare different models for the same likelihood and data[79]. The company reported that no meaningful differences in DIC between random effects and fixed effect models were observed. Random effects models were considered more appropriate due to the stated concerns in heterogeneity between the studies and the ERG considers that this decision was appropriate.

Where there were sufficient sample data, conventional reference prior distributions were used, however for certain endpoints there were too few studies to estimate the between study variance from the sample data alone and weakly informative priors were used. Although prior distributions should not be used without reasonable justification, the company considered *"the assumption of no heterogeneity made in the fixed effect model to be unrealistic."* (CS, [4] page 131). In the absence of further information on which to base the choice of prior, these were based on the recommendations of Turner *et al.,*[80], with details provided in CS,[4] Appendix 12. The ERG considers the company's choice of model and priors to be appropriate.

Statistical heterogeneity was assessed by presenting I^2 statistics from pairwise comparisons. Estimates of between study standard deviation from the conducted NMA were not reported in the original CS however the company provided this information upon response to clarification question A34, for key outcomes informing the economic model: all-cause mortality at 52 weeks; all-cause mortality at 72 weeks; PFS at 52 weeks; PFS at 72 weeks; IPF-related mortality at 52 weeks; IPF-related mortality at 72 weeks, and acute exacerbations.

Despite describing PANTHER as a multi-arm trial, it was treated as two separate placebo controlled trials for the statistical analyses. This was justified by describing PANTHER as an "atypical multi-arm trial", in which the "correlations between the arms will be less than those in a regular multi-arm trial" (see CS, Appendix 12). The ERG does not believe the issue is of importance, given that the interventions considered in PANTHER are not of relevance to the decision problem, but notes that if it is to be included, appropriate methods including correction for multi-arm trials should be used.

Reporting of results for the NMA

For continuous outcomes (FVC, 6MWD, SGRQ, UCSD SOBQ) the mean difference in the change from baseline is reported; for binary outcomes (acute exacerbations, discontinuation, categorical

decline in FVC) ORs are reported, and for survival outcomes (all-cause mortality, PFS, IPF related mortality) HRs are reported.

Results were summarised using posterior medians and 95% credible intervals (CrI). In the presence of heterogeneity, it is recommend that the predictive distribution, rather than the distribution of the mean treatment effect, better represents uncertainty about comparative effectiveness for a future rollout of a particular intervention.[81] The 95% predictive intervals (PrI) for key outcomes were provided by the company following a request for clarification (see clarification response,[10] question A34). The predictive intervals from the ERG analyses reported in Section 4.8 are used to inform the ERG base-case model in Section 5.

Implementation

Analyses were conducted using JAGS version 3.3.0 [82] and R version 3.0.1 or above.[83] The 'R2JAGS'[84] package was used to run JAGS from within R. The company stated that *"an appropriate burn-in period and number of iterations were allowed for."* (CS, Appendix 12).

Main results of NMA

Input data for the company's base-case network is provided for all outcomes in CS Appendix 11. In the original submission, pooled results for the two INPULSIS studies were used to inform the NMA for all survival outcomes (all-cause mortality, PFS, IPF related mortality). In response to clarification question A35 the company provided results using the individual study HR for all-cause mortality. The updated data used for the all-cause mortality NMA are presented in Table 40.

A full summary of the NMA results from the company's base-case network is provided in Table 39. The treatment effects for pirfenidone are broadly similar to those for nintedanib for all outcomes, with the pairwise treatment effects indicating that neither treatment is statistically significantly more effective.

For change from baseline in absolute (litres) and percent predicted FVC/VC, both pirfenidone and nintedanib were associated with beneficial effects relative to placebo. Pirfenidone was also associated with beneficial effects relative to placebo for all three time to event outcomes (all-cause mortality, PFS and IPF-related mortality). For nintedanib, the direction of the treatment effect favoured the active treatment, however the results were not statistically significant relative to placebo. For acute exacerbations, the treatment effects were not statistically significant for either treatment. For all-cause discontinuation of treatment, nintedanib was associated with increased odds of all-cause discontinuation relative to placebo, however the treatment effect was not statistically significant for pirfenidone.

The heterogeneity in treatment effects between studies is summarised for key outcomes in Table 41. The estimate of between-study standard deviation is mild-moderate for all outcomes, but with considerable uncertainty for IPF-related mortality and acute exacerbations. The network for IPF-related mortality contains fewer studies than that for all-cause mortality with no outcome data provided by SP3, and only pooled results were available for the INPULSIS trials. The NMA for acute exacerbations utilised a weakly informative prior for the between-study heterogeneity, as described above.

Sensitivity analyses conducted using fixed effects models were consistent with those reported from the random effects models. Results were also consistent across the company's base-case and restricted network.

For all-cause mortality, PFS and IPF-related mortality, the company's principal analyses use data from CAPACITY 1 and 2 evaluated at 52 weeks, rather than the full trial duration of 72 weeks. The main rationale behind this choice was to provide a comparison of data across similar timeframes for all studies. Other factors discussed by the company to justify this decision are that full follow up data were available for the majority of patients at this time point, that clinical data from ASCEND and CAPACITY was pre-specified to be pooled at 52 weeks, and that are no data available to support an assumption of proportional hazards in the longer term for nintedanib versus placebo (see clarification response,[10] question A37). The ERG considers that the use of the 52 week data would be appropriate if the purpose of analysis was to estimate the treatment effects at the specified time point, and there was reason to believe that treatment effects may not be consistent over the extended follow up period (and therefore bias results). However, the purpose of the analysis is to estimate the population mean survival time, and for the cost effectiveness modelling it was considered appropriate by the company to extrapolate the treatment effects over the full lifetime. The ERG therefore considers that the full evidence base with 72 week follow up should be used. Consequently, the 72 week data have been used in the additional ERG analyses presented in Section 4.8 and in health economic model .The ERG notes that the use of a constant HR in the economic model is appropriate only if the assumption of proportional hazards can be justified over both the observed and unobserved time period. The company's observation that there are no data available to support an assumption of proportional hazards in the longer term for nintedanib versus placebo therefore raises concerns over the reliability of the results based on extrapolated HR.

In response to clarification question A36 the company performed additional NMAs to justify the assumption that treatment effects are constant over time by including a covariate for trial duration through meta-regression, as described in the NICE TSD [81]. Analyses were conducted for the three

time-to-event outcomes (all-cause mortality, IPF related mortality, PFS) only. Results of the company's meta-regressions (see clarification response addendum,[30] question A36, pages 17 - 21) showed that including a covariate for study duration did not improve model fit, as judged using the DIC, and resulted in higher estimates for the between trial standard deviation. However these analyses were limited by the small number of studies and effect estimates at different trial durations were available only for the pirfenidone studies. The results should therefore be interpreted with caution and not viewed as robust evidence for a lack of treatment by time interaction.

Table 40: Input data for all-cause mortality NMA, company's base-case network (adapted from clarification response,[10] question A35 Table 29)

Study	Treatment	Comparator	HR	logHR	SE	N	n
CAPACITY 2	Pirfenidone	Placebo	0.37	-0.9942523	0.5304795	NA	NA
CAPACITY 1	Pirfenidone	Placebo	0.66	-0.4155154	0.5237481	NA	NA
ASCEND	Pirfenidone	Placebo	0.55	-0.597837	0.3793018	NA	NA
INPULSIS 1	Nintedanib	Placebo	0.63	-0.4620355	0.3942242	NA	NA
INPULSIS 2	Nintedanib	Placebo	0.74	-0.3011051	0.3103049	NA	NA
PANTHER	Triple therapy	Placebo	9.26	2.225704	1.0604775	NA	NA
PANTHER	NAC	Placebo	1.995622	0.6909556	0.6666667	NA	NA
SP3	Pirfenidone	Placebo	NA	NA	NA	110	3
SP3	Placebo	Placebo	NA	NA	NA	109	6
TOMORROW	Nintedanib	Placebo	NA	NA	NA	85	7
TOMORROW	Placebo	Placebo	NA	NA	NA	85	9

Table 41: Summary of results from company's base-case NMA, random effects model (adapted from CS Section 4.10 and clarification response,[10] question A34)

Outcome	Base-case network, RE model					
	Number of trials*		Treatment effect			Between study heterogeneity
	PFN	NTB	PFN vs placebo	NTB vs placebo	PFN vs NTB	
Lung Capacity		-				
Change from baseline in Percent Predicted FVC/VC (%)	4	3	3.39 (1.94,4.84)	3.33(2.34,4.5)	0.05 (-0.81,1.80)	NR
Change from baseline in FVC/VC (L)	4	3	0.12 (0.04,0.20)	0.12 (0.04,0.21)	0.00 (-0.11,0.12)	NR
FVC decline ≥10% Percent Predicted (OR)	3	2	0.58 (0.40,0.88)	0.65(0.42,1.02)	1.12(0.60,2.01)	NR
Physical Functioning and HRQoL		-				
Change in 6MWD	3	1	22.70 (8.82,36.31)	6.00 (-28.25,40.66)	16.63 (-20.83,53.81)	NR
SGRQ	2	3	-1.24(-4.94,2.39)	-2.11 (-5.48,0.37)	0.88 (-3.45,5.94)	NR
UCSD SOBQ	3	0	-3.19 (-6.24, -0.17)	NA	NA	NR
Time to event outcomes		-				
All-Cause Mortality up to 52 wks (HR)	4	3	0.52 (0.30, 0.88)	0.71 (0.43,1.16)	0.73 (0.35,1.50)	0.11 (0.03,0.54)
All-Cause Mortality up to 72 wks (HR)			0.62 (0.38, 0.99)	0.71 (0.43, 1.16)	0.87 (0.44, 1.72)	0.11(0.03,0.53)
PFS HR up to 52 wks (HR)	4	2**	0.63 (0.50, 0.80)	0.74(0.51,1.08)	0.85 (0.55,1.34)	0.09 (0.02,0.45)
PFS HR up to 72 wks (HR)			0.63 (0.50, 0.78)	0.74 (0.51,1.07)	0.85(0.55,1.31)	0.09 (0.02, 0.43)
IPF-Related Mortality up to 52 wks (HR)	3	3**	0.36(0.14, 0.90)	0.60 (0.22,1.33)	0.61 (0.18,2.34)	0.19 (0.03,1.44)
IPF-Related Mortality up to 72 wks (HR)			0.48 (0.22, 1.01)	0.60 (0.23, 1.28)	0.80 (0.27,2.63)	0.18 (0.03,1.29)
Other		-				
Acute Exacerbations (OR)	4	3	0.62 (0.29,1.39)	0.55 (0.26,1.09)	1.14 (0.41,3.44)	0.29(0.04,1.07)
All-cause Discontinuation of Treatment (OR)	4	3	1.28 (0.91,1.78)	1.42 (1.01,2.01)	0.90 (0.55,1.44)	NR

PFN, pirfenidone; NTB, nintedanib; HR, hazard ratio; OR, odds ratio

* number of trials are summarised for interventions relevant to the decision problem only. Network also includes NAC and triple therapy trials (PANTHER)

** uses pooled HR for INPULSIS

4.7.2 Safety

In response to clarification question A39, the company performed additional NMA to compare the treatment effects of pirfenidone, nintedanib, NAC, triple therapy and placebo for four key adverse events outcomes; diarrhoea, rash, discontinuation due to adverse event and serious cardiac adverse events. The results of these NMA were used to inform the updated economic model.

The base-case NMAs included all Phase II and III trials (eight trials in total), however not all trials reported data that could contribute to all AE outcomes. Table 42 summarises data available in each trial, for each AE outcome. As with the NMA of efficacy outcomes, the company also performed sensitivity analyses using a restricted network. A random effects model was specified for the principal analysis and sensitivity analyses were performed using a fixed effects model. Weakly informative priors, based on the recommendations of Turner *et al.,*[80] were used for the between study variance.

As with the data for the NMA of efficacy outcomes, there were differences in follow up time between studies. Data for the CAPACITY trials was collected at 72 weeks, rather than using intermediate follow up data (as was done for the NMA of efficacy outcomes) providing a greater range of follow up times. The CS page 126 states *"It is difficult to justify whether treatment effects will be stable over this longer time period"* and acknowledge that the difference in follow up time may lead to bias in the results.

A full summary of the NMA results and the number of studies included by scenario is provided in Table 43 for the company's base-case network, random effects model. Additional analyses are presented in the clarification response appendix D.[10]

Pirfenidone was associated with reduced odds of diarrhoea compared to nintedanib, and increased odds of rash as compared to nintedanib. For discontinuation due to adverse events and serious cardiac adverse events, the treatment effects for pirfenidone are broadly similar to those for nintedanib, with the pairwise treatment effects indicating that neither treatment is associated with more adverse events.

Table 42: Summary of evidence for the company's base-case adverse event NMAs (adapted clarification response,[10] appendix D, Table 137)

Trials	Study duration (weeks)	Diarrhoea	Rash	Discontinuation of treatment due to AE	Serious cardiac events
CAPACITY1	72	√	√	√*	√*
CAPACITY 2	72	√	√		
ASCEND	52	√	√	√	√
SP3	52	-	-	√	-
PANTHER (NAC)	60	√	-	√	√
PANTHER (Triple)	32	√	√	-	√
INPULSIS1	52	√	√	√	√*
INPULSIS2	52	√	-	√	
TOMORROW	52	√	√	√	√

** Pooled trials*

Table 43: Summary of results from the company's base-case AE NMAs, random effects model (adapted clarification response,[10] appendix D, Table 137)

Outcome	Number of trials*		Base-case network, RE model Treatment effect; OR (95% CrI)		
	PFN	NTB	PFN vs placebo	NTB vs placebo	PFN vs NTB
Diarrhoea	3	3	1.39 (0.94, 2.11)	7.32 (4.82, 11.13)	0.19 (0.11, 0.35)
Rash	3	2	3.85 (2.38, 6.29)	1.29 (0.49, 3.35)	2.99 (1.03, 8.88)
Discontinuation due to adverse event	4**	3	1.58 (1.04, 2.39)	1.52 (1.01, 2.29)	1.04 (0.58, 1.85)
Serious cardiac events	3**	3**	1.36 (0.54, 3.46)	0.64 (0.17, 1.49)	2.11 (0.65, 11.34)

PFN, pirfenidone; NTB, nintedanib; HR, hazard ratio; OR, odds ratio

** number of trials are summarised for interventions relevant to the decision problem only. network also includes NAC and triple therapy trials (PANTHER)*

*** uses pooled HR*

4.8 Additional work on clinical effectiveness undertaken by the ERG

4.8.1 Network meta-analysis

Additional analyses were conducted by the ERG, using the ERG base-case network described in Table 38. NMAs were conducted using random effects models for the following key outcomes used to inform the company's health economic model: all-cause mortality up to 72 weeks; PFS up to 72 weeks and acute exacerbations.

Analyses were conducted in the freely available software package WinBUGS[85] and R,[83] using the R2Winbugs[86] interface package. For all-cause mortality, there was evidence of poor convergence and so a weakly informative half-normal prior with variance 0.32^2 was used. WinBUGS code using this prior was provided by the company (see CS, Appendix 15). Under this prior, the between-study SD has a mean of 0.26. For all outcomes, a burn-in of 300,000 iterations of the Markov chain was used with a further 100,000 iterations retained to estimate parameters. Samples from the posterior distributions exhibited moderate correlation between successive iterations of the Markov chain and so were thinned by retaining every 10th sample.

All-cause mortality

Six trials were included in the network for all-cause mortality (Figure 27). The treatment effects are summarised in Figure 28.

Figure 27: Network of evidence for all-cause mortality, acute exacerbations and all-cause discontinuation, ERG base-case

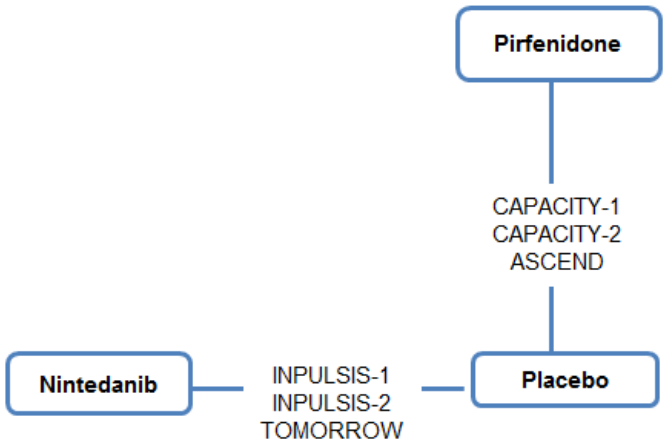
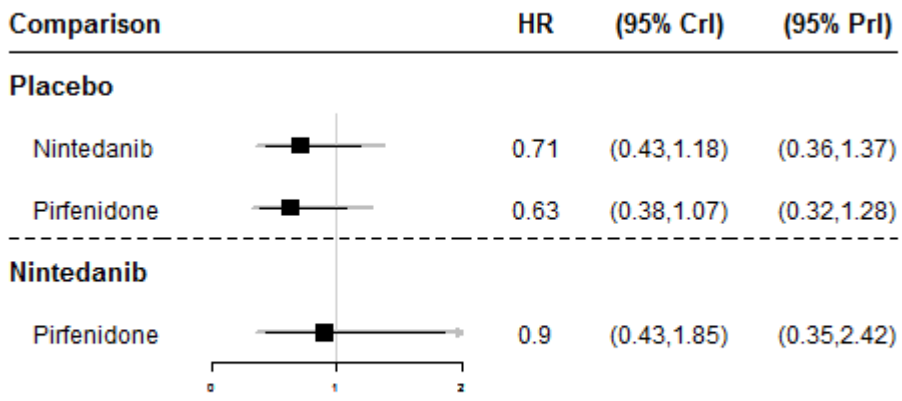


Figure 28: All-cause mortality, ERG base-case network - HR, 95% CrI and 95% PrI



Heterogeneity: between-study variance is 0.14 (95% CrI; 0.01, 0.52)

PFS

Five trials were included in the network for PFS (Figure 29), but a pooled HR was used for the INPULSIS trials since the individual study-level treatment effects were not available. The results of the NMA are summarised in

Figure 30.

Figure 29: Network of evidence for PFS, ERG base-case

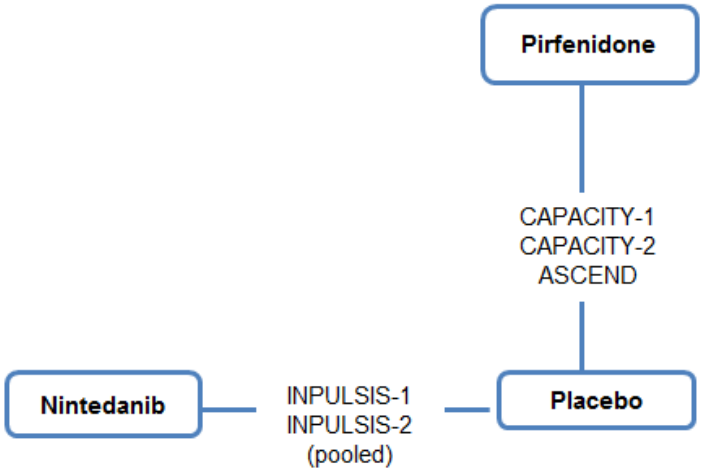
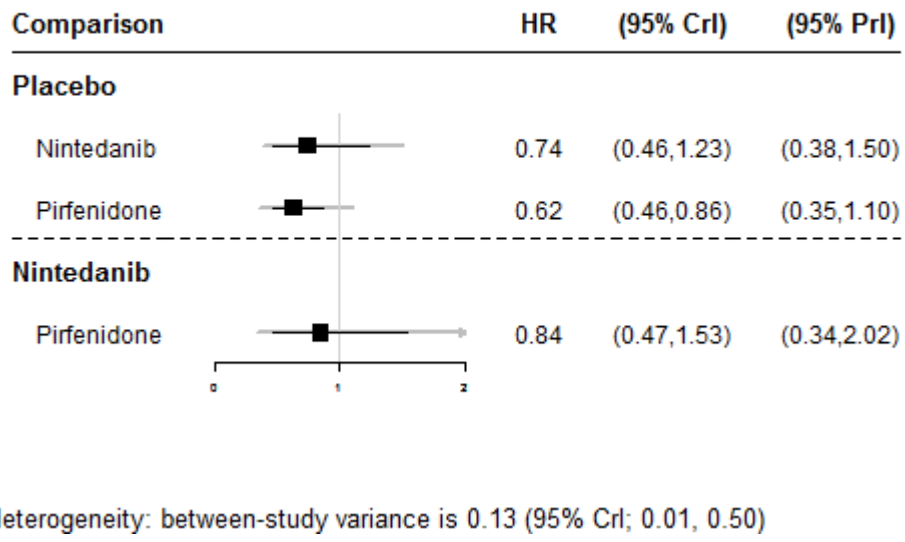


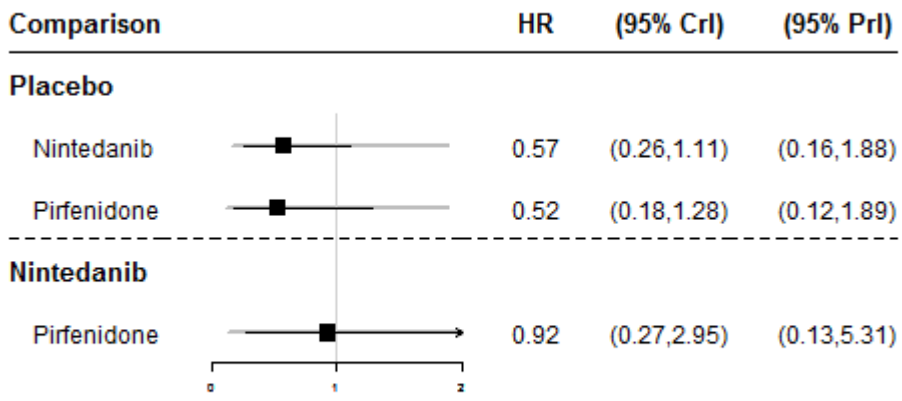
Figure 30: PFS, ERG base-case network - HR, 95% CrI and 95% PrI



Exacerbations

Six trials were included in the network for acute exacerbations (Figure 27). The pooled treatment effects are summarised in Figure 31.

Figure 31: Acute exacerbations, ERG base-case network - HR, 95% CrI and 95% PrI



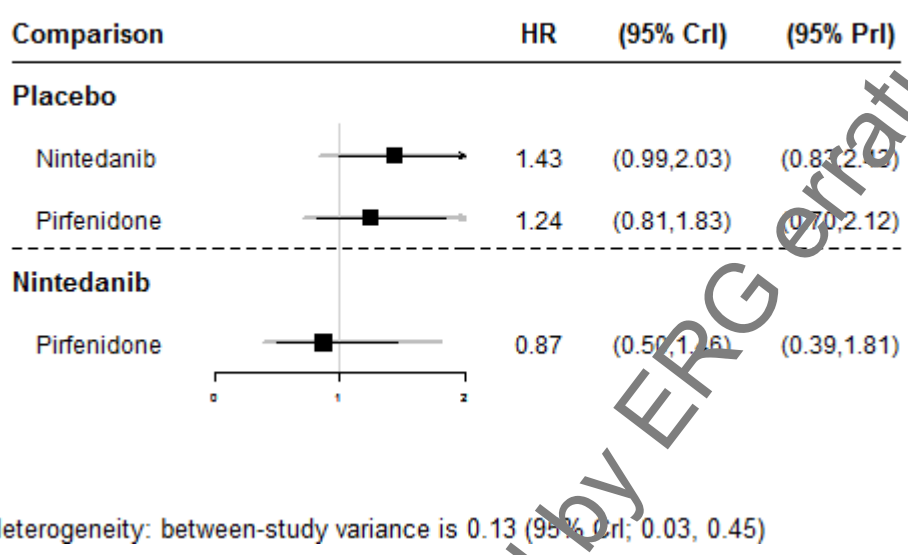
Heterogeneity: between-study variance is 0.29 (95% CrI; 0.05, 1.16)

All cause discontinuation

Six trials were included in the network for acute exacerbations (Figure 27). The pooled treatment effects (odds ratios) are summarised in Figure 32.

Treatment effects are estimated as odds ratios (OR), and then converted to relative risks (RR) using the average rate in the placebo arms over all studies in the NMA for use in the cost effectiveness model (clarification response,[10] Appendix D). For the ERG base-case network the average rate of all-cause discontinuation for placebo was 0.17. The estimated treatment effect for nintedanib vs pirfenidone on the odds ratio scale was OR: 1.14 (1/0.87) which equates to a relative risk of RR: 1.11.

Figure 32: All cause discontinuation, ERG base-case network - HR, 95% CrI and 95% PrI



4.9 Conclusions of the clinical efficacy section

Five RCTs compared pirfenidone at various doses with placebo in adults with mild or moderate IPF: ASCEND (Phase III),[34] CAPACITY 1 & CAPACITY 2 (Phase III),[49] SP3 (Phase III),[38] and SP2 (Phase II).[39] Three trials were international and multicentre (ASCEND and CAPACITY 1 & 2[49]), although only CAPACITY 2[49] included any UK centres (three of 110 centres across both CAPACITY trials). One RCT compared pirfenidone plus N-acetylcysteine (NAC) with placebo plus NAC in Chinese adults with mild or moderate IPF: Huang *et al.* 2015.[48]

Overall, the ERG assessed the potential risk of bias in ASCEND[34] and CAPACITY 1 & 2[49] to be low across most domains, with the exception of reporting bias and "other bias", which were judged to be "moderate" on account of inconsistencies between some outcomes and analyses presented in the trial protocols, those presented in published manuscripts and those reported in the CS,[4] and the possible influence of uncontrolled variables such as rate of disease progression.

The SP3,[38] SP2[39] and Huang *et al.* (2015) trials[48] were at a higher or more unclear risk of bias across many domains than the ASCEND[34] and CAPACITY[49] trials. These trials all evaluate lower, unlicensed doses of pirfenidone, apply different eligibility criteria and present noticeable differences from the other three trials in some baseline characteristics of participants.

The final selection of three trials (ASCEND,[34] CAPACITY 1 and CAPACITY 2[49]) for the main clinical efficacy review was considered to be appropriate by the ERG. However, there are some

between-trial differences across some baseline characteristics, such as mean FVC or 6MWD at baseline, but subgroup analyses suggested that these and other variables did not influence treatment effect. A *post hoc* pooled analysis of ASCEND[34] and CAPACITY 1 & 2[49] found no evidence of interaction between treatment for those patients with baseline FVC ≥ 80% predicted and those with FVC < 80% predicted.

The CS[4] reported three measures of lung function based on FVC: change from baseline in percent predicted FVC/VC; change from baseline in FVC/VC (ml); and relative proportions in each trial arm with FVC categorical decline of ≥10% percent predicted (this latter outcome measure included "death" in some analyses). The findings were not consistently statistically significant across trials for these outcome measures: ASCEND (52 weeks) [34] and CAPACITY 2 (72 weeks) [49] found statistically significant benefits for those on pirfenidone compared with those on placebo for mean change from baseline in percent predicted FVC (mean difference 4.78%; $p<0.001$ and mean difference 4.4%; relative difference 35.3%; 95% CI 0·7 to 9·1 $p=0.001$, respectively); but CAPACITY 1[49] found no statistically significant benefit for those on pirfenidone compared with those on placebo (absolute difference: 0.6%; relative difference: 6.5%; 95% CI -3·5 to 4·7 $p=0.501$). Pooled analyses of the CAPACITY trials[49] found statistically significant benefits for those on pirfenidone compared with those on placebo (absolute difference: 2.5%; relative difference: 22.8%; $p=0.005$). SP3[38] also reported statistically significant benefits for those on pirfenidone for change from baseline in percent predicted VC at 52 weeks ($p=0.044$); and change from baseline in VC (ml) ($p=0.042$). Huang *et al.* (2015)[48] reported a statistically significant mean change in FVC from baseline in favour of pirfenidone plus NAC compared with placebo plus NAC at 24 weeks ($p=0.02$) but not at 48 weeks ($p=0.11$). Meta-analyses of change in percent predicted FVC for CAPACITY 1 & 2[49] and ASCEND[34] and change in percent predicted VC for SP3,[38] suggested that pirfenidone reduces the decline in percentage predicted FVC compared with placebo up to 52 weeks (MD: 3.4, 95% CI: 1.87 to 4.94, *p*-value not reported). The meta-analysis also suggested that pirfenidone slows the rate of decline in FVC (MD: 0.12, 95% CI: 0.05 to 0.19, *p*-value not reported) up to 52 weeks.

In terms of decline in FVC by ≥10%, or death, ASCEND[34] reported a statistically significant difference in favour of pirfenidone compared with placebo at week 52 (absolute difference: 15.3 [95% CI not reported], $p<0.001$). For CAPACITY 1[49] the treatment effect at week 72 favoured pirfenidone but was not statistically significant (absolute difference: 3.8 [95% CI: -2.7 to 10.2], $p=0.440$), whilst CAPACITY 2[49] did report a statistically significant difference in favour of pirfenidone compared with placebo at week 72 (absolute difference: 14.4 [95% CI: 7.4 to 21.3], $p=0.001$). ASCEND also reported a significantly higher proportion of patients with no decline in percent predicted FVC (22.7% for pirfenidone versus 9.7% for placebo, $p<0.000001$), but CAPACITY 1[49] reported no difference between pirfenidone and placebo on this outcome measure (25.8% versus 22%, *p*-value not reported).

CAPACITY 2[49] reported a higher proportion of patients with no decline in percent predicted FVC for pirfenidone compared with placebo (24.1% versus 13.8%), but did not report a *p*-value. A meta-analysis of the ASCEND trial (52 weeks)[34] and the CAPACITY trials (48 weeks) [49] suggested that, compared with placebo, pirfenidone lowers the proportion of patients experiencing decline in FVC percent predicted of ≥10% (OR: 0.50, 95% CI: 0.31 to 0.82, *p*-value not reported).

There were fewer overall deaths or treatment-emergent IPF-related deaths in the pirfenidone than the placebo arms of the ASCEND[34] and CAPACITY[49] trials. However, these differences were not statistically significant in the ASCEND trial[34] at 52 weeks (for all-cause mortality or treatment-emergent IPF-related deaths, *p*=0.105 and *p*=0.226, respectively). The differences were significant in the pooled analyses for the CAPACITY trials[49] at 52 weeks (for all-cause mortality or treatment-emergent IPF-related deaths, *p*=0.047 and *p*=0.012, respectively) and in the pooled ASCEND[34] and CAPACITY[49] trials at 52 weeks (for all-cause mortality or treatment-emergent IPF-related deaths, *p*=0.011 and *p*=0.006, respectively). However, these differences were no longer significant at 72 weeks in the pooled CAPACITY trials (for all-cause mortality, *p*=0.315, IPF related mortality, *p*=0.117, or treatment-emergent all-cause mortality, *p*=0.141). There was only a significant difference between groups for treatment-emergent IPF-related mortality in the pooled CAPACITY trials[49] at 72 weeks (*p*=0.03). There appears to be a markedly increased rate of mortality in the CAPACITY trials[49] between the data reported for 52 weeks and for 72 weeks, the reasons for which are unclear. SP3, SP2 and Huang *et al.* (2015)[48] all reported all-cause mortality and found no statistically significant differences between pirfenidone and placebo arms. Meta-analysis of CAPACITY 1 & 2[49] and ASCEND[34] for pirfenidone compared with placebo, at 52 weeks, suggests that pirfenidone reduces all-cause mortality (HR: 0.52, 95% CI: 0.31 to 0.88, *p*-value not reported) and IPF-related mortality (HR: 0.37, 95% CI: 0.18 to 0.76, *p*-value not reported). Sensitivity analysis of the three trials at 72 weeks gave similar outcomes in favour of pirfenidone for both all cause-cause mortality and IPF-related mortality (HR: 0.64, 95% CI: 0.41 to 0.99, *p*-value not reported) and (HR: 0.49, CI: 0.27 to 0.87, *p*-value not reported). However, the reduction in mortality was lower at 72 weeks compared with 52 weeks.

Four of the key trials reported data for PFS: ASCEND, [34] CAPACITY 1 & 2[49] and SP3.[38] The definitions of PFS varied across the trials, but with a common element of a confirmed ≥10% decline from baseline in percent predicted FVC or VC. As with the findings for FVC outcomes, ASCEND at 52 weeks (HR 0.57; 95% CI, 0.43 to 0.77, *p*=0.0001) and CAPACITY 2[49] at 72 weeks (HR 0.64; 95% CI, 0.44 to 0.95, *p*=0.023) found statistically significant benefits in terms of PFS for those on pirfenidone compared with those on placebo; whilst for CAPACITY 1[49] the treatment effect was not statistically significant (HR: 0.84; 95% CI, 0.58 to 1.22, *p*=0.355). *Post hoc* pooled analyses of the ASCEND[34] and CAPACITY trials, [49] found statistically significant benefits for those on pirfenidone

compared with those on placebo (HR: 0.62; 95% CI: 0.51 to 0.76; $p<0.0001$). Huang *et al.* (2015)[48] also reported a significant treatment benefit for pirfenidone plus NAC compared with placebo plus NAC for PFS (HR=1.88, 95% CI: 1.09 to 3.24, $p=0.02$). Meta-analysis of the four trials showed pirfenidone improves PFS at 52 weeks compared with placebo (HR 0.63 95% CI, 0.53 to 0.74, *p*-value not reported). A sensitivity analysis based on CAPACITY trials[49] at 72 weeks, and ASCEND at 52 weeks,[34] with the assumption that the proportional hazards assumption holds up to 72 weeks, gave the same results.

All five included trials reported outcome data on acute exacerbations but used different definitions. The rates of acute exacerbation were much higher in the ASCEND trial[34] than in the CAPACITY trials, [49] with higher incidence in the placebo than the pirfenidone arms in the ASCEND[34] and CAPACITY 2[49] trials: no p values were reported. None of these three trials showed any statistically significant treatment effects compared to placebo for this outcome measure. SP2 t[39] did find a statistically significant difference in favour of the 1,800mg per day dose of pirfenidone for this outcome, but there was no consistency in the frequency of acute exacerbation reported across trials. This might be explained by the different definitions used. A meta-analysis of ASCEND,[34] CAPACITY 1 & 2[49] and SP3[38] also showed that pirfenidone is associated with a reduced risk of acute exacerbation of IPF with a HR of 0.64 (95% CI: 0.38 to 1.06, *p*-value not reported) compared with placebo, however the treatment effect was not statistically significant for the random effects model. CAPACITY 1 & 2[49] and SP2[39] also reported similarities in rates of hospitalisation (due to respiratory or non-respiratory causes) between pirfenidone and placebo arms.

Patient-reported outcomes were evaluated using the UCSD SOBQ and the SGRQ in the ASCEND and CAPACITY[49] trials. The treatment effects were not statistically significant for any of the individual trials , however results of the meta-analysis (using data from the CAPACITY trials at 48 weeks) suggest that pirfenidone is associated with a statistically significant reduction in in USCD SOBQ compared with placebo (Mean difference: -3.19 (95% CI: -5.74 to -0.63, *p*-value not reported).

The CS[4] reported the findings from two sets of analyses for 6MWD. The ASCEND[34] and CAPACITY[49] trials all reported findings on the pre-specified outcome of mean change from baseline in 6MWD for pirfenidone 2,403mg per day compared with placebo. ASCEND[34] at 52 weeks (absolute difference: 26.7m; relative reduction: 44.2%; $p=0.036$) and CAPACITY 1[49] at 72 weeks (absolute difference: 31.8m; relative difference: not reported; $p<0.001$) both reported a statistically significant and clinically important difference between pirfenidone and placebo on this outcome, but the treatment effect in CAPACITY 2[49] was not statistically significant (absolute difference: 16.4m; relative difference: not reported; $p=0.171$). A pooled analysis of the CAPACITY trials[49] at 72 weeks (absolute difference: 24m; relative difference: 31.2%; $p=0.0009$) also reported a statistically

significant and clinically important difference between pirfenidone and placebo on this outcome. Huang *et al.* (2015)[48] reported no difference between the pirfenidone and placebo arms in the 6MWT. Meta-analysis of CAPACITY 1 & 2 (data from week 48) [49] and ASCEND (data from week 52)[34] suggested that pirfenidone reduces the decline in 6MWD (MD: 22.9, 95% CI 10.58 to 35.23, *p*-value not reported).

A *post hoc* categorical analysis based on a mean decline ≥50 m in 6MWD from baseline, or death, in ASCEND[34] and CAPACITY 1 & 2,[49] also found that there was a statistically significant difference between pirfenidone and placebo in ASCEND trial (52 weeks: absolute difference: 9.8%; relative reduction: 27.5%; p=0.04)[34] The treatment effect was not statistically significant in CAPACITY 1 (p=0.10),[49] but was statistically significant for CAPACITY 2 (p=0.049).[49] A pooled analysis of the CAPACITY trials (72 weeks: absolute difference: 12.2%; relative reduction: 26%; p=0.001)[49] also reported a statistically significant effect for pirfenidone compared with placebo for this categorical outcome.

Four trials (CAPACITY 1 & 2,[49] SP3,[38] SP2[39]) reported data on the change from baseline in DLco. The CAPACITY trials[49] reported the change in percent predicted DLco, while SP2[39] and SP3[38] reported the mean decline (mL/min/mmHG). None of the trials reported statistically significant treatment effect for this outcome measure.

It is unclear why CAPACITY 1[49] reports different findings from ASCEND[34] and CAPACITY 2[49] in terms of FVC,PFS and 6MWD. For CAPACITY 1[49] the treatment effect is not statistically significant for FVC or PFS outcomes, unlike ASCEND[34] and CAPACITY 2,[49] but reports a positive statistically significant effect on one measure of 6MWD, which is not found to be statistically significant in CAPACITY 2[49]. An additional, small RCT of pirfenidone in combination with NAC in adults with mild and moderate IPF was identified by the ERG[48] and also reported no statistically significant effect on FVC, 6MWD or mortality outcomes.

The effect of the, "intrinsic variability in rates of FVC decline"[49], acknowledged as an issue in the CAPACITY trials' publication, and expanded on by the company in response to a request for clarification of this issue by the ERG (see clarification response,[10] question A26), might explain differences in outcomes across trials. Clinical advice received by the ERG suggested that there is currently no accepted single criterion by which to identify speed of progression of IPF. Participants in the trials included in the CS were not stratified by rate of progression, so it is possible, for example, that the placebo arm might have had more participants with more rapidly progressing disease than the intervention arm. As a result, the true treatment effect of the intervention relative to placebo might be uncertain. This could work either for or against the intervention.

In response to a clarification request from the ERG (see clarification response,[10] question A31), the company also provided results on OS and PFS from the ASCEND[34] and CAPACITY[49] trials for groups with a baseline percent predicted FVC of ≤80% (moderate IPF) and >80% (mild IPF), although exact, numbers within each subgroup in each trial arm were not reported. The findings did not suggest differential treatment effects according to disease severity for either outcome (as judged by the reported HR and 95% CI), however a treatment-by-subgroup interaction test was not reported so it is unclear if the difference between these subgroups was statistically significant.

The CS[4] also reported findings from non-randomised and non-controlled studies. First, the RECAP study (PIPF-012),[40] a non-randomised, non-controlled, open-label extension of the ASCEND and CAPACITY trials, which was principally designed to assess the long-term safety of pirfenidone 2,403mg/day in patients with IPF who received ≥80% of scheduled doses and completed the week 72 final study visit in CAPACITY 1 or CAPACITY 2. The RECAP study is ongoing. The most recent data-cut was performed in June 2015 and the next data-cut is planned in June 2016. The publication by Kreuter *et al*[63] found that discontinuation rates were highest in those enrolled patients who had originally received placebo, and especially in those who did not meet the ASCEND or CAPACITY entry criteria. Survival data and time-on-treatment data were reported in the CS,[4] (pages 159-161) and were presented for patients who received pirfenidone 2,403mg per day from baseline onwards in CAPACITY and ASCEND, and through the RECAP extension period, for whom data are available through to 8.8 years. Information on survival of patients with IPF was also presented from six registries to explore the relative survival rates of trial patients receiving pirfenidone compared with these "matched" real-world patients receiving best supportive care. The CS[4] stated that results were similar to the comparisons reported for the trials.

Based on the NMA, the treatment effects for pirfenidone were broadly similar to those for nintedanib for all outcomes, with the pairwise treatment effects indicating that neither treatment is statistically significantly more effective. For change from baseline in absolute (litres) and percent predicted FVC/VC, both pirfenidone and nintedanib were associated with beneficial effects compared with placebo. Pirfenidone was also associated with beneficial effects relative to placebo for all three time-to-event outcomes (all-cause mortality, PFS and IPF-related mortality). For nintedanib, the direction of the treatment effect favoured the active treatment, however the results were not statistically significant relative to placebo. For acute exacerbations, the treatment effects were not statistically significant for either treatment. For all-cause discontinuation of treatment, nintedanib was associated with beneficial effects relative to placebo; however the treatment effect was not statistically significant for pirfenidone.

The ERG noted that, overall, some adverse events (AEs) were frequent, especially nausea, rash, dizziness, dyspepsia, anorexia and photosensitivity, but that these were generally mild or moderate in severity. The ERG requested from the company more detailed data on serious adverse events and the adverse events leading to discontinuation. The most frequently-reported serious adverse events in the pirfenidone arms of the ASCEND[34] and CAPACITY[49] trials, other than worsening of IPF, were pneumonia, prostate cancer, angina pectoris, coronary artery disease, congestive cardiac failure, atrial fibrillation and pneumothorax. The AEs leading to discontinuation of treatment in ≥1% of patients in pirfenidone groups were pneumonia, rash, raised hepatic enzyme levels and decreased weight (in ASCEND), [34] photosensitivity, rash and respiratory failure (in CAPACITY 1)[49] and bladder cancer, nausea and rash (in CAPACITY 2). [49] The majority of safety data were from trials with a follow-up of no more than 72 weeks, but the CS[4] did present analyses that included more than 300 patients who had received pirfenidone for more than four years. However, the results for these patients were not presented separately. The ERG noted that the two ongoing studies to evaluate safety would address some outstanding issues: the non-randomised, non-controlled, OLE study that included a set of patients who completed either ASCEND, CAPACITY 1 or 2 (RECAP) and PIPF-002,[65] an ongoing open-label compassionate-use study in US patients with either IPF or secondary pulmonary fibrosis.

Meta-analyses of treatment-emergent serious adverse events using data from ASCEND,[34] CAPACITY 1&2[49] and SP3[38] at week 52 showed no difference between the pirfenidone and placebo group (OR: 0.90, 95% CI: 0.70 to 1.15, *p*-value not reported).

NMA of safety data indicated that pirfenidone is associated with reduced odds of diarrhoea compared to nintedanib, and increased odds of rash as compared to nintedanib. For discontinuation due to adverse events and serious cardiac adverse events, the treatment effects for pirfenidone are broadly similar to those for nintedanib, with the pairwise treatment effects indicating that neither treatment is associated with more adverse events.

There are two ongoing studies to evaluate safety: the non-randomised, non-controlled, open-label extension study that included a set of patients who completed either ASCEND, CAPACITY 1 or 2 (RECAP)[40] and PIPF-002,[65] an ongoing open-label compassionate-use study in US patients with either IPF or secondary pulmonary fibrosis.

Limitations

The ERG notes that the main limitations of the company's meta-analysis relate to the following:

- Combining the 48-week outcome data from the CAPACITY trials[49] with the 52 week data from ASCEND[34] and SP3 trials.[38] Although the direction of effect for all analysed outcomes were the same for the 52 week and 72 week data, the magnitude of effect of pirfenidone was generally less at 72 weeks than 52 weeks.
- Inclusion of the SP3 trial[38] to assess the following outcomes: lung capacity (FVC/VC percentage predicted, FVC/VC (L)); PFS; acute exacerbation; and serious adverse events. SP3[38] used a lower unlicensed dose (1,800mg/day) of pirfenidone and included only Japanese patients. In contrast, the CAPACITY 1 & 2[49] and ASCEND[34] studies used licence doses of pirfenidone (2,403mg/day) and included people from Europe and the USA.
- Variation in outcome definitions used across the included trials for PFS, acute exacerbation, 6MWT, lung function and combining data of FVC with VC for lung function.

The NMA included trials were of different durations. CAPACITY 1 and 2[49] presented data at 72 weeks whilst the maximum follow up for the other studies (of interventions relevant to the scope) was at 52 weeks. Trials with a shorter follow-up might be expected to observe fewer negative outcomes and so in order to facilitate synthesis across trials, the NMA used data from CAPACITY 1 and 2[49] evaluated at an earlier follow up time of either 48 or 52 weeks (depending on the outcome). This is a valid approach for evaluating the treatment effects at a specific time point but means that the analyses did not make use of the full follow-up data available. Alternative methods that allow the incorporation of trials of different durations, whilst accounting for time effects, could have been used.

For time-to-event outcomes (all-cause mortality, PFS, IPF related mortality) the treatment effects are reported as HRs, which are time averaged estimates of treatment effect and under the assumption of proportional hazards should be constant over time. The CS[4] provided evidence to support the assumption of proportional hazards but, despite this, data at 52 weeks were used in the company's base-case NMAs rather than the full 72-week data. Although there is not enough evidence to reject the assumption of proportional hazards for the presented pirfenidone data, the ERG notes that treatment effects at 72 weeks were often substantially lower than those at 52 weeks. The company[4] reported that there was no evidence to support that proportional hazards hold for nintedanib in the long-term.

The company also described other potential sources of heterogeneity between trials, in terms of differences in outcome definitions and handling of missing data. Due to the limited number of studies contributing to each network, a pragmatic approach was adopted, whereby trials were included regardless of these differences.

Despite including all available evidence in the NMAs, there were still a limited number of studies for certain outcomes. For binomial outcomes, there were too few studies to estimate the between-study variance from the sample data alone and weakly informative priors were used.

For the INPULSIS studies,[72] trial-level treatment effects were not available for two outcomes (PFS and IPF related mortality). Pooled HRs were therefore used to inform the NMA for these outcomes.

5 COST EFFECTIVENESS

5.1 ERG comment on company's review of cost-effectiveness evidence

5.1.1 Search strategy

A single search strategy was conducted in November 2015 to identify cost-effectiveness studies, HRQoL and resource use data. The ERG notes that the search was sufficiently comprehensive and sensitive and the ERG obtained a similar result when reproducing the searches. The structure of the search strategy was: Idiopathic pulmonary fibrosis AND (resource use OR cost-effectiveness OR utilities).

The following sources were searched:

- MEDLINE 1946 to 2015 November 16
- MEDLINE In-Process
- EMBASE 1974 to 2015 November 16
- Econlit 1886 to October 2015
- The Cochrane Database of Systematic Reviews
- The Health Technology Assessment (HTA) Database Issue 4 of 4, October 2015 (Cochrane Library)
- NHS Economic Evaluation Database (NHS EED): Issue 2 of 4, April 2015 (Cochrane Library
- Cost-Effectiveness Analysis registry
- PROQOLID
- ScHARRHUD
- EuroQol database

Supplementary searching included searching key HTA websites (NICE; the Pharmaceutical Benefits Scheme [PBS]; the Canadian Agency for Drugs and Technologies in Health [CADTH], and the Scottish Medicines Consortium [SMC]). NICE submissions were hand-searched and Google Scholar and conference posters and abstracts were also searched over the period 2014 to 2016. The CS states that a 'recent systematic review' was also hand-searched but does not report the citation details of the particular review.

The CS does not provide a reference to any published filter used in the search. However, the utilities search filter appears to have been directly derived (with no variation) from Arber et al,[87] whilst the cost-effectiveness filter appears to be a slightly modified version of the NHS EED search filter.[88] The company reported in their clarification response (see clarification response,[10] question B5) that this was amended in order to increase the sensitivity.[44]

No date or language restrictions were applied to the searches; however, the CS states that only studies that were published after 2010 were screened. The date limit was applied because although NHS EED was omitted from the original submission, the ERG report relating to the previous pirfenidone appraisal stated that this database was checked and no additional studies were identified.[89]

The ERG agrees that it was not appropriate to apply filters to the searches run on databases with an economic focus including Econlit and NHS EED, as these databases have a specific economic focus. The ERG notes that the reporting of the searches is very thorough and includes screenshots of the searches conducted on Google Scholar and conference abstracts.

5.1.2 Inclusion / exclusion criteria for the review of published cost-effectiveness studies

The CS (page 189) reports that study selection followed a two-stage process involving: (a) the assessment of titles and abstracts of potentially relevant studies by a single reviewer, checked independently by a second reviewer, followed by: (b) re-assessment of full texts of potentially includable studies against what the company refers to as the "*systematic review eligibility criteria.*"

The inclusion and exclusion criteria adopted are not clearly reported within the CS or accompanying appendices. The CS did not provide an explicit list of inclusion/exclusion criteria for the review of published cost-effectiveness studies. The CS states that the aim of the review was to identify cost-effectiveness studies of pirfenidone for adult patients with mild to moderate IPF in England. It also states that full economic evaluations were included as well as relevant economic data reported in technology assessments. The CS states that obviously irrelevant records (such as animal studies and studies about ineligible populations) were removed. Excluded studies are tabulated in Appendix 18 and the most common reasons for exclusion were either an ineligible population or the reporting of ineligible outcomes, however the appropriateness of these exclusions cannot be assessed without knowing explicitly which populations and outcomes were deemed relevant.

Included studies were assessed using the checklist reported by Drummond and Jefferson[90] by one reviewer and checked by a second reviewer. Studies were not selected or excluded from the review based on quality assessment.

5.1.3 *Studies included in the review of published cost-effectiveness studies*

The company's electronic searches yielded 3,474 potentially relevant unique citations for the single search to identify cost-effectiveness studies, HRQoL studies and resource use data. Of these, 4 studies (reported across 5 references according to the company) were included in the review of cost-effectiveness studies.[42, 91-94] The CS justifies the exclusion of the cost-effectiveness model used in the 2015 nintedanib NICE submission[26] on the basis that the model was for "*all patients with IPF and not*

just those patients with mild to moderate disease". The ERG disagrees with this exclusion because the modelled population in the nintedanib appraisal related to patients with a percent predicted FVC above 50%, even though this was a narrower population than that covered by the nintedanib licensed indication. The ERG considers the exclusion of this study to be inappropriate as it addressed a similar decision problem to that considered within the current pirfenidone appraisal.

The ERG notes that a total of 6 references are presented by the company instead of five (corresponding to 4 studies). This includes the model used in the previous submission to NICE reported in two references,[42, 89] the model developed by Loveman *et al.* (2014) for a health technology assessment of all available treatments for IPF reported in two references,[93, 94] and two separate Common Drug Review (CDR) reports published by the CADTH for nintedanib[91] and pirfenidone.[92] A table of reasons for exclusion of studies is presented in CS Appendix 18. The ERG notes some inconsistencies in that the CDR for pirfenidone published in 2015 included in the company's review is a re-submission and that an initial assessment was conducted in 2013; the original submission is not included in the company's systematic review.

The ERG notes that the included studies vary in terms of modelling approach. The model submitted to NICE by the company during the previous appraisal of pirfenidone (TA282), used a micro-simulation approach whereby surrogate outcomes (FVC and 6MWD) are used to estimate the risk of IPF-related mortality.[42, 89] In contrast, the model developed by Loveman *et al.* (2014) used a cohort state transition approach whereby OS is modelled as a function of PFS.[93, 94] The modelling approach used in the CDRs for nintedanib and pirfenidone are less clear given the lack of details provided in these brief reports.[91, 92] Effectiveness data and sources for utility values also vary between these studies. Data from the ASCEND trial[34] were not available during the previous submission to NICE[42, 89] or HTA by Loveman et al[93, 94] and therefore are only included in the two CDRs.[91, 92] Utility values in the previous model submitted to NICE were taken from the CAPACITY trials[49] based on the SGRQ scores mapped onto EQ-5D utilities based on an algorithm developed in COPD by Starkie *et al.* (2011).[95] The model developed by Loveman et al[93, 94] used utility values from two studies conducted under the auspices of the IPFCRN[75, 96] in the US. The pirfenidone model previously submitted to NICE[42, 89] took discontinuation rates from the trials and did not include a stopping rule. Although unclear, it also appears that no stopping rule was applied in the analyses submitted to the CDR for nintedanib[91] and pirfenidone.[92] In contrast, Loveman *et al.* (2014)[93, 94] assumed that treatments are discontinued following progression. ICERs reported also varied between studies with some ICERs only being available after the application of confidential price discounts. The previous pirfenidone model submitted to NICE[42, 89] reported an ICER for pirfenidone versus BSC for patients with percent predicted FVC ≤80% of £25,969 per QALY gained following a confidential price reduction. The ICER for pirfenidone was CAN$78,024 per QALY gained against BSC in the CDR for pirfenidone

following price reduction.[92] Pirfenidone was dominated by nintedanib in the CDR for nintedanib (assumption of equal efficacy but nintedanib was less costly).[91] Finally, Loveman *et al.* (2014) reported that, at the list price, pirfenidone was dominated by inhaled NAC.[93, 94]

Quality assessment tables are presented in CS Appendix 19. Following quality assessment, the company reports that "*the CDRs provide only a brief summary of the cost effectiveness results and therefore score poorly against most areas of the Drummond quality assessment check list*" (see CS page 194) and have limited relevance to the UK. The ERG considers this to be justified but raises attention to particular comments expressed during these assessments [91, 92] that are relevant for this appraisal including: (a) the uncertainty around the duration of the treatment effect for pirfenidone and nintedanib against BSC; (b) the uncertainty around the relative effectiveness between pirfenidone and nintedanib, and; (c) concerns regarding the discontinuation rate and the assumption that the treatment effect remains following discontinuation.

The CS does not report results from the quality assessment for the previous model submitted to NICE[42] but does summarise some of the concerns expressed by the ERG[89] including the appropriateness of the model structure, comparators included and uncertainty around the clinical effectiveness of pirfenidone versus BSC. In Appendix 19 of the CS, the ERG observes that according to the company, the model that was previously submitted to NICE performed poorly against most areas of the Drummond quality assessment checklist[90] (did not conform to 17 criteria, conformed to 15 criteria and 4 criteria were non-applicable).

Finally, the company considered the Loveman study[93, 94] to be of high quality when assessed against the Drummond quality assessment checklist but that the relevance to the UK is limited given: (a) the study did not include data from the ASCEND and IMPULSIS trials; (b) the inclusion of a trial in severe IPF; (c) utility values were taken from a non-UK source; (d) efficacy data were taken from studies outside the UK, and; (e) "*for pirfenidone the data were taken from two Japanese studies and two multi-national studies (of which the UK was one country).*" The ERG notes that whilst the company appears to suggest that the inclusion of Japanese studies is a limitation in its systematic review, as described in Section 4.6, despite a request from the ERG, the company refused to exclude Japanese studies from the NMA.

5.1.4 *Conclusions of the review of published cost-effectiveness studies*

The CS draws some conclusions regarding the quality of the included studies, comments on the applicability of the studies to the decision problem for this appraisal and tabulates the ICERs reported. Whilst the ERG is generally satisfied with the cost-effectiveness review presented by the company, the ERG considers the decision to exclude the model used for the nintedanib submission[26] from the cost-effectiveness review to be questionable. The ERG observes that the population entering the model resembles the population included in the IMPULSIS and TOMORROW trials which consisted of people with a percent predicted FVC >50% at baseline and therefore consists of people considered to have mild to moderate IPF which is relevant for this submission. The ERG further notes that whilst people included in the nintedanib trials had milder disease compared with the population included in the pirfenidone trials (approximately 45% had a FVC >80% compared with approximately 25% in the pirfenidone trials), an analysis is conducted for an ASCEND-like population (defined as FVC 50-90% predicted, $FEV_1/FVC \geq 0.8$).[12, 26] The ERG considers that this study should have been included in the company's systematic review in addition to the original CDR for pirfenidone for consistency. The nintedanib model uses a cohort state transition approach whereby people entering the model progress through a series of health states defined by roughly 10 point percent predicted FVC intervals. EQ-5D scores were taken directly from the IMPULSIS trials. In this assessment, pirfenidone was dominated by nintedanib when the stopping rule was applied to both or none of the interventions in people with a percent predicted FVC <80% at baseline (including the price discount for both interventions).

The ERG further notes that the CS does not provide any conclusions regarding the cost-effectiveness of pirfenidone compared with BSC or nintedanib based on this review of published cost-effectiveness analyses.

In summary, the ERG notes some inconsistencies in the company's review and considers that it is challenging to compare results from the different models given the differences in model structure, assumptions, data used and the existence of confidential price discounts.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

This section presents a summary description of the model submitted as part of the CS. ERG comments are provided directly after each aspect of the model is described.

5.2.1. Consistency of the CS with the requirements set out in the NICE reference case

As part of its submission to NICE, the company submitted a fully executable health economic model programmed in Microsoft Excel alongside a written description of the methods and results. A revised version of the model was submitted in response to the clarification questions from the ERG. The

original model and the changes made in the revised model are both summarised here, however the results are presented only for the revised model.

The company's economic evaluation (described in Table 44) assesses the cost-effectiveness of pirfenidone versus BSC from the perspective of the UK NHS and Personal Social Services (PSS) in three populations:

(i) the ITT trial population of the ASCEND/CAPACITY/RECAP trials,[34, 40, 49] comprising of adults with mild to moderate IPF at baseline;

(ii) a subgroup of people with a percent predicted FVC >80% at baseline (considered by the company to be mild IPF);

(iii) a subgroup of people with a percent predicted FVC > 50% and ≤ 80% at baseline (considered by the company to be moderate IPF)

Within the percent predicted FVC of 50 - 80% subgroup, a comparison of pirfenidone against both BSC and nintedanib is evaluated.

The company's model uses a lifetime horizon. All costs and health outcomes are discounted at a rate of 3.5% per annum.

Table 44: Scope of the company's health economic analysis

Population	(i) ITT - trial population – people with Mild to Moderate IPF (ii) People with a percent predicted FVC >80% at baseline (considered by the company to be mild IPF) (iii) People with a percent predicted FVC of 50 - 80% at baseline (considered by the company to be moderate IPF).
Interventions and comparators	For the ITT-trial population, the base-case analysis compares: • pirfenidone versus BSC[a] For people with a percent predicted FVC >80% at baseline (considered to be mild IPF), the base-case analysis compares: • pirfenidone versus BSC[a] For people with a percent predicted FVC of 50 - 80% at baseline (considered to be moderate IPF), the base-case analysis compares: • pirfenidone versus (i) nintedanib or (ii) BSC[a]
Primary health economic outcome	Incremental cost per QALY gained
Synthesis of health effects	The majority of clinical effectiveness and safety estimates included in the model are based on a systematic review of the literature and results are taken from NMAs.
Measuring and valuing health effects	The utility values for the main model health states (progression-free and progressed) were derived by mapping from a disease specific HRQoL instrument (SGRQ) measured in people with IPF to the EQ-5D-3L. The mapping algorithm between the SGRQ and EQ-5D-3L was estimated in a population with IPF from England. The utility decrements for AEs were based on the submission made by the company for nintedanib during TA379.[26]
Perspective	NHS and PSS for costs Direct health impact on patients only for outcomes (i.e. no carer QALYs are included)

Evidence on resource use and cost	Resource use estimates for routine management are based on telephone discussion with UK clinical experts. Hospitalisation data are based on estimates from pirfenidone trials. Unit costs are taken from NHS reference costs. Drug costs in the main CS are based on list prices (results which incorporated the PAS for nintedanib are reported in a confidential appendix). Costs of end of life care were taken from the literature.
Time horizon	Lifetime
Discount rate	3.5% per year for both costs and QALYs
Equality considerations	No weighting has been applied to QALYs

BSC – best supportive care; ITT – intention to treat; FVC – Forced vital capacity; QALY – quality-adjusted life year; IPF- idiopathic pulmonary fibrosis

[a] defined in the trial as symptom relief, pulmonary rehabilitation, management of comorbidities, and end of life care, including oxygen therapy

The population entering the company's model reflects the population included in the CAPACITY[49] and ASCEND trials.[34] Similarly, the intervention and associated treatment regimen assumed in the economic model reflects the regimens used in the Phase III trials.[34,49] The intervention consists of pirfenidone (267mg capsules, given orally), given as three 267mg capsules, three times a day, giving a total of 2403mg/day; before adjustments for dose reductions and interruptions. In the company's base-case, people initiating pirfenidone are assumed to discontinue treatment at the rate observed in the RECAP extension trial; therefore, no stopping rule is applied in the base-case. The stopping rule defined by NICE which formed the basis for the positive recommendation for pirfenidone[2] and nintedanib[12] is however applied to nintedanib in the company's base-case and only in a scenario analysis for pirfenidone.

5.2.1.1. ERG comments on the population described in the CS and included in the company's model

The ERG is satisfied that the population and subgroups addressed by the company are largely in line with the final NICE scope.[3] In the CAPACITY/ASCEND trials,[34,49] which formed the main basis of the evidence used in the economic model, individuals were eligible if they had a percent predicted FVC ≥50% and predicted diffusing capacity of the lungs for carbon monoxide (DLco) ≥ 35% (≥ 30% in the ASCEND trial). This is largely in line with the definition provided by NICE in the final scope[3] for mild-to-moderate IPF; defined as "*a FVC greater than or equal to 50% predicted and a diffusing capacity for carbon monoxide greater than or equal to 35%.*" Clinical experts to the ERG indicated

that it is challenging to assess the severity in IPF but considered the population included in the clinical trials and, by extension, in the model, to be consistent with the definition of mild to moderate IPF used in clinical practice.

In addition to the ITT population (adults with mild to moderate IPF), the company reports results for people with a percent predicted FVC > 80% and 50 - 80% at baseline, and considers these populations to be people with mild and moderate IPF, respectively. Clinical experts to the ERG reiterated that it is challenging to assess the severity in IPF and that percent predicted FVC alone may not be a sufficient surrogate marker and that DLco may be a better indicator of the severity in IPF. The final NICE scope[3] suggests that "*if evidence allows, subgroup analysis by disease severity, defined by FVC (such as above and below or 80% FVC) and/or diffusing capacity for carbon monoxide, will be considered.*" The ERG notes that an analysis by DLco is not presented by the company. The ERG further notes that InterMune (pirfenidone's company at the time of the previous NICE appraisal) in their original submission to NICE considered that "*in clinical practice a FVC of 70% or 80% predicted is often considered to represent mild IPF, whilst a FVC >50% and <70% predicted is considered indicative of moderate IPF (Nathan, 2011) although formal definitions within guidelines have not been made.*"[97] The ERG accepts the challenges in defining the severity in IPF, and considers the subgroups defined by the company to be clinically reasonable and broadly consistent with the final NICE scope.[3] Nevertheless, the ERG would have liked to see an analysis by DLco. The direction of the ICER for any subgroups using DLco as a stratification factor is unclear.

The company's model also reflects the population included in the ASCEND,[34] CAPACITY,[49] and RECAP extension trials.[40] As described in Section 3.1, the ERG observes that the populations recruited in those trials may not be fully reflective of a typical clinical population, notably;

- The majority of individuals recruited in the trials (approximately 75%) had a percent predicted FVC of 50 - 80% at baseline but the proportion with mild IPF may be higher in the UK;
- The majority of trial participants were not recruited in the UK and BSC may vary internationally particularly in countries without universal access to healthcare;
- Patients with comorbidities, particularly emphysema, were excluded from the trials but these patients may be offered treatment in current practice if their FVC is in the range of 50% to 80%.

Furthermore, the ERG notes that people included in the RECAP OLE study were pre-selected in that only people who were compliant to the drug (defined as compliance of ≥ 80% of dose) were included.

Finally, the CS reports results from the ITT population, a combination of people with a percent predicted FVC of 50 - 80% and >80% at baseline; as suggested in the final NICE scope.[3] The comparators specified in the final scope are different within these two populations. Nintedanib is a comparator in people with a percent predicted FVC of 50 - 80% (which composed the majority of people included in the trials) but not >80% at baseline. The correct interpretation of the results for the ITT population is therefore problematic, as the comparison is made only against BSC. The ERG advises that the subgroups of people with a percent predicted FVC of 50 - 80% and >80% at baseline should be interpreted separately for this reason.

5.2.1.2. ERG's comments on the treatment regimen assumed for the intervention

The ERG is largely satisfied with the treatment regimen for the intervention (pirfenidone) assumed in the company's model. The ERG notes that according to the SmPC,[1] the dose should be titrated over a 14-day period when initiating pirfenidone treatment according to the following schedule; one capsule, three times a day (801mg/day) in the first week and two capsules, three times a day (1,602mg/day) in the second week of initiating treatment. Individuals receive three capsules, three times a day (2,403mg/day) from week 2 onwards. The ERG notes that dose titrations have not been explicitly included in the company's model. Instead the average dose over the trial period following titration has been applied in the model.

In the company's base-case, people initiating pirfenidone discontinue at the rate observed in the Phase III trials.[34, 49] The appropriateness of the company's decision to not include a stopping rule is questionable. The ERG notes that the licensing of pirfenidone[1] does not specify a stopping rule. However, NICE issued a stopping rule for the use for pirfenidone[98] (TA379) and nintedanib[78] (TA282) in England and recommends that both treatments should be discontinued if there is evidence of disease progression (defined as a decline in predicted FVC of 10% or more within any 12 month period). The company justifies the exclusion of the stopping rule on the basis of: (i) the high unmet need for people with IPF; (ii) evidence that pirfenidone may benefit people with or without disease progression, and; (iii) references to arguments regarding the difficulty of imposing such a stopping rule from the nintedanib submission, and diverse comments received at the scoping consultation for this appraisal and during the consultation on the Appraisal Consultation Document (ACD) for nintedanib.

The ERG recognises that this issue may be open to debate; nevertheless, the ERG considers that an analysis including the stopping rule for pirfenidone and nintedanib should represent the base-case as this reflects current clinical practice in England. Clinical advice received by the ERG confirmed that the stopping rule defined by NICE has been implemented successfully in practice and that audits are

regularly conducted to ensure that clinics comply with these rules. The ERG further notes that the NICE Appraisal Committee considered the views expressed regarding the difficulty of implementing the stopping rule during the appraisal for nintedanib and concluded in the Final Appraisal Determination (FAD)[78] that: "*The Committee recognised the limitations of FVC but understood that in clinical practice the wider patient characteristics would be taken into account in interpreting percent predicted FVC. Clinical experts noted that they follow the stopping rule in NICE's technology appraisal guidance on pirfenidone for treating idiopathic pulmonary fibrosis, but explained that before withdrawing treatment they retest FVC to confirm that the 10% drop is not temporary, which might happen with an infection. The Committee concluded that, although it has some limitations, percent predicted FVC is the most reliable and widely used measure of lung function in clinical practice.*" The ERG further notes that the approach used by the company is somewhat inconsistent in that an identical stopping rule has been included in the NICE guidance for nintedanib (TA282) and pirfenidone (TA379) but the stopping rule is applied for nintedanib in the base-case but not for pirfenidone.

The ERG notes that whilst a scenario analysis is presented by the company including a stopping rule for both pirfenidone and nintedanib, the implementation of the stopping rule within the model lacks validity. This issue is further described in Section 5.2.2.2.

Finally, in the company's base-case analysis (assuming no stopping rule for pirfenidone), the duration and dosage of treatment is based on the discontinuation rate and dosage observed in the clinical trials.[34, 40, 49] The ERG is unclear whether the dosage received is representative of clinical practice and whether people would be treated for a shorter or longer duration than that assumed within the model. Nevertheless, the ERG considers that using the dose intensity and discontinuation rates from the same trials were used to generate the effectiveness estimates, could be considered reasonable as this ensures consistency in the extrapolated costs and benefits.

5.2.1.3. ERG's comments on the comparators included within the CS and company's model

In people with a percent predicted FVC of 50- 80% at baseline (considered to be moderate IPF), pirfenidone is compared with BSC (defined in the trial as symptom relief, pulmonary rehabilitation, management of comorbidities, and end of life care, including oxygen therapy) and nintedanib. The ERG considers the comparators included in the company's model for this subgroup to be appropriate as this is in line with the recent NICE recommendation regarding the use of nintedanib in adults with a percent predicted FVC of 50 - 80% at baseline[78] and the marketing authorisation for nintedanib.[17]

In people with a percent predicted FVC > 80% at baseline (considered to be mild IPF), pirfenidone is compared with BSC only. No analysis is presented against nintedanib. The ERG considers the

comparators included for this subgroup to be appropriate. Whilst nintedanib is licensed in this population,[17] NICE did not issue a positive recommendation for nintedanib in this subgroup.[78]

For the ITT-trial population, a combination of people with mild to moderate IPF, the only comparator considered is BSC. This is justified by the company on the basis that nintedanib has not been recommended by NICE for the treatment of people with a percent predicted FVC > 80% at baseline (see CS on page 207). The ITT-trial population represents a combination of those people with a percent predicted FVC of 50 - 80% or >80% at baseline; a proportion of these people are clearly suitable for treatment with nintedanib, which is not a comparator in the ITT analysis. The ERG further observes that a large majority of people (approx. 75% - see Table 16 in CS in page 198) included in the ASCEND/CAPACITY trials[34, 49] had a percent predicted FVC of 50 – 80% at baseline. The ERG advises that the subgroups of people with a percent predicted FVC of 50% to 80% and >80% at baseline should be interpreted separately.

Finally, within the company's model, the efficacy for BSC reflects the mix of therapies used in the ASCEND/CAPACITY trials[34,49] and includes interventions aiming to relieve symptoms, pulmonary rehabilitation, management of comorbidities, and end of life care, including oxygen therapy. The ERG notes that people in the ASCEND[34] and CAPACITY[49] trials were recruited from a large number of centres worldwide (127 sites in Australia, Brazil, Croatia, Israel, Mexico, New Zealand, Peru, Singapore, and the US for ASCEND and 110 centres in Australia, Europe, and North America for CAPACITY), with potentially varying clinical practice. The generalisability to the UK of treatments received as part of BSC within the ASCEND[34] and CAPACITY[49] trial populations is unclear, particularly for patients in those countries without universal access to healthcare.

5.2.1.4. ERG's comments on the perspective, discounting and time horizon used in company's base-case

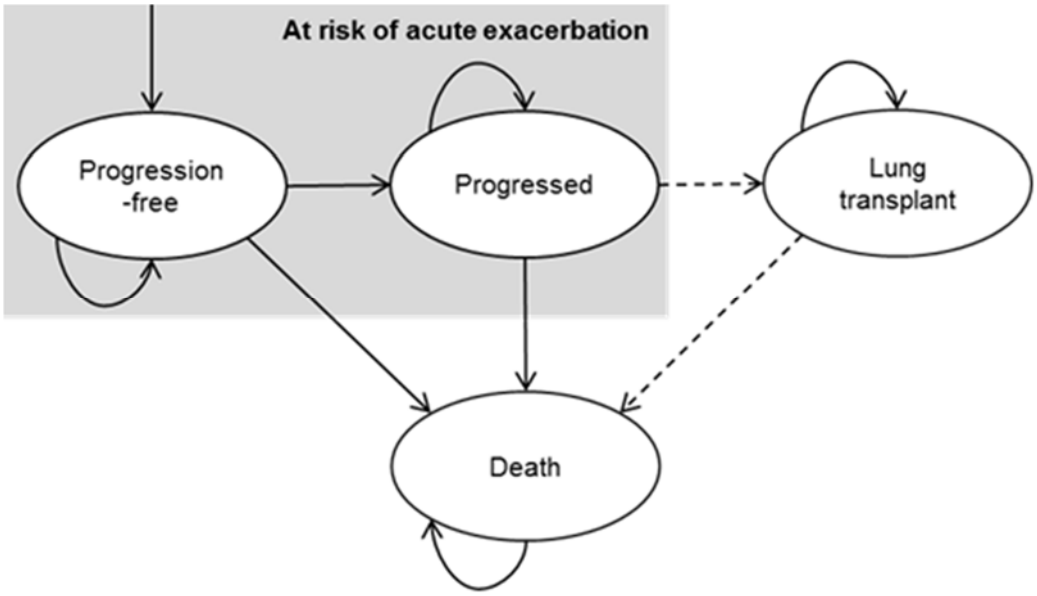
The company's base-case assesses costs and benefits over a lifetime horizon and adopts a UK NHS and PSS perspective. All costs and health outcomes are half-cycle corrected and discounted at a rate of 3.5% per annum. The ERG considers these to be appropriate and in line with the NICE Reference Case.[43]

5.2.2. Description and critique of the company's health economic model structure and logic

The description of the model's logic is based on information contained within the CS, and the ERG's assessment of the economic model. A simplified representation of the company's model structure is shown graphically in Figure 33. In summary, the model structure presented in the CS is based on three main health states; progression-free, progressed disease and death. Health states for progression-free

and progressed disease are further sub-divided into 'on-treatment' and 'off-treatment' periods (not shown in Figure 33). The model uses a 3-monthly cycle length.

Figure 33: Model structure (reproduced from CS,[4] Figure 42, page 205)



The company's model adopts a cohort-based partitioned survival approach whereby the OS, PFS and discontinuation curves from the Phase III trials[34,40,49] for pirfenidone are extrapolated over a lifetime horizon using parametric functions. These parametric functions are used to calculate the proportion of individuals in each health state over time. The time in the progressed disease health state is derived as the difference between the extrapolated OS and PFS curves. Consequently, movement between health states is not modelled using transitions probabilities, so this is not a traditional transition-state (Markov) model.

Treatment effects (HRs/RRs) estimated from the NMAs for BSC and nintedanib versus pirfenidone (with pirfenidone representing the baseline) are subsequently applied to the baseline hazards to estimate the hazards in people initiating nintedanib and BSC (see Section 5.2.4). The HRs/RRs are applied over the entire time horizon in the company's base-case, thereby assuming constant proportional hazards. Scenario analyses were conducted by the company whereby the treatment effects was assumed to stop after 7, 10 and 14 years. People initiating pirfenidone and nintedanib are assumed to receive BSC following treatment cessation.

In addition to the three main health states (progression-free, progressed disease and death), lung transplantation is included as a separate health state which is not used in the base-case. The model also includes the impact of acute exacerbations on HRQoL and resource use; these are not modelled

as separate health states, but are instead assumed to be treatment-specific and are applied within each model cycle.

QALYs are calculated as a function of time spent in the pre-/post-progression states with different utilities applied in each state. Cost components include drug acquisition, costs associated with the management of the condition, adverse events, acute exacerbation and end of life.

It should be noted that within its submission, the company makes reference to three modelling approaches that have been used in IPF: (i) the micro-simulation model submitted during the first appraisal of pirfenidone[2] (submitted by InterMune); (ii) the state transition approach based on percent predicted FVC categories submitted as part of the nintedanib NICE appraisal,[12,26] and; (iii) the state transition approach published by Loveman *et al.* (2014)[93,94] which is based on three main health states (progression-free, progressed disease and death). The company considers that the micro-simulation approach used in the previous NICE submission[97] and the approach used in the nintedanib NICE appraisal[26] add complexity and are difficult to parameterise and therefore are not appropriate.

5.2.2.1. ERG's comments on conceptual representation of the condition

The ERG has a number of concerns regarding the structure and logic of the company's model. These can be separated into four sets of issues: (i) the conceptual representation of the condition; (ii) the representation of the treatment pathway in IPF; (iii) the use of a partitioned survival model approach and HR, and; (iv) questionable structural assumptions.

The ERG considers that the company's model ignores a key facet of the disease: specifically that IPF is a progressive condition characterised by irreversible loss of lung function. The company's justification to use PFS in the model relies on three key sets of arguments: (i) findings from a review by Albera et al[99] which concluded that PFS could be deemed to be an appropriate endpoint in IPF trials; (ii) that this approach has been used in a previous economic evaluation,[94] and; (iii) the difficulty in parameterising a model based on percent predicted FVC (as used in the nintedanib appraisal[12,26]).

The ERG considers that whilst PFS could be considered as an appropriate endpoint in trials when evaluating the effect of an intervention in IPF, separating the natural history of IPF into two distinct consecutive phases (the presence/absence of progression) is overly simplistic and does not reflect the natural history of the condition or its progressive nature. This limitation is recognised in the CS (page 278) when results are compared against those generated during the original submission to NICE.[2] The company states that "*the impact on patient quality of life has been conservatively included for one progression alone in the updated model*" (see CS,[4] page 278). The CS therefore acknowledges that this simplification has the potential to bias the QALY gains estimated by the model. However,

contrary to the company's argument, the ERG considers that this simplification has the potential to overestimate the lifetime QALYs gained as the impact of subsequent progression on HRQoL is not captured. This overestimation could be favourable to pirfenidone as any survival gain for pirfenidone will translate into a larger QALY gain if subsequent declines in HRQoL after progression are ignored. Whilst the company's model structure made it difficult for the ERG to directly estimate the impact of this simplification on the incremental QALYs and ICER, an exploratory analysis conducted by the ERG (see Section 6) adjusting utility by age (and therefore assuming some form of progression – although with limitations) led to an increase in the ICERs of pirfenidone versus BSC.

Furthermore, within the company's model, all disease progression is assumed to be equally detrimental. Clinical advisors to the ERG considered that a 10% drop in percent predicted FVC would impact on HRQoL differently according to the baseline percent predicted FVC and therefore the clinical impact of disease progression, as defined in the model, would be different across individuals. The ERG notes that the model used in the nintedanib appraisal provides a better representation of the natural history in IPF, whereby individuals transit through multiple health states with different levels of percent predicted FVC (rather than just two), as their disease progresses. This structure allows for different HRQoL and cost estimates to be attached according to the individual's percent predicted FVC level. The model structure used in the nintedanib company submission was also considered by the clinical advisors to the ERG to be more representative of the progressive nature of IPF than the pre/post progression model presented by the company for pirfenidone.

In addition to the three main health states (progression-free, progressed disease and death), the company attempts to includes two key features of IPF; the impact on costs and health outcomes of acute exacerbations in the base-case and lung transplantations in a scenario analysis. The ERG considers the approach taken by the company to include lung transplantations as a scenario analysis to be appropriate given the uncertainty in the data available and the potential difficulty in incorporating lung transplantation within a cohort model. The ERG notes that the company's inclusion of lung transplantations relies on a series of assumptions and adjustments but this scenario analysis has a minimal impact on the ICER (an increase from [redacted] per QALY gained in the ITT population for the comparison of pirfenidone versus BSC).

The ERG considers the inclusion of acute exacerbations in the base-case to be appropriate given that exacerbations are considered to be an important clinical event in IPF.[78] Within the company's model, the impact of acute exacerbations is applied as a cost and HRQoL decrement during each model cycle and individuals could remain in the progression-free health state following an exacerbation. Clinical advisors to the ERG noted that the diagnosis of acute exacerbations is challenging and that it is often difficult to distinguish between an exacerbation and progression. Clinical advisors to the ERG

suggested that people who have experienced an exacerbation would usually be considered to have progressed. The ERG further notes from discussions held during the nintedanib appraisal that exacerbations are associated with high morbidity and mortality and therefore delaying/preventing exacerbations is an important aspect of maintaining quality of life.[78] Nevertheless, the ERG notes that the inclusion of acute exacerbations (as implemented by the company) has a minimal impact on the ICER (an increase from [redacted] per QALY gained in the ITT population for the comparison of pirfenidone versus BSC excluding acute exacerbations). The ERG considers the lack of impact associated with the inclusion of exacerbations in the model to be an artefact of the company's chosen model structure rather than a reflection on the relevance of exacerbations in IPF. This is because acute exacerbations are disconnected from the outcomes of progression and survival and are instead included as a simple cost and utility decrement during each model cycle.

The ERG further notes that within the company's model, the impact of exacerbations on costs and outcomes is modelled inconsistently and relies on a series of strong assumptions which are often not adequately supported by the evidence (especially over the long-term). The ERG notes that the impact of exacerbations on health outcomes is modelled by estimating the risk of exacerbations whilst on a particular treatment and applying utility decrements to those individuals having an exacerbation. In contrast, the impact of exacerbations on costs is included separately as a cost of hospitalisation specific to the treatment received (independent of the rate of exacerbations). It should be noted that in response to a request for clarification from the ERG (see clarification response,[10] question B15), the company confirmed that hospitalisation costs included in the model are not specific to acute exacerbations.

5.2.2.2. ERG's comments on the general modelling approach

The company's model adopts a partitioned survival approach and the CS (page 203) refers to the model published by Loveman *et al.* (2015).[94] The ERG notes that whilst both the Loveman *et al.* model and the company's model are based on PFS (although different definitions are used), each uses a different analytical approach (partitioned survival or state transition).

In the company's model, the OS, PFS and discontinuation curves from the trials are extrapolated using parametric functions and modelled independently from each other; these are used to determine the health state occupancy within the model. Within the company's model, individuals could also remain on treatment following progression. In contrast, in Loveman *et al.* (2015), a state transition approach is used and OS is estimated indirectly by assuming a relationship between OS and PFS. In the model described by Loveman *et al.,* treatment is assumed to be discontinued following progression. The ERG notes that both state-transition and partitioned survival approaches are used in the evaluation of cancer treatments and that both approaches have advantages and limitations. The

choice between approaches is often not straightforward and needs to be considered with respect to the quality and quantity of data available and whether the resulting model structure has face validity given the characteristics of the disease being modelled.

The ERG considers that whilst the partitioned-survival modelling approach is commonly used, the implementation of this approach in the company's model means that the outcomes of OS, PFS and discontinuation are modelled independently of each other. In simple terms, in the company's model, a change in either PFS or time to discontinuation has no impact on OS. To illustrate this, the ERG compared outcomes estimated when assuming no stopping rule (scenario 1 –company's base-case) with those estimated when assuming the stopping rule (scenario 2; as programmed by the company – company's scenario analysis). As can be seen from Table 45, different assumptions relating to the time on treatment have no impact on the mean life years, but impact treatment costs, and therefore the ICERs for pirfenidone.

Table 45: Impact of the stopping rule on health outcomes & treatment costs for pirfenidone and ICER against BSC for the ITT-trial population (results are discounted and half-cycle corrected)

	Scenario 1 - no stopping rule (company's base-case)	**Scenario 2 - Stopping rule (company's scenario analysis)**
Mean time on treatment (in years)	3.29	2.08
Mean time in PFS (in years)	2.05	2.05
Mean time in progressed disease (in years)	6.62	6.62
Mean life years	8.67	8.67
Treatment costs	[redacted]	[redacted]
ICER (vs. BSC)	[redacted]	[redacted]

During clarification, the ERG asked the company to provide evidence to support the assumption that time on treatment is independent from PFS and OS (see clarification response,[10] question B6). In response, the company stated that "*The ERG is correct that the model was constructed utilising the simplifying assumption that time on treatment, OS and PFS are independent of each other. This is a common practice in NICE submissions using time to event data (such as oncology submissions where disease is similar in severity and impact to IPF). To accurately quantify the relationship between time*

on treatment, OS and PFS, additional data would be required which are not publically available for nintedanib. Recent studies comparing the state-transition method (i.e. modelling time on treatment, PFS and OS separately) and area-under-the-curve (AUC) partitioned survival models show that the two methods produce similar results, and that either approach may be considered appropriate to a given decision problem, depending on the available data and scope of the evaluation [Briggs 2015]. We consider our approach the most appropriate given the data available".

The ERG considers the response from the company to be misleading. The company makes reference to a single case study conducted in advanced melanoma showing that the two methods provide similar results and could be appropriate in this particular case. However, the ERG is aware that different analytical approaches could lead to different estimates in other conditions, as shown in TA 257.[100] When deciding between modelling approaches it is important to consider the face validity of the model structure and any assumptions inherent within the structure as well as the amount and quality of the data available to parameterise the model.

Importantly, the ERG considers the modelling approach used by the company to be reasonable when the stopping rule is excluded; but inadequate when implementing a stopping rule given that treatment duration and treatment outcomes are disconnected from each other. The company's implementation of the stopping rule using tunnel states was also cumbersome and was not well described in the original submission but additional details were provided following the clarification request by NICE (see clarification response,[10] question B8. The company identified errors in the implementation which were corrected following the clarification request (see clarification response,[10] questions B8 – B10 and B23).

The ERG acknowledges that the implementation of a stopping rule, which was not implemented in the clinical trials, will usually be reliant on some assumptions to estimate treatment outcomes in those that discontinue due to the stopping rule, irrespective of the modelling approach chosen. Clinical advisors to the ERG commented that it hard to understand the relationship between treatment discontinuation and clinical outcomes such as disease progression and all-cause mortality because IPF is a heterogeneous condition with natural variability in the rates of decline in percent predicted FVC and the mechanism of action of pirfenidone has not been fully established. However, the ERG does not believe that the company's assumption that there is no relationship between treatment duration and treatment outcomes, such as PFS and OS, to be plausible. [redacted]

[redacted]

The ERG considers that the ICERs presented by the company using the stopping rule could represent a lower bound of the true ICER when the stopping rule is implemented in clinical practice, as the life-time costs of treatment are reduced when the stopping rule is applied in the model, but the incremental QALYs are not reduced by the shorter duration of treatment.

The ERG further notes that the CS includes a long description of the relationship between percent predicted FVC and OS to justify the definition of progression used in the model, but given that PFS and OS are modelled separately, no relationship is modelled between outcomes and therefore the definition of progression used in the model has no impact on OS.

5.2.2.3. ERG comments on the use HR for the comparators

The company estimates the baseline hazards of death, progression and discontinuation in people initiating pirfenidone from individual IPD from the CAPACITY,[49] ASCEND,[34] and RECAP trials[40] for all three populations evaluated in the model; i.e. the ITT population, and the subgroups of people with a percent predicted FVC of 50% - 80% and >80% at baseline. HRs taken from the NMAs are then applied to the hazards from the pirfenidone arms to estimate the hazards in people initiating nintedanib and BSC. Alternatives for OS for people initiating BSC are explored in scenario analyses such as using the Kaplan Meier (KM) curve up to the end of the observed period followed by extrapolation using HRs.

The ERG considers the use of HRs to capture the treatment effect to be reasonable and pragmatic with respect to the data available and the limited duration of follow-up in the evidence base for both nintedanib and BSC. Nevertheless, the ERG has a number of concerns with the values used and the duration over which the treatment effect is assumed to be constant in the company's base-case analysis. These issues are described in Section 5.2.2.5 and 5.2.4.1 respectively.

5.2.2.4. ERG's comments regarding the representation of the treatment pathway

The company's model assumes that people initiating pirfenidone and nintedanib receive BSC upon treatment discontinuation. The ERG considers that the treatment pathway assumed by the company is questionable. Nintedanib received a positive NICE recommendation in people with a percent predicted FVC of 50 - 80% at baseline; therefore it is possible that nintedanib could be used following the discontinuation of pirfenidone if individuals maintain a percent predicted FVC > 50%. Similarly, in principle, pirfenidone could be used following the discontinuation of nintedanib. Clinical advisors

to the ERG suggested that in practice, people initiating pirfenidone may switch to nintedanib upon discontinuation, and vice versa. This was acknowledged in the company's clarification response (see clarification response,[10] question B7), but the absence of sequences was justified by the company on the basis that a similar approach was used in the nintedanib appraisal.[10, 12]

The ERG notes that including treatment sequences within the economic model would require a complete restructuring of the model and the impact of their inclusion on the ICER is unclear.

5.2.2.5. ERG's comments regarding the assumption of proportional hazards

A key structural assumption in the company's model is that the treatment effect estimated at week 52 holds for the entire duration of the model (34 years) for both nintedanib and BSC against pirfenidone. The ERG considers the assumption of proportional hazards over the entire model duration to be overly optimistic and inadequately supported by the evidence for either pirfenidone against BSC or nintedanib.

The assumption of proportional hazards is somewhat justified by the company for the treatment effect between BSC and pirfenidone based on: (i) *post hoc* analyses conducted by the company (see CS,[4] Appendix 20) in the CAPACITY/ASCEND trials[34,49] which did not show a significant interaction between the treatment effect and time (see CS,[4] page 207 and clarification response,[10] question B12) for OS and PFS, and; (ii) inspection of the log-cumulative hazard between people initiating pirfenidone in the ASCEND/CAPACITY/RECAP trials[34,40,49] and (iii) data from three long-terms registries (Edinburgh, INOVA and EuroIPF).

The ERG has a number of concerns, which are discussed in turn:

- Despite there being no statistical evidence to contradict the assumption of proportional hazards between pirfenidone and BSC (up to 72 weeks for PFS and last follow-up for OS), the ERG notes that evidence from the CAPACITY-trials[49] reported a smaller treatment effect for OS between week 52 (HR: 0.49; 95% CI: 0.24 – 1.01) and week 72 (HR: 0.77; 95% CI: 0.47 – 1.28). Whilst the difference is not statistically significant, the ERG considers that the strong assumption of proportional hazards remains questionable. The ERG further observes a discrepancy in the company's argument in that the HRs estimated using data at 52 weeks are used in the company's base-case. As discussed in Section 4.7, the ERG considers that if the assumption of proportional hazards was valid, then the HR estimated at 72 weeks would be a more appropriate estimate as it incorporates more of the available data.
- The ERG also re-plotted the log-cumulative hazard plots for OS (using KM data available in the company's model) based on data from the ASCEND/CAPACITY/RECAP trials

(Figure 34). A parallel plot of the log-cumulative hazards for BSC and pirfenidone would suggest that the assumption of proportional hazards is reasonable within the trial period. Upon inspection of Figure 34, this assumption is questionable.

- Finally, the ERG advises considerable caution in the interpretation of any comparisons made by the company between the pirfenidone arm of the CAPACITY/ASCEND/RECAP trials[34, 40, 49] and data from registries. The ERG considers that such analyses are inherently subject to considerable bias. In brief:
  a. Despite the attempt by the company to select and match individuals from registries to people enrolled in the ASCEND and CAPACITY trials,[34, 49] the survival of individuals from the registries is inconsistent with the OS of people initiating BSC observed in the clinical trials (see Figure 35). The ERG notes that the company does not comment on the discrepancies between the OS in people enrolled in the CAPACITY/ASCEND trials[34, 49] and people enrolled in registries who were treated with BSC.
  b. The long-term survival for pirfenidone is based on the RECAP trial (OLE study of ongoing pirfenidone treatment) which enrolled people with IPF who completed the final follow-up visit of the CAPACITY trials and received ≥80% of the assigned study treatment. Clarifications were requested from the company regarding the rationale for excluding people from RECAP who received less than 80% of the assigned study treatment (see clarification response,[10] question B2). In response, the company stated that "*Patients using less than 80% of drug are considered to be non-compliant (a standard cut-off being used in many trials), and for this reason were not included in RECAP. Although RECAP was an open-label extension study, the standard compliance considerations were still applied.*" Consequently, the ERG considers that the exclusion of people who received less than 80% of the assigned study treatment could overestimate the survival for pirfenidone as only people that are considered to be compliant have been included in RECAP, thereby making comparison with long-term registries less relevant.
  c. Finally, whilst individuals from the registries were matched to people included in the clinical trials, the ERG notes some potential discrepancies in the inclusion criteria applied to the registry data which may bias the estimate in favour of pirfenidone. For instance, the company excluded individuals with a percent predicted FVC ≥90% (if DLco≥90%). However, according to data included in the company's model, approximately 8% of people in the ASCEND/CAPACITY trials had a percent predicted FVC ≥90%. Throughout the CS, the company discuss a potential link between FVC and mortality; thus, excluding people with a percent predicted FVC≥90% could underestimate the survival in individuals included in the registries.

Figure 34: Log-cumulative hazard plots for OS within the ASCEND/CAPACITY/RECAP trials (Plot drawn by the ERG)

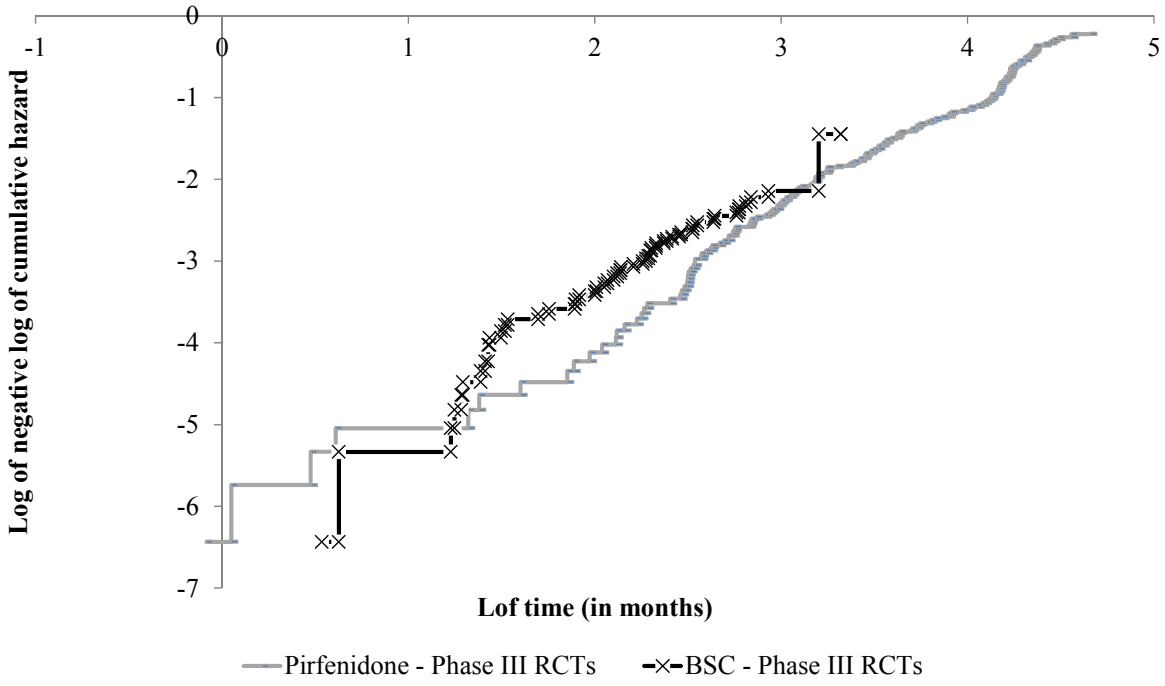
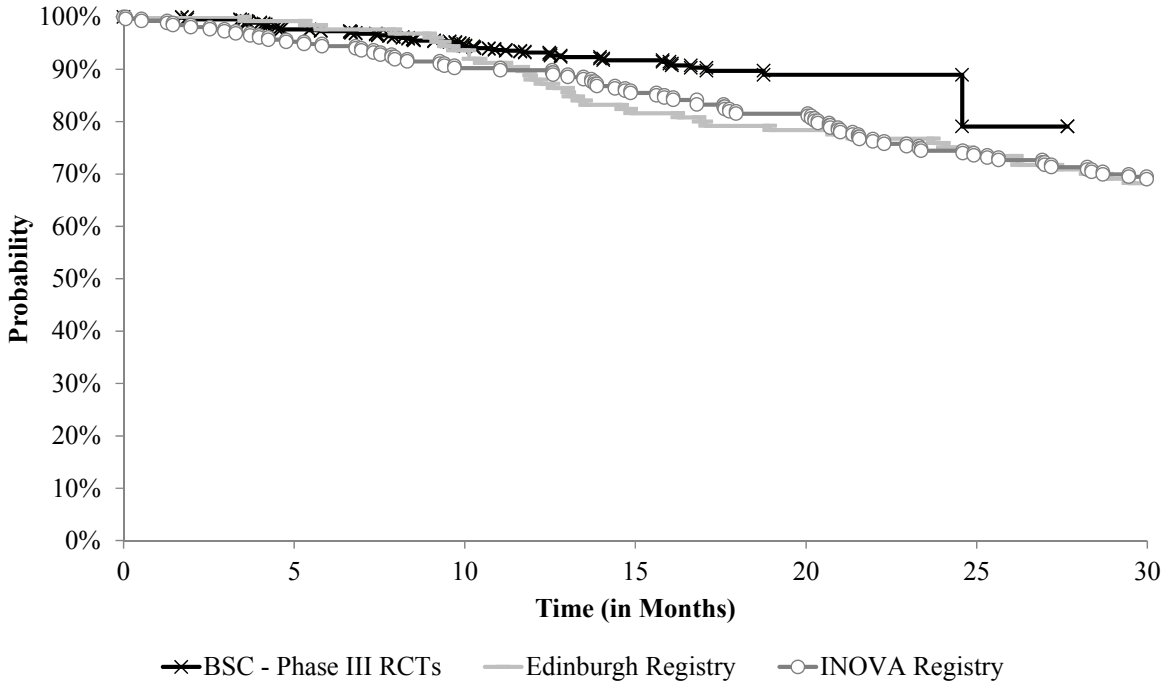


Figure 35: Plot of the OS for BSC from the ASCEND/CAPACITY trials and registries (Plot drawn by the ERG)



As a result, the ERG considers that the evidence presented by the company to support the assumption of proportional hazards for OS between BSC and pirfenidone in the long-term is inconclusive. The ERG notes however that the assumption of proportional hazards for PFS between BSC and pirfenidone appears more conclusive.

Clinical advisors to the ERG commented that it is possible that if a drug fundamentally alters the fibrosis pathway over the duration of a clinical trial, then with continued treatment it may be able to prevent declines over longer time periods. The ERG considers that this statement supports the possibility of continued effectiveness with long-term treatment but does not necessarily support a treatment effect for OS that is constant over the entire model duration.

As acknowledged by the company, the assumption of proportional hazard between pirfenidone and nintedanib is unclear. The ERG considers that assuming the treatment effect to hold for the entire model's duration is overly optimistic. The ERG notes that whilst indirect comparisons conducted by the ERG suggested (see Section 4.7) a slightly greater median treatment effect for pirfenidone using data up to 72 weeks (and excluding SP3) compared with nintedanib for OS (HR: 0.90; 95% CrI: 0.43 – 1.85), the differences were not statistically significant suggesting that the efficacy between nintedanib and pirfenidone could be similar. Results are also uncertain given the considerable heterogeneity between the population included in the trials for pirfenidone and nintedanib. As highlighted during the assessment for nintedanib by the CADTH, "*The two INPULSIS trials did not exclude people with normal lung function, while the ASCEND trial comparing pirfenidone against placebo imposed an upper limit on FVC. This resulted in a clinically meaningful difference in baseline per cent predicted FVC between the INPULSIS and ASCEND trials and suggested that patients in ASCEND may have had more advanced disease. This difference in baseline disease severity may have influenced the number of mortality events in the trials and impacted the ability to observe a mortality benefit with nintedanib.*"[91]

Consequently, the ERG considers that the company's base-case scenario provides a favourable estimate of the plausible ICERs for pirfenidone. Scenario analyses are presented by the company whereby the treatment effect is assumed to stop after 7, 10 and 14 years. The ERG notes that the ICER for pirfenidone compared with BSC for the ITT population increases from [redacted] when the treatment effect is assumed to stop after 7 years. The ERG considers that assuming the treatment effect to stop after 7 years is also arbitrary. The ERG notes that the treatment effect could stop earlier or later than 7 years, and therefore the ERG's preferred base-case are provided, in Section 6, using an optimistic (lifetime) and pessimistic assumption (treatment effect to stop at 2 years approximately at the end of the clinical evidence) regarding the duration of the treatment effect (lifetime to 2 year). This has been done because whilst the clinical advisors to the

ERG considered it possible that there may be continued effectiveness with long-term treatment, the duration of persistence for any long-term treatment effect is currently highly uncertain, particularly given that this is a heterogeneous condition and the mechanism of treatment is not fully understood at this time.

5.2.2.6. ERG's comments regarding the discontinuation with respect to progression

Within the company's model, people initiating pirfenidone and nintedanib could remain on treatment irrespective of progression status. Another structural assumption in the company' model is that the proportion of people who discontinue treatment would be the same irrespective of the progression status. The ERG considers that this is not adequately supported by the evidence. Nevertheless, the ERG notes that given the approach chosen by the company whereby OS, PFS and discontinuation are modelled separately, no impact is expected from this assumption as discontinuation is only used to calculate the treatment costs and treatment discontinuation has no impact on health outcomes.

5.2.3. Derivation of the baseline hazards of death, progression and discontinuation

This section focuses on the estimation of the baseline hazards of death, progression and discontinuation in people initiating pirfenidone. HRs/RRs are subsequently applied to the hazards from the pirfenidone arm to estimate the hazards of death, progression and discontinuation in people initiating BSC or nintedanib. These are discussed in Section 5.2.4. The source of data informing the KM curves for OS, PFS and time to discontinuation are summarised in Table 46.

Table 46: **Source of data informing KM curves (reproduced from clarification response, question B21,[10] Table 24)**

KM data	CAPACITY 1 & 2 (13 Jan 2009)*	ASCEND (14 Feb 2014)*	RECAP (30 June 2015)*
OS			
Pirfenidone – all	√	√	√
Pirfenidone - percent predicted FVC of 50 – 80%	√	√	√
Pirfenidone - percent predicted FVC >80%	√	√	√
BSC	√	√	
PFS			
Pirfenidone – all	√	√	
Pirfenidone - percent predicted FVC of 50 – 80%	√	√	
Pirfenidone - percent predicted FVC >80%	√	√	
BSC	√	√	
TTOT			
Pirfenidone – all	√	√	√
Pirfenidone - percent predicted FVC of 50 – 80%	√	√	√
Pirfenidone - percent predicted FVC >80%	√	√	√

* Date of data-cut

The baseline hazards of death and discontinuation (for reasons other than death and lung transplantations) in people initiating pirfenidone are estimated from IPD from the CAPACITY, ASCEND and RECAP trials[34,40,49] for all three modelled populations. Data from the ASCEND[34] and CAPACITY[49] trials, but not RECAP, are used to estimate the baseline hazards of progression in the company's model. The company justifies the exclusion of RECAP on the basis that progression data were not collected in this trial.

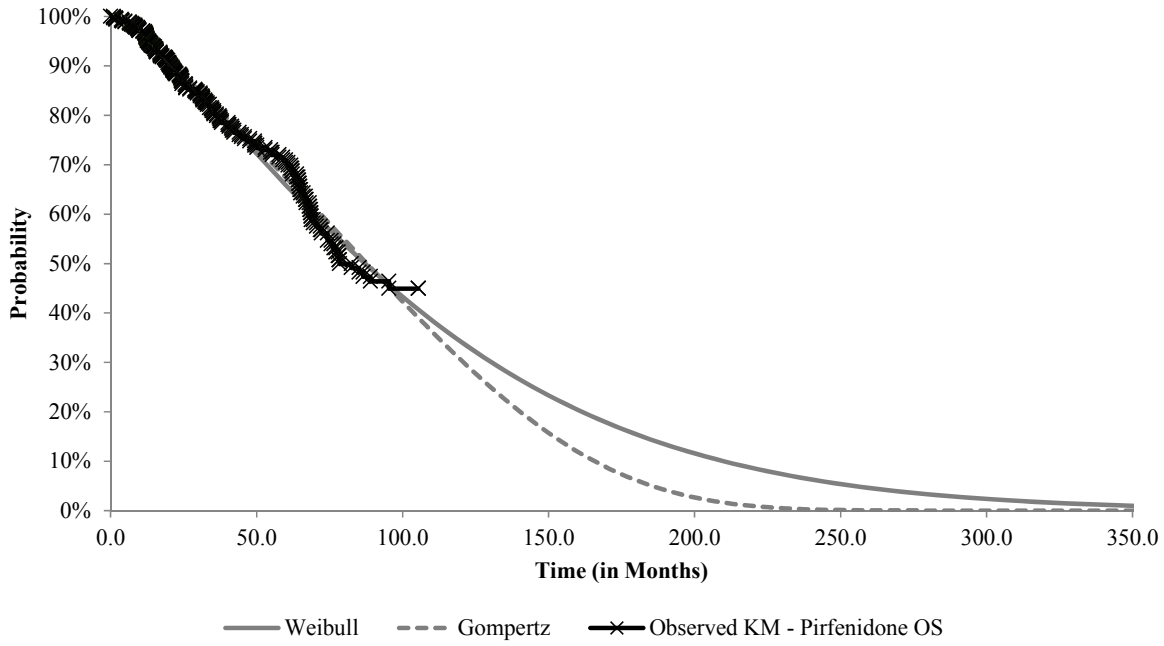
Progression is defined as per the ASCEND trial[34] definition and consists of confirmed ≥10% absolute decline in percent predicted FVC or confirmed ≥50m decline in 6MWD or death. This is principally justified by the company by the lack of data from the ASCEND trial on DLco.

A total of six single parametric functions were fitted to the observed KM curves: exponential, Weibull, Gompertz, log logistic, log normal and gamma. The Weibull distribution was selected for the base-case for the ITT population for all outcomes. This was justified in the CS based upon: (i) visual inspection of the fit during the observed period; (ii) statistical goodness of fit during the observed period (as measured by the Akaike Information Criterion [AIC] and Bayesian Information Criterion [BIC]), and; (iii) plausibility of the long-term extrapolation. Alternative parametric functions are examined in the sensitivity analyses.

5.2.3.1. ERG's comments regarding the estimating of overall survival

The ERG considers the process (i.e. assessing the fit to the observed data and assessing the plausibility of the long-term extrapolation) used by the company to select the most appropriate parametric distribution for OS to be generally appropriate. Nevertheless, the ERG considers the choice of the Weibull distribution in the company's base-case to be questionable and notes that the Gompertz distribution may provide a more clinically plausible extrapolation for OS.

Figure 36: Comparison of the observed KM for OS in people initiating pirfenidone against extrapolation using parametric distributions for the ITT population (Plot drawn by the ERG)



Of the six single distributions examined, the ERG considered the Weibull and Gompertz distributions to provide reasonable fits to both the observed period and a plausible long-term extrapolation in either the ITT population or people with a percent predicted FVC of 50 – 80% or >80% at baseline. Therefore these are the focus of comment in this section. The plot of the observed KM and Weibull

and Gompertz distributions are presented in Figure 36 for the ITT population (and Figure 46 and **Figure 47** in Appendix 1 in people with a percent predicted FVC > 80% at baseline and people with a percent predicted FVC of 50 - 80% at baseline, respectively). The ERG notes that only single parametric distributions are examined by the company and that the use of a piecewise distribution could potentially improve the fit.

5.2.3.1.1. Visual inspection of the fit to the observed period and goodness of fit.

The ERG considers that both the Weibull and Gompertz distributions provide a similar fit to the observed period and that it is difficult to differentiate between the two. The ERG notes that both curves provided a very similar visual fit to the observed period and had broadly similar BIC values (861.89 for Weibull vs. 869.44 for the Gompertz for the ITT population – see CS, Table 72, page 212). The ERG reiterates that goodness of fit criteria only provide an indication of the goodness of fit during the observed period and do not categorically indicate that one distribution should be preferred over alternative distributions.

5.2.3.1.2. Plausibility of the long-term extrapolation

Whilst the Weibull and Gompertz distributions provided a relatively similar fit during the observed period, these distributions provided different long-term extrapolations (Table 36). Contrary to the company, the ERG considers the Gompertz distribution to provide a more realistic long-term extrapolation for the following reasons:

i. A key argument from the company regarding the plausibility of the long-term extrapolation using the Weibull distribution relies on a comparison of the model prediction for BSC and registry data from the INOVA and Edinburgh cohorts. As described in Section 5.2.2.5, the ERG has a number of concerns with the survival observed in these registries compared with people initiating BSC that were enrolled in the ASCEND/CAPACITY trials. As shown in Figure 35, the survival from the registries did not validate the survival observed in people initiating BSC in the ASCEND/CAPACITY trials.[34, 49] The ERG further notes that the HR which is used to model the survival from BSC is taken from results from the NMA which uses data from the ASCEND/CAPACITY trials, and therefore, validating the model against registries is inconsistent when the registry data do not match the control data from the trials. The ERG considers that making inferences about the plausibility of the long-term extrapolation based on the modelled OS for BSC against registry data has limited relevance given the OS data from the registry do not match the placebo arm of the trials. The ERG further notes that both the modelled OS for BSC using the Weibull and Gompertz distribution provided a reasonable fit to the OS from the registries (Figure **37**). Therefore the

argument made by the company does not categorically indicate that one distribution should be preferred over the other one upon inspection of the fit of the modelled OS for BSC with registries.

ii. Second, the ERG notes that the OS curve from the trial and predicted in the model includes death from any cause (IPF and other causes) and that the Weibull distribution has a longer tail compared with the Gompertz distribution; consequently, the hazards of deaths at older ages may be underestimated. The UK life tables provide an estimate of the survival in the general population, in whom the average survival is expected to be greater than the survival observed in people with IPF who have a chronic progressive illness. For the Weibull or Gompertz distributions to be considered appropriate, a higher, or at least, equal hazard of death (compared with the general population life table estimates) should be observed. It can be seen from Figure 38 for the ITT population (and in appendix 2 for the subgroups in Figure 48 and **Figure 49**) that the use of the Weibull distribution in the model leads in some occasions to lower probabilities of death in people with IPF initiating pirfenidone compared with the probability of death from the general population. This is not considered by the ERG to be plausible. In contrast, the Gompertz distribution generates consistently greater probabilities of death when compared with the life tables in England.

As a result, the ERG considers that the Gompertz distribution provides a more plausible extrapolation of OS than the Weibull distribution.

Figure 37: Plot of the KM for OS from registries and modelled survival for BSC using the Weibull and Gompertz distribution (Plot drawn by the ERG)

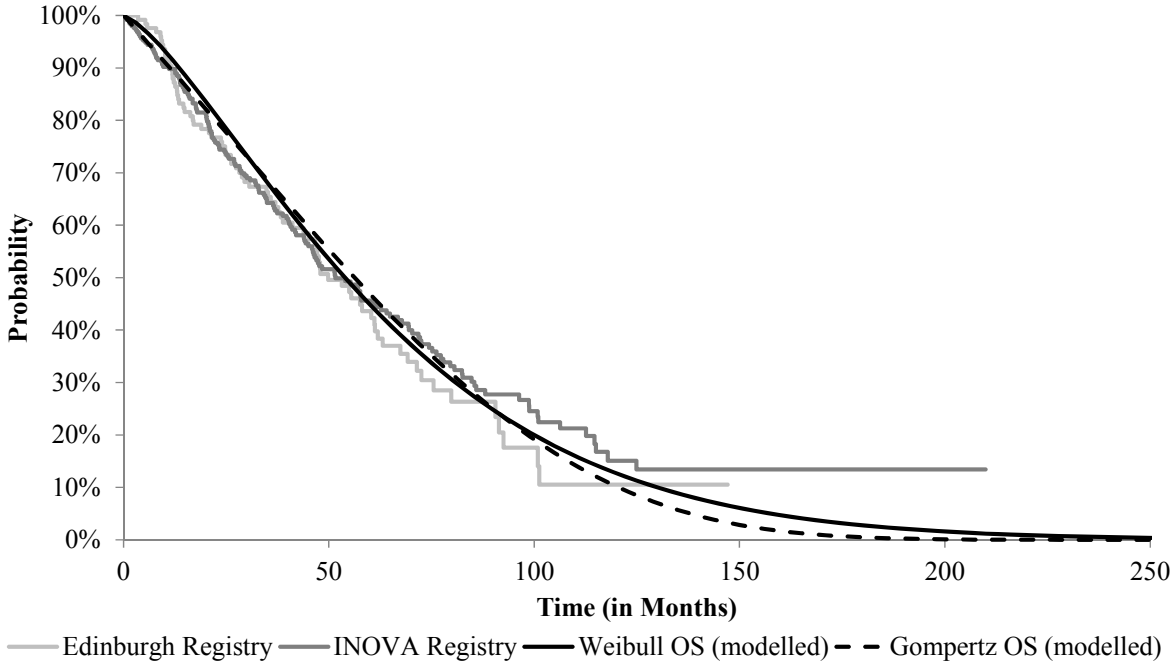
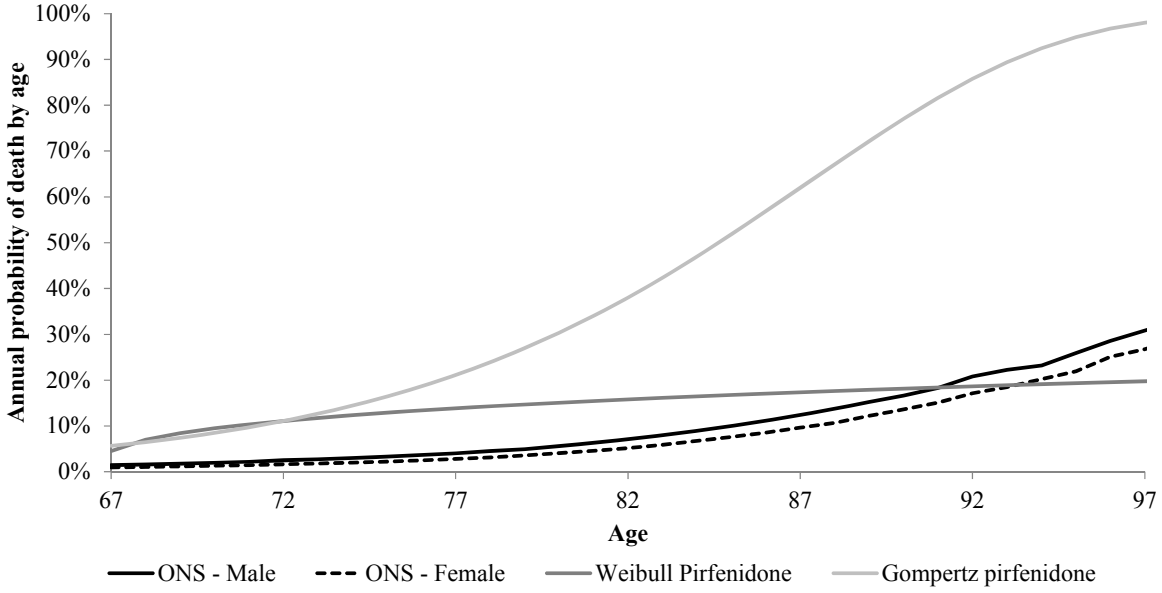


Figure 38: Plot of the annual hazard of death of modelled OS for BSC using the Weibull and Gompertz distribution and life tables in the UK in the ITT population (Plot drawn by the ERG)



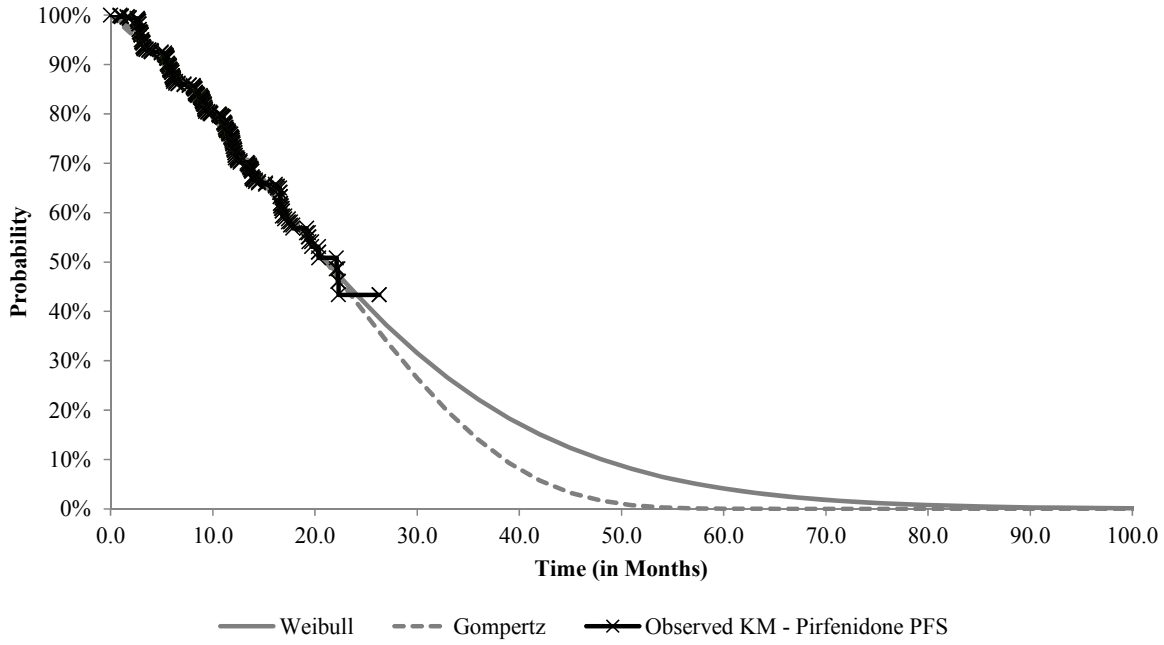
5.2.3.2. ERG's comments regarding the estimation of PFS

The ERG considers the definition of PFS used in the economic model to be largely appropriate. Discussion with clinical experts indicated that DLco is also considered to be clinically important but this is not as well accepted as a clinical trial endpoint (see Section 3) and was not included in the ASCEND trial.[34] Therefore, the ERG considered the definition used by the company based on the ASCEND trial[34] to be largely appropriate. Nevertheless, the ERG observes that PFS and its definition have only a minimal impact in the model and that the key driver of cost-effectiveness is OS.

The choice of the Weibull over the Gompertz distribution in the base-case is again questionable (Figure **39**). However, the impact on the ICER is minimal (increase from [redacted] per QALY gained using the Gompertz distribution – ITT population, against BSC), so any bias is likely to be small given the current model structure. However, the ERG notes that they would expect PFS to have a larger impact on the ICER if the relationship between disease progression, treatment discontinuation and treatment effect following discontinuation had been modelled in a more realistic manner.

The ERG further identified some inconsistencies in the approach used to estimate the hazard of progression for the subgroups. In response to a request for clarification from the ERG (see clarification response,[10] question B20), the company provided additional analyses and options in the economic model to use a more consistent methodology for PFS for the subgroups. Whilst the ERG expected the change to affect the subgroup analyses, the ERG is unclear why this also affects the results for the ITT population.

Figure 39: Comparison of the observed KM for PFS in people initiating pirfenidone against extrapolation using parametric distributions for the ITT population (Plot drawn by the ERG)

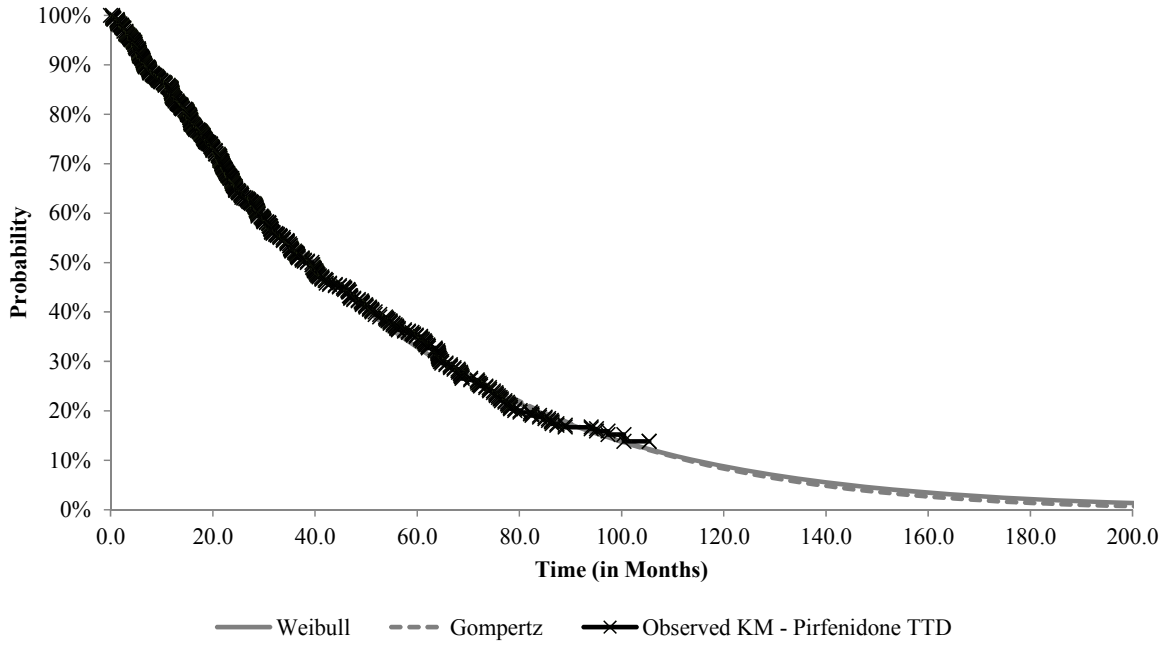


5.2.3.3. ERG's comments regarding the estimation of time to discontinuation

The ERG is satisfied with the approach used by the company to censor death and lung transplantation when estimating the time to discontinuation. However, the ERG notes that the censoring of lung transplantation may introduce bias as lung transplantation was not included in the base-case model structure; however, the impact is likely to be minimal given the small numbers discontinuing due to lung transplantation (see CS, Figures 4 to 6) in the ASCEND and CAPACITY trials.

The company's base-case uses the Weibull distribution for time to discontinuation based upon both visual and statistical goodness of fit to the observed portion of the pirfenidone curve. Alternative curve fits are explored as sensitivity analyses. The ERG considers that the choice between the Gompertz and Weibull distribution is questionable (Figure 40), but also that the impact on the ICER is again minimal (reduction in the ICER from [redacted] per QALY gained for ITT population for the comparison between pirfenidone versus BSC). However, as with PFS, the ERG would expect treatment discontinuation to have a larger impact on the ICER if the relationship between treatment discontinuation and treatment effect following discontinuation had been modelled in a more realistic manner. As detailed in Section 5.2.2.2, health outcomes are disconnected from costs, and therefore increasing the discontinuation rate leads to similar health outcomes at a lower costs and therefore an improved ICER for pirfenidone.

Figure 40: Comparison of the observed KM for discontinuation in people initiating pirfenidone against extrapolation using parametric distributions for the ITT population (Plot drawn by the ERG)



5.2.4. Treatment effects used in the company's base case for OS, PFS and time to discontinuation for pirfenidone vs. BSC and nintedanib

Treatment effects for pirfenidone against nintedanib and BSC are summarised in Table 47. The company's base-case analysis uses the treatment effects (HR) for pirfenidone against nintedanib and BSC (applied as inverse HR to the baseline pirfenidone curve) for the outcomes of OS and PFS reported in the Section 5.3 of the CS. The treatment effects are estimated from a random effects model which included all Phase II and Phase III trials using data up to 52 weeks (with the exception of SP2). Whilst only the OR are presented in the CS within the clinical section for the relative increase in discontinuation for nintedanib (compared with pirfenidone), ORs from the NMA are transformed into relative risks and used in the model subsequently. The company uses alternative models in scenario analyses including fixed effect models and data up to 72 weeks.

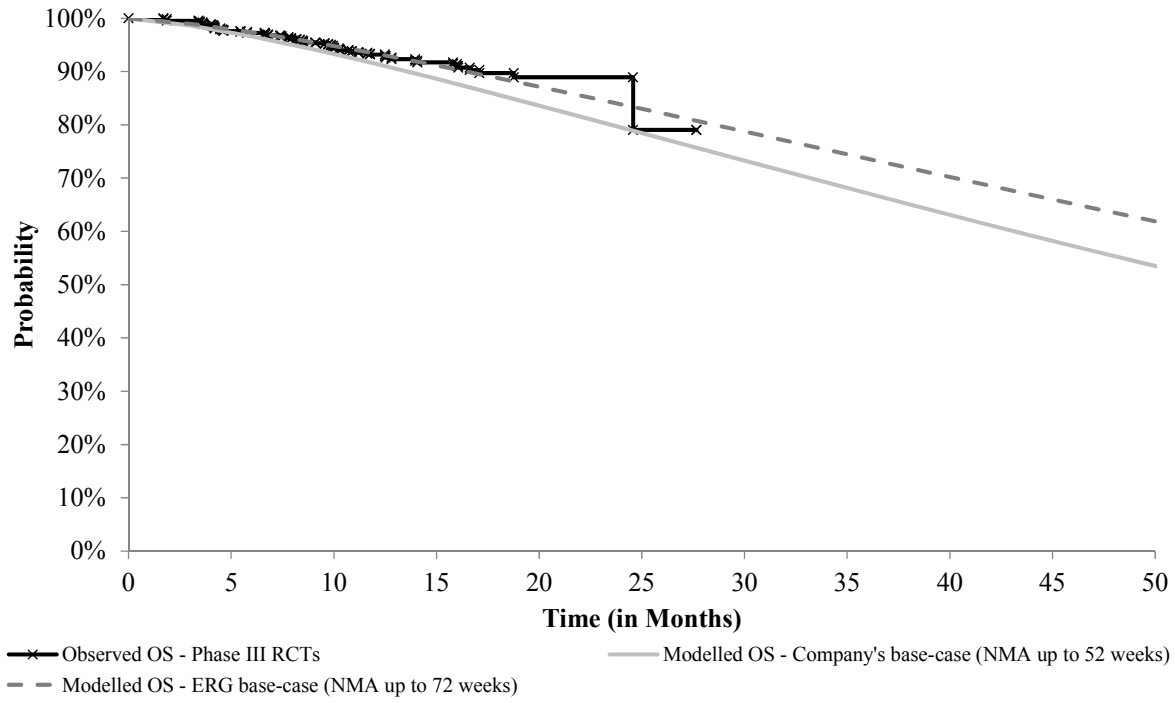
Table 47: Treatment effects used in the company's base-case

Treatment	Base-case HR (pirfenidone vs comparator)		Base-case RR (pirfenidone vs comparator)
	OS	**PFS**	**TTD**
Nintedanib	0.72	0.85	1.08
Best supportive care	0.52	0.63	NA

5.2.4.1. ERG's comments regarding the treatment effects used for OS in the company's model

As described in Sections 4.6 and 4.7, the CS reports results from a series of NMAs with varying strength and weaknesses, which are subsequently used in the company's model. In the company's base-case, the treatment effects (median) are estimated from a random effects model including all Phase II and III trials (referred to as the expanded network by the company) using data up to 52 weeks. As discussed in Section 4.7, the ERG considers that: (i) the treatment effects estimated using data up to 72 weeks are more appropriate and consistent with the company's assumption of proportional hazards; (ii) SP3 should be excluded from the base-case as this is a different (Japanese) population with a different dose and statistical adjustments were required as HRs were not reported, and; (iii) using the treatment effect at 52 weeks does not provide a reasonable fit to the observed KM for BSC (see Figure 41).

Figure 41: Fit of the modelled BSC using results from the NMA at 52 week (company's base-case) and 72 weeks (estimated by the ERG used in exploratory analyses)



The ERG further notes that the treatment effects taken from the NMA reported in the clinical section and subsequently used in the economic model, use posterior medians as point estimates, and associated 95% CrI. The ERG considers the use of the median in the economic model to be inappropriate and considers that the CODA samples (from the predictive distribution) should be used for the purpose of the modelling. As shown in section 6.1, using the median or mean point estimate could lead to inconsistent results.

5.2.4.2. ERG's comments regarding the treatment effects used for PFS in the company's model

The ERG notes that the treatment effect for PFS taken from the company's NMA using data up to 52 and 72 weeks are broadly the same between pirfenidone and BSC (HR: approximately 0.63) or nintedanib (HR: approx. 0.74), and therefore the company's decision to use the 52 week data instead of the 72 week data is unlikely to have a significant impact on the ICERs for pirfenidone. The ERG also notes that different definitions of PFS are used between trials included in the NMA. This is acknowledged in the CS (page 143). The ERG considers that this is likely to introduce some biases between pirfenidone and nintedanib but reiterates that PFS has a minimal impact on the ICER in the company's model. Although, as stated previously, it is expected that it would have a greater impact if progression was linked to treatment discontinuation and treatment effects were allowed to differ after discontinuation. Finally, as described above, the ERG considers that the CODA samples should be used in the model.

5.2.4.3. ERG's comments regarding the relative difference in treatment discontinuation in the company's model for pirfenidone and nintedanib

Although unclear from the CS, the OR estimated from the NMA were transformed into RR. The ERG notes that the relative risk for discontinuation (RR) is calculated for discontinuations for any reason but is applied in the model to people who discontinued treatment from reasons other than death and lung transplants. This may introduce bias if the rates of death or lung transplant differ between the trial arms compared in the NMA. The ERG further considers that the CODA samples should be used in the model.

5.2.5. Inclusion of costs associated with IPF-related mortality

The company's original base-case analysis assumes that 57.89% of deaths occurring in people initiating pirfenidone are IPF-related, based on the data from the CAPACITY/ASCEND trial (see CS, page 214), with the remaining deaths occurring due to causes unrelated to IPF. The proportion of deaths related to IPF in people initiating BSC and nintedanib was assumed to be greater compared with people initiating pirfenidone (72.22% and 68.57%, respectively). These figures were reported as being derived by applying data from the NMA, although the CS does not describe exactly how this was done.

In response to a request for clarification (question B10) regarding the source of the estimates for the proportion of deaths which are IPF-related, the company amended the methodology in the revised economic model by using the proportion of observed deaths that are IPF-related for each treatment, according to the company at 52 weeks from their respective trials (Table 48). The ERG observed that compared with the statement from the company, data at 72 weeks from the CAPACITY trials are used.

Table 48: Revised IPF-related mortality figures (reproduced from the clarification response,[10] question B10)

Intervention	Time point	n of IPF-related deaths	N of all-cause deaths	Proportion of death IPF-related
Pirfenidone	52 weeks*	17	32	53.13%
Placebo		35	50	70.00%
Nintedanib		26	42	61.90%
* contrary to the statement from the company data at week 72 from the CAPACITY trials are used				

A one-off cost of £9,996 is assigned in the model only for deaths attributable to IPF. This cost was taken from estimates provided in a report from the National Audit Office (2008)[101] and inflated to 2014 prices. This data source was also used in the nintedanib submission.[26] No costs are applied to deaths that are unrelated to IPF. Little justification is provided in the CS with the exception of the following statement "*costs associated with IPF are greatly increased in the last year of life due to increased resource use, home care and length of stay in hospital*" (see CS,[4] page 2014).

5.2.5.1. ERG's comments on the impact of IPF related-mortality on costs

The ERG has a number of concerns with the approach used by the company which included costs associated with end of life only in people dying from IPF-related causes.

The ERG considers the approach used by the company to estimate the proportion of death that are IPF-related in the revised economic model to be questionable and that ideally results from the NMA should be used. The ERG notes that whilst an NMA was used in the original submission to NICE, it was unclear on how this was done. The baseline proportion of death (whilst on BSC) was also unclear

in the original submission. As a result, both of the approaches presented by the company are considered questionable.

No details are included within the CS on the source used to represent the costs associated with death from IPF-related causes. In brief, Hatziandreu *et al.* (2008)[101] is a modelling study which aimed to estimate the total costs of care provided to people in their last year of life for both cancer and organ failure (heart and pulmonary disease) in three settings; hospital, hospice and home. As with any modelling study, this economic analysis relied on a series of assumptions. Due to time and resource constraints, the ERG is not able to provide a complete assessment of this study. Nevertheless, the ERG notes that the company assumes the costs associated with end of life care in the last year of life in people dying from organ failure to be a reasonable approximation of the cost of care provided in the last year of life for people dying of causes related to IPF.

The ERG considers that deaths from causes other than IPF are also likely to be associated with costs that fall on the NHS. The exclusion of costs associated with death from other causes is likely to bias the cost-effectiveness estimate in favour of pirfenidone given the lower IPF-related mortality assumed in the company's model. The ERG asked the company to provide evidence which demonstrates that the costs in the last year of life for IPF-related deaths are higher than the costs in the last year of life for deaths from other causes (see clarification response,[10] question B14). In response, the company stated: "*The purpose of the submission was to assess the cost and clinical implications associated with pirfenidone for the treatment of IPF compared with current care. As a result of this, only costs borne by the condition have been considered in the analysis, with other costs deemed out of the scope of this analysis and unrelated to the decision problem. Consequently for end of life care, costs attributable to death from other causes than IPF are not included.*" The ERG considers the response from the company to be unsatisfactory and considers that the assumption of zero costs for people with IPF who die from other causes to be inadequately supported by evidence. The ERG further notes that Hatziandreu *et al.* (2008) states that "*It is undisputable that patients with organ failure who are at the end of their life have palliative care needs as severe and distressing as those with cancer. Patients suffering from other conditions of equal importance in terms of prevalence, and economic burden, such as dementia or renal failure are also subject to end of life care services*."[101]

Consequently, in the absence of evidence relating to the differential impact of deaths on resource use, the ERG considers that the cost associated with end of life should be applied to all deaths irrespective of the cause.

5.2.6. Incorporation of the impact of acute exacerbations on costs and quality of life in the company's model

As described in Section 5.2.2, the company's model includes the impact of acute exacerbations as a cost and HRQoL decrement during each model cycle which is applied according to the treatment received (pirfenidone, BSC and nintedanib). Table 49 summarises the costs and utility decrements per model cycle assumed in the company's base-case in people receiving pirfenidone, BSC and nintedanib.

Table 49: Management costs and utility decrements (per model cycle) associated with acute exacerbations assumed in the company's model

	Pirfenidone	**Best supportive care**	**Nintedanib**
Decrement in utilities assumed per model cycle	- 0.00103	- 0.00165	- 0.00091
Management costs assumed per model cycle	£114.08	£226.34	£114.08

The utilities decrements associated with acute exacerbations are calculated from three components:
The incidence of acute exacerbations per model-cycle (3 months) by treatment. The rate for people treated with BSC is taken from the nintedanib submission.[26] The incidence is then adjusted in individuals receiving pirfenidone and nintedanib using ORs estimated from the NMA (see Table 50).

(i) The decrement in utilities associated with an acute exacerbation. The decrement is also taken from the nintedanib submission[26] and is calculated as a weighted average between the decrement in the first and subsequent months (see Table 51).

(ii) Assumptions regarding the duration over which HRQoL is reduced due to an exacerbation. This duration is assumed to be 3 months.

Table 50: Incidence of acute exacerbations (adapted from the CS, page 219, Table 80)

Treatment	Base-case OR vs BSC	Incidence of acute exacerbation assumed in the company's model
BSC	Baseline risk of 1.46% per model cycle	
PFN	0.62	0.91%
NTB	0.55	0.81%

Key: BSC, best supportive care; comp, comparator; NTB, nintedanib; OR, odds ratio; PFN, pirfenidone.

Table 51: Decrement in utility associated with an acute exacerbation (reproduced from CS,[4] Table 90, page 233)

Time frame	Utility [SE]	Reference
First month	-0.274 [0.059]	Nintedanib NICE company submission[26]
Subsequent months	-0.033 [0.053]	
Per model cycle	-0.113	First month + 2 * Subsequent months

Key: Dist, distribution; N/A, not applicable; NICE, National Institute for Health and Care Excellence; SE, standard error.

In contrast, the costs associated with the management of acute exacerbations per model cycle (see Table 52), are calculated from two components:

(i) The probability and duration of hospitalisations in people initiating BSC and pirfenidone. These are calculated using data from the CAPACITY trials.

(ii) The average cost associated with a hospital bed day. This is taken from the NHS Reference Costs for hospitalisations due to respiratory failure (HRG code: DZ27S, DZ27T, DZ27U) and is assumed to be £266.71 per day.

In the absence of comparable data, people treated with nintedanib are assumed to incur the same hospitalisation costs as people treated with pirfenidone.

Table 52: Calculation of hospitalisation cost (reproduced from CS,[4] Table 99, page 243)

	PFN	BSC
Number of cycle-length intervals observed [a]	3768	3771
Number of subjects with hospitalisation [b]	195	202
Rate of hospitalisation per cycle [c = a/b]	0.052	0.054
Probability of hospitalisation per cycle [d = 1-exp[c]]	0.050	0.052
Average length of stay in hospital [e]	8.48	16.27
Total cost of hospitalisation [f = e * cost of bed day]	£2,261.70	£4,339.37
Hospitalisation cost applied per cycle [g = d * f]	£114.08	£226.34

5.2.6.1. ERG's comments regarding the estimation of the costs and decrement in utilities associated with acute exacerbations

As described in Section 5.2.2.1, the inclusion of acute exacerbations (as implemented by the company in the economic model) has a minimal impact on the ICER, and therefore, only a brief critique is presented here. The ERG considered the lack of impact of acute exacerbations in the economic model to be an artefact of the model structure chosen by the company and not a reflection of the relevance of exacerbations in IPF.

The ERG also notes that the approach used by the company to include the impact of acute exacerbations and costs and HRQoL is inconsistent as different trial outcomes are used to estimate the impact of acute exacerbations on costs and QALYs. As a consequence of this inconsistency, pirfenidone is assumed to have a greater decrement in HRQoL associated with acute exacerbations compared with nintedanib during each model cycle, whilst the per cycle cost is assumed to be identical. The ERG considers the inputs for the decrement in utilities and costs for acute exacerbations to be broadly reasonable and notes that the majority of inputs are taken from the nintedanib submission and have a limited impact on the ICER.

The ERG further notes that data on hospitalisations used to represent the costs associated with acute exacerbations are not specific to hospitalisations due to acute exacerbations. The ERG sought clarification from the company on the inconsistencies in the approach to include the impact of acute exacerbations on costs and HRQoL (see clarification response,[10] question B15). In response, the company confirmed that data include hospitalisations from any causes and therefore the hospitalisation costs is broader than just the cost associated with acute exacerbations. The CS provides limited detail on the data used to inform the incidence and length of hospital stay. The ERG is unclear whether data on only respiratory-related, IPF-related or hospitalisations from any cause were used.

5.2.7. Incorporation of the impact of AEs on costs and HRQoL in the company's model

The impact of AEs on costs and HRQoL is applied during each model cycle according to the treatment currently received.

Table 53 summarises the costs and utility decrements per model cycle assumed in the company's revised base-case model in people receiving pirfenidone, BSC and nintedanib. In response to the ERG's clarification requests (see clarification response,[10] question A39 and the summary of model changes on page 38), the company amended the calculation of the incidence of AEs which led to a number of new errors being introduced into the model (described below). As can be seen from

Table 53, in the revised economic model the errors led to people on BSC experiencing greater costs and QALY impacts compared with people on pirfenidone or nintedanib.

Table 53: Costs and utility decrements (per model cycle) associated with AEs when averaged across the treated cohort (as applied in the revised company model) - prior to correction of errors by the ERG

	Pirfenidone	**Best supportive**	**Nintedanib**
Management costs	£93.79	£109.47	£32.18
Decrement in utility	-0.0040	-0.0052	-0.0015

The decrements in utilities and management costs associated with AEs per model cycle are calculated from:

(i) The incidence of AEs by treatment
(ii) The management costs associated with each AE
(iii) The utility decrement associated with AE
(iv) The assumed duration of the decrement for each AE.

Inputs and assumptions for each AE are summarised in Table 54. The company considered the same AEs included in the nintedanib submission;[26] namely: serious cardiac events, serious gastrointestinal event (which is subsequently replaced by the company by diarrhoea in the revised economic model), gastrointestinal perforation (nintedanib only), photosensitivity reaction (pirfenidone only) and rash (pirfenidone only in the original economic model and both pirfenidone and BSC in the revised economic model).

Table 54: Incidence, costs and decrement in utilities associated with each AE (per individual experiencing the event) included in the revised company model (prior to correction of errors by the ERG)

	Incidence			Decrement in utilities		Costs
	PFN	**BSC**	**NTB**	**Disutilities**	**Duration**	
Serious cardiac event	0.79%	1.05%	0.23%	-0.198	3 months	£2,200.15
Diarrhea	3.48%	4.52%	1.36%	-0.068		£1,910.91
Gastrointestinal perforation	-	-	0.08%	-0.118		£1,583.03
Photosensitivity reaction	2.32%	-	-	-0.032	15 days	£467.62
Rash	0.00%	1.82%	0	-0.03		£428.63

The incidence of AEs for rash, serious gastrointestinal events (assumed to be diarrhoea in the revised economic model) and serious cardiac events for BSC are calculated based on the average incidence for the placebo arm across the Phase III trials. Relative risks from the NMA submitted as part of the company's clarification response (see clarification response,[10] question A39) are then used to derive the incidence in people treated with pirfenidone and nintedanib. The incidence of AEs for gastrointestinal perforation and photosensitivity were taken from the nintedanib submission.[26]

The utility decrement associated with serious cardiac events, gastrointestinal events (replaced by diarrhoea in the revised economic model) and perforation were taken from the nintedanib submission[26] and are assumed to last 3 months. The utilities decrement associated with photosensitivity and rash were taken from Handorf *et al.* (2012)[102] and the NICE Centre for Clinical Practice, respectively, and are assumed to last 15 days. Costs associated with the management of AEs were taken from the NHS Reference Costs (2014-15)[103] using a similar approach to that used in the nintedanib submission.[26]

5.2.7.1. ERG comments on the inclusion of AEs in the economic model

As with acute exacerbations, the ERG notes that the inclusion of AEs (as implemented by the company in the economic model) has a minimal impact on the ICER, and therefore, only a brief critique is provided here.

The company revised their estimates of the incidence of AEs using results from the NMA in response to clarification on a separate issue (see clarification response,[10] question A39). The ERG considers this revised approach to be more appropriate. Nevertheless, the ERG identified a series of errors in the implementation of this within the company's revised model. The ERG notes that results from the NMA suggest that pirfenidone has a greater incidence of AEs (serious cardiac events, rash, diarrhoea) compared with placebo. However, in the economic model, the incidence of AEs used for pirfenidone is lower compared with BSC. The ERG notes that this is because the RRs from the NMA are applied incorrectly in the model. Furthermore, the ERG notes some discrepancies between results from the NMA reported in the clarification response (see clarification response,[10] question A39) and the data from the NMA used in the economic model.

The ERG further notes that in the original economic model, the company included serious gastrointestinal events; which was subsequently replaced by diarrhoea in the revised economic model. The ERG considers the inclusion of diarrhoea to be appropriate. However, whilst data on the incidence of diarrhoea appear to be used, the costs and utility decrements associated with serious gastrointestinal events are still used. If the costs and utility decrements associated with diarrhoea are lower than those associated with serious gastrointestinal events, then the approach used by the company is likely to be unfavourable to nintedanib which is considered to be associated with a greater incidence of diarrhoea compared with BSC and pirfenidone.

5.2.8. HRQoL

Table 55 summarises the health state utility values assumed within the company's model.

Table 55: Summary of health state utility values used in the company's model for the base-case and sensitivity analyses

Health state	Pirfenidone (Base-case)	Pirfenidone Trial (Alternative mapping)[95]	Panther & ACE	Ofev STA[26]
Progression-Free	0.847	0.791	0.82	0.777
Progression	0.7818	0.744	0.74	0.744
Transplant	Assumed to be the same as progression-free			

The CS includes details of a systematic review of studies which provide estimates of HRQoL for adult people with mild to moderate IPF (see CS, Section 5.4). As described in Section 5.1.1, the company undertook a single search in November 2015 to identify cost-effectiveness studies, HRQoL studies and resource use data. A total of 22 references were included in the HRQoL review, of which 5

references (corresponding to 4 studies) reported EQ-5D data which are briefly described within the CS (CS, pages 228 to 230) with the remaining studies described in the CS Appendix 22. This included EQ-5D data collected in Richeldi *et al.* (2014)[72] and used in the 2015 nintedanib submission,[26] EQ-5D (measure using time trade off) data from a registry,[104] SF-36 and EQ-5D data from a RCT comparing bosentan and placebo,[105] and EQ-5D data from a RCT comparing sildenafil and placebo.[96]

In the base-case, the company obtained EQ-5D utility scores for the progression-free and progressed health states based on the mean SGRQ score collected in the CAPACITY trials for each health state. The mean SGRQ in people who are progression-free (37.31) and progressed (42.40) was estimated from the CAPACITY trials[49] using a generalised estimating equation (GEE) regression model to account for the correlation between measurements from the same individual at different time points. These mean scores were then mapped onto the EQ-5D using a mapping algorithm published by Freemantle *et al.* (2015) which was developed specifically for people with IPF.[106] A scenario analysis is also presented using a mapping algorithm published by Starkie et al;[95] this algorithm was estimated in people with Chronic Obstructive Pulmonary Disease (COPD).

Alternative utility values, identified from the company's systematic review, are used in additional scenario analyses. For the sensitivity analysis including lung transplantation, the utility value associated with lung transplantation was assumed to be similar to the value for the progression-free health state. The limitation of making such an assumption is acknowledged by the company.

5.2.8.1. ERG's comments regarding the estimation of health state quality of life

Limited detail is included on the methods for the systematic review of HRQoL data. Whilst the search strategies and list of excluded studies are provided in the CS Appendices 17 and 18, no information is provided on the outcomes that were eligible for this review, thereby making it difficult to assess why some studies were excluded. The ERG notes that the link between the systematic review of HRQoL literature and the evidence selected for use in the model is unclear. For example, the mapping study used to map from the SGRQ data from the CAPACITY trials to the EQ-5D[106] was excluded from the systematic review as it was deemed to include an ineligible patient population. However, the study by Freemantle was conducted in people with IPF (as described on page 225 of the CS) and therefore the reason for exclusion is inconsistent. The study by Nathan *et al.* (2015) was also excluded (see CS,[4] Appendix 18) due to the inclusion of an "ineligible population", however the population relates to people with diagnosed IPF and therefore seems relevant. The company also did not identify the study by Raghu *et al.* (2013)[75] which reported EQ-5D data. This study was reported in Loveman *et al.* (2014)[93, 94] which was included in the review of cost-effectiveness studies but appears to have been excluded from the HRQoL review. Therefore the ERG cannot be certain that all relevant HRQoL data have been identified and presented in the CS.

In the company base-case, EQ-5D utility scores are estimated from the trial. The ERG notes that the mean SGRQs are transformed into an EQ-5D score based on a linear mapping function published by Freemantle *et al.*[106] The ERG considers the general method of using aggregate data for the mean SGRQ to be appropriate in principle given the use of a linear mapping function. Nevertheless, the ERG notes that the mapping algorithm is unconstrained and therefore could predict values greater than 1.0 when the SGRQ is below 26 when applied at the individual level. In response to a request for clarification from the ERG (see clarification response,[10] question B18a), the company reported that 25.5% of individuals at baseline and 19.5% based upon the last observation had a SGRQ score below 25. The ERG considers that given the large proportion of people at the individual level with a SGRQ score below 25, utility values estimated at the aggregate level could be biased and that utility values should be estimated at the individual level and capped at 1.0. Within the company's clarification response (see clarification response,[10] question B18c), these estimates were provided by the company and led to a decrease in utility value in people without progression (from 0.8485 to 0.8185) and people who progressed (from 0.7835 to 0.7597). The impact is examined by the ERG in Section 6.

The company further assumes that HRQoL is constant within the respective health states. The ERG considers that HRQoL is likely to reduce over time within each health state, or at least, within the progressed disease health state given the progressive nature of the condition. However, as described, given the model structure chosen by the company, including a HRQoL decrement for further disease progression in the progressed state is challenging. The ERG considers this implication to likely overestimate the number of QALYs but the size and direction of the effect of incremental QALYs and therefore the ICER is unclear.

The ERG observes that alternative utility values used in scenario analyses have a moderate impact on the ICER. The ERG considers this to be an artefact of the model structure chosen by the company in that HRQoL is assumed to remain constant within health states and not a reflection of the impact of HRQoL in IPF. The ERG further observes that no adjustment for utility by age is assumed and therefore, people are assumed to maintain the same level of HRQoL irrespective of age. The ERG considers this implication to likely overestimate the number of QALYs and potentially the incremental QALY gain associated with pirfenidone. This is examined by the ERG in Section 6.

5.2.9. Drug acquisition costs and resource use associated with the management of IPF

This section discusses the drug acquisition and administration costs and the resources use and costs associated with the management of IPF included in the company's model.

Drug acquisition costs

Before accounting for dose reductions and interruptions, the company's model assumes that people treated with pirfenidone receive a total of nine 267mg capsules per day, at a daily cost of £71.70 (or a cost per 3-month model cycle of £5,730.62) using the list price.

The company calculates that in the CAPACITY and ASCEND trials,[34, 49] a fewer number of pills were given daily, with an average of 7.88 capsules per day. This is used in the company's base-case, leading to a daily cost of £62.80 (or a cost per 3-month model cycle of £5,730.62) using the list price.

The daily cost in people treated with nintedanib is assumed to be at parity with the daily cost calculated for pirfenidone. This is justified on the basis that this assumption was made in the nintedanib submission.[26]

No drug acquisition costs are assumed for people receiving BSC. This is justified by the company on the basis that BSC represents the placebo arm of the trials from which efficacy data were derived. Similarly, no concomitant medications are assumed in the company's model.

Administration costs

No administration costs were assumed for either treatment as both nintedanib and pirfenidone are taken orally.

Costs associated with the management of IPF and lung transplant

The CS presents the methods and results of a systematic review of studies with the aim to identify data on resource use and costs for adult patients with mild to moderate IPF. As described in Section 5.1.1, a single search strategy was conducted in November 2015 to identify cost-effectiveness studies, HRQoL studies and resource use data. A total of 7 references were included in its review for resource use, of which 2 were cohort studies,[107, 108] one was an economic evaluation alongside an RCT in patients with IPF,[109] one was a cost-effectiveness analysis in patients with cystic fibrosis[110] and three were technology appraisals by national bodies (NICE and the SMC).[111,42, 112]

The company separates treatment-specific monitoring costs (resource use at initiation and liver function tests) and the costs associated with the progression status. These are summarised in Table 56.

Table 56: Cost of resource use per model cycle (reproduced from CS,[4] Table 96)

Cost type		**PFN**	**BSC**	**NTB***
Cycle specific costs	Cycle 1	£969.38	£964.71	£969.38
	Cycle 2	£5.61	£0.94	£5.61
	Cycle ≥ 3	£1.87	£0.56	£1.87
Progression status specific costs	Pre-progression	£513.22	£513.22	£513.22
	Post-progression	£525.44	£525.44	£525.44

Key: BSC, best supportive care; NTB, nintedanib; PFN, pirfenidone.

* only for people with moderate IPF

Cycle : 3 month

Resource use by progression status, which is not specific to the treatment received is summarised in Table 57. This was estimated from discussion with a panel of UK clinicians (see CS, page 239). As shown in Table **57**, the difference in resource use between health states is assumed to be dependent on the individual's percent predicted FVC for oxygen, healthcare professional visits and GP visits. Therefore, the only difference in management costs between health states is due to the differences in FVC between people who are progression-free and those with progression. The frequency of liver function tests (that are specific to treatments) was taken from the pirfenidone and nintedanib SmPC.

The cost associated with lung transplantation (which is used only in a sensitivity analysis) was taken as the average cost of lung/heart transplant reported in a report published by the NHS Blood and Transplant (2013) uplifted to 2014/2015.

Table 57: Resource use assumed in the company's model (based on CS, Tables 94 and 95)

Resource use item	At treatmen	Subsequent MRU	Unit cost
Liver function test	TRUE	*	£1.87
Gas transfer	TRUE	every 4 months	£202.08
Lung volume study	TRUE	None	£170.54
Full pulmonary (covers	TRUE	every 4 months	£165.85
Field exercise test	TRUE	every 6 months	£177.13
Oxygen	FALSE	for all patients with <80% FVC	£206.08
Healthcare professional visit	TRUE	every 4 months if FVC >60%, every 3 months if FVC<60%	£248.17 £177.53
GP visit	FALSE	based upon FVC	£37.00
Key: FVC, forced vital capacity; GP, general practitioner; MRU, medical resource use; SmPC, summary of product			
** Liver function tests were administered as per the pirfenidone SmPC for pirfenidone and nintedanib patients (every month for the first 6 months of treatment, then every 3 months), and for BSC are administered according to clinician opinion (every 1.5 months for the first 6 months of treatment, then 0.3 times per model cycle).*			

5.2.9.1. ERG's comments

As with the HRQoL searches, few details are included in the CS regarding the methods for the systematic review of resource use. Whilst the search strategies and list of excluded studies are provided in the CS (Appendices 17 and 18), no information is provided on the outcomes that were eligible for this review, thereby making it difficult to assess why certain studies were excluded. The ERG notes that the link between the systematic review of resource use and the evidence selected for use in the model is also unclear. For example, the CS states that the estimates of unit costs from the nintedanib submission are preferred by the company over those identified from the literature (see CS, page 237). However, the ERG notes that the nintedanib submission is not included in the company's resource use review. The CS also does not present whether the costs are consistent across the different sources, thus it is unclear whether the data in the nintedanib submission reflect the data from other published sources. The ERG further notes that the company identified a study reporting the impact of pirfenidone in a real-world setting through the UK Named Patient Programme using a retrospective study design. Findings from this study are not used or compared with the resource used in the model. The study by Loveman et al[93, 94] which was included in the review of cost-effectiveness studies, is also not included in the company's resource use review. The ERG considers this to be inappropriate

particularly given the similarities between health states between the Loveman et al[93, 94] model and the company's model. Therefore, the ERG cannot be certain that all relevant resource use data have been identified and presented in the CS.

The ERG is generally satisfied with the inclusion of drug acquisition costs in the company's model but notes following clarification that; dose interruptions and reductions for pirfenidone are calculated after titration and therefore exclude the first 2 weeks. The ERG considers that a more appropriate approach would have been to separate the costs for the first model cycle from those for subsequent cycles. This is amended in the ERG preferred-base-case.

The ERG notes that the daily cost of pirfenidone and nintedanib is equivalent when assuming the full indicated dose is taken (after the titration period for pirfenidone) and when using the current list price.[16] However, assuming the same daily costs for pirfenidone and nintedanib based on the average dose used in the pirfenidone trials implies the same impact of dose reductions/interruptions for pirfenidone and nintedanib. The ERG notes that the price structure for pirfenidone and nintedanib is different and that a dose reduction with nintedanib (for instance, from 150mg to 100 mg) would not be associated with a reduction in costs. The ERG observes that the IMPULSIS trial[113] reported a compliance with nintedanib of 96.4 % whereas the mean dose applied in the model for pirfenidone is 87.6% of the indicated dose. Therefore, the ERG considers that assuming the same cost for pirfenidone and nintedanib is likely to favour nintedanib.

The company's base-case assumes no drug acquisition costs for BSC and/or concomitant medications. The ERG considers this to be inappropriate as within the trials, individuals received concomitant medications as part of BSC. This was included in the nintedanib submission at a cost of approximately £25 per model cycle calculated from the trial for both nintedanib and BSC.[26] However, the ERG notes that the impact of the ICER is likely to be minimal given that the cost will be applied to all arms.

The CS also reports that resources use estimates were derived from discussion with a panel of clinicians, although no details were provided in the CS. In response to a request for clarification (clarification response,[10] question B16), the company provided further details, stating that: "*One-to-one telephone interviews were conducted with the panel of UK clinical experts. Content of the earlier NICE manufacturer submission was discussed, along with how the approach employed to assess resource use in the earlier submission matched current clinical practice in IPF. Discussions accounted for the revised descriptions of the NHS Reference Cost list for 2014-15 compared to earlier years (e.g. revision of 'simple lung exercise function test' to 'field exercise test').*" Despite this additional clarification, the ERG considers the process used by the company to elicit resource use has

not been reported in a sufficiently transparent manner. It is also unclear how any potentially divergent views between clinicians were accounted for.

Contrary to the statement from the company, oxygen was not included in the company's original model; this was amended in the revised model. However, the ERG notes that resources use by progression status (notably oxygen, healthcare professional visits and GP visits) are driven by the level of percent predicted FVC. The ERG observes that given the structure chosen by the company, which is based on progression status and not percent predicted FVC level, the implementation of these percent predicted FVC dependent costs in the model relies on a series of assumptions. Notably, the company uses the percent predicted FVC distribution at baseline (divided into 10% bands) to represent the distribution of percent predicted FVC in people without progression and assumes a shift of one band in percent predicted FVC in people with progressive disease. This is arbitrary and an artefact of the chosen model structure.

The ERG further observes some double-counting in the first cycle for the costs associated with the management of the condition. However, as this has been done consistently in the pirfenidone, BSC and nintedanib arms, the effect is cancelled out across the treatment arms and therefore has no impact on the ICER.

In addition, as with the modelling of HRQoL by health state, the company assumes that resource use is constant within the respective health states. The ERG considers that resource use is likely to increase over time within each health states, or at least, within the progressed disease health state given the progressive nature of the condition. The ERG considers this implication to likely underestimate the management costs associated with IPF. The size of the effect and direction of the ICER is unclear.

5.2.10. Summary of data used for the subgroups of people with a percent predicted FVC of 50% to 80% and >80%

The only data that are subgroup-specific are the baseline OS, PFS and discontinuation curves. For the subgroup analyses (percent predicted FVC of 50 – 80% or >80% at baseline), parametric functions were fitted to the observed KM using percent predicted FVC as covariates (percent predicted FVC<50%, 50%≤FVC<80%, FVC≥80%) for both OS and time on treatment. However, a different approach was taken for PFS. In this case, the KM data for each subgroup were used separately to fit the parametric distributions. Whilst unclear from the CS, the Weibull distribution was also selected for the subgroups of people with a percent predicted FVC of 50 to 80% or >80. The other model parameters applied in the subgroup analyses were the same as those applied in the ITT analysis.

5.2.10.1. ERG's comments

The company uses data by subgroup to derive the baseline hazards of death, discontinuation and progression in people initiating pirfenidone. The ERG considers this to be appropriate. As described in Section 5.2.3.1, the ERG considers the Gompertz distribution to provide a more plausible long-term extrapolation for OS compared with the Weibull distribution in both the ITT population and in the subgroups of people with a percent predicted FVC of 50% to 80% and >80% (see [redacted] and [redacted]). [redacted]

The ERG further identified that a different approach was used for PFS compared with OS and time to discontinuation for the subgroup analyses. In their clarification response (see clarification response,[10] question B20), the company acknowledged the inconsistency and stated that using a consistent approach with OS and time to discontinuation had a minimal impact on the ICER (increase from [redacted] per QALY gained). Nevertheless, the ERG identified an error when the Gompertz and gamma distribution was used for PFS; this is corrected in the ERG's exploratory analyses (see Section 6).

The treatment effects from the ITT population are used to represent the treatment effect for the subgroup. The ERG considers this to be appropriate given the lack of stratification by subgroup in the trial, and the *post hoc* nature of the subgroup analysis. The ERG notes that approximately 75% of people enrolled in the trial had a percent predicted FVC ≤80%. The ERG notes that *post hoc* analyses provided by the company in response to clarification questions (question A31) showed numerically different treatment effects for OS, although it is unclear if the differences are real. The ERG further notes some apparent typographical errors in some of the values reported in Table 23 for OS in people with a percent predicted FVC > 80% from the ASCEND trial. As it is not possible to rule out with certainty a different treatment effect by subgroup, the ERG's exploratory analyses examine the impact of using the treatment effects by subgroups from the *post hoc* analyses.

The company further assumed that the impact of acute exacerbations, IPF-related mortality and AEs are the same between the subgroups of people with a percent predicted FVC of 50% to 80% and >80%. As described in Sections 5.2.5 to 5.2.7, the impact of different assumptions on the ICER is minimal and thus, the ERG is satisfied with using the same data by subgroup for the purpose of the model but highlights that the impact on costs and HRQoL could be different according to the subgroup examined.

Finally, the company assumed no differences in HRQoL and resource use for the progression-free and progression states in people with a percent predicted FVC of 50% to 80% and >80%. The ERG considers this to be inappropriate given that both HRQoL and resource use are a function of percent

predicted FVC for the progression-free and progressed states. The impact on the ICER of making this assumption is uncertain given the chosen structure as progression is not modelled.

5.2.11. Cost-effectiveness results

Results for the ITT population and the subgroup of people with a percent predicted FVC >80% are presented in this report both using the list price and with the PAS for pirfenidone. Results for the subgroup of people with a percent predicted FVC of 50 – 80% are presented using the list price only in this report with results using the PAS for pirfenidone and the PAS for nintedanib available in a confidential appendix. Whenever possible, results reported here are taken either from the results provided at the clarification stage following amendments made by the company, or in the case of results incorporating the PAS, from the PAS submission template. On some occasions results had to be re-run by the ERG where there existed discrepancies between the model and values reported by the company. These are highlighted as being taken from ERG analysis. Finally, it should be noted that the base-case results presentd by the company and reproduced here exclude the stopping rule for pirfenidone, but include a stopping rule for nintedanib as this was the assumption used in the company's base-case.

- **ITT population – Mild to Moderate IPF**

Table 58 summarises the estimated health gains and costs for each strategy for the ITT population. Pirfenidone is estimated to result in an additional 1.87 QALYs at an incremental cost of [redacted] (using the list price) compared with BSC, over a life-time horizon. This corresponds to an ICER for pirfenidone versus BSC of [redacted] per QALY gained. It can be seen from Table 59 that when applying the PAS for pirfenidone, the incremental cost for pirfenidone versus BSC is £40,010 and the ICER is £21,387.

Table 58: Central estimates (based on point estimates of parameters) of cost-effectiveness for the ITT population – discounted results (list price)

	Costs	LY	QALYs	Inc Costs	Inc LY		Inc QALYs	ICER
BSC	[redacted]	5.38	3.80					
pirfenidone	[redacted]	8.67	5.67	[redacted]	3.29		1.87	[redacted]

BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life

Table 59 Summary of definitions used for progression-free survival

	Costs	LY	QALYs	Inc Costs	Inc LY	Inc QALYs	ICER
BSC	£26,627	5.38	3.80				
pirfenidone	£66,638	8.67	5.67	£40,010	3.29	1.87	£21,387

BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

- **People with a percent predicted FVC > 80% at baseline**

Table 60 summarises the estimated health gains and costs for each strategy in people with a percent predicted FVC > 80% at baseline. Pirfenidone is estimated to generate an additional 2.17 QALYs at an incremental cost of [redacted] (using the list price) compared with BSC, over a life-time horizon. This corresponds to an ICER for pirfenidone versus BSC of [redacted] per QALY gained.

It can be seen from Table 61 that when the PAS for pirfenidone is applied, the incremental costs for pirfenidone versus BSC are £52,480 and the ICER is £24,187.

Table 60: Central estimates (based on point estimates of parameters) of cost-effectiveness in people with a percent predicted FVC > 80% – discounted results (list price)

	Costs	LY	QALYs	Inc Costs	Inc LY	Inc QALYs	ICER
BSC	[redacted]	7.11	4.82				
Pirfenidone	[redacted]	11.26	6.99	[redacted]	4.15	2.17	[redacted]

Table 61: Central estimates (based on point estimates of parameters) of cost-effectiveness in people with a percent predicted FVC > 80% – discounted results (PAS price)

	Costs	LY	QALYs	Inc Costs	Inc LY	Inc QALYs	ICER
BSC	£31,729	7.11	4.82				
Pirfenidone	£84,209	11.26	6.99	£52,480	4.15	2.17	£24,187

- **People with a percent predicted FVC of 50 - 80% at baseline**

Table 62 summarises the results for the subgroup of people with a percent predicted FVC of 50 - 80% at baseline. The company's model suggests that, at the list price, pirfenidone is the most effective and most expensive option and BSC is the least effective and least expensive option. Based on an incremental analysis of the three options, nintedanib is expected to be ruled out due to extended dominance at the list price. When compared with BSC, pirfenidone is estimated to generate an additional 1.696 QALYs at an incremental cost of [redacted]. The corresponding ICER for pirfenidone versus BSC is estimated to be [redacted] per QALY gained at the list price. When applying the pirfenidone PAS, the ICER for pirfenidone versus BSC was £21,318. The results for pirfenidone versus nintedanib are in the confidential appendix.

Table 62: Central estimates (based on point estimates of parameters) of cost-effectiveness in people with a percent predicted FVC of 50 - 80% – discounted results (list price)

	Costs	QALYs	Incremental results versus BSC			ICER (incremental analysis)
			Costs	QALYs	ICER	
BSC	[redacted]	3.443	-	-	-	-
Nintedanib	£65,065	4.226	£40,197	0.783	£51,331	Extendedly dominated
Pirfenidone	[redacted]	5.138	[redacted]	1.696	[redacted]	[redacted]

Table 63: Central estimates (based on point estimates of parameters) of cost-effectiveness in people with a percent predicted FVC of 50 - 80% – discounted results (with PAS)

	Costs	QALYs	Incremental results versus BSC			ICER (incremental analysis)
			Costs	QALYs	ICER	
BSC	£24,868	3.44	-	-	-	-
Nintedanib	See confidential appendix					Extendedly dominated
Pirfenidone	£61,012	5.14	£36,145	1.70	£21,318	£21,318

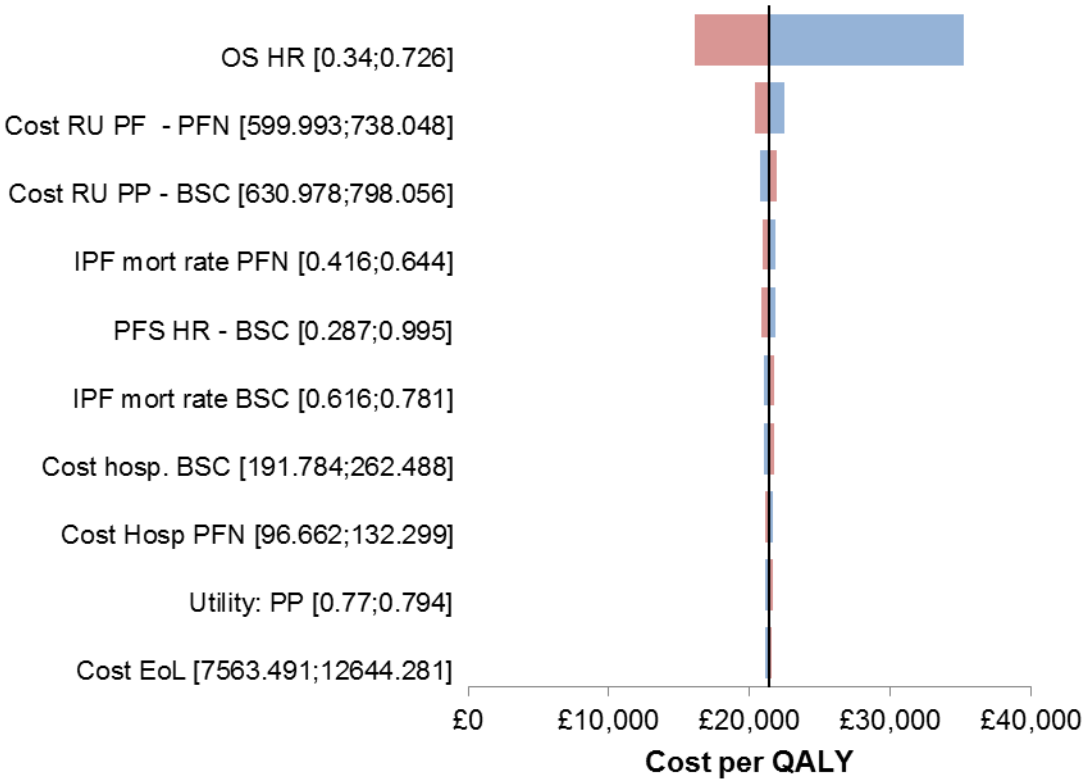
5.2.12. Sensitivity analyses

The company conducted a range of uncertainty analyses including probabilistic sensitivity analysis (PSA), deterministic univariate sensitivity analyses (DSA) and scenario analyses. Results for the ITT population for the DSA and PSA are reported here (taken from the company's clarification response[10] and the PAS submission template). Findings for scenario analyses and the subgroups of people with a percent predicted FVC of 50% to 80% and >80% are similar and therefore not reported here, but are available in the clarification responses[10] and the PAS submission template.

- **Deterministic one-way sensitivity analysis**

Figure 42Figure 42 shows the one-way DSA conducted by the company for the ITT population with the PAS applied. As recognised by the company, the ICER is most sensitive to the HR for OS.

Figure 42: Tornado diagram – ITT population, PAS price (reproduced from Figure 1 of the PAS submission template)



Key: BSC, best supportive care; EoL, end of life; Hosp, hospitalisation; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; mort, mortality; PF, progression-free; PFN, pirfenidone; PP, post-progression; QALY, quality-adjusted life year; RU, resource use.

- **Probabilistic sensitivity analysis**

The company reported results from the PSA for the ITT population and the subgroup of people with a percent predicted FVC of 50% to 80% and >80% in its clarification response.[10] These are summarised below in Table 64 when using the list price. Results when incorporating the PAS are presented in Table 65 for pirfenidone versus BSC but results for pirfenidone versus nintedanib for the moderate population are reported in the confidential appendix.

The company report PSA results which are close to the deterministic results for the ITT population. The ERG notes that results for the mild and moderate IPF subgroups for the deterministic and probabilistic analyses are also similar.

The cost-effectiveness acceptability curves (CEACs) and cost-effectiveness plane for the ITT population when incorporating the PAS are presented in Figure 43 and Figure 44, respectively.

Table 64: PSA results (list price) for ITT population and people with a percent FVC >80% at baseline

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER	Probability that pirfenidone is optimal at £20,000 per QALY gained	Probability that pirfenidone is optimal at £30,000 per QALY gained
ITT – mild to moderate IPF (see clarification response, Table 179)							
BSC	[redacted]	3.765					
Pirfenidone	[redacted]	5.68	[redacted]	1.91	[redacted]	[redacted]	[redacted]
People with a percent predicted FVC of 50 - 80% at baseline (see clarification response, Table 208)							
BSC	[redacted]	3.42					
Pirfenidone	[redacted]	5.15	[redacted]	1.74	[redacted]	[redacted]	[redacted]
People with a percent predicted FVC > 80% at baseline (see clarification response, Table 190)							
BSC	[redacted]	4.799					
Pirfenidone	[redacted]	7.00	[redacted]	2.21	[redacted]	[redacted]	[redacted]

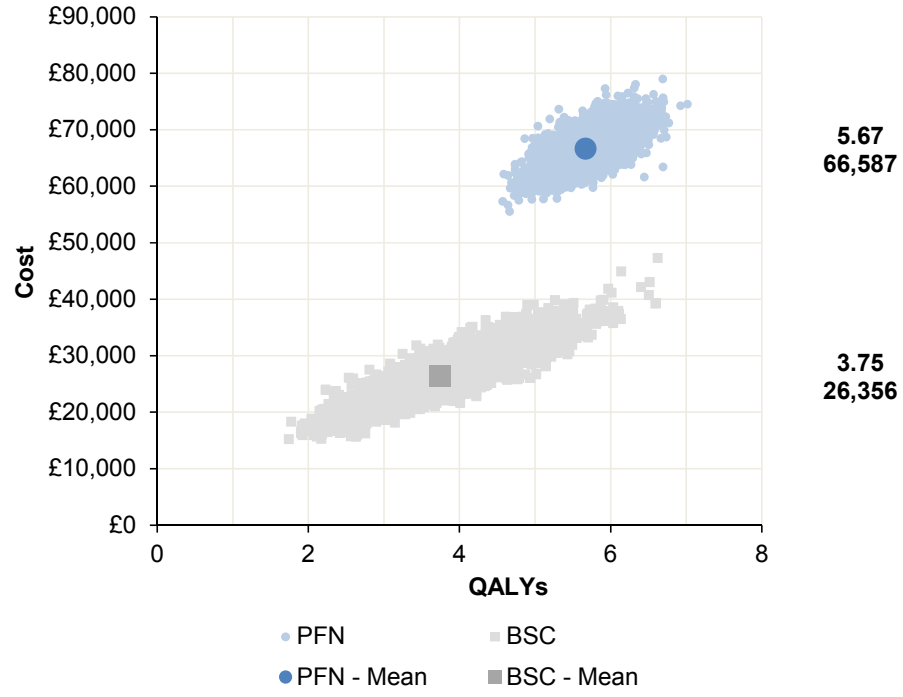
BSC – best supportive care

Table 65 PSA results (with PAS) for ITT population and people with a percent FVC >80% at baseline

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER	Probability that pirfenidone is optimal at £20,000 per QALY gained	Probability that pirfenidone is optimal at £30,000 per QALY gained
ITT – mild to moderate IPF (generated by ERG using company model submitted following clarification)							
BSC	£26,356	3.748					
Pirfenidone	£66,587	5.670	£40,231	1.92	£20,928	0.44	0.85
People with a percent predicted FVC of 50 – 80% at base-line (reproduced from PAS template, Table 15)							
BSC	£24,651	3.40					
Pirfenidone	£61,029	5.14	£36,378	1.74	£20,863	0.34	0.47
People with a percent predicted FVC > 80% at base-line (reproduced from PAS template, Table 14)							
BSC	£31,448	4.758					
Pirfenidone	£84,283	7.01	£52,835	2.25	£23,476	0.27	0.76

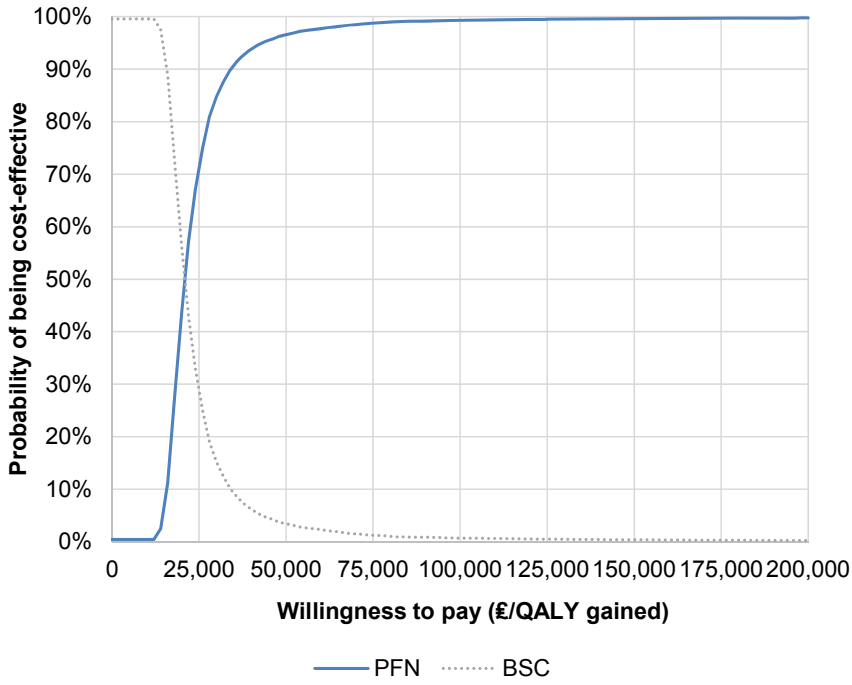
BSC – best supportive care

Figure 43: PSA scatterplot – ITT population, PAS price (generated by ERG)



Key: *BSC, best supportive care; PFN, pirfenidone; QALY, quality-adjusted life year*

Figure 44: Cost-effectiveness acceptability curve – ITT population, PAS price (generated by ERG)



Key: *BSC, best supportive care; PFN, pirfenidone; QALY, quality-adjusted life year*

ERG comments on PSA

The ERG has a number of concerns with the company's PSA. Notably, the majority of distributions appear to be arbitrary.

- The treatment effects from the NMA are varied using a gamma distribution based on the confidence interval assuming the HR to be normally distributed around the median. The ERG considers this to be inappropriate and that the CODA samples, using estimates from the predictive distributions should be used in the PSA. The ERG considers that the approach used by the company tend to underestimate the uncertainty in the treatment effects.
- PFS and OS are also modelled independently from each other and therefore no correlation is included. It is also possible in theory within the company's model, for the PFS curve to be greater than the OS curve as no constraint is added. However, the ERG notes that within the company's base-case assumptions, OS is consistently greater than PFS.
- The ERG considers the sampling of health utility values to be questionable and may underestimate uncertainty. The ERG observes that the mean SGRQ scores in people who are progression-free and with progression are sampled independently from each other assuming a normal distribution based on the mean score and standard errors estimated from the GEE model. This approach ignores the correlation between health states; the ERG considers that the variance-covariance matrix from the GEE should be used instead. The ERG further notes that the uncertainty in the mapping algorithm used to estimate the EQ-5D score is not accounted for within the company's model.
- The ERG further notes that the majority of distributions used to sample costs (resource use, management of AE, hospitalisation costs, and end of life) appear to be arbitrary. The company arbitrarily varied costs from a gamma distribution assuming an arbitrary variance of 20% around the mean despite having sufficient information to estimate the precision around some of these parameters.

- **Scenario analysis**

In addition to the DSA and PSA, the company reports cost-effectiveness results across nine groups of scenarios; these involved altering the model time horizon (10-year to 30-year), utility values, duration of the treatment effect, baseline hazard of death, progression and time to discontinuation, the studies included in the NMA for OS, PFS and exacerbations, implementation of the stopping rule and resource use. Results are available in the CS and clarification response. In brief, the ICERs were sensitive to the time horizon, assumption relating to the duration over which the treatment effect is assumed to remain constant, the parametric distributions for OS in people initiating pirfenidone, the treatment effects taken from the NMAs for OS only, and the inclusion of the stopping rule.

5.2.13. Model validation

The company reports two main methods of model validation:

- Comparison of the model predictions with results from previous evaluations,
- Validation of the long-term prediction of survival.

The CS provides a comparison of the model outcomes from its model with those from the company's submission, in the nintedanib appraisal,[26] and the company's submission in the previous appraisal of pirfenidone[2] (see Table 66).

Table 66: Comparison of LYs and QALYs – moderate population (reproduced from CS,[4] Table 122)

Outcome	NTB submission[26]			This submission			TA282	
	BSC	NTB	PFN	BSC	NTB	PFN	BSC	PFN
Total QALYs	3.27	3.67	3.62	2.15	3.77	4.46	3.18	4.30
Total LYs	4.36	4.86	4.86	4.33	5.30	6.47	4.40	5.96

Key: IPF, idiopathic pulmonary fibrosis; LY, life year; NTB, nintedanib; PFN, pirfenidone; QALY, quality-adjusted life year

The CS also provides a comparison of OS from their model compared with two studies (see Table 67) which uses observational data (both sources are described further in Table 59 of the CS). Fisher al (2015)[64] reports OS from a modelling study whereby the OS in patients initiating BSC is modelled from a log-normal distribution which is fitted to data from the National Jewish Health Interstitial Lung Disease database and not the US strand registry as suggested by the company. The OS in patients initiating pirfenidone is modelled from a log-normal distribution which is fitted to data from the RECAP trial. The Roskell *et al.* study[66] is also a modelling study and uses data from the RECAP OLE for pirfenidone (Weibull distribution fitted to the KM). The survival in patients initiating BSC was taken the CPRD and included patients with a FVC > 50% only. A Weibull distribution was fitted to the CPRD data.

Table 67: Comparison of OS and PFS – ITT population (reproduced from CS,[4] Table 123)

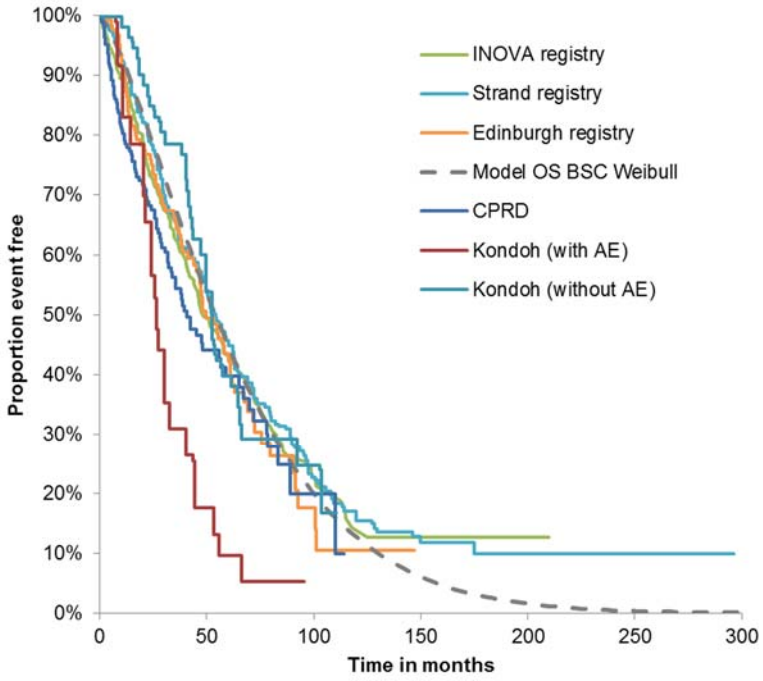
Outcome	This submission		Fisher *et al.*[64]		Roskell *et al.*[66]	
	BSC	PFN	BSC	PFN	BSC	PFN
Mean OS	5.38	8.67	6.10	9.29	5.25	9.26
Mean PFS	1.50	2.05	1.28	3.28	NR	NR

Key: IPF, idiopathic pulmonary fibrosis; LY, life year; NR, not reported; NTB, nintedanib; PFN, pirfenidone; QALY, quality-adjusted life year.

Overall, the company considers that their model generates estimates for OS and PFS (Table 67) and QALYs (Table 66) for both people initiating pirfenidone and BSC which are comparable to previous economic evaluations.

In addition, the company reports a comparison of the modelled OS in people initiating BSC predicted by the model against long-term registry data (see Figure 45). The company considers that the predicted OS for BSC in the model is consistent with long-term registries.

Figure 45: Long-term overall survival for BSC IPF people – ITT population



Number at risk	0	50	100	150	200	250	300
Strand	321	174	72	41	31	31	0
Edinburgh	125	44	6	0	0	0	0
CPRD	193	31	2	0	0	0	0
Kondoh with AE	23	4	0	0	0	0	0
Kondoh without AE	51	27	12	0	0	0	0
INOVA	234	82	22	8	8	0	0

Key: AE, acute exacerbation; BSC, best supportive care; OS, overall survival.

5.2.13.1. ERG's comments regarding the model validation of the company's model

The ERG observes that whilst the CS presents information regarding the external validity of the model, the CS does not describe any other forms of quality assurance such as:

- Validation of the model structure and key structural assumptions using clinical experts to ensure face validity;
- Peer review of the model by an independent health economist, or;
- Verification of the calculations within the model by an independent modeller.

As described in Section 5.2, the ERG has a number of concerns with the company's model regarding the conceptual representation of the condition, the representation of the treatment pathway in IPF, the implementation of the stopping rule and questionable structural assumptions including the assumption of a constant treatment effect over time. Based on these concerns, the ERG considers the company's model to lack face validity.

As part of its critical appraisal, the ERG checked the calculation to identify any programming errors and/or inconsistencies in the economic model. Inputs were also varied to establish if changes in inputs resulted in expected changes to the model outputs. Checks were also performed to ensure that the parameters presented in the CS and the company response to clarification correspond to those used in the economic model. No major programming errors were identified in the company's model during this process. The ERG identified however some minor programming errors and discrepancies, some of which were rectified in the revised economic model submitted by the company following responses to clarifications. These included:

- Lack of discounting for the cost of end of life (rectified in the revised model),
- Inclusion of the cost oxygen (rectified in the revised model following ERG comment),
- Double-counting of resource use in the first cycle,
- Discrepancies between results from the NMA presented in the clarifications responses for AEs and treatment effects used in the economic model,
- Miscalculation of the incidence of AEs –NMA outputs applied incorrectly in the model,
- Minor programming error for the PSA for cost for health professional visits for the progression-free health state,

The ERG further observes that in the revised model, an additional error was introduced by the company which was not present in the original model submitted to NICE. In brief, in the original

submission to NICE the cost associated with progression was correctly applied to patients in the progression health state irrespective of treatments. In the revised model, the company applied the costs for the progression-free health state to patient in the progressed health state, but only for the pirfenidone treatment arm. This change between the original model and the model submitted following clarification was not mentioned by the company.

The ERG also notes that whilst the PAS for pirfenidone was implemented correctly in the revised company model, the ERG had to correct the implementation of the PAS for nintedanib, in addition to setting the discount to its true confidential value, before generating the results presented in the confidential appendix. This was because the discount for nintedanib was being applied in addition to the discount for pirfenidone when calculating the nintedanib drug costs.

The company compares the model prediction for OS with estimates from two survival modelling studies reported in a power point presentation by Fisher *et al.* (2015) and a poster presented by Roskell *et al.* (2014). The ERG considers the comparison to be of limited relevance given that the same source of data is used for pirfenidone (RECAP OLE) and there are potential biases associated with the use of registry data as described in section 5.2.

The company also justifies its structure based on a previous economic evaluation conducted by Loveman et al[94] but does not provide a comparison of the results with this study. The ERG notes that Loveman *et al.* estimated the number of total QALYs to be 2.98 in people initiating BSC and 3.34 in people initiating pirfenidone compared with 3.15 and 4.46 in the company's model. These differences are not discussed by the company.

Finally, a key argument from the company regarding the model's validity relies on a comparison between the modelled OS for BSC and the OS observed in three long-term registries. As noted in Section 5.2, the ERG has several concerns with the survival observed in these registries compared with people initiating BSC that were enrolled in the ASCEND/CAPACITY trials in that the survival from the registry did not validate the survival from the BSC arm of the trial.

In conclusion, the ERG considers the validation undertaken by the company to be misleading and considers that the company's base-case may overestimate the benefit of pirfenidone compared with BSC. The ERG further considers the lack of reporting on the assessment of face validity for the model using clinical experts to be a matter of concern.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

This section summarises additional analyses undertaken by the ERG using the company's model as well as the development of an ERG-preferred base-case.

The ERG expressed a number of concerns regarding the model structure and parameterisation of the company's model. A key concern related to the lack of ability of the model to capture the progressive nature of IPF and inflexibility associated with the modelling approach chosen by the company (partitioned survival model) which meant that correlations between outcomes are not captured in the model. This is a concern as the modelled stopping rule impacts on costs but not health outcomes. Importantly, the company's model also relies on a strong assumption that the treatment effect estimated within the trials (up to 52 weeks) is maintained over the entire model's duration (34 years). Such extrapolation is questionable and subject to considerable uncertainty. This leads to discrepancies between the model-predicted OS and observed OS in people initiating BSC from the ASCEND/CAPACITY trial (see Figure 35).

Unfortunately, a number of the issues identified cannot be addressed by the ERG without major restructuring of the economic model. It should also be noted that changes to the model are challenging given the structure of the model whereby outcomes are disconnected from each other. The ERG is not able to adequately amend the implementation of the stopping rule within the company's existing model structure and thus, considers that any ICER generated in the scenarios using the stopping rule need to be interpreted with caution as they are likely to provide ICERs that are favourable to pirfenidone when compared against BSC.

The following analyses were undertaken by the ERG to inform its base-case:

1. **Using the ERG's preferred estimate of the treatment effect, which uses data up to 72 weeks, excludes SP3, and uses the CODA samples from the predictive distribution.** As described in Section 3.2.4.1, the ERG considered the treatment effect estimated at 72 weeks to be more appropriate and more consistent with the company's assumption of proportional hazards. Furthermore, the ERG considered that SP3 should be removed from the network as this trial was conducted in a Japanese population, an unlicensed dose was given and the HR was not directly available which could introduce a bias. Finally, the ERG considered that the CODA samples (from the predictive distribution) should be used instead of the median HR in order to properly capture the joint uncertainty in the effectiveness estimates, and therefore the results for this scenario are run probabilistically.

2. **Use of the Gompertz distribution for OS (rather than the Weibull).** As described in Section 5.2.3.1, the ERG considered the Gompertz distribution to provide a more plausible long-term extrapolation compared with the Weibull distribution.
3. **Stopping the treatment effect after 2 years (approximately the end of follow-up of the clinical evidence for pirfenidone vs. BSC) compared with extrapolating the treatment effect to the entire model duration.** As described in Section 5.2.2.5, there is considerable uncertainty regarding the duration over which the treatment effects observed in the trials could be reasonably expected to persist. Consequently, the ERG present results using an optimistic scenario (treatment effect assumed to be constant over the entire lifetime – as assumed in the company's base-case) and a pessimistic scenario (treatment effect stop approximately after the end of follow-up of the clinical evidence at 2 years).
4. **Capping of utility estimates for individuals at 1.0 in the IPD used to derive average utilities for the progression-free and progressed state.** As described in Section 5.2.8.1, the company's base-case utility values were estimated from the aggregate mean SGRQ mapped onto EQ-5D using a mapping function. The ERG noted that when applied at the individual-level, the mapping function predicted values over 1.0. The ERG considers that the utility values estimated at the individual-level and capped at 1.0 is more appropriate leading to a mean utility value of 0.82 for progression-free and 0.76 for progressed disease (compared with 0.85 and 0.78 in the company's base-case).
5. **Adjustment of utility by age.** As described in Section 5.2.8.1, the company's base-case assumes utility values to be constant with respect to age or time. This has the effect of over-estimating the total number of QALYs. Whilst it is not possible within the company's model to explore the impact of progression in quality of life with respect to time, the ERG considers that utility values should at least be adjusted by age to avoid over-estimating QALYs. Health utilities were adjusted by age by the ERG based on the ratio of the change in utility values observed in the general population taken from an analysis conducted by Ara *et al.* (2010)[114] using data from the Health Survey for England (HSE).
6. **Including costs associated with end of life for all people irrespective of the cause of death.** The ERG considers that the company's assumption that only IPF-related mortality is associated with end of life costs is inadequately justified by the evidence. In the absence of evidence on the differential costs according to the cause of death, the one year cost assumed by the company for end of life care (£9,996) is applied to all deaths, irrespective of cause. The ERG notes that the impact will be slightly different between treatment arms due to discounting.
7. **Including titration in the first cycle based on data provided by the company at the clarification stage.** A different dose intensity is used between the first (3 months) and subsequent cycles. In response to clarification, the company provided the ICER for this

analysis. It should be noted that the ERG was not able to replicate the ICER provided by the company and therefore the ICER for this analysis presented by the ERG are inconsistent with those reported by the company in Table 19 of the response to clarification (see addendum to clarification response).

8. **Using compliance from IMPULSIS for nintedanib.** Given the different price structure, the ERG considered that assuming the same impact of dose reductions/interruptions between pirfenidone and nintedanib is likely to be unfavourable to pirfenidone. Consequently, an analysis is conducted assuming a compliance of 96.4% for nintedanib based on data from the IMPULSIS trial.[113]
9. **Corrections of errors in the economic model.** As part of the critical appraisal of the model, the ERG identified a series of minor programming errors which have been corrected. These are described in appendix 4.

The impact of each individual change is reported in section 6 in addition to the ERG-preferred base-case which combines all these changes. For consistency, results are reported with and without the stopping rule (same assumption for both treatments). It should also be noted that the ERG-preferred base-case is presented as a range (most optimistic to most pessimistic scenario) given the uncertainty surrounding the extrapolation of the treatment effect.

5.4 Conclusions of the cost-effectiveness section

The company submitted a fully executable economic model as part of their submission to NICE. The analysis was undertaken from the perspective of the UK NHS and PSS over a lifetime horizon. The company's analysis is presented for three populations: (1) the ITT trial population, which is comprised of adults with mild to moderate IPF; (2) people with a percent predicted FVC > 80% at baseline (considered to be mild IPF), and; (3) people with a percent predicted FVC of 50 - 80% at baseline (considered to be moderate IPF). All three analyses include BSC as a comparator (defined in the trial as symptom relief, pulmonary rehabilitation, management of comorbidities, and end of life care, including oxygen therapy). Nintedanib is included as a comparator only in the analysis of people with a percent predicted FVC of 50 - 80% at baseline.

The analysis in the ITT population does not include nintedanib as a comparator as nintedanib is only a valid comparator for the subgroup of the ITT population with moderate IPF (percent predicted FVC of 50 - 80%). The ERG considers that it is more appropriate to conduct an economic analysis separately within the mild and moderate subgroups as the comparators vary by subgroup.

In the company's base-case, people initiating pirfenidone are assumed to discontinue treatment at the rate observed in the trials, hence no stopping rule is applied. The stopping rule defined by NICE

which formed the basis for the positive recommendation for pirfenidone[2] and nintedanib[12] is however applied to nintedanib in the company's base-case. A scenario analysis is presented in which the stopping rule is applied for both pirfenidone and nintedanib.

Within the ITT trial population (adults with mild to moderate IPF), the company's model estimates the ICER for pirfenidone against BSC to be [redacted] per QALY gained (probabilistic ICER: [redacted]) using the list price for pirfenidone and £21,387 per QALY gained (probabilistic ICER: £20,928) when incorporating the PAS. Within the subgroup of people with a percent predicted > 80% at baseline (considered to be mild IPF), the company's model estimates the ICER for pirfenidone against BSC to be [redacted] per QALY gained (probabilistic ICER: [redacted]) using the list price for pirfenidone and £24,187 per QALY (probabilistic ICER: £23,476) when incorporating the PAS. Within the subgroup of people with a percent predicted FVC of 50 - 80% at baseline (considered to be moderate IPF), nintedanib is ruled out due to extended dominance at the list price; the resulting ICER for pirfenidone versus BSC is estimated to be to be [redacted] per QALY gained (probabilistic ICER: [redacted]) using the list price and £21,318 (probabilistic ICER: £20,863) when incorporating the PAS. The results for pirfenidone versus nintedanib when incorporating the PAS for the moderate subgroup cannot be reported here and can be found in the confidential appendix.

The company presented a series of scenario analyses. The ICERs were mostly sensitive to the time horizon, assumptions regarding the duration over which the treatment effect is assumed to remain constant, the parametric distributions for OS in people initiating pirfenidone, the treatment effects taken from the NMAs for OS only, and the inclusion of the stopping rule.

The ERG critically appraised the company's health economic analysis and the model upon which this analysis was based. The ERG has a number of concerns regarding the model structure and parameterisation of the company's model. These include: (a) the inability of the model to capture the progressive nature of IPF; (b) the absence of stopping rule in the company's base-case; (c) the inadequacy of the partitioned survival approach when implementing a stopping rule; (d) the assumption that treatment effect is constant over the entire model duration, and; (e) estimation of the treatment effect. The ERG further observes that under the company's base-case assumption, there are discrepancies between the model's predictions of OS and the observed trial data in people initiating BSC. The company does not comment on these discrepancies and focus instead on a comparison with registry data which does not match the trial data in people initiating BSC.

Whilst a number of issues identified could not be addressed by the ERG without major restructuring of the economic model (particularly amending the implementation of the stopping rule), a number of analyses were undertaken by the ERG which informed the ERG preferred base-case. The main

changes include: (a) using the ERG's preferred estimate of the treatment effect, which uses data up to 72 weeks, excludes SP3, and uses the CODA samples from the predictive distribution; (b) use of the Gompertz distribution for OS (rather than the Weibull); (c) exploring different durations over which the treatment effect is assumed to be maintained (ranging between 2 years to the entire time horizon); (d) capping utility estimates for individuals at 1.0; (e) adjusting utility by age; (g) including costs associated with end of life for all people irrespective of cause of death; (h) amending dose reductions/interruptions for pirfenidone and nintedanib, and; (i) correcting of errors in the economic model. The results of these exploratory analyses are summarised in Section 6.

6. IMPACT ON THE ICER OF EXPLORATORY ANALYSES UNDERTAKEN BY THE ERG

Section 6.1.1 summarises the impact of each individual change which forms part of the ERG preferred base-case. Section 6.1.2 presents the ERG preferred base-case.

6.1.1. Impact of each individual change which forms the ERG-preferred base-case assumptions

Table 68 and Table 69 presents the impact on the ICER of each individual change for the ITT population and subgroups of people with a percent predicted FVC of 50 – 80% and > 80% using the list price. Table 68 presents the results for the deterministic model, whilst Table 69 presents the results for the probabilistic model for the scenario where the efficacy estimates were updated to use the CODA samples. Results using the PAS are presented in Table 69 and Table 71 for the deterministic and probabilistic analysis. The ICERs for nintedanib versus BSC are not included in Table 68 as nintedanib is always extendedly dominated by pirfenidone and BSC at the list price. It can be seen from Table 68 and Table 69 that the ICERs are sensitive to four key assumptions: (i) the duration of extrapolation of the treatment effect; (ii) inclusion of the stopping rule; (iii) the treatment effect assumed and; (iv) the use of the Gompertz rather than the Weibull distribution for OS.

As expected, the ICERs are the most sensitive to the assumption around the extrapolation of the treatment effect. Assuming that the treatment effect does not persist beyond 2 years (compared with the company's base-case whereby the treatment effect is extrapolated over the entire model's duration) has the effect of increasing the ICER for pirfenidone against BSC from approximately [redacted] per QALY gained for the ITT population when using the list price. The ICER changed from approximately £21,000 to £73,000 when incorporating the PAS. This is because people initiating pirfenidone experience a shorter duration of benefits (2 years).

In contrast, the implementation of the stopping rule for pirfenidone has the effect of reducing the ICER for pirfenidone versus BSC from approximately [redacted] per QALY gained for the ITT population. The ICER changed from approximately £21,000 to £15,000 when incorporating the PAS. However, as described in Section 5.2.2.2, this is an artefact of the model structure whereby treatment discontinuation limits the cots but is disconnected from health outcomes. The ERG reiterates that the analyses using the stopping rule lacks face validity and provides a lower bound of the plausible ICER (i.e. most optimistic scenario).

The ERG's preferred estimate of the treatment effect, which uses data up to 72 weeks, excludes SP3, and uses the CODA samples from the predictive distribution, has a moderate to large effect on the

ICER for pirfenidone versus BSC with an increase from approximately [redacted] per QALY gained for the ITT population (see Table 70). The ICER changed from approximately £21,000 to £29,000 when incorporating the PAS. This is because a lower treatment effect (higher HRs) leads to fewer health gains whilst on pirfenidone.

Finally, the use of the Gompertz distribution to represent the baseline hazard of death in people initiating pirfenidone leads to a moderate increase in the ICER for pirfenidone against BSC from approximately [redacted] per QALY gained for the ITT population. The ICER changed from £21,000 to £25,000 when incorporating the PAS. This is attributable to the fact that the Gompertz distribution has a shorter tail compared with the Weibull distribution, reducing the period over which treatment benefits can be accrued.

The impact of each change on the mild and moderate populations was similar to that for the ITT population, with the exception of the assumption that treatment effect stops after 2 years, where the impact was greater for the mild population (see Table 69).

Full incremental results for the four changes which had the biggest effect on the ICERs for pirfenidone versus BSC are presented in Appendix 3 for the analysis incorporating the PAS.

Table 68: Summary of the impact of individual changes to the ICER for pirfenidone versus BSC[a] using the list price and mean parameter inputs (deterministic model)

	ITT population	People with a percent predicted FVC of 50 – 80%[a]	People with a percent predicted FVC > 80%
Company base-case	[redacted]	[redacted]	[redacted]
No stopping rule for nintedanib	[redacted]	[redacted]	[redacted]
Inclusion of stopping rule for pirfenidone	[redacted]	[redacted]	[redacted]
Treatment effect assumed to stop after 2 years	[redacted]	[redacted]	[redacted]
Gompertz distribution for OS	[redacted]	[redacted]	[redacted]
HRQoL capped at 1.0	[redacted]	[redacted]	[redacted]
Adjustment of HRQoL by age	[redacted]	[redacted]	[redacted]
End of life costs applied to death irrespective of causes	[redacted]	[redacted]	[redacted]
Pirfenidone dose titration	[redacted]	[redacted]	[redacted]
Nintedanib compliance taken from IMPULSIS	[redacted]	[redacted]	[redacted]
Correction of errors	[redacted]	[redacted]	[redacted]

[a] nintedanib is extendedly dominated

Table 69: **Summary of the impact of individual changes to the ICER for pirfenidone versus BSC[a], using the PAS price and mean parameter inputs (deterministic model)**

	ITT population	**People with a percent predicted FVC of 50 - 80%**	**People with a percent predicted FVC > 80%**
Company base-case	£21,387	£21,331[b]	£24,187
Inclusion of stopping rule for pirfenidone	£14,847	£15,197	£15,707
Treatment effect assumed to stop after 2 years	£72,599	£66,503	£112,214
Gompertz distribution for OS	£25,360	£24,855	£31,379
HRQoL capped at 1.0	£22,041	£21,983	£24,928
Adjustment of HRQoL by age	£22,716	£22,487	£26,129
End of life costs applied to death irrespective of causes	£21,957	£22,000	£24,606
Pirfenidone dose titration	£21,120	£21,060	£23,893
Correction of errors	£22,574	£22,501	£25,519

[a] results for pirfenidone versus nintedanib are presented in the confidential appendix

[b] generated by ERG after correcting error in calculation of days within drug costs

Table 70 Summary of the impact of individual changes to the ICER for pirfenidone versus BSC, using the list price for the probabilistic model [b]

	ITT population	**People with a percent predicted FVC of 50 - 80%**			**People with a percent predicted FVC > 80%**
	Pirfenidone vs. BSC	**Nintedanib vs. BSC**	**Pirfenidone vs. nintedanib**	**Pirfenidone vs. BSC**	**Pirfenidone vs. BSC**
Company base-case	[redacted]	**Nintedanib extendedly dominated by pirfenidone**		[redacted]	[redacted]
Treatment effect at 72 weeks (CODA sample)[b]	[redacted]	£40,436	[redacted]	[redacted]	[redacted]

[b] Run probabilistically in order to incorporate the CODA sample

Table 71: Summary of the impact of individual changes to the ICER for pirfenidone versus BSC, using the PAS price for the probabilistic model [b]

	ITT population	**People with a percent predicted FVC of 50- 80%**			**People with a percent predicted FVC > 80%**
	Pirfenidone vs. BSC	**Nintedanib vs. BSC**	**Pirfenidone vs. nintedanib**	**Pirfenidone vs. BSC**	**Pirfenidone vs. BSC**
Company base-case	£20,928	**See confidential appendix**		£20,863	£23,476
Treatment effect at 72 weeks (CODA sample)[b]	£28,922			£28,766	£33,060

[b] Run probabilistically in order to incorporate the CODA sample

6.1.2. ERG-preferred base-case ICERs

The ERG's preferred base-case, which combines individual changes detailed in Section 6.1, is presented in Table 72 assuming no stopping rule for either treatment and in Table 74 assuming the stopping rule for both treatments, for the list price. Equivalent results when incorporating PAS are reported in Table 73 and Table 75 (with the results for pirfenidone versus nintedanib in the moderate subgroup reported in the confidential appendix). The ERG's preferred base-case is presented using an optimistic and pessimistic assumption regarding the duration of the treatment effect (lifetime to 2 year). This has been done because whilst the clinical advisors to the ERG considered it possible that there may be continued effectiveness with long-term treatment, the duration of persistence for any long-term treatment effect is currently highly uncertain, particularly given that this is a heterogeneous condition and the mechanism of treatment is not fully understood at this time. Results are run probabilistically (5,000 iterations) to incorporate the CODA sample.

Based on the ERG's preferred base-case assumptions, no stopping rule and the pirfenidone list price, within the ITT population (adults with mild to moderate IPF), the ICER for pirfenidone versus BSC is expected to be in the range [redacted] per QALY gained. The inclusion of the stopping rule results in ICERs for pirfenidone versus BSC of [redacted] per QALY gained. When incorporating the PAS the ICERs range from £39,895 to £115,751 per QALY gained without the stopping rule and £27,124 to £75,121 per QALY gained with the stopping rule.

Based on the ERG's preferred base case assumptions, no stopping rule and the pirfenidone list price, within people with a percent predicted FVC > 80% (considered to be mild IPF), the ICER for pirfenidone versus BSC is expected to be in the range of [redacted] to [redacted] per QALY gained. When the stopping rule is assumed, the ICER for pirfenidone versus BSC is expected to be in the range of [redacted] to [redacted] per QALY gained. When incorporating the PAS the ICERs range from £49,921 to £186,260 per QALY gained without the stopping rule and £31,722 to £113,365 per QALY gained with the stopping rule.

Based on the list price, within people with a percent predicted FVC of 50 - 80% (considered to be moderate IPF), pirfenidone consistently produced greater QALYs compared with nintedanib at a lower cost, and therefore nintedanib was dominated by pirfenidone, irrespective of whether the stopping rule is included. Excluding the stopping rule, the expected ICER for pirfenidone versus BSC is expected to be in the range of [redacted] per QALY gained. When the stopping rule is assumed, the ICER for pirfenidone versus BSC is expected to be in the range of

[redacted] per QALY gained. When incorporating the PAS the ICERs range from £39,166 to £104,915 per QALY gained without the stopping rule and £27,432 to £70,234 per QALY gained with the stopping rule. The results for pirfenidone versus nintedanib when incorporating the PAS are reported in the confidential appendix.

Full incremental results for the scenairos presented in Tables 73 and 75 are provided in Appendix 3.

Table 72: ERG-preferred base-case assuming no stopping rule (ICER for pirfenidone versus BSC), analyses conducted, using the list price

	ITT population	People with a percent predicted FVC of 50 - 80%	People with a percent predicted FVC > 80%
Optimistic ERG base-case (life-time treatment effect) – probabilistic	[redacted]	[redacted][a]	[redacted]
Pessimistic ERG base-case (2 years of treatment effect) - probabilistic	[redacted]	[redacted][a]	[redacted]

[a] Nintedanib dominated by pirfenidone in ERG preferred base-case

Table 73: ERG-preferred base-case assuming no stopping rule (ICER for pirfenidone versus BSC)[a], analyses conducted, using the PAS price

	ITT population	People with a percent predicted FVC of 50 – 80%	People with a percent predicted FVC > 80%
Optimistic ERG base-case (life-time treatment effect) – probabilistic	£39,895	£39,166	£49,921
Pessimistic ERG base-case (2 years of treatment effect) – probabilistic	£115,751	£104,915	£186,260

[a] results for pirfenidone versus nintedanib are presented in the confidential appendix

Table 74: ERG-preferred base-case assuming the stopping rule to apply (ICER for pirfenidone versus BSC), analyses conducted using the list price

	ITT population	People with a percent predicted FVC of 50 - 80%	People with a percent predicted FVC > 80%
Optimistic ERG base-case (life-time treatment effect)	[redacted]	[redacted][a]	[redacted]
Pessimistic ERG base-case (2 years of treatment effect)	[redacted]	[redacted][a]	[redacted]

[a] Nintedanib dominated by pirfenidone in ERG preferred base-case

Table 75: ERG-preferred base-case assuming the stopping rule to apply (ICER for pirfenidone versus BSC)[a], analyses conducted using the PAS price

	ITT population	People with a percent predicted FVC of 50 - 80%	People with a percent predicted FVC > 80%
Optimistic ERG base-case (life-time treatment effect)	£27,124	£27,432	£31,722
Pessimistic ERG base-case (2 years of treatment effect)	£75,121	£70,234	£113,365

[a] results for pirfenidone versus nintedanib when incorporating the PAS are in the confidential appendix

7 END OF LIFE

The CS states that life-expectancy in people with IPF is 3 years from the time of diagnosis (CS, page 43). The ERG therefore does not consider that pirfenidone for the treatment of IPF meets the criteria laid out in the NICE methods guide for a 'life-extending treatment at the end of life', which is that the treatment is indicated for patients with a short life expectancy, normally less than 24 months.[43]

8 OVERALL CONCLUSIONS

The ERG had some concerns regarding the generalisability of the trial population to patients with IPF and comorbid obstructive airway disease. These patients were excluded from the three main RCTs comparing pirfenidone with placebo (ASCEND, CAPACIY 1 and CAPACITY 2), but according to clinical advisors to the ERG, these patients would be considered for treatment in current practice provided they have a percent predicted FVC of between 50% and 80%.

The meta-analysis of trial data for the outcome of PFS is considered to be subject to some uncertainty due the combination of data from trials which used different definitions of PFS and should be interpreted with caution.

The three main pirfenidone RCTs (ASCEND, CAPACITY 1 and CAPACITY 2) were considered by the ERG to be at low to moderate risk of bias, on account of inconsistencies between some protocol-specified outcomes and analyses and those reported in the CS, and the possible influence of uncontrolled variables such as rate of disease progression.

The ERG considers the data from SP2 and SP3 to be less relevant to the decision problem due to the use of a non-licensed dose of pirfenidone, and differences in the population, which was exclusively Japanese and was therefore considered to be less relevant to the population likely to be treated in England. These two studies were also assessed to be at higher risk of bias than the three main pirfenidone RCTs.

The ERG concludes that whilst the available evidence suggests that there is a statistically significant reduction in all-cause mortality for pirfenidone compared with placebo, there remains uncertainty regarding whether the size of the treatment benefit for overall survival is constant over time due to variation in the treatment effect estimated using data from 52 weeks and 72 weeks.

The ERG concludes that there is some evidence to support a statistically significant reduction in the decline in percent predicted FVC compared with placebo, but notes that a statistically significant treatment effect was not demonstrated in one of the RCTs (CAPACITY 1), which weakens the strength of the evidence for this outcome.

The ERG concludes that pirfenidone does not appear to have a significant effect in individual trials on other outcomes that are important to patients, such as disease specific health-related quality of life measures (SGRQ, UCSD SOBQ). The evidence for a statistically significant treatment effect on 6MWD was not consistent in the CAPACITY trials, and therefore the effect of pirfenidone on

physical function, which is understood by the ERG to be an important driver of HRQoL, remains uncertain.

The ERG concludes that the AEs from the trials are consistent with those listed in the SmPC and that pirfenidone is generally well tolerated with most AEs experienced being mild to moderate.

A *post hoc* pooled analysis of ASCEND and CAPACITY 1 & 2 found no evidence for differential treatment effects according to disease severity, as assessed using three key efficacy outcomes; absolute ≥10% FVC decline, ≥50m 6MWD decline, and ≥20-point worsening of dyspnoea as measured by UCSD SOBQ. For these analyses disease severity was categorised according to baseline percent predicted FVC of 50 - 80% (moderate IPF) and >80% (mild IPF). In response to a clarification request from the ERG, the company also provided subgroup analyses according to disease severity for OS and PFS from the ASCEND and CAPACITY trials, although exact numbers within each subgroup in each trial arm were not reported. The findings did not suggest differential treatment effects according to disease severity for either outcome (as judged by the reported HR and 95% CI), however a treatment-by-subgroup interaction test was not reported so it is unclear if the difference between these subgroups was statistically significant. This subgroup analysis is particularly relevant to the decision problem as these groups had different comparator treatments and therefore separate analyses have been presented in the economic section for these subgroups. The ERG concludes that the evidence presented in the CS is not sufficient to support the use of subgroup specific treatment effects for these two groups.

Based on the NMA, the treatment effects for pirfenidone were broadly similar to those for nintedanib for all outcomes, with the pairwise treatment effects indicating that neither treatment is statistically significantly more effective.

Based on the company model when using the list price, the ICER for pirfenidone against BSC is [redacted] per QALY gained within the ITT-trial population in adults with mild to moderate IPF. The ICER for people with a percent predicted FVC >80% at baseline was [redacted] per QALY against BSC when using the list price. The ICER for people with a percent predicted FVC of 50 - 80% at baseline is [redacted] per QALY gained against BSC when using the list price. Nintedanib was extendedly dominated when using the list price in patients with a percent predicted FVC of 50 - 80% at baseline.

Based on the company model when incorporating the PAS for pirfenidone, the ICER for pirfenidone versus BSC was £21,387 per QALY in the ITT population and £24,187 per QALY in the mild subgroup (percent predicted FVC >80% at baseline) and £21,318 per QALY in the moderate

subgroup (percent predicted FVC of 50 - 80% at baseline). The results for pirfenidone versus nintedanib when incorporating the nintedanib and pirfenidone PAS (moderate subgroup) are reported in the confidential appendix.

The analysis in the ITT population does not include nintedanib as a comparator as nintedanib is only a valid comparator for the subgroup of the ITT population with moderate IPF (percent predicted FVC of 50 - 80%). The ERG considers that it is more appropriate to conduct an economic analysis separately within the mild and moderate subgroups as the comparators vary by subgroup.

The ERG identified a number of concerns regarding the model structure and parameterisation of the company's model including (a) the inability of the model to capture the progressive nature of IPF, (b) the absence of stopping rule in the company's base-case; (c) the inadequacy of the partition-survival approach when implementing a stopping rule, (d) the assumption that treatment effect is constant over the entire duration of the model, and (e) estimation of the treatment effect. The ERG further observes that under the company's base-case assumption, there are discrepancies between the model's prediction for OS and observed trial data in people initiating BSC.

A number of analyses were undertaken by the ERG which informed the ERG preferred base-case. The ERG's exploratory analysis led to consistently higher ICERs for pirfenidone against BSC for all three populations (ITT, mild [percent predicted FVC>80%] and moderate [percent predicted FVC of 50 - 80%] subgroups), even under the company's optimistic base-case assumption that the treatment effect is assumed to hold for the entire duration of the model. Using the list price, the ICER for pirfenidone ranged from [redacted] per QALY against BSC in the ITT-trial population. In the mild subgroup, the ICER ranged from [redacted] per QALY against BSC when using the list price. In the moderate subgroup the ICERs ranged from [redacted] per QALY against BSC when using the list price. When incorporating the PAS the ICERs for pirfenidone versus BSC were above £27,000 per QALY in the ITT population, above £31,000 in the mild subgroup and above £27,000 in the moderate subgroup. Results for pirfenidone versus nintedanib in the moderate subgroup when incorporating the PAS are presented in the confidential appendix.

A key uncertainty is around the duration of the extrapolation of the treatment effect. In the company's base-case, the treatment effect is assumed to be constant over the model's entire duration. The ERG considered this to be overly optimistic and inadequately supported by the evidence and believes that the treatment effect could reduce over time; although there is a lack of data to support either assumption. Assuming a shorter duration of extrapolation for the treatment effect led to an increase in the ICERs.

An important limitation in the company's model implementation regards the implementation of the stopping rule. Despite the fact that the NICE recommendations for pirfenidone (TA379) and nintedanib (TA282) include identical stopping rules, the company's model structure does not accommodate the robust exploration of the impact of this stopping rule on the ICERs. Whilst a scenario analysis including the stopping rule for both treatments is presented in the CS, the ERG has a number of concerns with this analysis as stopping treatment earlier led to a reduction in treatment costs, but left the gain in life years and QALYs unchanged. The ERG considers that results from these analyses need to be interpreted with caution and that no robust ICERs have been presented by the company when the stopping rule is implemented. The ERG considers that the ICERs presented by the company using the stopping rule could represent a lower bound of the true ICER when the stopping rule is implemented in clinical practice, as the life-time costs of treatment are reduced when the stopping rule is applied in the model, but the incremental QALYs are not reduced by the shorter duration of treatment.

8.1 Implications for research

IPF is a heterogeneous condition and there is natural variability in the rates of decline in percent predicted FVC. It is therefore difficult for clinicians to know if treatment is benefiting an individual patient as a patient who experiences stability on the drug may have had a low rate of decline in FVC without treatment and a patient who experiences a moderate rate of decline in FVC on treatment may have experienced a more rapid decline without treatment. Further research into biomarkers which predict the rate of disease progression or which predict response to treatment would be beneficial.

9 REFERENCES

1. Roche Products Ltd. Esbriet 267 mg Hard Capsules, Summary of Product Characteristics: eMC; 2015.
2. National Institute for Health and Care Excellence. Pirfenidone for treating idiopathic pulmonary fibrosis (TA282). London, UK: The National Institute for Health and Care Excellence; 2013.
3. National Institute for Health and Care Excellence (NICE). Pirfenidone for the treatment of idiopathic pulmonary fibrosis (review of TA282) [ID837]) - Final scope. London: NICE; 2015.
4. Roche Product Limited. Pirfenidone for the treatment of adults with idiopathic pulmonary fibrosis: Company evidence submission. London, UK; 2016.
5. Kim DS, Collard HR, King TE, Jr. Classification and natural history of the idiopathic interstitial pneumonias. *Proc Am Thorac Soc* 2006;3:285-92. http://dx.doi.org/10.1513/pats.200601-005TK
6. Schmidt SL, Tayob N, Han MK, Zappala C, Kervitsky D, Murray S*, et al.* Predicting pulmonary fibrosis disease course from past trends in pulmonary function. *Chest* 2014;145:579-85. http://dx.doi.org/10.1378/chest.13-0844
7. Behr J, Bendstrup E, Crestani B, Gnther A, Olschewski H, Skld M*, et al.*: Safety and tolerability of N-acetylcysteine (NAC) with pirfenidone in idiopathic pulmonary fibrosis (IPF): PANORAMA. *European Respiratory Journal* 2015;46. http://dx.doi.org/10.1183/13993003.congress-2015.OA3477
8. Behr J, Bendstrup E, Crestani B, Gunther A, Olschewski H, Skold M*, et al.* Pirfenidone and N-acetylcysteine (NAC) for the Treatment of Idiopathic Pulmonary Fibrosis (IPF): Exploratory Efficacy Results From the PANORAMA Study. *Chest* 2015;148. http://dx.doi.org/10.1378/chest.2335639
9. Raghu G, Rochwerg B, Zhang Y, Garcia CA, Azuma A, Behr J*, et al.* An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. *Am J Respir Crit Care Med* 2015;192:e3-19. http://dx.doi.org/10.1164/rccm.201506-1063ST
10. Roche Products Ltd. Response to Clarification Letter.; 2016.
11. Nathan. Effect of Continued Treatment With Pirfenidone Following a Clinically Meaningful Decline in Percent Predicted Forced Vital Capacity in Patients With Idiopathic Pulmonary Fibrosis. ATS 2015 Annual Congress; Denver, USA, abstract no. 1.
12. National Institute for Health and Care Excellence. Nintedanib for treating idiopathic pulmonary fibrosis (TA379). London, UK; 2016.
13. Harari S, Caminati A. Idiopathic pulmonary fibrosis: from clinical trials to real-life experiences. *Eur Respir Rev* 2015;24:420-7. http://dx.doi.org/10.1183/16000617.0042-2015
14. Chaudhuri N, Duck A, Frank R, Holme J, Leonard C. Real world experiences: pirfenidone is well tolerated in patients with idiopathic pulmonary fibrosis. *Respir Med* 2014;108:224-6. http://dx.doi.org/10.1016/j.rmed.2013.11.005
15. Koschel D, Cottin V, Skold M, Tomassetti S, Azuma A, Giot C*, et al.* Pirfenidone post-authorization safety registry (PASSPORT)–Interim analysis of IPF treatment. *European Respiratory Journal* 2014;44:1904.
16. MIMS. Esbriet. Haymarket Media Group Ltd; 2016. URL: http://www.mims.co.uk/drugs/respiratory-system/cough-congestion-respiratory-distress/esbriet (Accessed 28/03/2016, 2016).
17. Boehringer Ingelheim Limited. Ofev 150 mg soft capsules, Summary of Product Characteristics; 2015.
18. Saketkoo LA, Mittoo S, Huscher D, Khanna D, Dellaripa PF, Distler O*, et al.* Connective tissue disease related interstitial lung diseases and idiopathic pulmonary fibrosis: provisional core sets of domains and instruments for use in clinical trials. *Thorax* 2013:thoraxjnl-2013-204202.
19. Nathan SD, Meyer KC. IPF clinical trial design and endpoints. *Current opinion in pulmonary medicine* 2014;20:463-71.

20. du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, *et al.* Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. *American journal of respiratory and critical care medicine* 2011;184:1382-9.
21. Swigris JJ, Kuschner WG, Jacobs SS, Wilson SR, Gould MK. Health-related quality of life in patients with idiopathic pulmonary fibrosis: a systematic review. *Thorax* 2005;60:588-94. http://dx.doi.org/10.1136/thx.2004.035220
22. Swigris JJ, Esser D, Conoscenti CS, Brown KK. The psychometric properties of the St George's Respiratory Questionnaire (SGRQ) in patients with idiopathic pulmonary fibrosis: a literature review. *Health Qual Life Out* 2014;12. http://dx.doi.org/10.1186/s12955-014-0124-1
23. du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, *et al.* Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011;184:459-66. http://dx.doi.org/10.1164/rccm.201011-1790OC
24. Nathan SD, Shlobin OA, Weir N, Ahmad S, Kaldjob JM, Battle E, *et al.* Long-term Course and Prognosis of Idiopathic Pulmonary Fibrosis in the New Millennium. *Chest* 2011;140:221-9. http://dx.doi.org/10.1378/chest.10-2572
25. du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, *et al.* Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference. *American journal of respiratory and critical care medicine* 2011;183:1231.
26. Boehringer Ingelheim Limited. Nintedanib for the treatment of idiopathic pulmonary fibrosis: Company evidence submission. The National Institute for Health and Care Excellence; 2015.
27. InterMune. Protocol PIPF-016: InterMune Inc; 2012.
28. InterMune. Clinical Study Report No. PIPF-006: Appendix 16.1.1. Protocol and Protocol Amendments: InterMune Inc; 2006.
29. InterMune. Clinical Study Report No. PIPF-004: Appendix 16.1.1. Protocol and Protocol Amendments: InterMune Inc; 2006.
30. Roche Products Ltd. Addendum to the response to Clarification Letter. 2016.
31. Swigris JJ, Han M, Vij R, Noth I, Eisenstein EL, Anstrom KJ, *et al.* The UCSD shortness of breath questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. *Respiratory medicine* 2012;106:1447-55.
32. Moher D, Liberati A, Tetzlaff J, Altman DG, Group. TP. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 2009;6:e1000097. http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000097
33. ASCEND Clinicaltrials.gov. Efficacy and Safety of Pirfenidone in Patients With Idiopathic Pulmonary Fibrosis (IPF). In; 2016.
34. King TE Jr, Bradford WZ, Castro-Bernardini S, Fagan EA, Glaspole I, Glassberg MK, *et al.* A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2083-92. http://dx.doi.org/doi: 10.1056/NEJMoa1402582.
35. CAPACITY 2 ClinicalTrials.gov. Three-Arm Study of the Safety and Efficacy of Pirfenidone in Patients With Idiopathic Pulmonary Fibrosis. In; 2016.
36. CAPACITY 1 ClinicalTrials.gov. Safety and Efficacy of Pirfenidone in Patients With Idiopathic Pulmonary Fibrosis. In; 2016.
37. Noble PW, Albera C, Bradford WZ, Costabel U, du Bois RM, Fagan EA, *et al.* Pirfenidone for idiopathic pulmonary fibrosis: analysis of pooled data from three multinational phase 3 trials. *Eur Respir J* 2016;47:243-53. http://dx.doi.org/doi: 10.1183/13993003.00026-2015
38. Taniguchi H, Ebina M, Kondoh Y, Ogura T, Azuma A, Suga M, *et al.* Pirfenidone in idiopathic pulmonary fibrosis. *Eur Respir J* 2010;35:821-9. http://dx.doi.org/doi: 10.1183/09031936.00005209.
39. Azuma A, Nukiwa T, Tsuboi E, Suga M, Abe S, Nakata K, *et al.* Double-blind, placebo-controlled trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2005;171:1040-7.

40. Costabel U, Albera C, Bradford W, Hormel P, *et al.* Analysis of lung function and survival in RECAP: An open-label extension study of pirfenidone in patients with idiopathic pulmonary fibrosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2014;31:198-205.
41. National Jewish Health. Participation Program for Pulmonary Fibrosis - P_3F Registry. In; 2016.
42. InterMune. Single technology appraisal: Pirfenidone for the treatment of idiopathic pulmonary fibrosis: Manufacturer's submission. Brisbane, CA, USA; 2011.
43. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013; 2013.
44. Centre for Reviews Dissemination. Systematic review: CRD's guidance for undertaking reviews in health care. CRD, University of York; 2009. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]; 2011.
45. Aravena C, Labarca G, Venegas C, Arenas A, Rada G. Pirfenidone for Idiopathic Pulmonary Fibrosis: A Systematic Review and Meta-Analysis. *Plos ONE* 2015;10:e0136160. http://dx.doi.org/10.1371/journal.pone.0136160
46. Atkins CP, Loke YK, Wilson AM. Outcomes in idiopathic pulmonary fibrosis: A meta-analysis from placebo controlled trials. *Respir Med* 2014;108:376-87.
47. Jiang C, Huang H, Liu J, Wang Y, Lu Z, Xu Z. Adverse Events of Pirfenidone for the Treatment of Pulmonary Fibrosis: A Meta-Analysis of Randomized Controlled Trial. *PLoS ONE* 2012;7:e47024.,
48. Huang H, Dai HP, Kang J, Chen BY, Sun TY, Xu ZJ. Double-Blind Randomized Trial of Pirfenidone in Chinese Idiopathic Pulmonary Fibrosis Patients. *Medicine (Baltimore)* 2015;94:e1600. http://dx.doi.org/doi: 10.1097/MD.0000000000001600.
49. Noble PW, Albera C, Bradford WZ, Costabel U, Glassberg MK, Kardatzke D, *et al.* Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. *Lancet* 2011;377:1760-9. http://dx.doi.org/doi:10.1016/S0140-6736(11)60405-4.
50. Kondoh Y, Taniguchi H, Katsuta T, Kataoka K, Kimura T, Nishiyama O, *et al.* Risk factors of acute exacerbation of idiopathic pulmonary fibrosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2010;27:103-10.
51. European Medicines Agency. Guideline on Missing Data in Confirmatory Clinical Trials. 2011. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/09/WC500096793.pdf (Accessed 9 March, 2016).
52. Shinogi. SP3 study protocol. Date of preparation: 16 October 2006.
53. Shinogi. SP2 study protocol. Date of preparation: 9 October 2001.
54. American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). *Am J Respir Crit Care Med* 2000;161:646-64.
55. King TE Jr, Bradford WZ, Castro-Bernardini S, Fagan EA, Glaspole I, Glassberg MK, *et al.* A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis [Supplementary Appendix]. *N Engl J Med* 2014;370:2083-92.
56. Roche Products Ltd. Cost-effectiveness / cost-utility AUC model of Esbriet (pirfendione) compared to BSC and Ofev (nintedanib) in mild to moderate IPF. 2016.
57. FDA. U.S. Food and Drug Administration. Pirfenidone Capsules [NDA 22-535]. Sponsor's Briefing Document for the FDA Pulmonary-Allergy Drugs Advisory Committee Meeting Scheduled for 09 March 2010. 2010. URL: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/PulmonaryAllergyDrugsAdvisoryCommittee/UCM203083.pdf (Accessed 9 March, 2016).
58. Woods BS, Hawkins N, Scott DA. Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. *BMC Med Res Methodol* 2010;10.
59. Sahn SA, Albera C, Du Bois RM, Bradford W, Costabel U, King TE. The effect of treatment with pirfenidone on progression-free survival in patients with idiopathic pulmonary fibrosis (IPF): Exploratory analysis of outcomes using novel criteria for disease progression.

American Journal of Respiratory and Critical Care Medicine 2011;183:(Meeting Abstracts 1):[A3810].
60. Kondoh Y, Taniguchi H, Ebina M, Azuma A, Ogura T, Taguchi Y*, et al.* Risk factors for acute exacerbation of idiopathic pulmonary fibrosis-Extended analysis of pirfenidone trial in Japan. *Respir Investig* 2015;53:271-8. http://dx.doi.org/doi: 10.1016/j.resinv.2015.04.005.
61. Nathan SD. Effect of Continued Treatment With Pirfenidone Following a Clinically Meaningful Decline in Percent Predicted Forced Vital Capacity in Patients With Idiopathic Pulmonary Fibrosis. ATS 2015 Annual Congress; 15-20 May; Denver, USA, abstract no. 22.
62. Nathan SD, Albera C, Bradford WZ. Effect of continued treatment with pirfenidone following clinically meaningful declines in forced vital capacity: Analysis of data from three phase 3 trials in patients with idiopathic pulmonary fibrosis. In; 2016 (Unpublished, AiC).
63. Kreuter M. Pirfenidone: an update on clinical trial data and insights from everyday practice. *Eur Respir Rev* 2014;23:111-7. http://dx.doi.org/doi:10.1183/09059180.00008513.
64. Fisher M. Disease-progression modeling in idiopathic pulmonary fibrosis: Prediction of time to disease progression and life expectancy with pirfenidone. Paper presented at: BTS 2015 Annual Congress2-4 Dec, London, UK.
65. PIPF-002 ClinicalTrials.gov. Safety Study of Oral Pirfenidone in Patients With Pulmonary Fibrosis/Idiopathic Pulmonary Fibrosis. In; 2016.
66. Roskell N, Saunders O, Lee D, Fisher M. Long-term survival analysis: Pirfenidone compared to standard care for the treatment of patients with idiopathic pulmonary fibrosis. *European Respiratory Journal* 2014;44:A1905.
67. Navaratnam V, Fleming KM, West J, Smith CJ, Jenkins RG, Fogarty A*, et al.* The rising incidence of idiopathic pulmonary fibrosis in the U.K. *Thorax* 2011;66:462-7. http://dx.doi.org/doi: 10.1136/thx.2010.148031
68. Genentech I. Safety Study of Oral Pirfenidone in Patients With Pulmonary Fibrosis/Idiopathic Pulmonary Fibrosis. 2015, https://clinicaltrials.gov/ct2/show/NCT00080223.
69. Albera C. Pirfenidone Is Efficacious in Patients With Idiopathic Pulmonary Fibrosis (IPF) and Mild Restrictive Disease: Benefit of Early Intervention. Paper presented at: ATS 2015 Annual Congress; Denver, USA, 15-20 May.
70. Behr J, Demedts M, Buhl R, Costabel U, Dekhuijzen RP, Jansen HM*, et al.* Lung function in idiopathic pulmonary fibrosis--extended analyses of the IFIGENIA trial. *Respir Res* 2009;10. http://dx.doi.org/doi: 10.1186/1465-9921-10-101
71. Demedts M, Behr J, Buhl R, Costabel U, Dekhuijzen R, Jansen HM*, et al.* High-dose acetylcysteine in idiopathic pulmonary fibrosis. *N Engl J Med* 2005;353:2229-42.,
72. Richeldi L, du Bois RM, Raghu G, Azuma A, Brown KK, Costabel U*, et al.* Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2071-82.
73. Richeldi L, Costabel U, Selman M, Kim DS, Hansell DM, Nicholson AG*, et al.* Efficacy of a tyrosine kinase inhibitor in idiopathic pulmonary fibrosis. *N Engl J Med* 2011;365:1079-87. http://dx.doi.org/doi: 10.1056/NEJMoa1103690
74. Martinez FJ, de Andrade JA, Anstrom KJ, King TE, Raghu G. Randomized Trial of N-acetylcysteine in Idiopathic Pulmonary Fibrosis. *The New England journal of medicine* 2014;370:2093-101. http://dx.doi.org/10.1056/NEJMoa1401739
75. Raghu. The Idiopathic Pulmonary Fibrosis Clinical Research Network. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. *N Engl J Med* 2012;366:1968-77.
76. Chhabra SK. Forced vital capacity, slow vital capacity, or inspiratory vital capacity: which is the best measure of vital capacity? *J Asthma* 1998;35:361-5.
77. The Idiopathic Pulmonary Fibrosis Clinical Research Network. Prednisone, Azathioprine, and N-Acetylcysteine for Pulmonary Fibrosis. *The New England journal of medicine* 2012;366:1968-77. http://dx.doi.org/10.1056/NEJMoa1113354
78. National Institute for Health and Care Excellence. Idiopathic pulmonary fibrosis - nintedanib. Final appraisal determination. December 2015, Available on: http://www.nice.org.uk/guidance/indevelopment/gid-tag491/documents [date accessed: Jan 2016].

79. Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A. Bayesian measures of model complexity and fit. *J Roy Stat Soc B* 2002;64:583-616. http://dx.doi.org/Doi 10.1111/1467-9868.00353
80. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *Int J Epidemiol* 2012;41:818-27. http://dx.doi.org/10.1093/ije/dys041
81. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 5: the baseline natural history model. *Med Decis Making* 2013;33:657-70. http://dx.doi.org/10.1177/0272989X13485155
82. Plummer M. JAGS User Manual. In. 3.3.0 ed; 2012.
83. R Developement Core Team. R: A language and environment for statistical computing. In: R Foundation for Statistical Computing; 2013.
84. Su Y-S, Yajima M. R2jags: Using R to Run 'JAGS'. In; 2015.
85. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Stat Comput* 2000;10:325-37. http://dx.doi.org/Doi 10.1023/A:1008929526011
86. Sturtz S, Ligges U, Gelman A. R2WinBUGS: A package for running WinBUGS from R. *J Stat Softw* 2005;12:1-16. <Go to ISI>://WOS:000232806500001
87. Arber M, Garcia S, Veale T, Edwards M, Shaw A, Glanville J. Sensitivity of a Search Filter Designed to Identify Studies Reporting Health State Utility Values; 2015.
88. Centre for Reviews of Dissemination. Search strategies. York; 2014.
89. Cooper K, Mendes D, Picot J, Loveman E. Pirfenidone for the treatment of idiopathic pulmonary fibrosis : A Single Technology Appraisal; 2012.
90. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;313:275-83. http://www.ncbi.nlm.nih.gov/pubmed/8704542
91. Canadian Agency for Drugs and Technologies in Health. CDEC Final Recommendation - Nintedanib. Ottwa, Canadia: Canadian Agency for Drugs and Technologies in Health; 2015.
92. Canadian Agency for Drugs and Technologies in Health. CDEC Final Recommendation - Pirfenidone Resubmission. Ottwa: Canadian Agency for Drugs and Technologies in Health. 2015, Available from: https://www.cadth.ca/pirfenidone-12. .
93. Loveman E, Copley VR, Colquitt JL, Scott DA, Clegg AJ, Jones J*, et al.* The effectiveness and cost-effectiveness of treatments for idiopathic pulmonary fibrosis: systematic review, network meta-analysis and health economic evaluation. *BMC Pharmacol Toxicol* 2014;15:63. http://dx.doi.org/10.1186/2050-6511-15-63
94. Loveman E CV, Colquitt J, Scott DA, Clegg A, Jones J, et al. The clinical effectiveness and cost-effectiveness of treatments for idiopathic pulmonary fibrosis: a systematic review and economic evaluation. *Health Technology Assessment* 2015;19:1-336.,
95. Starkie HJ, Briggs AH, Chambers MG, et al. Predicting EQ-5D Values Using the SGRQ. *Value in Health* March–April 2011;Volume 14:Pages 354–60.
96. Zisman DA SM, Anstrom KJ, Collard HR, Flaherty KR, Hunninghake GW. . A controlled trial of sildenafil in advanced idiopathic pulmonary fibrosis. *New England Journal of Medicine* 2010;363:620-8.,
97. National Institute for Health and Clinical Excellence. Updated information and analyses to support the Single Technology Appraisal submission of pirfenidone for the treatment of idiopathic pulmonary fibrosis (IPF) - TA282. 2012, Available at: https://www.nice.org.uk/guidance/ta282/resources/idiopathic-pulmonary-fibrosis-pirfenidone-intermune-additional-analyses2.
98. National Institute for Health and Care Excellence. Idiopathic pulmonary fibrosis - Pirfenidone. Final appraisal determination. *NICE* April 2013, Available at: https://www.nice.org.uk/guidance/ta282/documents/idiopathic-pulmonary-fibrosis-pirfenidone-final-appraisal-determination-document2.
99. Albera C. Challenges in idiopathic pulmonary fibrosis trials: the point on end-points. . *EurRespirRev* 2011;20.

100. National Institute for Health and Care Excellence (NICE). Assessment group report (TA257) - Lapatinib or trastuzumab in combination with an aromatase inhibitor for the first-line treatment of metastatic hormone-receptor-positive breast cancer that overexpresses HER2. June 2012, Available at http://publications.nice.org.uk/lapatinib-or-trastuzumab-in-combination-with-an-aromataseinhibitor-for-the-first-line-treatment-of-ta257.
101. Hatziandreu E. The potential cost savings of greater use of home- and hospice based end of life care in England (2008). *NAO* 2008, Available at: https://www.nao.org.uk/wp-content/uploads/2008/11/07081043_potentialsavings.pdf.
102. Handorf EA, McElligott S, Vachani A, et al. Cost Effectiveness of Personalized Therapy for First-Line Treatment of Stage IV and Recurrent Incurable Adenocarcinoma of the Lung. *Journal of Oncology Practice* 2012;8:267-74.
103. Department of Health. NHS reference costs 2013 to 2014. 2015. URL: https://www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014 (Accessed 1 March, 2016).
104. Pittrow D KJ, Kreuter M, Hoeper MM, Wirtz H, Koschel D, et al. . Symptom burden and health related quality of life in patients with idiopathic pulmonary fibrosis in clinical practice: Insights-IPF registry. *Value in Health* 2014;17:A600.
105. King TE J, Brown KK, Raghu G, du Bois RM, Lynch DA, Martinez F, et al. . BUILD-3: a randomized, controlled trial of bosentan in idiopathic pulmonary fibrosis. *American Journal of Respiratory and Critical Care Medicine* 2011;184:92-9.
106. Freemantle N, Wilson A, Fisher M. Mapping The St George's Respiratory Questionnaire To The Euroqol 5 Dimensions: A Study In Patients With Idiopathic Pulmonary Fibrosis. *Value in Health* 2015;18:A503.
107. Parfrey H LC, Gibbons MA, Armstrong E, Harris E, Frank R, et al. . Healthcare utilisation by patients with idiopathic pulmonary fibrosis; observations from the UK pirfenidone named patient programme. *Thorax* 2013;68:A164-A65.,
108. Navaratnam V FA, Glendening R et al. . The Increasing Secondary Care Burden of Idiopathic Pulmonary Fibrosis: Hospital Admission Trends in England From 1998 to 2010. *Chest* 2013;143:1078-84.
109. Wilson E, Shulgina L, Cahn A, Chilvers E, Parfrey H, Clark. AB*., et al.* Treating idiopathic pulmonary fibrosis with the addition of co-trimoxazole: an economic evaluation alongside a randomised controlled trial. *Pharmacoeconomics* 2014;31:87-99.
110. Whiting P AM, Burgers L, Westwood M, Ryder S, Hoogendoorn M, et al. . Ivacaftor for the treatment of patients with cystic fibrosis and the G551D mutation: a systematic review and cost-effectiveness analysis. *Health Technology Assessment* 2014;18:1-106.
111. National Institute for Health and Care Excellence. Costing report: Idiopathic pulmonary fibrosis. London: NICE; 2013. 2013, Available from: https://www.nice.org.uk/guidance/cg163/resources/costing-report-190080397. .
112. Scottish Medicines Consortium. Detailed advice for nintedanib (Ofev®) [SMC no. 1076/15]. 2015, Available on: http://www.scottishmedicines.org.uk/SMC_Advice/Advice/1076_15_nintedanib_Ofev/nintedanib_Ofev [date accessed: Dec 2015].
113. Corte T, Bonella F, Crestani B, Demedts M, Richeldi L, Coeck C*, et al.* Safety, tolerability and appropriate use of nintedanib in idiopathic pulmonary fibrosis. *Respir Res* 2015 Sep 24;16:doi: 10.1186/s12931-015-0276-5.,
114. Ara R, Brazier J. Populating an Economic Model with Health State Utility Values: Moving toward Better Practice. *Value in Health* 2010;13:509-18.

10. APPENDICES

Appendix 1: Comparison of the observed KM for OS in people initiating pirfenidone against extrapolation using parametric distributions for people with a percent FVC > and 50 - 80% at baseline

Figure 46: Comparison of the observed KM for OS in people initiating pirfenidone against extrapolation using parametric distributions for people with a percent FVC of 50 - 80% at baseline (Plot drawn by the ERG)

Figure 47: Comparison of the observed KM for OS in people initiating pirfenidone against extrapolation using parametric distributions for people with a percent FVC of 50 - 80% at baseline (Plot drawn by the ERG)

Appendix 2: Plot of the annual hazard of death of modelled OS for BSC using the Weibull and Gompertz distribution and life tables in UK in people with a percent FVC of 50 – 80% and > 80% at baseline

Figure 48: Plot of the annual hazard of death of modelled OS for BSC using the Weibull and Gompertz distribution and life tables in UK in people with a percent FVC >80% at baseline (Plot drawn by the ERG)

Figure 49: Plot of the annual hazard of death of modelled OS for BSC using the Weibull and Gompertz distribution and life tables in UK in people with a percent FVC of 50 - 80% at baseline (Plot drawn by the ERG)

Appendix 3: Full incremental analysis for key sensitivity analyses and ERG base-case

Table 76 to Table 79 present full incremental results for the four key sensitivity analyses identified in Table 69 on the main report. These results are all deterministic and incorporate the PAS. The results for pirfenidone versus nintedanib can be found in the confidential appendix.

Table 76: Treatment effect assumed to stop after 2 years – deterministic incorporating PAS

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£33,798	5.215			
Pirfenidone	£66,638	5.667	£32,840	0.452	£72,599
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£31,180	4.690			
Nintedanib	See confidential appendix				
Pirfenidone	£61,035	5.138	£29,854	0.449	£66,503
Mild population (percent predicted FVC >80%)					
BSC	£40,671	6.606			
Pirfenidone	£84,209	6.994	£43,539	0.388	£112,214

Table 77: Gompertz distribution for OS – deterministic incorporating PAS

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£25,996	3.687			
Pirfenidone	£64,362	5.200	£38,366	1.513	£25,360
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£24,430	3.374			
Nintedanib	See confidential appendix				
Pirfenidone	£59,276	4.776	£34,846	1.402	£24,855
Mild population (percent predicted FVC >80%)					
BSC	£30,124	4.520			
Pirfenidone	£79,543	6.094	£49,420	1.575	£31,379

Table 78: Inclusion of stopping rule for pirfenidone – deterministic incorporating PAS

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£26,627	3.797			
Pirfenidone	£54,360	5.664	£27,733	1.868	£14,847
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£24,868	3.443			
Nintedanib	See confidential appendix				
Pirfenidone	£50,596	5.136	£25,728	1.693	£15,197
Mild population (percent predicted FVC >80%)					
BSC	£31,729	4.824			
Pirfenidone	£65,740	6.989	£34,011	2.165	£15,707

Table 79: Treatment effect at 72 weeks (incorporating CODA samples) – probabilistic incorporating PAS

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£29,694	4.393			
Pirfenidone	£66,685	5.672	£36,991	1.279	£28,922
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£27,683	3.995			
Nintedanib	See confidential appendix				
Pirfenidone	£61,097	5.157	£33,414	1.162	£28,766
Mild population (percent predicted FVC >80%)					
BSC	£35,220	5.520			
Pirfenidone	£84,133	6.999	£48,913	1.480	£33,060

Table 80 to Table 83 below present full incremental results for the ERG-preferred base-case under both optimistic and pessimistic assumptions regarding the duration of treatment effect, both with and without the stopping rule applied, when incorporating the PAS.

Table 80: ERG-preferred base-case assuming no stopping rule with optimistic assumption regarding duration of treatment effect (life-time effect) – probabilistic incorporating PAS

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£30,972	3.968			
Pirfenidone	£69,560	4.935	£38,589	0.967	£39,895
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£29,220	3.64	-	-	-
Nintedanib	See confidential appendix				
Pirfenidone	£64,325	4.53	£35,106	0.90	£39,166
Mild population (percent predicted FVC >80%)					
BSC	£35,053	4.747			
Pirfenidone	£84,735	5.742	£49,682	0.995	£49,921

Table 81: ERG-preferred base-case assuming no stopping rule with pessimistic assumption regarding duration of treatment effect (2 years) – probabilistic incorporating PAS

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£34,439	4.617			
Pirfenidone	£69,352	4.918	£34,913	0.302	£115,751
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£32,032	4.18	-	-	-
Nintedanib	See confidential appendix				
Pirfenidone	£63,604	4.48	£31,571	0.30	£104,915
Mild population (percent predicted FVC >80%)					
BSC	£39,060	5.498			
Pirfenidone	£84,712	5.743	£45,652	0.245	£186,260

Table 82: ERG-preferred base-case assuming stopping rule with optimistic assumption regarding duration of treatment effect (life-time) – probabilistic incorporating PAS

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£30,947	3.964			
Pirfenidone	£57,216	4.932	£26,269	0.968	£27,124
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£29,225	3.64	-	-	-
Nintedanib	See confidential appendix				
Pirfenidone	£53,790	4.53	£24,565	0.90	£27,432
Mild population (percent predicted FVC >80%)					
BSC	£35,035	4.757			
Pirfenidone	£66,796	5.759	£31,761	1.001	£31,722

Table 83: ERG-preferred base-case assuming stopping rule with pessimistic assumption regarding duration of treatment effect (2 years) – probabilistic incorporating PAS

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£34,430	4.610			
Pirfenidone	£57,048	4.911	£22,618	0.301	£75,121
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£32,081	4.20	-	-	-
Nintedanib	See confidential appendix				
Pirfenidone	£53,249	4.50	£21,169	0.30	£70,234
Mild population (percent predicted FVC >80%)					
BSC	£39,063	5.501			
Pirfenidone	£66,794	5.745	£27,731	0.245	£113,365

Appendix 4: Errors identified and corrected by the ERG

Error identified	Description of the error	Change by the ERG
Discrepancies in the NMA AE outputs between those used in the model and those reported in clarification response	Result from Table 16 in clarification response does not match NMA outputs in sheet "NMA" in the economic model in cells E108-110 & Cells K108-110.	Values reported in Table 16 in clarification response are used in the economic model by the ERG
Incorrect application of outputs of NMA for AEs	In sheet "Model Inputs", RR are applied to the incidence of AEs on BSC. However, the inverse of RR are calculated in Sheet "NMA" and used.	The inverse of the RR used in sheet "NMA" are used for AE in sheet "Model Inputs"
Parameters for PFS using the Gompertz and Gamma for the subgroup analyses	In sheet "PFS Parameters", cells H25:26 for Gompertz are linked to incorrect Cells. Same for Gamma distribution.	In sheet "PFS Parameters", in cell H25, replace D138 by D115. In sheet "PFS Parameters", in cell H26, replace D140 by D116.
Calculation of the drug acquisition cost per cycle.	In sheet "Model Inputs" in Cell E200, assume 365 days instead of 365.25 (as used throughout the rest of the model).	Replace 365 by 365.25
Calculation of the cost for healthcare professional visits in PSA	In sheet "Costs" in Cell G64. Calculation use the deterministic cost for health care professional (Cell F53)	In sheet "Costs" in Cell G64. Replace F53 by H53
Use of the cost for progression-free for the progressive health state in people initiating pirfenidone	In sheet "Esbriet" column BJ, use of "c_dm_pir_pre" instead of cost for progression health state	In sheet "Esbriet" column BJ, replace "c_dm_pir_pre" by "c_dm_pir_post"
Implementation on nintedanib discount in PAS analyses	Nintedanib discount applied in addition to pirfenidone discount	Nintedanib PAS discount applied to nintedanib list

		price. Corrected as follow: [([2151.11/60] x 2) x (365.25/4)] x (1 – PAS discount) Also corrected error in days (from 365 by 365.25) for both drugs

Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282): A Single Technology Appraisal

Produced by School of Health and Related Research (ScHARR), The University of Sheffield

Authors Sarah Davis, Senior Lecturer in Health Economics, ScHARR, University of Sheffield, Sheffield, UK

Rachid Rafia, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Christopher Carroll, Reader in Systematic Review and Evidence Synthesis, ScHARR, University of Sheffield, Sheffield, UK

Munira Essat, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Jean Sanderson, Research Associate in Statistics, ScHARR, University of Sheffield, Sheffield, UK

Naila Dracup, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK

Dr Stephen Bianchi, Consultant in Respiratory Medicine, Sheffield Teaching Hospitals NHS Foundation

Professor David Thickett, Consultant in Respiratory Medicine, University Hospitals of Birmingham NHS Foundation Trust and Chair in Respiratory Medicine, University of Birmingham

Correspondence Author Sarah Davis, Senior Lecturer in Health Economics, ScHARR, University of Sheffield, Sheffield, UK

Date completed Date completed (05/03/2016)

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 142/06/02.

ASCEND, CAPACITY 1 & 2 and SP3 at week 52 showed no difference between the pirfenidone and placebo group (OR: 0.90, 95% CI: 0.70 to 1.15, *p*-value not reported).

In the absence of head-to-head RCTs evaluating nintedanib against pirfenidone the company conducted a Bayesian NMA to perform an indirect treatment comparison. NMAs were conducted for 11 outcomes relevant to the decision problem and the results of four of these outcomes (overall survival [OS], PFS, time to treatment discontinuation and acute exacerbations) were used to inform the economic model. Based on the NMA, the treatment effects for pirfenidone were broadly similar to those for nintedanib for all outcomes, with the pairwise treatment effects indicating that neither treatment is statistically significantly more effective.

NMA of safety data indicated that pirfenidone is associated with reduced odds of diarrhoea compared to nintedanib, and increased odds of rash as compared to nintedanib. For discontinuation due to adverse events and serious cardiac adverse events, the treatment effects for pirfenidone are broadly similar to those for nintedanib, with the pairwise treatment effects indicating that neither treatment is associated with more adverse events.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The final selection of three trials (ASCEND, CAPACITY 1 and CAPACITY 2) for the main clinical efficacy review was considered to be appropriate by the ERG, as was the inclusion of the trials from Japan, SP3 and SP2, as supporting evidence. An additional relevant trial was also identified by the ERG and included as supporting evidence: this was a multicenter Chinese trial, which compared pirfenidone plus N-acetylcysteine (NAC) with placebo plus NAC in adult patients with mild or moderate IPF (Huang 2015). The ERG noted that there were between-trial differences across some baseline characteristics in the three key trials (ASCEND, CAPACITY 1 & 2), such as mean FVC or 6MWD at baseline, but subgroup analyses suggested that these and other variables did not influence treatment effect.

Overall, the ERG assessed the potential risk of bias in ASCEND and CAPACITY 1 & 2 to be low across most domains, with the exception of reporting bias and "other bias", which were judged to be "moderate", on account of inconsistency between some of the outcomes and analyses specified in the trial protocols and those presented in the CS, and the possible influence of uncontrolled variables such as rate of disease progression. The SP3, SP2 and Huang *et al.* (2015) trials were at a higher or more unclear risk of bias across many domains than the ASCEND and CAPACITY trials. These trials all evaluated lower doses of pirfenidone, which are licensed in Japan but not in the UK, applied different eligibility criteria and presented noticeable differences from the other three trials in some baseline characteristics of participants.

- the inadequacy of the partitioned survival approach when implementing the stopping rule
- the assumption that treatment effect is constant over the entire model duration
- the estimation of the treatment effect

The ERG further observes that under the company's base-case assumptions, there are discrepancies between the model's prediction of OS for people initiating BSC and the observed trial data for OS in patients who were randomised to placebo. The CS does not comment on these discrepancies and instead focuses on a comparison of the model prediction with registry data for patients receiving BSC, even though the registry data does not match the trial data for people randomised to placebo.

1.6 ERG commentary on the robustness of evidence submitted by the company

The ERG notes the following strengths and weaknesses in the evidence submitted by the company.

1.6.1 Strengths

- The CS reports a generally good quality systematic review of the RCT evidence.
- The three principal RCTs are generally at a low risk of bias.
- Generally, there are no major safety concerns, and some long-term safety evidence is available.
- *Evidence in the model for pirfenidone is based upon long-term data for people included in RECAP.*
- *Results from NMAs are used to inform the relative treatment effects for the comparators.*
- Whilst EQ-5D data were not directly available in the trials, SGRQ data from the trials were mapped onto the EQ-5D using a mapping algorithm developed in people with IPF.

1.6.2 Weaknesses and areas of uncertainty

- There is a moderate risk of reporting bias in the three key RCTs and unclear, moderate or high risk of bias across some domains in the three supporting RCTs.
- There are difficulties in controlling for the rate of disease progression among IPF trial participants, which might moderate outcomes, however the extent of this is unclear.
- The efficacy findings are not consistent across individual trials; one of the key trials reports no statistically significant treatment effect for pirfenidone compared with placebo on the primary outcomes measures relating to FVC or the secondary outcome of PFS.
- Individual trials do not report any statistically significant treatment effect compared to placebo for all-cause mortality; a statistically significant treatment effect is only observed when pooling or meta-analysing studies.

at baseline, FVC >80% at baseline). For the ITT population the ICERs incorporating the PAS ranged from £27,124 to £115,751. For the mild population (percent predicted FVC >80%) the ICERs incorporating the PAS ranged from £31,722 to £186,260. Fort the moderate population (percent predicted FVC 50 – 80%), the ICERs for pirfenidone versus BSC ranged from £27,432 - £104,915 when incorporating the PAS. Results incorporating the PAS for pirfenidone versus nintedanib in the moderate population are presented in the confidential appendix.

A key uncertainty in the company's model concerns the duration of the extrapolation of the treatment effect. As reported in the company's scenario analyses and the ERG's exploratory analyses, truncating the duration over which the treatment effect applies increases the ICERs for pirfenidone versus BSC. A further important limitation in the company's model relates to the implementation of stopping rules for pirfenidone and nintedanib. The inclusion of the stopping rule in the economic model lacks validity in that the modelled stopping rule impacts on costs but not health outcomes. The ERG considers that the analysis incorporating the stopping rule as implemented in the economic model provides a lower bound of the plausible ICER (i.e. most optimistic scenario).

4.1.5 *Evidence synthesis*

The synthesis for the review of clinical efficacy was a basic descriptive summary of the evidence from the five pirfenidone trials for the following outcomes: change from baseline in percent predicted FVC; all-cause and IPF-related mortality; PFS; acute exacerbation; hospitalisation; changes from baseline in 6MWD, the UCSD SOBQ and the SGRQ. This approach to evidence synthesis was neither described nor justified in the CS.[4]

Meta-analyses using both fixed and random effects models comparing pirfenidone with placebo were performed for selected outcomes and time-points, based on available trial data, and the methods used were described in the CS[4] (Section 4.9 and Appendix 9). Data were combined from CAPACITY 1 & 2[33, 36, 49] and ASCEND[33, 34] using the UK licence dosage (2,403mg/day) and from SP3[38] which uses a lower dosage (1,800 mg/day which is the licensed dose in Japan but is not a licensed dose in the UK). The company considered this to be appropriate as the dose by weight would be similar for all studies given the lower body weight of the Japanese population compared with the North American and European population. An NMA comparing effects across all treatments was also performed by the company. This is critiqued in Sections 4.6 and 4.7 of this report.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 *Review of clinical efficacy (relevant pirfenidone RCT evidence)*

The CS[4] provides a detailed description of trials identified by the company as satisfying the requirements of the final NICE scope,[3] i.e. pirfenidone compared with placebo or nintedanib (see Table 1). No trial compared pirfenidone with nintedanib. Five RCTs compared pirfenidone at various doses with placebo: ASCEND (Phase III),[34] CAPACITY 1 & CAPACITY 2 (Phase III),[49] SP3 (Phase III),[38] and SP2 (Phase II).[39] Three trials were international and multicentre (ASCEND[34] and CAPACITY 1 & 2[33, 36, 49]), although only CAPACITY 2 included any UK centres[35] (three of 110 centres across both CAPACITY trials).[49] The inclusion criteria in all three trials were adult patients with mild or moderate IPF based on percentage predicted FVC of ≥50% (in ASCEND[34] this had an upper limit of ≤90%). Two trials were conducted exclusively in Japan (SP3[38] and SP2[39]) and did not report baseline levels of FVC or VC. One trial was conducted in China and evaluated pirfenidone in combination with N-acetylcysteine (NAC). The trials varied in criteria relating to lung function, concomitant medications permitted for IPF, and the investigated doses of pirfenidone (for the purposes of this appraisal, ASCEND,[34] and CAPACITY 1 & 2[49] all evaluated the efficacy of the licensed dose of 2,403mg/d; the SP2,[39] SP3[38] and the Huang *et al.* [48] trial evaluated lower doses; the applicability of these lower doses to clinical practice in England and Wales is unclear.

Table 1: Characteristics of included pirfenidone RCTs (reproduced in part from CS,[4] Tables 10 and 15, pages 59 and 82)

Trial No. of patients	Location	Inclusion criteria		Exclusion criteria	Intervention and co-interventions (No. of patients)	Comparator (No. of patients)	Follow-up
		IPF diagnosis	Lung function parameters	Patient factors			
ASCEND (PIPF-016)[33, 34] n=555	International multi-centre	– Confident clinical and radiographic diagnosis of IPF, confirmed centrally with diagnosis of IPF >6 months but <48 months. – No improvement of IPF in preceding year. –	– FVC (% predicted value) 50-90% – DLco 30-90% – 6MWT ≥150 m	– Abnormal lab parameters – Obstructive airway disease – History of unstable /deteriorating cardiac or pulmonary disease – History of severe hepatic impairment/ end-stage liver disease/end-stage renal disease requiring dialysis	Pirfenidone 2,403mg/day (n=278) Concomitant treatment with any investigational drug for the treatment of IPF was prohibited. However, concomitant medications used in the treatment of IPF were permitted if given for a non-IPF indication and there was no clinically acceptable alternative.	Placebo (n=277)	52 weeks
CAPACITY 1 (PIPF-006)[36, 49] n=344	International multi-centre	– Confident clinical and radiographic diagnosis of IPF, confirmed locally (diagnosis previous 48 months) – No improvement of IPF in preceding year	– FVC (% predicted value) ≥ 50% – DLco ≥35% – FVC or DLco ≤90% – 6MWT ≥150 m		Pirfenidone 2,403mg/day (n=171) Concomitant treatments for IPF were prohibited, with exceptions of short courses of azathioprine, cyclophosphamide, corticosteroids, or acetylcysteine for protocol-defined acute exacerbation of IPF, acute respiratory decompensation, or progression of disease.	Placebo (n=173)	72 weeks
CAPACITY 2 (PIPF-004)[35, 49] n=435	International multi-centre		– FVC (% predicted value) ≥50% – DLco ≥35% – FVC or DLco ≤90%		Pirfenidone 2,403mg/day (n=174) Pirfenidone 1,197mg/day (n=87) As CAPACITY 1	Placebo (n=174)	72 weeks

IPF: Idiopathic Pulmonary Fibrosis; FVC: Forced Vital Capacity; DLco: Diffusing capacity of the lungs for carbon monoxide; 6MWT: 6-minute walking test

The exclusion of certain patients otherwise eligible for pirfenidone, based on co-morbidities, such as obstructive airways disease, must also be taken into account when judging the generalisability of the trials' findings.

The outcomes reported in the CS[4] are generally consistent with those that are listed in the final NICE scope.[3] The ASCEND,[34] CAPACITY[49] and Huang *et al*[48] trials use change from baseline in percent predicted FVC as an endpoint, while SP3[38] and SP2[39] use VC. The CS states that the decision to use VC in the SP3[38] and SP2[39] trials was dictated by the ATS international consensus statement published in 2000, which recommended measurement of VC.[54] The CS[4] did not state when the recommended measurement changed to FVC or provide any reference to substantiate the change. The CS[4] states that VC and FVC should be treated as comparable endpoints as there is little difference between VC and FVC in subjects without obstructive pathology. Whilst the clinical advisors to the ERG agreed with this statement, the ERG noted that the exclusion criteria for SP3[38] were not as explicit regarding the exclusion of patients with emphysema as the exclusion criteria for the other pirfenidone trials. Therefore, the ERG considers that the synthesis of VC data from SP3[38] with FVC data from the ASCEND[34] and CAPACITY trials[49] is questionable.

The outcomes listed in the trial protocols publicly-available from the clinical trials register (https://clinicaltrials.gov/ct2/home) are not entirely consistent with those reported in the CS.[4] For example, the principal efficacy outcome of "percent predicted FVC or death" does not appear in any protocol as a trial outcome but appears to describe the method used by the company in order to impute a FVC measurement for patients who have died (see clarification response[10], questions, A11 and A13). Furthermore, neither of the secondary outcomes of "treatment-emergent IPF-related mortality" nor the SGRQ was listed in the protocols, but both appear *post hoc* as outcomes in the CS[4] (as well as in the ASCEND[34] publication, but not in the CAPACITY trials' publication,[49] see Table 3).

The following outcome was listed in protocols but was not reported in the results for the CAPACITY 1 & 2[49] and SP3[38] trials: Change in Worst Oxygen Saturation by Pulse Oximetry (SpO_2) measurement observed during the 6-Minute Walk Test. The CAPACITY trial protocols[35, 37, 51] also listed lung transplantation as a secondary outcome, but this is not included as an outcome in the CS[4] (pages 53 and 66). The CS[4] lists fibrosis by use of high resolution computed tomography (HRCT) (see CS, Table 12, page 68) as an outcome, but this only appears to be used as a diagnostic criterion for IPF or as part of the definitions of acute exacerbations (see CS,[4] pages 104-105).

Definitions of outcomes are first provided under the trial results section of the CS[4] (Section 4.7, pages 90-113). The outcomes, and the definitions applied in each of the trials, taken from the CS and the

Table 2: Primary efficacy outcomes and measures in ASCEND, CAPACITY 1, CAPACITY 2, SP3 and SP2

Outcome	**ASCEND**[33, 34]	**CAPACITY 1**[36, 49]	**CAPACITY 2**[35, 49]	**SP3**[38]	**SP2**[39]
Protocol-listed outcome	Change in percent predicted FVC from baseline to week 52†	Mean and absolute change in percent predicted FVC from baseline to week 72		No protocols available	
Reported outcomes	Change in percent predicted FVC from baseline to week 52	Change in percent predicted FVC **or death** from baseline to week 52		Change in VC from baseline to week 52	Change in the lowest SpO_2 during 6MWT. Full definition given in Azuma, page 1041 Change in VC from baseline was listed as a secondary outcome
	Categorical decline of ≥10% in percent predicted FVC or death	Categorical decline of ≥10% in percent predicted FVC. This was listed as a secondary outcome in the protocols and publication, defined as "Categorical Assessment of Absolute Change in Percent Predicted Forced Vital Capacity (FVC) based on the change in baseline percent predicted FVC at week 72, patients were assigned to 1 of 5 categories: mild decline (<10% but >=0% decline), moderate decline (<20% but >=10% decline), severe decline (>=20% decline), mild improvement (>0% but <10% improvement), or moderate improvement (>=10% improvement). Those who died or had a lung transplant before week 72 were included in the severe decline category. The results indicate the number of patients who experienced a Categorical Change in Percent Predicted Forced Vital Capacity" [35, 36, 49]			
Magnitude of treatment effect	The magnitude of the treatment effect was estimated by comparing the distribution of patients in the pirfenidone group with those in the placebo group across two thresholds of change at week 52: an absolute decline of 10 percentage points in the percentage of the predicted **FVC or death**, or no decline in the percentage of the predicted FVC (King 2014, page 2085)[34]	Estimated by use of differences in treatment group means and categorical change in FVC (page 1763, Noble 2011)[49]			

† This outcome was not reported in the ASCEND publication; the data were only made available by Roche in the CS,[4] Table 20 and pages 93-94.

Table 3: Secondary efficacy outcomes and measures in ASCEND, CAPACITY 1, CAPACITY 2, SP3 and SP2

Outcome	ASCEND[33, 34]	CAPACITY 1[36, 49]	CAPACITY 2[35, 49]	SP3[38]	SP2[39]
All-cause mortality	Yes				
IPF-related death	Yes	Yes*		No	
Treatment-emergent IPF mortality	Yes. Defined as death occurring after randomisation and within 28 days of the last dose of the study drug (CS, page 96).[4] Listed only in the ASCEND NEJM protocol but reported for all mortality outcomes in ASCEND publication and separately, applied and not-applied, to all-cause and IPF-related mortality in the CAPACITY publication: appears to be a *post hoc* outcome measure.				
Progression-free Survival (PFS)	Defined in the CS (page 99)[4] as a confirmed ≥10% decline from baseline in %FVC, confirmed ≥50 m decline from baseline in 6MWD, or death	PFS is defined as the first occurrence of a 10% absolute decline from baseline in percent predicted Forced Vital Capacity, a 15% absolute decline from baseline in percent predicted hemoglobin(Hgb)-corrected carbon monoxide diffusing capacity (DLco), or death		Defined as VC decline of ≥10% or death. When the VC data could not be obtained due to worsening of respiratory symptoms, including acute exacerbation, the case was also classified as disease progression. (Taniguchi, page 822)[38]	No
Acute Exacerbations	Identified via a *post hoc* analysis of adverse events based on the MedDRA lower level term "acute exacerbation of IPF".(CS, page 104)	Definition not provided in clinical trials register protocols (where it is reported only as part of a composite measure*). CS (page 104)[4] defines this outcome as requiring all of the following within a 4-week interval: Worsening of PaO_2 (≥8 mm Hg drop from the most recent value); clinically significant worsening of dyspnoea; new, superimposed ground-glass opacities on HRCT in one or more lobes; all other cardiac, thromboembolic, aspiration, infectious processes ruled out		†Definition not provided in protocols. CS (page 104)[4] defines this outcome as requiring all of the following within a month: increase in dyspnoea; new, ground-glass opacities on HRCT in addition to previous honeycomb lesion; all oxygen partial pressure in resting arterial blood (PaO2) is lower by more than 10 Torr than previous one; exclusion of obvious causes, such as infection, pneumothorax, cancer, pulmonary embolism or congestive heart failure; the serum levels of CRP, LDH are usually elevated as well as serum markers of interstitial pneumonias, such as KL-6, Sp-A or Sp-D	†Definition not provided in protocols. CS (page 104)[4] defines this outcome as requiring all of the following: worsening, otherwise unexplained clinical features within 1 month; progression of dyspnoea over a few days to less than 5 weeks; new radiographic/HRCT parenchymal abnormalities without pneumothorax or pleural effusion (e.g., new, superimposed ground-glass opacities); a decrease in the PaO_2 by 10 mm Hg or more; exclusion of apparent infection based on absence of Aspergillus and pneumococcus antibodies in blood, urine for Legionella pneumophila, and sputum cultures

4.2.2 Results

Participants' baseline characteristics

More than 620 participants received the licensed 2,403mg/day dose during the three international RCTs compared with more than 620 control patients who received placebo in these trials. Another 322 participants received lower doses of pirfenidone in the CAPACITY 2,[49] SP2[39] and SP3[38] trials.

The final selection of three trials (ASCEND,[34] CAPACITY 1[36, 49] and CAPACITY 2[35, 49]) for the main clinical efficacy review was considered to be appropriate by the ERG. However, there are some between-trial differences across some baseline characteristics (see **Error! Reference source not found.**). The ASCEND trial[34] participants had a lower mean percentage predicted FVC (range across arms of 67.8-68.6) than the CAPACITY trials[49] (range across arms of 73.1-76.4) and lower pre-enrollment corticosteroid use (range across arms of 0.7%-2.2%) than the CAPACITY trials[49] (range across arms of 5.2%-12.9%). CAPACITY 1[49] participants had a lower mean 6MWD (range across arms of 378.0-399.1) than in ASCEND[34] and CAPACITY 2[35] (range across arms of 410.0-420.7), and there was a relatively lower proportion of patients in CAPACITY 2[35] requiring supplemental oxygen use (range across arms 14.0%-17.0%) than in ASCEND[34] and CAPACITY 1[49] (range across arms of 27.4%-28.1%). All of these variables, with the exception of corticosteroid use, are accepted potential treatment effect modifiers and therefore were the subject of subgroup analyses in the CS,[4] (Section 4.8, pages 114-117).

The ERG considers the relevance of the smaller SP3[38] and SP2[39] trials, which were conducted exclusively in Japan, to be more questionable. These trials evaluate lower doses of pirfenidone which are licensed in Japan but not in the UK, apply different eligibility criteria and present noticeable differences from the other three trials in some baseline characteristics of participants (see **Error! Reference source not found.**), for example, higher proportions of male participants (range across arms of 78%-94% for SP2[39] and SP3[38] compared with 68%-80% for ASCEND[33, 34] and CAPACITY 1 and 2[49]) and smokers (60%-86% compared with 58%-66%); higher mean percentages of predicted DLco compared with ASCEND[34] and the CAPACITY trials[49] (52.1-57.7 compared with 43.7-47.8), lower trial corticosteroid use (SP3[38] only, 4.8-10.9 compared with 21.0-36.5 in the ASCEND[34] and CAPACITY trials[49]), and smaller proportions having received surgical lung biopsies (21.0%-29.1% compared with 28.5%-55%, see **Error! Reference source not found.**).

Baseline data from participants on patient-reported outcome measures, such as the SGRQ and UCSD SOBQ, were not reported in the CS.[4]

The Huang *et al.* trial[48] comparing pirfenidone plus NAC with placebo plus NAC reported comparability between arms across all baseline characteristics except for smoking status.[48]

Table 4: Categorical analysis of change from baseline in percent predicted FVC or death (reproduced from CS,[4] Table 18)

Study	Time point	Treatment group	Decline ≥10% FVC or death, n (%)	No decline* in FVC, n (%)	*p*-value†
ASCEND[34]	52 weeks	PFN 2,403mg/day (N=278)	46 (16.5)	63 (22.7)	*p*<0.000001
		PBO (N=277)	88 (31.8)	27 (9.7)	
CAPACITY 1[49] §	72 weeks	PFN 2,403mg/day (N=171)	39 (22.8)	44 (25.8)	*p*=0.440
		PBO (N=173)	46 (26.6)	38 (22.0)	
CAPACITY 2[49] §	72 weeks	PFN 2,403mg/day (N=174)	35 (20.1)	42 (24.1)	*p*=0.001
		PBO (N=174)	60 (34.5)	24 (13.8)	
Pooled CAPACITY 1 & 2[49]	72 weeks	PFN 2,403mg/day (N=345)	74 (21)	86 (24.9)	*p*=0.003
		PBO (N=347)	106 (31)	62 (17.9)	

PFN: pirfenidone; PBO: placebo
** CAPACITY trials data not reported in original publication (Noble 2011[49]) but taken from respective CSRs*
†Rank ANCOVA (pirfenidone 2,403mg/day vs placebo). It is unclear if this p value relates to the "Decline or death" or the "No decline" comparison: the numbers in the CS, Table 18 refer to the "No decline" comparison in ASCEND (King 2014[34]), but the "Decline or death" comparison for the CAPACITY trials (Noble 2011[49])
§ Note: these data are from the original publication (Noble 2011[49]), which only reports decline of >10% FVC and not decline of >10% or death

A pooled analysis of ASCEND[34] (week 52) and CAPACITY 1 & 2[49] (week 48) reported a statistically significant difference in favour of pirfenidone compared with placebo in terms of those who had experienced a decline in FVC by ≥10% or had died (absolute difference: 10.0 [95% CI not reported], *p*<0.003), and reported a higher proportion of patients with no decline in percent predicted FVC for pirfenidone compared with placebo (24.9% versus 17.9%, *p*-value not reported). This analysis is described as "pre-specified" in the CS[4] (page 91), but this is inaccurate: there is no reference to this analysis for this outcome in the any of the ASCEND protocols,[33, 55] which only refer to pooling of these trials for mortality (see Section 5.4.2.3.2 in the protocols). The protocol that accompanied the ASCEND publication (Section 13.2, page 29) stated that, "*The clinical study protocol (dated 16 March 2011, section 5.4.2.1) describes a supportive analysis of FVC as the change from baseline to Week 52 in FVC volume (in mL) … A categorical analysis of relative change from baseline has been added*".[55]

The ASCEND trial[34] reported that at 52 weeks there were fewer overall deaths and TE IPF-related deaths in the pirfenidone group than the placebo group, but these differences were not statistically significant (p=0.105 and p=0.226 respectively).

In the pooled analysis of CAPACITY 1 & 2[49] at 52 weeks, there were fewer overall deaths and TE IPF-related deaths in the pirfenidone groups compared with the placebo groups and this difference was statistically significant in both groups (p=0.047 and p=0.012 respectively).

In the pooled analysis of CAPACITY 1 & 2[49] at 72 weeks, there were fewer overall deaths and TE IPF-related deaths in the pirfenidone groups compared with the placebo groups. Overall, there was a 23% reduction in all-cause mortality versus placebo among patients treated with pirfenidone 2,403mg/day (HR=0.77; 95% CI: 0.47 to 1.28; p=0.315), a 38% reduction in IPF-related mortality (HR=0.62; 95% CI: 0.35 to 1.13; p=0.117) and a 35% reduction in TE all-cause mortality (HR=0.65; 95% CI: 0.36 to 1.16; p=0.141). However, none of these differences were statistically significant.

For TE IPF-related mortality, the HR between the pirfenidone and placebo groups at week 72 also favoured pirfenidone and was statistically significant (HR=0.48; 95% CI: 0.24 to 0.95; p=0.03, see **Error! Reference source not found.**).

There appears to be a markedly increased rate of mortality for the CAPACITY trials[49] between the data reported in the CS[4] for 52 weeks (Table 23, page 97) and the data reported in the publication for 72 weeks.[49] There is a substantial increase in all-cause mortality in the pirfenidone group, from 11 at 52 weeks to 27 at 72 weeks, compared with a much smaller increase in the placebo group from 22 at 52 weeks to 34 at 72 weeks (the p-values for the differences between groups are 0.047 and 0.315 for 52 weeks and 72 weeks, respectively). (see **Error! Reference source not found.**). In the same way, TE IPF-related mortality in the pirfenidone group increases from 4 deaths at 52 weeks to 12 deaths at 72 weeks in the pirfenidone group, and from 15 at 52 weeks to 25 at 72 weeks in the placebo group (p-values for the differences between groups are 0.012 and 0.030 for 52 and 72 weeks, respectively). No explanation is provided in the CS[4] for these relative increases in rates of mortality, particularly for the pirfenidone groups, between weeks 52 and 72 in the CAPACITY trials.[49]

In the pooled analysis of the data from 52 weeks for ASCEND[34] and CAPACITY 1 & 2[49] (required by the Food and Drug Administration (FDA)[57] and finalised as an analysis in the Statistical Analysis Plan only on 1st January 2014, according to the company's clarification response[10] (question A22), there were significantly fewer overall deaths (p=0.011) and TE IPF-related deaths (p=0.006) in the pirfenidone groups compared with the placebo groups.

Table 5: Post hoc analysis of data on hospitalisations in CAPACITY 1 & 2 (reproduced from CS, Table 28)[4]

Study arm	CAPACITY 1[49]		CAPACITY 2[49]		Pooled	
	PFN n=171	PBO n=173	PFN n=174	PBO n=174	PFN n=345	PBO n=347
Respiratory hospitalisations (RH)						
Number of patients with at least 1 RH	22 (12.9%)	23 (16.7%)	29 (16.7%)	29 (16.7%)	51 (14.8%)	52 (15.0%)
Number of RH	31	37	34	40	65	77
Mean length of RH (days)	8.5	17.3	7.6	12.1	8.0	14.6
Total number of days in hospital	264	640	259	484	522	1124
Average number of NRH days per patient	1.5	3.7	1.5	2.8	1.5	3.2
Non-respiratory hospitalisations (NRH)						
Number of patients with at least 1 NRH	37 (21.6%)	25 (14.5%)	35 (20.1%)	31 (17.8%)	72 (20.9%)	56 (16.1%)
Number of NRH	48	31	38	42	86	73
Mean length of NRH (days)	10.1	20.8	7.2	16.0	8.8	8.0
Total number of days in hospital	485	645	274	672	758	1317
Average number of NRH days per patient	2.8	3.7	1.6	3.9	2.2	3.8

PFN: pirfenidone 2,403mg/d; PBO: placebo

In SP2,[39] five patients in the placebo arm and none in the pirfenidone treatment were hospitalised due to exacerbations (Azuma 2005[39]). The company did not conduct a meta-analysis as data were only available for the CAPACITY trials.

4.2.2.6 Patient-Reported Outcomes (Quality of Life)

University of San Diego (UCSD) Shortness of Breath Questionnaire (SOBQ)

The ASCEND[34] and CAPACITY trials[39] reported this outcome. The CS[4] states (pages 111 and 112) that the SOBQ can be used to formulate clinically relevant inferences about IPF patients; that the total score in this questionnaire increases with increased dyspnoea, and an increment of 20 points is considered a clinically relevant threshold based on estimates of the minimal important difference for the USCD SOBQ that range from 5-11.[31] In ASCEND,[34] the proportion of patients with ≥20 point increase in shortness of breath as measured by SOBQ at week 52 was smaller in patients receiving pirfenidone than in those receiving placebo, but this difference was not statistically significant (p=0.1577, see **Error! Reference source not found.**).

Table 6: Serious treatment-emergent adverse events reported by ≥2 patients in CAPACITY 1 & 2 at 72 weeks[49]

	CAPACITY 1[49]		CAPACITY 2[49]		
Adverse event, n (%)	**PFN 2,403mg/d (n=171)**	**PBO (n=173)**	**PFN 2,403mg/d (n=174)**	**PFN 1,197mg/d (n=87)**	**PBO (n=174)**
Pneumonia	7 (4.1)	7 (4.0)	4 (2.3)	3 (3.4)	6 (3.4)
Respiratory failure	4 (2.3)	6 (3.5)	2 (1.1)	3 (3.4)	2 (1.1)
Angina pectoris			2 (1.1)	2 (2.3)	1 (0.6)
Atrial fibrillation	2 (1.1)	1 (0.6)	1 (0.6)	3 (3.4)	1 (0.6)
Coronary artery disease	6 (3.5)	0 (0)	0	3 (3.4)	2 (1.1)
Acute renal failure	2 (1.2)	2 (1.2)	1 (0.6)	2 (2.3)	0 (0)
Fall	2 (1.2)	1 (0.6)			
Hypotension	2 (1.2)	1 (0.6)			
Colitis	2 (1.2)	0 (0)			
Hip fracture	2 (1.2)	0 (0)			
Prostate cancer	2 (1.6)*	0 (0)			
Interverterbral disc profusion	2 (1.2)	0 (0)			
Liver test function abnormal	2 (1.2)	0 (0)			
Nephrolithiasis	2 (1.2)	0 (0)			
Sick sinus syndrome	2 (1.2)	0 (0)			
Pneumothorax			3 (1.7)	2 (2.3)	0
Pulmonary embolism			1 (0.6)	3 (3.4)	1 (0.6)
Syncope			3 (1.7)	1 (1.1)	1 (0.6)
Chest pain			3 (1.7)	0	0
Bladder cancer			2 (1.1)	0	0
Gastroesophageal reflux disease			2 (1.1)	0	0
Bronchitis	0 (0)	5 (2.9)†	2 (1.1)		2 (1.1)
Lobar pneumonia	-	-	2 (1.1)		2 (1.1)
Non-cardiac chest pain	-	-	2 (1.1)		2 (1.1)
Myocardial infarction	-	-	0 (0)		4 (2.3)†

* Male patients only †p<0.05

Adverse events leading to discontinuation of treatment

In the ASCEND trial,[34] the proportion of patients discontinuing treatment due to an AE was 14.4% (n=40) in the pirfenidone group and 10.8% (n=30) in the placebo group. The most common AE leading to treatment discontinuation was worsening IPF (1.1% [n=3] in the pirfenidone group versus 5.4% [n=15] in the placebo group), but again the caveats should be noted regarding the categorisation of this event as a safety outcome. The only other AEs leading to discontinuation of treatment in ≥1% of patients in the pirfenidone group were elevated hepatic enzymes levels, pneumonia, rash and decreased weight, which each occurred in 3 patients (1.1%).

In the CAPACITY trials,[49] treatment was discontinued due to AEs in 15% (n=51) of 345 patients in the pooled pirfenidone 2,403mg/day group compared with 9% (n=30) of 347 patients in the placebo group. The most common AE leading to discontinuation was worsening of IPF (3% in both groups). The other AEs leading to discontinuation of treatment in ≥1% of patients in the pirfenidone group were provided by the company in response to a request for clarification from the ERG (see clarification response addendum,[30] question A24). In CAPACITY 1,[49] these were elevated IPF (2.3% in each arm), photosensitivity, rash and respiratory failure, which each occurred in 2 patients (1.2%) in the pirfenidone trial arm but not at all in the placebo arm. In CAPACITY 2,[49] for the 2,403mg per day dose, these were elevated IPF (1.1% for pirfenidone versus 1.7% for placebo), bladder cancer (1.1% vs 0%), nausea (2.3% versus 0%) and rash (1.7% versus 0%). The following substantial laboratory abnormalities (Grade 4 or a shift of 3 grades e.g. from 0 to 3) occurred more frequently in the CAPACITY 1 and 2 pooled pirfenidone 2,403mg/day group compared with placebo: hyperglycaemia (1% [n=4] versus <1% [n=3], respectively); hyponatraemia (1% [n=5] versus 0%); hypophosphatemia (2% [n=6] versus <1% [n=3]); and lymphopenia (1% [n=5] versus 0). However, none were associated with clinically significant consequences. More patients in the pooled pirfenidone-treated group than in the pooled placebo group had elevations in alanine aminotransferase and aspartate aminotransferase of more than 3x the upper limit of normal (4% [n=14] versus <1% [n=2]). However, all reports were reversible and without clinical sequelae.

SP2[39] reported that 11 patients discontinued pirfenidone treatment, compared with 2 patients in the placebo arm, due to AEs.[39] The CS (page 172) stated that skin photosensitivity was the AE that was principally responsible for discontinuing or reducing pirfenidone dose; full data on AE discontinuations were provided in the publication:[39] the principal AEs affecting discontinuation from pirfenidone treatment were: photosensitivity (n=5); vomiting (n=1); fever (n=1); abnormality of hepatic function (n=1); dizziness (n=1); facial paralysis (n=1) and hepatoma (n=1). There were no instances of any of these events in the placebo arm.

Studies included in NMA

The company's systematic review identified 10 RCTs of reasonable methodological quality that compared pirfenidone, nintedanib, NAC, or triple therapy with placebo in patients with IPF. However, the company excluded two of the trials; SP2[39] (pirfenidone) and IFIGENIA[70, 71] (double and triple therapy) from the NMA. INFIGENIA[70, 71] was excluded from the NMA as the trial compares double and triple therapy, which are not comparators of interest for this appraisal. SP2[39] was excluded from the NMA as it was considered as an outlier by the NICE Appraisal Committee for the review of nintedanib (TA379)[12] and there was no useable data at one year as the trial was stopped early at 36 weeks. In addition, a non-valid primary end point, SpO_2, was used.

A total of eight studies were included in the company's NMA: ASCEND[34] (pirfenidone), CAPACITY 1[49] (pirfenidone), CAPACITY 2[49] (pirfenidone), SP3[38] (pirfenidone), INPULSIS 1[72] (nintedanib), INPULSIS 2[72] (nintedanib), TOMORROW[73] (nintedanib) and PANTHER[74, 75] (NAC and triple therapy). However, not all trials presented outcome data that could contribute to each NMA for all outcomes.

The ERG notes that although not in the final NICE scope,[3] the evidence network includes NAC and triple therapy. The trials of comparators contributing data to the NMA were all placebo-controlled RCTs and therefore all comparisons were made with placebo (see **Error! Reference source not found.**). The ERG therefore believes that PANTHER[74, 75] has little influence on the NMA results for nintedanib and pirfenidone, and therefore data from PANTHER[74, 75] have been excluded from the additional analyses performed by the ERG in Section 4.8. In this section, only data from the trials of relevance to the decision problem are summarised.

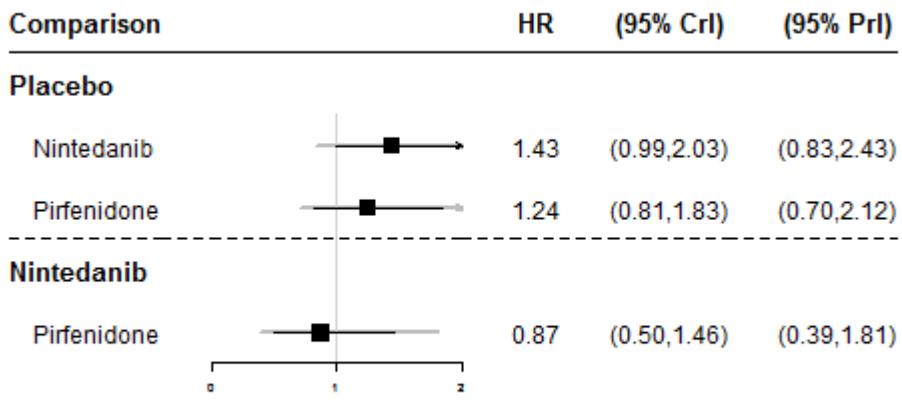
A summary of the design and study characteristics of the studies included in the NMA is provided in **Error! Reference source not found.**.

Table 7: Reported outcomes and definitions adapted from CS,[4] (including response from clarification question A14, and A17 and A32)[10]

Outcome	**ASCEND[34]**	***CAPACITY 1 & 2[49]**	**SP3[38]**	**INPULSIS 1&2[72]**	**TOMORROW[73]**	**PANTHER[74]**
Study duration**	52 weeks	72 weeks	52 weeks	52 weeks	52 weeks	60 weeks (NAC), 32 weeks (Triple therapy)
Lung function						
Change in percent predicted FVC	Yes	Yes	Reported change in % predicted VC	Yes	Yes	Yes (NAC only)
Change from baseline in FVC (L)	Yes	Yes	Reported change from baseline in VC (L)	Yes	Yes	Yes
Categorical decline of ≥ 10% in percent predicted FVC	Yes	Yes	No	Yes	Not clearly defined, therefore excluded	Yes (NAC only)
Survival						
All-cause mortality	Defined as rate of death from any cause	Defined as OS	Number of deaths	Defined as OS	Deaths from any cause	
IPF-related death	Reported as treatment-emergent -IPF-related mortality and defined as deaths occurring between randomisation and within 28 days of last dose of study drug	Reported as IPF-related mortality and defined as deaths occurring between randomisation and within 28 days of last dose of study drug	No	Defined as death from respiratory cause		
PFS	Defined as a confirmed ≥10%	Defined as confirmed ≥10% decline in percent predicted	Defined as VC decline of ≥10% or death.)	No	Excluded as only reported the	Defined as decline of

Treatment effects are estimated as odds ratios (OR), and then converted to relative risks (RR) using the average rate in the placebo arms over all studies in the NMA for use in the cost effectiveness model (clarification response,[10] Appendix D). For the ERG base-case network the average rate of all-cause discontinuation for placebo was 0.17. The estimated treatment effect for nintedanib vs pirfenidone on the odds ratio scale was OR: 1.14 (1/0.87) which equates to a relative risk of RR: 1.11.

Figure 1: All cause discontinuation, ERG base-case network - HR, 95% CrI and 95% PrI



Heterogeneity: between-study variance is 0.13 (95% CrI; 0.03, 0.45)

4.9 Conclusions of the clinical efficacy section

Five RCTs compared pirfenidone at various doses with placebo in adults with mild or moderate IPF: ASCEND (Phase III),[34] CAPACITY 1 & CAPACITY 2 (Phase III),[49] SP3 (Phase III),[38] and SP2 (Phase II).[39] Three trials were international and multicentre (ASCEND and CAPACITY 1 & 2[49]), although only CAPACITY 2[49] included any UK centres (three of 110 centres across both CAPACITY trials). One RCT compared pirfenidone plus N-acetylcysteine (NAC) with placebo plus NAC in Chinese adults with mild or moderate IPF: Huang *et al.* 2015.[48]

Overall, the ERG assessed the potential risk of bias in ASCEND[34] and CAPACITY 1 & 2[49] to be low across most domains, with the exception of reporting bias and "other bias", which were judged to be "moderate" on account of inconsistencies between some outcomes and analyses presented in the trial protocols, those presented in published manuscripts and those reported in the CS,[4] and the possible influence of uncontrolled variables such as rate of disease progression.

The SP3,[38] SP2[39] and Huang *et al.* (2015) trials[48] were at a higher or more unclear risk of bias across many domains than the ASCEND[34] and CAPACITY[49] trials. These trials all evaluate lower doses of

pirfenidone which are licensed in Japan but not in the UK, apply different eligibility criteria and present noticeable differences from the other three trials in some baseline characteristics of participants.

The final selection of three trials (ASCEND,[34] CAPACITY 1 and CAPACITY 2[49]) for the main clinical efficacy review was considered to be appropriate by the ERG. However, there are some between-trial differences across some baseline characteristics, such as mean FVC or 6MWD at baseline, but subgroup analyses suggested that these and other variables did not influence treatment effect. A *post hoc* pooled analysis of ASCEND[34] and CAPACITY 1 & 2[49] found no evidence of interaction between treatment for those patients with baseline FVC ≥ 80% predicted and those with FVC < 80% predicted.

The CS[4] reported three measures of lung function based on FVC: change from baseline in percent predicted FVC/VC; change from baseline in FVC/VC (ml); and relative proportions in each trial arm with FVC categorical decline of ≥10% percent predicted (this latter outcome measure included "death" in some analyses). The findings were not consistently statistically significant across trials for these outcome measures: ASCEND (52 weeks) [34] and CAPACITY 2 (72 weeks) [49] found statistically significant benefits for those on pirfenidone compared with those on placebo for mean change from baseline in percent predicted FVC (mean difference 4.78%; $p<0.001$ and mean difference 4.4%; relative difference 35.3%; 95% CI 0·7 to 9·1 $p=0.001$, respectively); but CAPACITY 1[49] found no statistically significant benefit for those on pirfenidone compared with those on placebo (absolute difference: 0.6%; relative difference: 6.5%; 95% CI -3·5 to 4·7 $p=0.501$). Pooled analyses of the CAPACITY trials[49] found statistically significant benefits for those on pirfenidone compared with those on placebo (absolute difference: 2.5%; relative difference: 22.8%; $p=0.005$). SP3[38] also reported statistically significant benefits for those on pirfenidone for change from baseline in percent predicted VC at 52 weeks ($p=0.044$); and change from baseline in VC (ml) ($p=0.042$). Huang *et al.* (2015)[48] reported a statistically significant mean change in FVC from baseline in favour of pirfenidone plus NAC compared with placebo plus NAC at 24 weeks ($p=0.02$) but not at 48 weeks ($p=0.11$). Meta-analyses of change in percent predicted FVC for CAPACITY 1 & 2[49] and ASCEND[34] and change in percent predicted VC for SP3,[38] suggested that pirfenidone reduces the decline in percentage predicted FVC compared with placebo up to 52 weeks (MD: 3.4, 95% CI: 1.87 to 4.94, *p*-value not reported). The meta-analysis also suggested that pirfenidone slows the rate of decline in FVC (MD: 0.12, 95% CI: 0.05 to 0.19, *p*-value not reported) up to 52 weeks.

In terms of decline in FVC by ≥10%, or death, ASCEND[34] reported a statistically significant difference in favour of pirfenidone compared with placebo at week 52 (absolute difference: 15.3 [95% CI not reported], $p<0.001$). For CAPACITY 1[49] the treatment effect at week 72 favoured pirfenidone

Limitations

The ERG notes that the main limitations of the company's meta-analysis relate to the following:

- Combining the 48-week outcome data from the CAPACITY trials[49] with the 52 week data from ASCEND[34] and SP3 trials.[38] Although the direction of effect for all analysed outcomes were the same for the 52 week and 72 week data, the magnitude of effect of pirfenidone was generally less at 72 weeks than 52 weeks.
- Inclusion of the SP3 trial[38] to assess the following outcomes: lung capacity (FVC/VC percentage predicted, FVC/VC (L)); PFS; acute exacerbation; and serious adverse events. SP3[38] used a lower dose (1,800mg/day) of pirfenidone, which is licensed in Japan but not in the UK, and included only Japanese patients. In contrast, the CAPACITY 1 & 2[49] and ASCEND[34] studies used licence doses of pirfenidone (2,403mg/day) and included people from Europe and the USA.
- Variation in outcome definitions used across the included trials for PFS, acute exacerbation, 6MWT, lung function and combining data of FVC with VC for lung function.

The NMA included trials were of different durations. CAPACITY 1 and 2[49] presented data at 72 weeks whilst the maximum follow up for the other studies (of interventions relevant to the scope) was at 52 weeks. Trials with a shorter follow-up might be expected to observe fewer negative outcomes and so in order to facilitate synthesis across trials, the NMA used data from CAPACITY 1 and 2[49] evaluated at an earlier follow up time of either 48 or 52 weeks (depending on the outcome). This is a valid approach for evaluating the treatment effects at a specific time point but means that the analyses did not make use of the full follow-up data available. Alternative methods that allow the incorporation of trials of different durations, whilst accounting for time effects, could have been used.

For time-to-event outcomes (all-cause mortality, PFS, IPF related mortality) the treatment effects are reported as HRs, which are time averaged estimates of treatment effect and under the assumption of proportional hazards should be constant over time. The CS[4] provided evidence to support the assumption of proportional hazards but, despite this, data at 52 weeks were used in the company's base-case NMAs rather than the full 72-week data. Although there is not enough evidence to reject the assumption of proportional hazards for the presented pirfenidone data, the ERG notes that treatment effects at 72 weeks were often substantially lower than those at 52 weeks. The company[4] reported that there was no evidence to support that proportional hazards hold for nintedanib in the long-term.

The company also described other potential sources of heterogeneity between trials, in terms of differences in outcome definitions and handling of missing data. Due to the limited number of studies

following price reduction.[92] Pirfenidone was dominated by nintedanib in the CDR for nintedanib (assumption of equal efficacy but nintedanib was less costly).[91] Finally, Loveman *et al.* (2014) reported that, at the list price, pirfenidone was dominated by inhaled NAC.[93, 94]

Quality assessment tables are presented in CS Appendix 19. Following quality assessment, the company reports that "*the CDRs provide only a brief summary of the cost effectiveness results and therefore score poorly against most areas of the Drummond quality assessment check list*" (see CS page 194) and have limited relevance to the UK. The ERG considers this to be justified but raises attention to particular comments expressed during these assessments [91, 92] that are relevant for this appraisal including: (a) the uncertainty around the duration of the treatment effect for pirfenidone and nintedanib against BSC; (b) the uncertainty around the relative effectiveness between pirfenidone and nintedanib, and; (c) concerns regarding the discontinuation rate and the assumption that the treatment effect remains following discontinuation.

The CS does not report results from the quality assessment for the previous model submitted to NICE[42] but does summarise some of the concerns expressed by the ERG[89] including the appropriateness of the model structure, comparators included and uncertainty around the clinical effectiveness of pirfenidone versus BSC. In Appendix 19 of the CS, the ERG observes that according to the company, the model that was previously submitted to NICE performed poorly against most areas of the Drummond quality assessment checklist[90] (did not conform to 17 criteria, conformed to 15 criteria and 4 criteria were non-applicable).

Finally, the company considered the Loveman study[93, 94] to be of high quality when assessed against the Drummond quality assessment checklist but that the relevance to the UK is limited given: (a) the study did not include data from the ASCEND and INPULSIS trials; (b) the inclusion of a trial in severe IPF; (c) utility values were taken from a non-UK source; (d) efficacy data were taken from studies outside the UK, and; (e) "*for pirfenidone the data were taken from two Japanese studies and two multi-national studies (of which the UK was one country).*" The ERG notes that whilst the company appears to suggest that the inclusion of Japanese studies is a limitation in its systematic review, as described in Section 4.6, despite a request from the ERG, the company refused to exclude Japanese studies from the NMA.

5.1.4 Conclusions of the review of published cost-effectiveness studies

The CS draws some conclusions regarding the quality of the included studies, comments on the applicability of the studies to the decision problem for this appraisal and tabulates the ICERs reported. Whilst the ERG is generally satisfied with the cost-effectiveness review presented by the company, the ERG considers the decision to exclude the model used for the nintedanib submission[26] from the cost-effectiveness review to be questionable. The ERG observes that the population entering the model resembles the population included in the INPULSIS and TOMORROW trials which consisted of people with a percent predicted FVC >50% at baseline and therefore consists of people considered to have mild to moderate IPF which is relevant for this submission. The ERG further notes that whilst people included in the nintedanib trials had milder disease compared with the population included in the pirfenidone trials (approximately 45% had a FVC >80% compared with approximately 25% in the pirfenidone trials), an analysis is conducted for an ASCEND-like population (defined as FVC 50-90% predicted, $FEV_1/FVC \geq 0.8$).[12, 26] The ERG considers that this study should have been included in the company's systematic review in addition to the original CDR for pirfenidone for consistency. The nintedanib model uses a cohort state transition approach whereby people entering the model progress through a series of health states defined by roughly 10 point percent predicted FVC intervals. EQ-5D scores were taken directly from the INPULSIS trials. In this assessment, pirfenidone was dominated by nintedanib when the stopping rule was applied to both or none of the interventions in people with a percent predicted FVC <80% at baseline (including the price discount for both interventions).

The ERG further notes that the CS does not provide any conclusions regarding the cost-effectiveness of pirfenidone compared with BSC or nintedanib based on this review of published cost-effectiveness analyses.

In summary, the ERG notes some inconsistencies in the company's review and considers that it is challenging to compare results from the different models given the differences in model structure, assumptions, data used and the existence of confidential price discounts.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

This section presents a summary description of the model submitted as part of the CS. ERG comments are provided directly after each aspect of the model is described.

5.2.1. Consistency of the CS with the requirements set out in the NICE reference case

As part of its submission to NICE, the company submitted a fully executable health economic model programmed in Microsoft Excel alongside a written description of the methods and results. A revised version of the model was submitted in response to the clarification questions from the ERG. The

Evidence on resource use and cost	Resource use estimates for routine management are based on telephone discussion with UK clinical experts. Hospitalisation data are based on estimates from pirfenidone trials. Unit costs are taken from NHS reference costs. Drug costs in the main CS are based on list prices (results which incorporated the PAS for nintedanib are reported in a confidential appendix). Costs of end of life care were taken from the literature.
Time horizon	Lifetime
Discount rate	3.5% per year for both costs and QALYs
Equality considerations	No weighting has been applied to QALYs
BSC – best supportive care; ITT – intention to treat; FVC – Forced vital capacity; QALY – quality-adjusted life year; IPF- idiopathic pulmonary fibrosis *[a] defined in the trial as symptom relief, pulmonary rehabilitation, management of comorbidities, and end of life care, including oxygen therapy*	

The population entering the company's model reflects the population included in the CAPACITY[49] and ASCEND trials.[34] Similarly, the intervention and associated treatment regimen assumed in the economic model reflects the regimens used in the Phase III trials.[34,49] The intervention consists of pirfenidone (267mg capsules, given orally), given as three 267mg capsules, three times a day, giving a total of 2403mg/day; before adjustments for dose reductions and interruptions. In the company's base-case, people initiating pirfenidone are assumed to discontinue treatment at the rate observed in the CAPACITY 1 & 2, ASCEND and RECAP studies; these discontinuation rates are not adjusted to reflect the implementation of the stopping rule in the base-case. The stopping rule defined by NICE which formed the basis for the positive recommendation for pirfenidone[2] and nintedanib[12] is however applied to nintedanib in the company's base-case and only in a scenario analysis for pirfenidone.

5.2.1.1. ERG comments on the population described in the CS and included in the company's model

The ERG is satisfied that the population and subgroups addressed by the company are largely in line with the final NICE scope.[3] In the CAPACITY/ASCEND trials,[34,49] which formed the main basis of the evidence used in the economic model, individuals were eligible if they had a percent predicted FVC ≥50% and predicted diffusing capacity of the lungs for carbon monoxide (DLco) ≥ 35% (≥ 30% in the ASCEND trial). This is largely in line with the definition provided by NICE in the final scope[3] for mild-to-moderate IPF; defined as "*a FVC greater than or equal to 50% predicted and a diffusing*

as separate health states, but are instead assumed to be treatment-specific and are applied within each model cycle.

QALYs are calculated as a function of time spent in the pre-/post-progression states with different utilities applied in each state. Cost components include drug acquisition, costs associated with the management of the condition, adverse events, acute exacerbation and end of life.

It should be noted that within its submission, the company makes reference to three modelling approaches that have been used in IPF: (i) the micro-simulation model submitted during the first appraisal of pirfenidone[2] (submitted by InterMune); (ii) the state transition approach based on percent predicted FVC categories submitted as part of the nintedanib NICE appraisal,[12,26] and; (iii) the state transition approach published by Loveman *et al.* (2014)[93,94] which is based on four main health states (unprogressed IPF, progressed IPF, lung-transplant and dead). The company considers that the micro-simulation approach used in the previous NICE submission[97] and the approach used in the nintedanib NICE appraisal[26] add complexity and are difficult to parameterise and therefore are not appropriate.

5.2.2.1. ERG's comments on conceptual representation of the condition

The ERG has a number of concerns regarding the structure and logic of the company's model. These can be separated into four sets of issues: (i) the conceptual representation of the condition; (ii) the representation of the treatment pathway in IPF; (iii) the use of a partitioned survival model approach and HR, and; (iv) questionable structural assumptions.

The ERG considers that the company's model ignores a key facet of the disease: specifically that IPF is a progressive condition characterised by irreversible loss of lung function. The company's justification to use PFS in the model relies on three key sets of arguments: (i) findings from a review by Albera et al[99] which concluded that PFS could be deemed to be an appropriate endpoint in IPF trials; (ii) that this approach has been used in a previous economic evaluation,[94] and; (iii) the difficulty in parameterising a model based on percent predicted FVC (as used in the nintedanib appraisal[12,26]).

The ERG considers that whilst PFS could be considered as an appropriate endpoint in trials when evaluating the effect of an intervention in IPF, separating the natural history of IPF into two distinct consecutive phases (the presence/absence of progression) is overly simplistic and does not reflect the natural history of the condition or its progressive nature. This limitation is recognised in the CS (page 278) when results are compared against those generated during the original submission to NICE.[2] The company states that "*the impact on patient quality of life has been conservatively included for one progression alone in the updated model*" (see CS,[4] page 278). The CS therefore acknowledges that this simplification has the potential to bias the QALY gains estimated by the model. However,

- (**Error! Reference source not found.**). A parallel plot of the log-cumulative hazards for BSC and pirfenidone would suggest that the assumption of proportional hazards is reasonable within the trial period. Upon inspection of **Error! Reference source not found.**, this assumption is questionable.
- Finally, the ERG advises considerable caution in the interpretation of any comparisons made by the company between the pirfenidone arm of the CAPACITY/ASCEND/RECAP trials[34, 40, 49] and data from registries. The ERG considers that such analyses are inherently subject to considerable bias. In brief:
  a. Despite the attempt by the company to select and match individuals from registries to people enrolled in the ASCEND and CAPACITY trials,[34, 49] the survival of individuals from the registries is inconsistent with the OS of people initiating BSC observed in the clinical trials (see **Error! Reference source not found.**). The ERG notes that the company does not comment on the discrepancies between the OS in people enrolled in the CAPACITY/ASCEND trials[34, 49] and people enrolled in registries who were treated with BSC.
  b. The long-term survival for pirfenidone is based on the RECAP trial (OLE study of ongoing pirfenidone treatment) which enrolled people with IPF who completed the final follow-up visit of the CAPACITY-trials and received ≥80% of the assigned study treatment. Clarifications were requested from the company regarding the rationale for excluding people from RECAP who received less than 80% of the assigned study treatment (see clarification response,[10] question B2). In response, the company stated that "*Patients using less than 80% of drug are considered to be non-compliant (a standard cut-off being used in many trials), and for this reason were not included in RECAP. Although RECAP was an open-label extension study, the standard compliance considerations were still applied.*" Consequently, the ERG considers that the exclusion of people who received less than 80% of the assigned study treatment could overestimate the survival for pirfenidone as only people that are considered to be compliant have been included in RECAP, thereby making comparison with long-term registries less relevant.
  c. Finally, whilst individuals from the registries were matched to people included in the clinical trials, the ERG notes some potential discrepancies in the inclusion criteria applied to the registry data which may bias the estimate in favour of pirfenidone. For instance, the company excluded individuals with a percent predicted FVC ≥90% (if DLco≥90%).Throughout the CS, the company discuss a potential link between FVC and mortality; thus, excluding people with a percent predicted FVC≥90% could underestimate the survival in individuals included in the registries. However, this

exclusion criteria only resulted in the exclusion of 1 patient from INOVA and 1 patient from Euro IPF, so any bias introduced is likely to be minimal.

particularly given the similarities between health states between the Loveman et al[93, 94] model and the company's model. Therefore, the ERG cannot be certain that all relevant resource use data have been identified and presented in the CS.

The ERG is generally satisfied with the inclusion of drug acquisition costs in the company's model but notes following clarification that; dose interruptions and reductions for pirfenidone are calculated after titration and therefore exclude the first 2 weeks. The ERG considers that a more appropriate approach would have been to separate the costs for the first model cycle from those for subsequent cycles. This is amended in the ERG preferred-base-case.

The ERG notes that the daily cost of pirfenidone and nintedanib is equivalent when assuming the full indicated dose is taken (after the titration period for pirfenidone) and when using the current list price.[16] However, assuming the same daily costs for pirfenidone and nintedanib based on the average dose used in the pirfenidone trials implies the same impact of dose reductions/interruptions for pirfenidone and nintedanib. The ERG notes that the price structure for pirfenidone and nintedanib is different and that a dose reduction with nintedanib (for instance, from 150mg to 100 mg) would not be associated with a reduction in costs. The ERG observes that the INPULSIS trial[113] reported a compliance with nintedanib of 96.4 % whereas the mean dose applied in the model for pirfenidone is 87.6% of the indicated dose. Therefore, the ERG considers that assuming the same cost for pirfenidone and nintedanib is likely to favour nintedanib.

The company's base-case assumes no drug acquisition costs for BSC and/or concomitant medications. The ERG considers this to be inappropriate as within the trials, individuals received concomitant medications as part of BSC. This was included in the nintedanib submission at a cost of approximately £25 per model cycle calculated from the trial for both nintedanib and BSC.[26] However, the ERG notes that the impact of the ICER is likely to be minimal given that the cost will be applied to all arms.

The CS also reports that resources use estimates were derived from discussion with a panel of clinicians, although no details were provided in the CS. In response to a request for clarification (clarification response,[10] question B16), the company provided further details, stating that: "*One-to-one telephone interviews were conducted with the panel of UK clinical experts. Content of the earlier NICE manufacturer submission was discussed, along with how the approach employed to assess resource use in the earlier submission matched current clinical practice in IPF. Discussions accounted for the revised descriptions of the NHS Reference Cost list for 2014-15 compared to earlier years (e.g. revision of 'simple lung exercise function test' to 'field exercise test')."* Despite this additional clarification, the ERG considers the process used by the company to elicit resource use has

5.2.13. Model validation

The company reports two main methods of model validation:

- Comparison of the model predictions with results from previous evaluations,
- Validation of the long-term prediction of survival.

The CS provides a comparison of the model outcomes from its model with those from the company's submission, in the nintedanib appraisal,[26] and the company's submission in the previous appraisal of pirfenidone[2] (see Table 8).

Table 8: Comparison of LYs and QALYs – moderate population (reproduced from CS,[4] Table 122)

Outcome	NTB submission[26]			This submission			TA282	
	BSC	NTB	PFN	BSC	NTB	PFN	BSC	PFN
Total QALYs	3.27	3.67	3.62	3.15	3.77	4.46	3.18	4.30
Total LYs	4.36	4.86	4.86	4.33	5.30	6.47	4.40	5.96

Key: IPF, idiopathic pulmonary fibrosis; LY, life year; NTB, nintedanib; PFN, pirfenidone; QALY, quality-adjusted life year

The CS also provides a comparison of OS from their model compared with two studies (see **Error! Reference source not found.**) which uses observational data (both sources are described further in Table 59 of the CS). Fisher al (2015)[64] reports OS from a modelling study whereby the OS in patients initiating BSC is modelled from a log-normal distribution which is fitted to data from the US strand registry. The OS in patients initiating pirfenidone is modelled from a log-normal distribution which is fitted to data from the RECAP trial. The Roskell *et al.* study[66] is also a modelling study and uses data from the RECAP OLE for pirfenidone (Weibull distribution fitted to the KM). The survival in patients initiating BSC was taken the CPRD and included patients with a FVC > 50% only. A Weibull distribution was fitted to the CPRD data.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

This section summarises additional analyses undertaken by the ERG using the company's model as well as the development of an ERG-preferred base-case.

The ERG expressed a number of concerns regarding the model structure and parameterisation of the company's model. A key concern related to the lack of ability of the model to capture the progressive nature of IPF and inflexibility associated with the modelling approach chosen by the company (partitioned survival model) which meant that correlations between outcomes are not captured in the model. This is a concern as the modelled stopping rule impacts on costs but not health outcomes. Importantly, the company's model also relies on a strong assumption that the treatment effect estimated within the trials (up to 52 weeks) is maintained over the entire model's duration (34 years). Such extrapolation is questionable and subject to considerable uncertainty. This leads to discrepancies between the model-predicted OS and observed OS in people initiating BSC from the ASCEND/CAPACITY trial (see **Error! Reference source not found.**).

Unfortunately, a number of the issues identified cannot be addressed by the ERG without major restructuring of the economic model. It should also be noted that changes to the model are challenging given the structure of the model whereby outcomes are disconnected from each other. The ERG is not able to adequately amend the implementation of the stopping rule within the company's existing model structure and thus, considers that any ICER generated in the scenarios using the stopping rule need to be interpreted with caution as they are likely to provide ICERs that are favourable to pirfenidone when compared against BSC.

The following analyses were undertaken by the ERG to inform its base-case:

1. **Using the ERG's preferred estimate of the treatment effect, which uses data up to 72 weeks, excludes SP3, and uses the CODA samples from the predictive distribution.** As described in Section **Error! Reference source not found.**, the ERG considered the treatment effect estimated at 72 weeks to be more appropriate and more consistent with the company's assumption of proportional hazards. Furthermore, the ERG considered that SP3 should be removed from the network as this trial was conducted in a Japanese population, a dose licensed in Japan but not in the UK was given and the HR was not directly available which could introduce a bias. Finally, the ERG considered that the CODA samples (from the predictive distribution) should be used instead of the median HR in order to properly capture the joint uncertainty in the effectiveness estimates, and therefore the results for this scenario are run probabilistically.

2. **Use of the Gompertz distribution for OS (rather than the Weibull).** As described in Section **Error! Reference source not found.**, the ERG considered the Gompertz distribution to provide a more plausible long-term extrapolation compared with the Weibull distribution.

7. those reported by the company in Table 19 of the response to clarification (see addendum to clarification response).
8. **Using compliance from INPULSIS for nintedanib.** Given the different price structure, the ERG considered that assuming the same impact of dose reductions/interruptions between pirfenidone and nintedanib is likely to be unfavourable to pirfenidone. Consequently, an analysis is conducted assuming a compliance of 96.4% for nintedanib based on data from the INPULSIS trial.[113]
9. **Corrections of errors in the economic model.** As part of the critical appraisal of the model, the ERG identified a series of minor programming errors which have been corrected. These are described in appendix 4.

The impact of each individual change is reported in Section 6 in addition to the ERG-preferred base-case which combines all these changes. For consistency, results are reported with and without the stopping rule (same assumption for both treatments). It should also be noted that the ERG-preferred base-case is presented as a range (most optimistic to most pessimistic scenario) given the uncertainty surrounding the extrapolation of the treatment effect.

5.4 Conclusions of the cost-effectiveness section

The company submitted a fully executable economic model as part of their submission to NICE. The analysis was undertaken from the perspective of the UK NHS and PSS over a lifetime horizon. The company's analysis is presented for three populations: (1) the ITT trial population, which is comprised of adults with mild to moderate IPF; (2) people with a percent predicted FVC > 80% at baseline (considered to be mild IPF), and; (3) people with a percent predicted FVC of 50 - 80% at baseline (considered to be moderate IPF). All three analyses include BSC as a comparator (defined in the trial as symptom relief, pulmonary rehabilitation, management of comorbidities, and end of life care, including oxygen therapy). Nintedanib is included as a comparator only in the analysis of people with a percent predicted FVC of 50 - 80% at baseline.

The analysis in the ITT population does not include nintedanib as a comparator as nintedanib is only a valid comparator for the subgroup of the ITT population with moderate IPF (percent predicted FVC of 50 - 80%). The ERG considers that it is more appropriate to conduct an economic analysis separately within the mild and moderate subgroups as the comparators vary by subgroup.

In the company's base-case, people initiating pirfenidone are assumed to discontinue treatment at the rate observed in the trials, hence no stopping rule is applied. The stopping rule defined by NICE which formed the basis for the positive recommendation for pirfenidone[2] and nintedanib[12] is however

Table 9: Summary of the impact of individual changes to the ICER for pirfenidone versus BSC[a] using the list price and mean parameter inputs (deterministic model)

	ITT population	People with a percent predicted FVC of 50 – 80%[a]	People with a percent predicted FVC > 80%
Company base-case	[redacted]	[redacted]	[redacted]
No stopping rule for nintedanib	[redacted]	[redacted]	[redacted]
Inclusion of stopping rule for pirfenidone	[redacted]	[redacted]	[redacted]
Treatment effect assumed to stop after 2 years	[redacted]	[redacted]	[redacted]
Gompertz distribution for OS	[redacted]	[redacted]	[redacted]
HRQoL capped at 1.0	[redacted]	[redacted]	[redacted]
Adjustment of HRQoL by age	[redacted]	[redacted]	[redacted]
End of life costs applied to death irrespective of causes	[redacted]	[redacted]	[redacted]
Pirfenidone dose titration	[redacted]	[redacted]	[redacted]
Nintedanib compliance taken from INPULSIS	[redacted]	[redacted]	[redacted]
Correction of errors	[redacted]	[redacted]	[redacted]

[redacted]

Addendum to Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282): A Single Technology Appraisal

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Sarah Davis, Senior Lecturer in Health Economics, ScHARR, University of Sheffield, Sheffield, UK Rachid Rafia, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Jean Sanderson, Research Associate, ScHARR, University of Sheffield, UK
Correspondence Author	Sarah Davis, Senior Lecturer in Health Economics, ScHARR, University of Sheffield, Sheffield, UK
Date completed	Addendum completed (28/04/2016)

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 142/06/02.

Declared competing interests of the authors

None of the authors have any conflicts of interest to declare.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Davis S, Rafia R, Sanderson J. Addendum to Pirfenidone for treating idiopathic pulmonary fibrosis: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2016.

Contributions of authors

Jean Sanderson conducted the network meta-analysis. Rachid Rafia conducted the cost-effectiveness analysis. All three authors contributed to the drafting of this addendum.

Abbreviations

BSC	Best supportive care
CS	Company's submission
ERG	Evidence Review Group
FVC	Forced vital capacity
ICER	Incremental cost-effectiveness ratio
ITT	Intention to treat
NMA	Network meta-analysis
QALY	Quality-adjusted life year

CONTENTS

TABLES

FIGURES

1 INTRODUCTION

Prior to the first Appraisal Committee meeting, NICE requested that the Evidence Review Group (ERG) provide cost-effectiveness results for two additional scenarios.

Firstly, NICE requested that the ERG provide results for a scenario in which the SP3 study was included in the network meta-analyses (NMAs) that inform the cost-effectiveness model. Methods for this scenario are reported in section 2.1, and results for each of the ERG base-scenarios are provided in section 3.1 for the intention-to-treat (ITT) population, the mild subgroup (percent predicted FVC >80%) and the moderate subgroup (percent predicted FVC of 50% to 80%).

Secondly, NICE requested that the ERG provide results with the stopping rule implemented for nintedanib, but not for pirfenidone. Methods for this scenario are provided in section 2.2 and results for the moderate subgroup (percent predicted FVC of 50 to 80%) are provided in section 3.1. Results for the mild subgroup and the intention to treat population are not required as results for pirfenidone when excluding the stopping rule for pirfenidone have already been provided within the main report, and nintedanib is not a comparator in the mild subgroup or the ITT population.

2 METHODS

2.1 Scenario including SP3 within the NMA

The NMA was re-run with SP3 added to the studies included in the ERG base-case network, which has been previously described in Table 38 of the main report. Revised results were generated for the outcomes that inform the model; overall survival, progression-free survival, acute exacerbation and all-cause discontinuation. Results of the NMA for this scenario are presented in section 3.2.1.

To generate incremental cost-effectiveness ratios (ICERs) for this scenario, the ERG used the ERG adapted model which was previously used to generate the ERG-preferred base-case ICERs presented in section 6.1.2 of the main report, but updated the efficacy estimates to reflect outputs of the NMA when including SP3. The ICERs for this scenario are presented in section 3.2.2.

2.2 Scenario with the stopping rule implemented for nintedanib only

The implantation of the stopping rule for nintedanib, but not for pirfenidone was presented in the company submission (CS) as the company's base-case scenario and so the ERG was able to use the options provided within the company's model to implement this scenario. The ERG used the version of the ERG adapted model which was previously used to generate the ERG-preferred base-case

ICERs presented in section 6.1.2 of the main report, with the only change being the selection of 'Ofev only' in cell E90 of the 'Model inputs' sheet which controls implementation of the stopping rule. The ERG notes their previous concerns regarding the implementation of the stopping rule within the company model which are described in detail section 5.2.2.2 of the main report. Briefly, the stopping rule, as implemented in the company model, restricts the life-time cost of treatment but does not have any impact on clinical outcomes and therefore the ERG considers that the lifetime benefits are likely to be overestimated when the stopping rule is implemented.

3 RESULTS

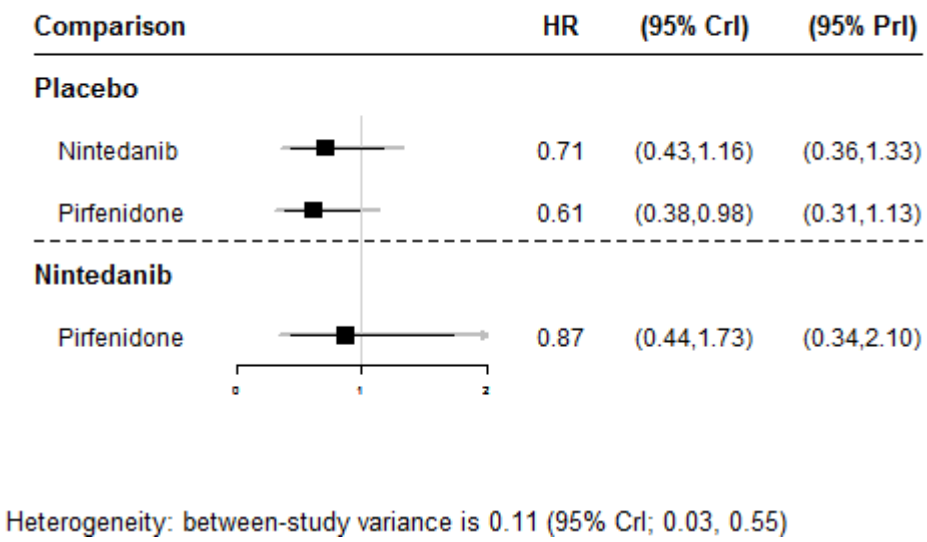
3.1 Scenario including SP3 within the NMA

3.1.1 NMA results with SP3 included

All-cause mortality

On addition of SP3, seven trials were included in the network for all-cause mortality. Two of these trials (SP3 and TOMORROW) reported results as the proportion of deaths, rather than the HR. The treatment effects are summarised in Figure 1.

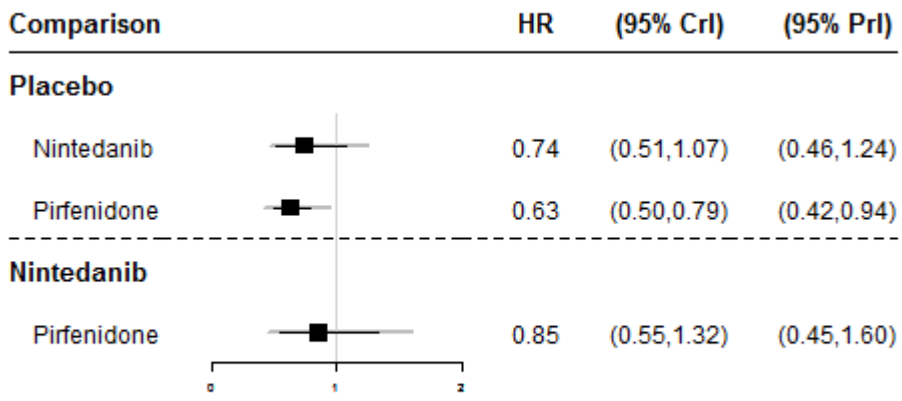
Figure 1: All-cause mortality, ERG base-case network plus SP3 - HR



PFS

On addition of SP3, six trials were included in the network for PFS, but a pooled HR was used for the INPULSIS trials since the individual study-level treatment effects were not available. All trials provided results as HR. The results of the NMA are summarised in Figure 2.

Figure 2: PFS, ERG base-case network plus SP3 - HR

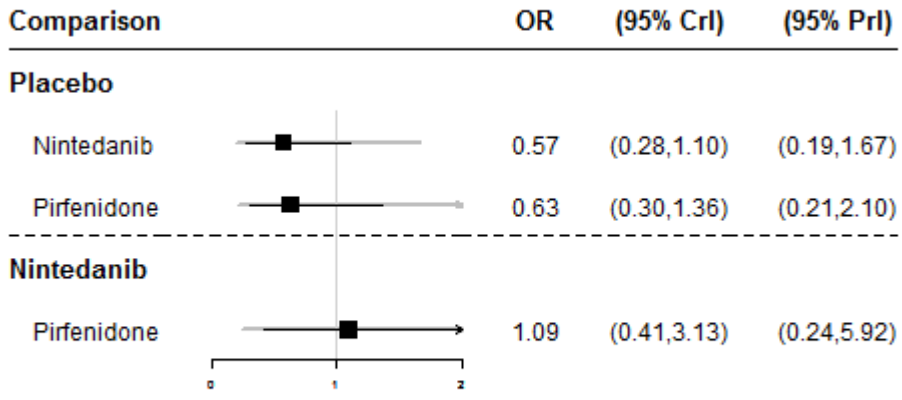


Heterogeneity: between-study variance is 0.09 (95% CrI; 0.02, 0.45)

Exacerbations

On addition of SP3, seven trials were included in the network for acute exacerbations. The pooled treatment effects are summarised in Figure 3.

Figure 3: Acute exacerbations, ERG base-case network plus SP3 - OR



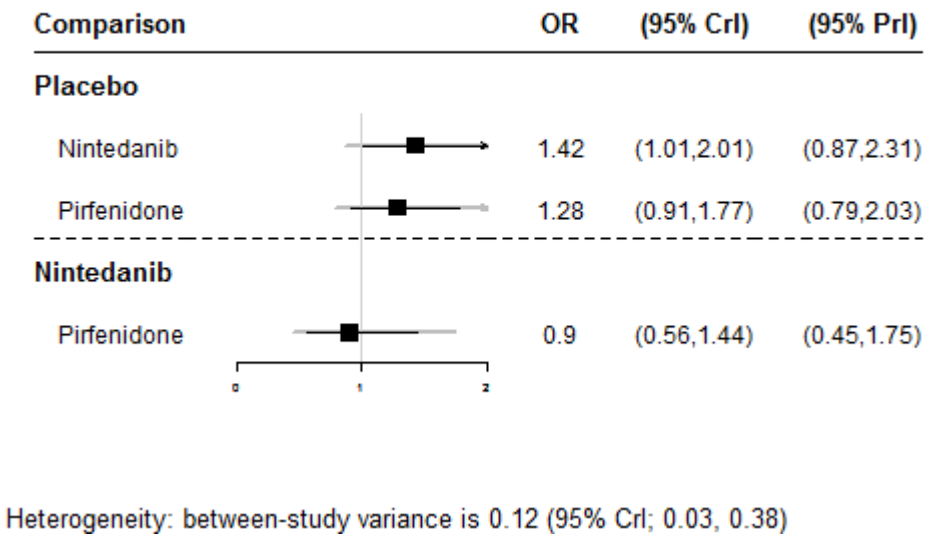
Heterogeneity: between-study variance is 0.26 (95% CrI; 0.04, 1.05)

All cause discontinuation

On addition of SP3, seven trials were included in the network for acute exacerbations. The pooled treatment effects (OR) are summarised in Figure 4.

Treatment effects are estimated as odds ratios (OR), and then converted to relative risks (RR) using the average rate in the placebo arms over all studies in the NMA for use in the cost effectiveness model (clarification response,[10] Appendix D). For the ERG base-case network plus SP3, the average rate of all-cause discontinuation for placebo was 0.19. The estimated treatment effect for nintedanib vs pirfenidone on the odds ratio scale was OR: 1.11 (1/0.90) which equates to a relative risk of RR: 1.08.

Figure 4: All cause discontinuation, ERG base-case network plus SP3 - OR



Heterogeneity: between-study variance is 0.12 (95% CrI; 0.03, 0.38)

3.1.2. Cost-effectiveness results with SP3 included

The results when including SP3 in the NMA are presented in Table 1 to Table 4 when using the list price and in Table 5 to Table 8 when incorporating the PAS for pirfenidone. The results when incorporating the PAS for nintedanib are provided in the confidential appendix.

For the analyses incorporating the PAS, the ICERs versus best supportive care (BSC) for the ITT population vary from £25,365 to £106,584, the ICERs vs BSC for the moderate population vary from £25,603 to £96,662 and the ICERs vs BSC for the mild population vary from £29,607 to £170,279. The addition of SP3 to the network appears to have had minimal impact on the cost-effectiveness estimates. The ICERs for pirfenidone versus BSC when incorporating the PAS have reduced by between 5 and 10% depending on the scenario.

Table 1: **ERG-preferred base-case assuming no stopping rule with pessimistic assumption regarding duration of treatment effect (2 years) – probabilistic list price – include SP3**

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	[redacted]	4.596			
Pirfenidone	[redacted]	4.926	[redacted]	0.330	[redacted]
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	[redacted]	4.208			
Nintedanib	[redacted]	4.397	[redacted]	0.189	[redacted]
Pirfenidone	[redacted]	4.540	[redacted]	0.332	[redacted]
Mild population (percent predicted FVC >80%)					
BSC	[redacted]	5.484			
Pirfenidone	[redacted]	5.754	[redacted]	0.270	[redacted]

Table 2: **ERG-preferred base-case assuming no stopping rule with optimistic assumption regarding duration of treatment effect (life-time effect) – probabilistic list price – include SP3**

	Costs	**QALYs**	**Inc Costs (vs. BSC)**	**Inc QALYs (vs. BSC)**	**ICER (vs. BSC)**
ITT population					
BSC	[redacted]	3.872			
Pirfenidone	[redacted]	4.926	[redacted]	1.054	[redacted]
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	[redacted]	3.557			
Nintedanib	[redacted]	4.266	[redacted]	0.708	[redacted]
Pirfenidone	[redacted]	4.538	[redacted]	0.980	[redacted]
Mild population (percent predicted FVC >80%)					
BSC	[redacted]	4.657			
Pirfenidone	[redacted]	5.747	[redacted]	1.090	[redacted]

Table 3: **ERG-preferred base-case assuming stopping rule with pessimistic assumption regarding duration of treatment effect (2 years) – probabilistic list price – include SP3**

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	[redacted]	4.599			
Pirfenidone	[redacted]	4.929	[redacted]	0.330	[redacted]
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	[redacted]	4.206			
Nintedanib	[redacted]	4.401	[redacted]	0.194	[redacted]
pirfenidone	[redacted]	4.539	[redacted]	0.332	[redacted]
Mild population (percent predicted FVC >80%)					
BSC	[redacted]	5.494			
Pirfenidone	[redacted]	5.761	[redacted]	0.267	[redacted]

Table 4: **ERG-preferred base-case assuming stopping rule with optimistic assumption regarding duration of treatment effect (life-time effect) – probabilistic list price – include SP3**

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	[redacted]	3.878			
Pirfenidone	[redacted]	4.939	[redacted]	1.060	[redacted]
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	[redacted]	3.555			
Nintedanib	[redacted]	4.267	[redacted]	0.711	[redacted]
pirfenidone	[redacted]	4.537	[redacted]	0.981	[redacted]
Mild population (percent predicted FVC >80%)					
BSC	[redacted]	4.671			
Pirfenidone	[redacted]	5.764	[redacted]	1.092	[redacted]

Table 5: **ERG-preferred base-case assuming no stopping rule with pessimistic assumption regarding duration of treatment effect (2 years) – probabilistic incorporating PAS – include SP3**

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£34,246	4.596			
Pirfenidone	£69,444	4.926	£35,198	0.330	£106,584
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£32,290	4.209			
Nintedanib	See confidential appendix				
Pirfenidone	£64,404	4.541	£32,114	0.332	£96,662
Mild population (percent predicted FVC >80%)					
BSC	£39,023	5.483			
Pirfenidone	£84,872	5.752	£45,849	0.269	£170,279

Table 6: **ERG-preferred base-case assuming no stopping rule with optimistic assumption regarding duration of treatment effect (life-time effect) – probabilistic incorporating PAS – include SP3**

	Costs	**QALYs**	**Inc Costs (vs. BSC)**	**Inc QALYs (vs. BSC)**	**ICER (vs. BSC)**
ITT population					
BSC	£30,456	3.868			
Pirfenidone	£69,501	4.922	£39,046	1.054	£37,033
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£28,810	3.555			
Nintedanib	See confidential appendix				
pirfenidone	£64,353	4.536	£35,542	0.981	£36,230
Mild population (percent predicted FVC >80%)					
BSC	£34,592	4.672			
Pirfenidone	£84,885	5.767	£50,293	1.095	£45,921

Table 7: ERG-preferred base-case assuming stopping rule with pessimistic assumption regarding duration of treatment effect (2 years) – probabilistic incorporating PAS – include SP3

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£34,322	4.598			
Pirfenidone	£57,223	4.929	£22,901	0.331	£69,250
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£32,248	4.207			
Nintedanib	See confidential appendix				
Pirfenidone	£53,851	4.540	£21,603	0.333	£64,949
Mild population (percent predicted FVC >80%)					
BSC	£38,971	5.493			
Pirfenidone	£66,885	5.762	£27,915	0.269	£103,893

Table 8: ERG-preferred base-case assuming stopping rule with optimistic assumption regarding duration of treatment effect (life-time effect) – probabilistic incorporating PAS – include SP3

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)
ITT population					
BSC	£30,458	3.873			
Pirfenidone	£57,175	4.926	£26,717	1.053	£25,365
Moderate population (percent predicted FVC of 50 – 80%)					
BSC	£28,713	3.552			
Nintedanib	See confidential appendix				
Pirfenidone	£53,754	4.530	£25,041	0.978	£25,603
Mild population (percent predicted FVC >80%)					
BSC	£34,644	4.665			
Pirfenidone	£66,894	5.754	£32,250	1.089	£29,607

3.2 Scenario with the stopping rule implemented for nintedanib only

Results are provided in Table 9 and Table 10 when applying the stopping rule for nintedanib but not for pirfenidone, for the ERG base-case scenarios incorporating pessimistic and optimistic assumptions regarding the duration of treatment effect. Results when incorporating the PAS are provided in a confidential appendix to this addendum. Results for pirfenidone versus BSC when incorporating the PAS are not reported here as they are unaffected by the change to the nintedanib stopping rule.

The incorporation of the stopping rule for nintedanib but not for pirfenidone improves the cost-effectiveness of nintedanib relative to pirfenidone as it limits the life-time costs of treatment with nintedanib, but does not have any impact on the life-time benefits accrued.

Table 9: Stopping rule for nintedanib only (other assumptions as per ERG preferred base-case) with treatment effect assumed to stop after 2 years – Moderate subgroup (probabilistic results using list price)

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)	ICER (incremental analysis)
BSC	[redacted]	4.237				
Nintedanib	[redacted]	4.426	[redacted]	0.189	[redacted]	[redacted]
Pirfenidone	[redacted]	4.541	[redacted]	0.303	[redacted]	[redacted]

Table 10: Stopping rule for nintedanib only (other assumptions as per ERG preferred base-case) with lifetime treatment effect – Moderate subgroup (probabilistic results using list price)

	Costs	QALYs	Inc Costs (vs. BSC)	Inc QALYs (vs. BSC)	ICER (vs. BSC)	ICER (incremental analysis)
BSC	[redacted]	3.643				
Nintedanib	[redacted]	4.366	[redacted]	0.723	[redacted]	[redacted]

Pirfenidone	[redacted]	4.536	[redacted]	0.894	[redacted]	[redacted]

4 CONCLUSIONS

The NMA amended to include SP3 suggests that the treatment effects for pirfenidone are broadly similar to those for nintedanib for all outcomes, with the pairwise treatment effects indicating that neither treatment is statistically significantly more effective. Compared to the ERG base case network, there is a minimal change in the median treatment effects that favours pirfenidone for all-cause mortality, and favours nintedanib for PFS, exacerbations and all-cause discontinuation. There is a reduction in the estimated between study heterogeneity for all outcomes.

The impact on the ICERs of including SP3 within the NMA, whilst maintain all other assumptions that make up the ERG's preferred base-case scenario, is minimal with the ICERs being broadly similar to the results for the ERG's preferred base-case scenario. This is expected given the minimal impact of adding SP3 on the results of the NMA for the outcomes that inform the cost-effectiveness model.

The ERG would advise caution in the interpretation of the results for the scenario in which SP3 was included in the network meta-analysis because the SP3 study was conducted in a population which is likely to be less representative of patients in the UK (Japanese patients), it used a dose licensed in Japan but not in the UK, and statistical adjustments were required to incorporate data from this study as hazard ratios were not provided for the outcome of all-cause mortality.

The ERG notes that because of the manner in which the stopping rule has been implemented within the company model, wherein treatment costs are affected, but clinical outcomes are not, the ERG does not believe that the company model provides an accurate estimation of the relative benefits of the two treatments under a scenario where the stopping rule is implanted for nintedanib but not for pirfenidone. The ERG therefore believes that the results for the scenario incorporating the stopping rule for nintedanib but not for pirfenidone should be interpreted with caution as they are likely to be biased in favour of nintedanib.

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

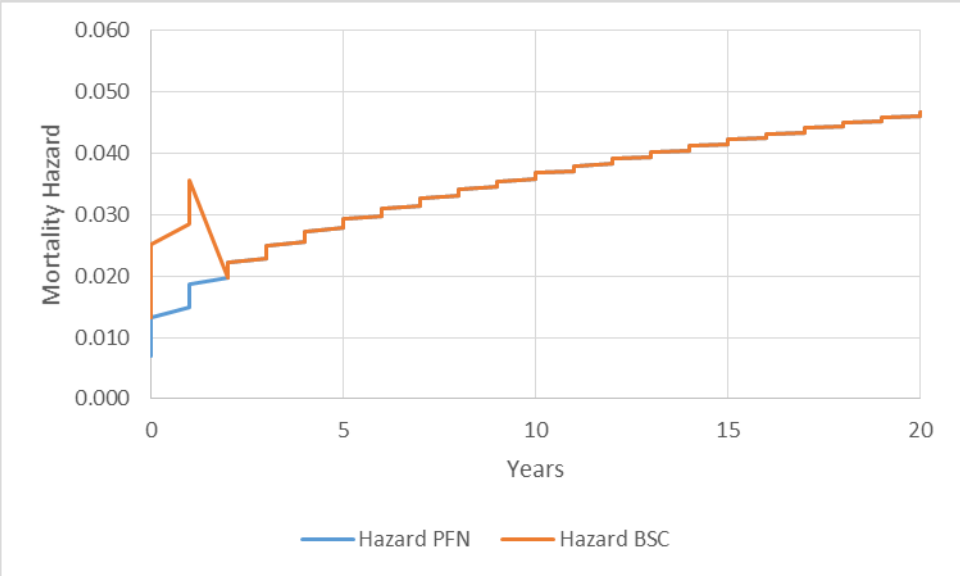
Pirfenidone for treating idiopathic pulmonary fibrosis (review of TA282) [ID837]

You are asked to check the ERG report from the School of Health and Related Research (ScHARR) to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **6pm, 19 April 2016** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 Duration over which treatment effect is applied in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 177 of the report, the ERG states: *"The ERG notes that the treatment effect could stop earlier or later than 7 years, and therefore the ERG's preferred base-case are provided, in Section 6, using an optimistic (lifetime) and pessimistic assumption (treatment effect to stop at 2 years approximately at the end of the clinical evidence) regarding the duration of the treatment effect (lifetime to 2 year)."* We do not consider a treatment effect of only two years to be clinically plausible, or appropriate in consideration of the	Discontinuation of treatment effect at the end of trial follow up for BSC patients (at approximately 27.7 months) has been chosen as an arbitrary cut off directly due to the end of trial follow up. Therefore, use of this time point at which to assume cessation of treatment effect has no clinical rationale, and we would therefore not consider this to be a reasonable cut off. We note that the scenario analysis with treatment effect limited to two years implies a steep drop in the hazard function for BSC (see below) which is not supported by the available data.  The assumption of proportional hazards supported using evidence from registry data has already been presented in Appendix 20 of the company submission. The	The duration over which the treatment effect is applied in the model is inherently a large driver of model results. We agree that there is uncertainty associated with this model parameter. However, we feel it is important to consider the wealth of data available for BSC patients from registries, which have been shown in response to Issue 15 to be suitable for comparison to pirfenidone trial data.	This is not a matter of factual accuracy. We clearly state on page 177 that the duration of persistence for any long-term treatment effect is currently highly uncertain. Results are presented for 2 years of treatment effect and life-time treatment effect to provide the Committee with information regarding how this uncertainty may affect the ICERs. We clearly state that the scenario using 2 years of treatment effect is pessimistic. No change has been made to the report in response to this issue.

data presented.	curves were shown to be approximately parallel, and therefore support the assumption of applying the treatment effect for the length of the trial data (7 years for pirfenidone, and 10-14 years for the registries). We agree that it is difficult to make assumptions regarding the treatment effect after the end of trial follow up data, which was why the scenarios regarding duration of treatment effect were explored as sensitivity analysis. Therefore, we would consider the pessimistic assumption should be amended to 7 years, as this is the maximum duration over which clinical data for pirfenidone are available, and consequently propose the following amendment: *"The ERG notes that the treatment effect could stop at any point later than 7 years, and therefore the ERG's preferred base-case are provided, in Section 6, using an optimistic (lifetime) and pessimistic assumption (treatment effect to stop at 7 years approximately at the end of the clinical evidence for pirfenidone, and equivalent registry data for patients on BSC) regarding the duration of the treatment effect (lifetime to 7 years)."* Additionally, we would request the ERG to re-consider all sections which allude to treatment effect not applying after 2 years to be adjusted to consider the actual duration of available clinical data: 7 years as opposed to 2 years.		

Issue 2 Increase in mortality from CAPACITY studies from Week 52 to 72 compared different outcomes

Description of problem	**Description of proposed amendment**	**Justification for amendment**	**ERG response**
On page 71 of the report, it appears that the ERG have assumed that all-cause mortality is the same as treatment-emergent all-cause mortality, and stated there was	Adjust page 71 text to refer to all-cause mortality data at Week 72 for CAPACITY studies. We would propose: "*there was an increase in death rates in the pirfenidone group, from 3.2% at 52 weeks 8% at 72 weeks; compared to placebo group from 6.3% at 52 weeks to 10% at 72 weeks.*"	On page 71 figures are not comparing like for like.	The data that are reported in the text are consistent with the Tables and, as such, are accurate, but some text needs revision. We have deleted the following: "If one assumes that the

<table>
<tr>
<td>increase in mortality from CAPACITY 1 & 2 from Week 52 to Week 72 by stating all-cause mortality data at Week 52 and treatment-emergent all-cause mortality data at Week 72.

In addition on page 16. the ERG note that “there appears to be a markedly increased rate of mortality in the CAPACITY trials between the data reported for 52 weeks and for 72 weeks, the reasons for which are unclear“</td>
<td>Clarify that increase in mortality appears to only impact CAPACITY-1 (not CAPACITY-2). It should be noted that the number of patients dying in each of the trials individually between 52 and 72 weeks is small therefore this increase in rate is most likely an artefact of low numbers of events. The EMA’s preferred analysis to assess mortality differences was to use pooled data from ASCEND and the CAPACITY studies in order to ensure a sufficient sample size to assess this endpoint. For this reason, data has been used throughout the model at 52 weeks as at this point data is available from the ASCEND study which provides substantial additional patient numbers enabling reliable analysis (see later Issues on this point).
<table>
<tr><th></th><th colspan="2">ASCEND</th><th colspan="2">CAPACITY-1</th><th colspan="2">CAPACITY-2</th></tr>
<tr><th></th><th>Pirfenidone</th><th>Placebo</th><th>Pirfenidone</th><th>Placebo</th><th>Pirfenidone</th><th>Placebo</th></tr>
<tr><td>n</td><td>278</td><td>277</td><td>171</td><td>173</td><td>174</td><td>174</td></tr>
<tr><td>Number of deaths at 52 wks</td><td>11</td><td>20</td><td>6</td><td>9</td><td>5</td><td>13</td></tr>
<tr><td>Number of deaths between 52 and 72 wks</td><td>N/A</td><td>N/A</td><td>7</td><td>6</td><td>3</td><td>2</td></tr>
</table>
</td>
<td>The statement on page 16 should refer only to CAPACITY-1; additionally it should be made clear that the EMA’s preferred analysis for the assessment of mortality was to use pooled data (which is only possible at 52 weeks) precisely in order to have a sufficient sample size with which to make informed comparison.</td>
<td>reported “all-cause mortality” is actually “treatment emergent all-cause mortality” … in the pirfenidone group, from 11 at 52 weeks to 19 at 72 weeks, … (the p-values for the differences between groups are 0.047 and 0.315 for 52 weeks and 72 weeks, respectively). The numbers for non-treatment emergent all-cause mortality are higher (see Error! Reference source not found.).”

We have added in replacement: “There is a substantial increase in all-cause mortality in the pirfenidone group, from 11 at 52 weeks to 27 at 72 weeks, compared with a much smaller increase in the placebo group from 22 at 52 weeks to 34 at 72 weeks (the p-values for the differences between groups are 0.047 and 0.315 for 52 weeks and 72 weeks, respectively). (see Error! Reference source not found.)”.

Regarding the statement on page 16: The rates of mortality do appear to be relatively higher in the week 52-72</td>
</tr>
</table>

			period in CAPACITY 1 than in CAPACITY 2, but the reasons for this are still unclear.

Issue 3 Choice of data cut off point for application of hazard ratio for OS

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 174 of the report, the ERG states: *"The ERG further observes a discrepancy in the company's argument in that the HRs estimated using data at 52 weeks are used in the company's base-case. As discussed in Section 4.7, the ERG considers that if the assumption of proportional hazards was valid, then the HR estimated at 72 weeks would be a more appropriate estimate as it incorporates more of the available data."* This statement is misleading, as the data presented to inform the hazard ratio should be considered in full ahead of making assumption regarding the appropriate cut off point.	As data were pooled to inform the OS for BSC patients, at around 52 weeks a large proportion of patients on BSC are lost to follow up due to the end of the ASCEND trial. This was presented in Figure 3 of the company submission (Section 4.4). Between 12 and 15 months, approximately half the total number of patients are no longer at risk (576 to 290), and therefore it is expected that the hazard ratio based on the first 52 weeks of trial data is likely to be more reflective of the true outcomes for patients on BSC as all trial data is taken into account at this point. This is important as similar drops are measured in the first four 5 month periods of the BSC KM curve, but the numbers at risk drop substantially at this later cut off time (please see the table below for details): **Time (months) / At risk / Alive / Difference** 0 / 624 / 100% / - 5 / 604 / 98% / 2.41% 10 / 576 / 94% / 3.09% 15 / 290 / 92% / 2.78% 20 / 8 / 89% / 2.78%	The impact of amending this error will clarify that the choice of cut off point for the HR derivation was made due to the time point at which as much data was available for the largest possible proportion of patients. 52 weeks is the last timepoint at which data from the ASCEND trial is available. Given this is the trial requested by the FDA to investigate the impact of pirfenidone on mortality we would recommend that the most relevant HR to use is the one in which data from this trial is still included.	This is not a matter of factual accuracy. The company's statement that "*the data presented to inform the hazard ratio should be considered in full ahead of making assumption regarding the appropriate cut off point*" is reasonable for an analysis of the treatment effect up to specific time point. However for use in the cost effectiveness model, the ERG does not agree with the decision to limit the data to the first 52 weeks of follow up and subsequently extrapolate those results. The ERG does not consider the fact that the ASCEND trial ends at 52 weeks provides justification to exclude 20 weeks of follow up data from the CAPACITY trials. The HRs are estimated separately within each study, and under the assumption of PH. The

	To address this feature of the data, we would propose the following amendment *"The ERG notes that the considered HRs estimated using data at 52 weeks are used in the company's base-case. The rationale the company gave for using these HRs was due to the loss of information from one of the 3 relevant clinical trials after week 52 resulting in later timepoints being unrepresentative of the entire population of the relevant clinical trials."*		company's statement that "*we would recommend that the most relevant HR to use is the one in which data from this [ASCEND] trial is still included*" is therefore not accurate, since using the 72 week data for the CAPACITY trials does not mean that data from the ASCEND trial will be excluded. No change has been made to the report in response to this issue.

Issue 4 Analyses are incorrectly described as *post-hoc*, but were pre-specified in the EMA ISE SAP developed for the type II variation

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG report frequently describes analyses conducted on pooled data from CAPACITY I & II, or CAPACITY I & II and ASCEND as *post hoc*. This is not correct: these analyses – including the assessment of mortality across the three trials – were pre-specified in the integrated summary of efficacy (ISE) SAP	Section 6.4 of the EMA SAP for ISE, states: "*The primary efficacy analysis of the pooled data is the rank ANCOVA model for the change from Baseline to Month 12 in %FVC between the 2403 mg/day pirfenidone and placebo groups. ITT population from Studies PIPF-004, PIPF-006 and PIPF-016 will be used.*" Secondary outcomes are also discussed in the ISE: "*Mortality data from the pirfenidone 2403 mg/d or 1800 mg/d and placebo groups in Studies PIPF-004, PIPF-006 and PIPF-016 will be pooled. ITT population with all randomized patients from these three studies will be used*." (Section 6.5.1). "*All-cause mortality through Month 12 will be the secondary efficacy endpoint*." (Section 6.5.2). Analyses for the 6MWT and PFS were also described in the Plan. The SAP ISE has been supplied to NICE alongside this	Correction	The ERG contends that the terminology "pre-specified" implies that the analyses were specified before the trial was undertaken; this was not the case. The SAP was dated 2014, which post-dates the completion of two of the three included trials. The ERG accepts that the trial protocol (dated 25th April 2012) specifies that data will be pooled across all 3 trials for mortality outcomes. However, it does not specify that data will be pooled for the other outcomes such as PFS and 6MWD.

developed for the EMA type II variation.	response. Based on these analyses being pre-specified in the EMA SAP for ISE, we request the ERG report be amended to reflect this.		No change has been made to the report in response to this issue.

Issue 5 Source of data for patients discontinuing treatment with pirfenidone

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 162 of the report, the ERG state: *"In the company's base-case, people initiating pirfenidone are assumed to discontinue treatment at the rate observed in the RECAP extension trial; therefore, no stopping rule is applied in the base-case."* This is incorrect as data were not solely taken from the RECAP extension trial, and the rationale for not applying the stopping rule was not restricted to this	In the model, patients are assumed to discontinue treatment at the rate observed in the Phase III studies (CAPACITY 1&2 and ASCEND) as well as the RECAP extension trial. We would propose the following amendment: *"In the company's base-case, people initiating pirfenidone are assumed to discontinue treatment at the rate observed in the pirfenidone clinical trials."* Additionally, the rationale for not including the stopping rule in the base case for pirfenidone extends further than discontinuation data observed in the trials. The stopping rule was also not applied as it was highly criticised by clinicians at a Boehringer Ingelheim advisory board and by clinicians during the Appraisal Committee meeting for nintedanib in August 2015, the consultation responses to the ACD, and the scoping consultation for this re-appraisal of pirfenidone. We acknowledge that the ERG states in the report: *"Clinical advisors to the ERG reported that to their knowledge the stopping rule is being rigorously applied in clinical practice, but they agreed that the stopping rule is clinically problematic as a prior decline in lung function does not predict a future decline, and periods of stability can*	The discontinuation rate is derived using data from all available trials, as stated on page 217 of the CS: *"In each cycle, the proportion of patients on and off treatment are calculated based upon curves fitted to patient level data from ASCEND/CAPACITY/RECAP."* The reason for not applying the stopping rule was not solely due to these data, but was also informed by previously deliberation through the nintedanib NICE appraisal process, as stated on page 218 of the CS: *"The same discontinuation rule was highly criticised by clinicians at an advisory board held by Boehringer Ingelheim on 23rd April 2014, as it was considered difficult to impose [Boehringer Ingelheim 2015]. These opinions were reiterated by clinicians during the Appraisal Committee meeting for nintedanib in August 2015, the consultation responses to the ACD, and the scoping consultation for this re-appraisal of*	We agree that it was incorrect to state that the rate of discontinuation was based on RECAP alone. This sentence has been amended to "*In the company's base-case, people initiating pirfenidone are assumed to discontinue treatment at the rate observed in the CAPACITY 1 & 2, ASCEND and RECAP studies; these discontinuation rates are not adjusted to reflect the implementation of the stopping rule in the base-case."*.

reason.	*sometimes only be identified retrospectively."* We would propose the following amendment: *"No stopping rule is applied in the base-case, as the company considered that people initiating pirfenidone are assumed to discontinue treatment at the rate observed, and the stopping rule was associated with strong clinical criticism in previous NICE committee meetings."*	*pirfenidone [NICE 2015d, NICE 2015e]."* The impact of amending this error will not affect cost-effectiveness results, but may improve understanding regarding the use of clinical trial data within the model, and reasoning why the stopping rule was not included in the base case.	

Issue 6 Issues with trial data for pirfenidone and placebo OS

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 186 of the report, the ERG states: "*As discussed in Section 4.7, the ERG considers that: (i) the treatment effects estimated using data up to 72 weeks are more appropriate and consistent with the company's assumption of proportional hazards*…" Figure 41 is also presented by the ERG to support this point.	Removal of the following text: "*The treatment effects estimated using data up to 72 weeks are more appropriate and consistent with the company's assumption of proportional hazards…*" Removal of Figure 41 and legend.	The hazard ratio used in the NMA for overall survival is stratified by the three studies and two regions. Therefore comparing this stratified hazard ratio with the "raw" Kaplan- Meier is incorrect as the two estimates are not comparable.	This is not a matter of factual accuracy. The ERG considers that it is reasonable to compare the survival in the BSC arm predicted by the model, with the raw Kaplan-Meier data for the pooled studies. No change has been made to the report.

Issue 7 Uncertainty whether SP3 protocol excluded obstructive lung diseases & implication for pooling of VC and FVC results

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The comparability of VC and FVC was questioned by the ERG, as	Removal of the following statements:	The NICE Committee have previously accepted the approach	We do not believe that the ERG report is factually

the inclusion / exclusion criteria for SP3 were not explicit regarding the patients with obstructive airway disease (pages 34, 54 and 119). The lack of clarity contributed to the ERG dismissing the relevance of all data from SP3.	• "*Therefore, the ERG considers that the combination of VC data from SP3 with FVC data from the ASCEND and CAPACITY trials is questionable*" (page 34) • "*Therefore, the ERG considers that the synthesis of VC data from SP3 with FVC data from the ASCEND and CAPACITY trials is questionable*" (page 54) • "*Therefore the ERG considers that the combination of VC data from SP3 with FVC data from the ASCEND and CAPACITY trials is questionable*" (page 119) Inclusion of SP3 in all analyses (in combination with Issue 8).	of pooling data from VC and FVC outcomes into one combined assessment as part of their review of the evidence informing TA282. Whilst a clarification question on this point was not raised by the ERG, following receipt of the ERG report, we have been able to confirm that clinical, radiological (pathological) diagnosis of obstructive disease was performed by central review, and evident COPD patients were to be excluded from the study. After a review of the patient lists, there was one patient in the placebo group of SP3 (n=55) with co-morbid COPD, and two with emphysema. Of the 108 patients in the pirfenidone 1800mg/day dose arm of SP3, three patients had co-morbid emphysema (none had COPD).	inaccurate given the information available to the ERG at the time the report was prepared. We welcome this additional clarification, but we still believe that the combination of data from these two outcomes is questionable given that some patients with emphysema were included in the SP3 study. We think that it is important that the Committee considers this limitation of the evidence explicitly and would therefore prefer to leave these statements in the report as they are not factually inaccurate. At the request of NICE we have provided the Committee with analyses that include data from SP3 for their consideration. No change has been made to the report in response to this issue. The additional analyses are presented in an addendum.

Issue 8 Referring to 1800mg dose as unlicensed

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
In many places the ERG report, the 1800mg/day dose is referred to as being "*unlicensed*" (pages: 15, 50, 59, 152, 216). On page 152, the ERG state that the 1800mg/day dose is a limitation of including SP3 to assess efficacy outcomes.	Clarify upfront that the 1800mg/day is licensed in Japan and state that the dose licensed in Europe and North America (2403mg/day) was derived from the licensed Japanese dose, based on mean weights of the North American and European populations. Delete "unlicensed dose/s" and replace with "lower dose/s" (pages 15, 50, 59, 152, 216). Based on the rationale for the differences in dose across geographical patient populations, remove this as a limitation of including SP3 to assess efficacy outcomes (page 152).	Stating that 1800mg/day is an unlicensed dose is inaccurate: the lower dose is licensed in Japan, and was the basis for derivation of the licensed dose in Europe and North America. The 2403mg/day dose was derived by normalisation of the 1800 mg/day dose used in the Japanese studies. This was based on a calculation that adjusted for the differences in mean weights in the North American and European patient population compared with the Japanese patient population. In general, North American and European are approximately 30% heavier than Japanese patients of similar age. Thus, the 1800 mg daily dose in Japanese IPF patients translates to approximately 2400 mg/day in North American and European IPF patients. There is no evidence to suggest a difference in effect between the doses in the two populations, a view which was accepted by the NICE Committee in their consideration of the evidence supporting TA282. The Recent nintedanib appraisal was also based on analyses which pooled data/outcomes across the two doses, and this was again accepted by the NICE Committee (TA379).	Whilst the ERG does not consider it factually inaccurate to describe the lower dose used in the SP3 trials as unlicensed in the context of a report on the use of pirfenidone in the UK, we feel that it is reasonable to make it clearer that this is the licensed does in Japan. Therefore for each of the six instances where we described this as an unlicensed dose the text has been amended to make it clear that the dose is licensed in Japan but not in the UK.

Issue 9 Appropriateness of absolute FVC based structure

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 168 of the report, the ERG state: *"Furthermore, within the company's model, all disease progression is assumed to be equally detrimental. Clinical advisors to the ERG considered that a 10% drop in percent predicted FVC would impact on HRQoL differently according to the baseline percent predicted FVC and therefore the clinical impact of disease progression, as defined in the model, would be different across individuals.* *The ERG notes that the model used in the nintedanib appraisal provides a better representation of the natural history in IPF, whereby individuals transit through multiple health states with different levels of percent predicted FVC (rather than just two), as their disease progresses. This structure allows for different HRQoL and cost estimates to be attached according to the individual's percent predicted FVC level.* *The model structure used in the nintedanib company submission was also considered by the clinical advisors to the ERG to be more representative of the progressive nature of IPF than the pre/post progression model presented by the company for pirfenidone."* Additionally, on page 179 of the report, the ERG state:	The first statement suggests that data are available to estimate the transitions between percent predicted FVC health states. We would consider any estimation of these transition probabilities to be fundamentally flawed on the basis of small patient numbers to inform them. Additionally, HRQoL across the percent predicted FVC model health states used in the nintedanib submission yielded very similar overall HRQoL values for the health states used in our modelling. The assumption made here was that patients "progress" by moving to the next category of percent predicted FVC (i.e. progression from 70 - < 80% goes to 60 - < 70%). This is shown below: - PF: 0.82 (current ERG model) - PF: 0.78 (estimate based on nintedanib model health states) - PP: 0.76 (current ERG model) - PP: 0.74 (estimate based on nintedanib model health states) We would argue that the analyses used within nintedanib model do not add much additional information in terms of change in patient quality of life over time given these extremely small	The FVC structure utilised in our model comprised of the best choice given the data available, transparency for external assessment. This is supported by prior use in published literature (Loveman *et al.* (2013)).	This is not a matter of factual accuracy. It is the opinion of the ERG and their clinical advisors that a model structure which distinguishes between different levels of percent predicted FVC has more face validity than one that categorises patients as either progressed or unprogressed. We note that whilst the model by Loveman et al. used similar health states to the company model, it was a state-transition model and not a partitioned survival model, which is an important difference. No changes have been made to the report in response to this issue.

"The ERG also notes that different definitions of PFS are used between trials included in the NMA. This is acknowledged in the CS (page 143). The ERG considers that this is likely to introduce some biases between pirfenidone and nintedanib but reiterates that PFS has a minimal impact on the ICER in the company's model. Although, as stated previously, it is expected that it would have a greater impact if progression was linked to treatment discontinuation and treatment effects were allowed to differ after discontinuation." These statements critique the use of the FVC-based structure undertaken in our modelling. Whilst we agree that the model structure used by Boehringer Ingelheim for the nintedanib appraisal has its merits, we would also consider it to comprise of some fundamental flaws and be inappropriate for the data available to us. Consequently, we consider both of these statements to contain misleading points about how percent predicted FVC was used in the model.	changes across health states. Therefore, we would propose the following addition to the end of the statement: *"In spite of these apparent complications, the model produces comparable estimates of HRQoL when considering an aggregated estimate across all percent predicted FVC subgroups when partitioned into the health states applied in the company model."* Regarding the second statement, our ability to model the clinical pathway of IPF using this structure is hindered by the lack of available data. As we do not have access to the same breakdown of FVC data, we are unable to produce a similar structure without imposing substantial uncertainty on transitions throughout the model. Furthermore, even if these data were available, due to the relatively small number of patients, transitions would already be associated with copious amounts of uncertainty. Therefore, we would not consider the model structure adopted by Boehringer Ingelheim to be appropriate for the purpose of this economic evaluation.		

Issue 10 Inclusion of acute exacerbations

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 170 of the report, the ERG state:	The ERG state within the report on page 169 that:	Acute exacerbations are captured within the model structure with progression	We do not believe that the statement in the ERG report is

"The ERG considered the lack of impact of acute exacerbations in the economic model to be an artefact of the model structure chosen by the company and not a reflection of the relevance of exacerbations in IPF." This statement is contradictory to another statement made earlier in the document.	*"Clinical advisors to the ERG noted that the diagnosis of acute exacerbations is challenging and that it is often difficult to distinguish between an exacerbation and progression."* Resultantly, the impact of acute exacerbations is likely captured within disease progression and thus the model structure captures the impact of acute exacerbations appropriately. Therefore, we would propose the following amendment to the statement: *"The ERG considered the impact of acute exacerbations in the economic model to be somewhat restricted by the model structure chosen by the company, and not necessarily a reflection of the relevance of exacerbations in IPF.* *However, it should be noted that the diagnosis of acute exacerbations is challenging and that it is often difficult to distinguish between an exacerbation and progression."*	(which is intrinsically linked), the cost of hospitalisation which includes acute exacerbation and the additional utility decrement assigned to patients. Difficulty accounting for these events separately within the model is driven primarily by the difficulty associated with diagnosing acute exacerbations and for this reason lack of comparable data to compare across trials. A similar point was also raised during the NICE Committee's assessment of the model supporting the appraisal of nintedanib, despite the use of a different model structure: "*However, taking into account the clinical experts comments about the substantial impact of exacerbations on quality of life (see section 4.2), the Committee was concerned that the results of the model were not sensitive to changes in the rate of exacerbations*." (page 20 of ACD supporting TA379).	factually inaccurate. As acknowledged by the ERG, it is challenging to distinguish between an exacerbation and progression. However, within the model, the company attempts to include the impact of acute exacerbation separately from progression. Whilst the ERG understand the challenge of including exacerbations, the point raised by the ERG refers to the fact that acute exacerbations as currently included in the economic model have limited impact on the results. The ERG considers that if acute exacerbations were modelled differently (for example by assuming that patients are deemed to have progressed following an exacerbation), the impact of acute exacerbations on the results would be greater. Therefore, no changes have been made to the report in response to this issue.

Issue 11 Clarification on external model validation

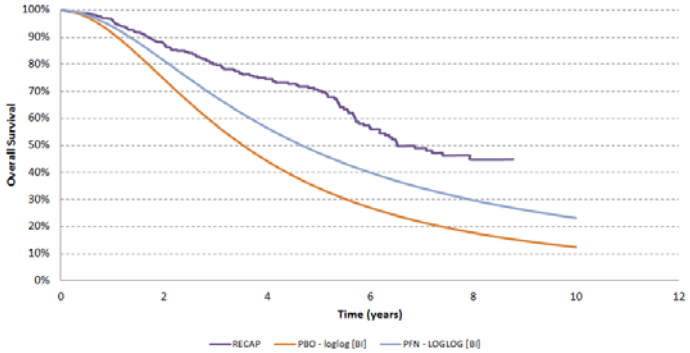
Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 214 of the report, the ERG state: "*The ERG observes that whilst the CS presents information regarding the external validity of the model, the CS does not describe any other forms of quality assurance such as:* • *Validation of the model structure and key structural assumptions using clinical experts to ensure face validity;* • *Peer review of the model by an independent health economist, or;* • *Verification of the calculations within the model by an independent modeller*." We would like to clarify that all forms of quality assurance mentioned above were undertaken during model development.	We would propose amendment of this statement to: "*The ERG observes that whilst the CS presents information regarding the external validity of the model, the CS does not describe any other forms of quality assurance. However, upon clarification with the manufacturer, the ERG was made aware that during external validation of the model the following forms of quality assurance were undertaken:* • *Validation of the model structure and key structural assumptions using clinical experts to ensure face validity;* • *Peer review of the model by an independent health economist, or;* • *Verification of the calculations within the model by an independent modeller."*	Whilst this information was not included within the submission, we note that the ERG did not request clarification on these additional forms of quality assurance of the model. These are, therefore, described below. Validation of the model structure was undertaken in March 2014 with two clinical experts in IPF: • Prof. Athol Wells (Consultant Physician, Royal Brompton) • Prof. Ron du Bois (Senior Research Investigator, Imperial College) Two independent 1:1 telephone calls were held with the experts. The objective was to gain clinical feedback on the plausibility / face validity of two proposed model structures: • A microsimulation model with updated data following the same structure as that submitted and accepted by NICE in TA282 • A denovo partitioned survival model model using parametric survival curves to determine	Section 5.10 of the template for company submissions provides the company with an opportunity to report information on the validation of the cost-effectiveness model. This information was not provided within the CS, so the text on page 214 is accurate given the information available to the ERG at the time the report was prepared. Therefore no change has been made to the report in response to this issue.

		transitions between alive and dead Both experts preferred the partitioned survival model approach for its simplicity. However, it was cautioned whether FVC alone could capture quality of life differences between patients. Following the publication by Loveman *et al*. it was decided to adopt a similar structure, which adopted survival curves and reflected quality of life differences through progression and lung transplant. During this time, the nintedanib model structure was unpublished, and analysis of the ASCEND/CAPACITY data found that there was a weak correlation between FVC change and quality of life; this has also been documented in the literature and confirmed by experts. During the model development, >3 independent health economists have peer-reviewed the model and verified the calculations. The model structure, although only focusing on survival outcomes and comparison with registry data, has been published at ATS 2015 and BTS 2015. Dr. Toby Maher (Consultant respiratory physician, Royal Brompton) is the second author on both publications, and has clinically validated the	

		estimations of survival and progression-free survival. We therefore feel that the statement regarding our validation of the model requires amendment.	

Issue 12 Implication that nintedanib model provides a better prediction of long term survival in IPF patients

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 169 of the report, it is stated that: "*The ERG notes that the model used in the nintedanib appraisal provides a better representation of the natural history in IPF, whereby individuals transit through multiple health states with different levels of percent predicted FVC (rather than just two), as their disease progresses. This structure allows for different HRQoL and cost estimates to be attached according to the individual's percent predicted FVC level. The model structure used in the nintedanib company submission was also considered by the clinical*	We propose that the statement is adjusted to account for this poor predictive ability, despite the use of a more complex characterisation of the disease.	As part of our response to the nintedanib ACD, it was stated that: "*The extrapolation for overall survival within the manufacturer's economic model poorly reflects the long term survival of IPF patients in the UK. Figure 2* [reproduced below] *presents the survival estimates for pirfenidone and placebo (extracted from the manufacturer's economic model), and compares this to the Kaplan Meier data from the RECAP study through overlaying the KM curve from this study. The RECAP study is the long term follow up of patients enrolled in the CAPACITY I & II trials.* *From Figure 2, it is clear that this important outcome of the model lacks face validity: the extrapolated pirfenidone estimate from the manufacturer's model [light blue] has diverged from the observed KM data by 1 year, with this gap growing over time: by year 3, the model under-predicts survival for patients initially receiving pirfenidone by over 10%*." **Figure 2: Manufacturer's model poorly reflects the long term survival of IPF patients**	This is not a matter of factual accuracy. As highlighted in response to Issue 9, it is the opinion of the ERG and their clinical advisors that a model structure which distinguishes between different levels of percent predicted FVC has more face validity than one that categorises patients as progressed or unprogressed. The ERG did not comment on whether the nintedanib model provides a better prediction of long term survival in IPF patients compared with the company's model. Therefore, no changes have been made to the report in response to this issue.

advisors to the ERG to be more representative of the progressive nature of IPF than the pre/post progression model presented by the company for pirfenidone." Based on information submitted as part of the consultation to the NICE ACD for nintedanib, we believe this statement is misleading: the nintedanib model does not accurately reflect long term survival of IPF patients in the UK.			

Issue 13 Implication that data from the UK Named Patient Programme has been ignored

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 201 of the report, the ERG state: *"The ERG further notes that the company identified a study reporting the impact of pirfenidone in a real-world setting through the UK Named Patient Programme using a retrospective study design. Findings from this study are not used or compared with the resource used in the model."* This statement suggests data were excluded from this study without reason.	The study via UK Named Patient Programme as reported by Parfrey *et al.* (2013) was a conference poster abstract, which did not contain any resource use estimates that could be used in the model. Therefore, we would propose removing the statement or amending to	The statement is not relevant as the reference is not an appropriate source of resource use data to inform the model.	Whilst we accept that the data from the UK Named Patient Programme was reported only as a conference abstract by Parfrey et al. (2013), it was listed as an included study. It presents the following information which we would consider to be relevant information on resource use and which could have been summarised in the CS. *"In the first 6 months from baseline, 11 patients had 15 IPF-related hospitalisations of which 6 were for an acute exacerbation. One patient was*

This is incorrect.	clarify that no relevant data was available.		*hospitalised in the 6–9 months period. The mean ± S.D. and median (IQR) hospital bed days in the first 6 months were 11.0 ± 7.5 and 11.0 (6.0–13.5) days respectively for hospitalised patients and mean ± S.D. 1.2 ± 4.2 days for all patients. One patient was admitted to the Intensive Care Unit, for 5 days. Eighteen patients had IPF-related Accident and Emergency department visits, 3 had an IPF-related day-case and 67 outpatient clinic visits, with a mean of 2.1 ± 2.0 outpatient clinic visits per patient in the first 6 months of the observation period."* No change has been made to the report in response to this issue.

Issue 14 Misinterpretation of RECAP bias

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 174 of the report, the ERG states that: "*Consequently, the ERG considers that the exclusion of people who received less than 80% of the assigned study treatment could overestimate the survival for pirfenidone as only people that are considered to be compliant have been included in RECAP, thereby making comparison with long-term registries less relevant*." This is misleading, as it is not stated that	Based on the high proportion of patients from the CAPACITY and ASCEND trials being enrolled into RECAP, we propose the following amendment: "I*nclusion into RECAP required that patients received 80% of the assigned study treatment during the CAPACITY and ASCEND studies. This may give scope for over-estimation of the survival for pirfenidone as only people that are considered to be compliant would be included in RECAP, thereby making comparison with long-term registries less relevant. However, as 96% and 94% of patients*	Clarification	This is not a matter of factual inaccuracy. Trial populations are already biased towards patient groups who are more likely to be compliant. Excluding non-compliant patients from the extension study will make the extension study less applicable to patients seen in clinical practice. No change has been made to the

96% and 94% of patients completing CAPACITY and ASCEND entered the RECAP study.	*completing the CAPACITY and ASCEND from these trials were enrolled into RECAP, the study is likely to be representative of the majority of patients receiving pirfenidone for 12 months or longer."*		report in response to this issue.

Issue 15 Appropriateness of registry data for estimating the long-term OS for BSC patients

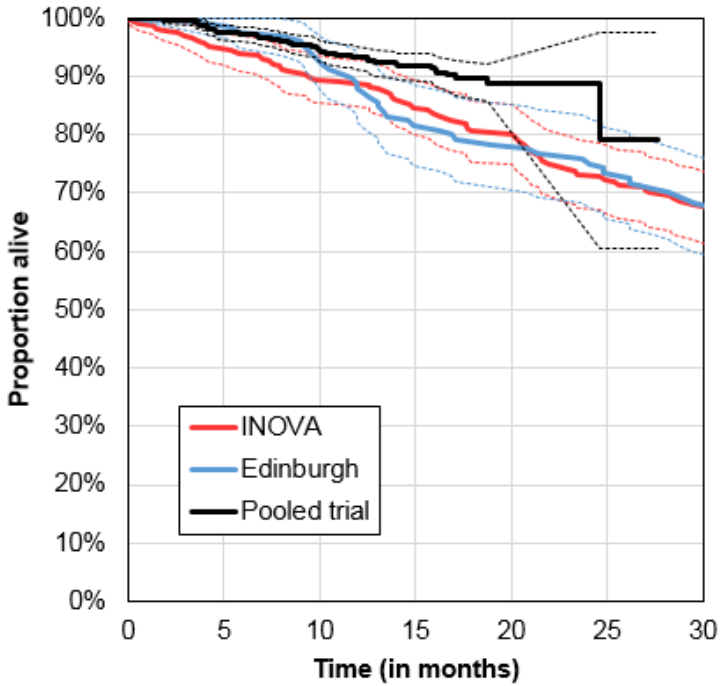
Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 175 of the report, the ERG states: *"Despite the attempt by the company to select and match individuals from registries to people enrolled in the ASCEND and CAPACITY trials, the survival of individuals from the registries is inconsistent with the OS of people initiating BSC observed in the clinical trials (see Figure 35)."* We consider the statement of inconsistency between the data sources to be misleading, in consideration of the evidence presented. Additionally on page 175 of the report, the ERG state: *"… the company excluded* [from consideration within the registries] *individuals with a percent predicted FVC ≥90%*	We would consider Figure 35 to demonstrate similar levels of consistency between registry data and the ASCEND and CAPACITY trials. We feel this is an appropriate assumption given that: - IPF is an orphan disease, and therefore available evidence for patients with IPF is sparse and often associated with caveats attributable to differences in baseline characteristics. - At the end of follow up across the ASCEND and CAPACITY trials, approximately 80% of patients were still event-free. This inherently causes issues in estimating the long-term outcomes associated with BSC. - Considering the length of follow up in ASCEND and CAPACITY, and the associated length of follow up in the registries, outcomes do not appear substantially dissimilar. At the end of the follow up period of the trials, the difference in OS between the registries and the trials is approximately 8%. A revision of Figure 35 from the ERG report showing the numbers at risk and associated 95% confidence intervals around the registry curves is shown in Figure 1 below.	The impact of amending these errors will not affect cost-effectiveness results, but aims to clearly demonstrate the full reasoning of why the differences in trial and registry OS appear inconsistent (within the context of short follow up).	This is not a matter of factual accuracy, but we accept that this point regarding excluded patients is less strong now that the company has clarified that only one patient with an FVC above 90 was excluded. We have amended the text to remove the statement regarding the potential exclusion of 8% of patients due to an FVC >90%, and added the following to the end of this bullet, "*However, this exclusion criteria only resulted in the exclusion of 1 patient from INOVA and 1 patient from Euro IPF, so any bias introduced is likely to be minimal."* However, the ERG does not believe that the statement on page 181 regarding *"making inferences about the*

(if DLco≥90%). However, according to data included in the company's model, approximately 8% of people in the ASCEND/ CAPACITY trials had a percent predicted FVC ≥90%. Throughout the CS, the company discuss a potential link between FVC and mortality; thus, excluding people with a percent predicted FVC≥90% could underestimate the survival in individuals included in the registries." This statement is misleading given that only patients with FVC≥90% and DLco≥90% were removed; and that this amounted to only 1 patient. Furthermore, on page 181 of the report, the ERG state: *"The ERG considers that making inferences about the plausibility of the long-term extrapolation based on the modelled OS for BSC against registry data has limited relevance given the OS data from the registry do not match the placebo arm of the trials."* We consider this statement to be contradictive.	Considering the large proportion of patients still to experience an event at the end of follow up in the pooled trial data, the differences in the numbers at risk (0 versus 74 and 141) as well as the overlapping of confidence intervals, we consider outcomes from the trials and registries to be consistent. We feel that the ERG should reconsider whether these data are in fact inconsistent, with the number of patients still at risk and lack of significant difference in mind. In regards to the statement regarding excluding patients with a percent predicted FVC ≥90% (if DLco≥90%), please see Figure 2 for a more detailed flow of patient inclusion within consideration for the registries. These flows show that for both registries, only 1 patient was removed for having a percent predicted FVC and DLco ≥90%. Therefore we would consider that outcomes from the registries are largely reflective of the expected IPF patient population. Furthermore, within the pirfenidone clinical trials, only 1 patient had a baseline percent predicted FVC and DLco ≥90%, as shown in Figure 3. Therefore, the exclusion of these patients from the registry data will have a neglible difference on the overall results. We would therefore propose the following ammendment: *"… the company excluded* [from consideration within the registries] *one patient from each registry with a percent predicted FVC ≥90% (if DLco≥90%) in line with the exclusion criteria applied to the ASCEND clinical trial."* Finally, the statement regarding use of the registry data in the long term implies that OS for BSC patients is underpredictive of the trial data, but is also inappropraite for estimating the long-term outcomes for patients on		*plausibility of the long-term extrapolation based on the modelled OS for* BSC against registry" is factually inaccurate. Therefore this statement has not been removed as suggested by the company.

	BSC. We consider this statement to be contradictive (i.e. the data cannot simulatenously be underpedictive and not appropraite in the long-term). Therefore, we would propose removing the following statement, along with others alluding to this contradiction: *"The ERG considers that making inferences about the plausibility of the long-term extrapolation based on the modelled OS for BSC against registry data has limited relevance given the OS data from the registry do not match the placebo arm of the trials."*		

Associated figures for Issue 15

Figure 1: Revision of Figure 35 from the ERG report including numbers at risk and 95% confidence intervals



INOVA	234	217	201	187	177	158	141
Edinburgh	125	123	115	101	95	85	74
Pooled trial	624	604	576	290	8	8	0

Figure 2: Flow charts for patient numbers included in the INOVA and Edinburgh registries

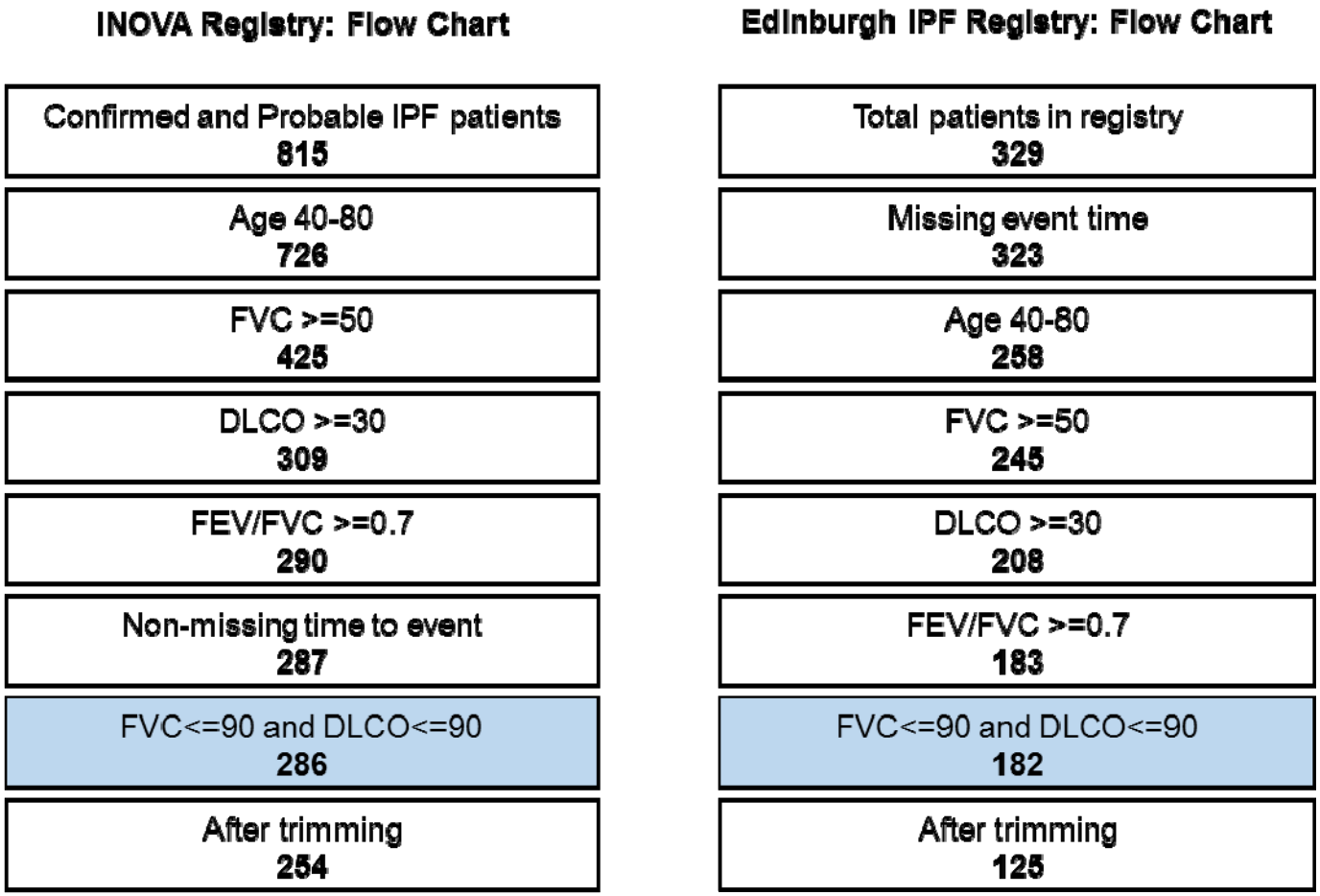


Figure 3: Patients in the pooled pirfenidone trials with baseline percent predicted FVC and DLco ≥90%

Pooled analysis
Listing of patients with both a baseline % predicted FVC and DLCO >=90 at baseline

Study Identifier	Unique Subject Identifier	Planned Arm Code	BL % Predicted Forced Vital Capacity	BL % Predicted Hgb-corrected DLco
PIPF-004	Censored	PLACEBO	96.4253	90.1397

Issue 16 Rationale for lack of inclusion of sequential active treatment following cessation

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 173 of the report, the ERG states: *"Clinical advisors to the ERG suggested that in practice, people initiating pirfenidone may switch to nintedanib upon discontinuation, and vice versa. This was acknowledged in the company's clarification response (see clarification response, question B7), but the absence of sequences was justified by the company on the basis that a similar approach was used in the nintedanib appraisal."* This is partially correct, but	In the model, patients are not permitted to receive pirfenidone (or nintedanib) following cessation of nintedanib (or pirfenidone). This modelling feature was incorporated as there is no clinical or safety evidence available on the impact of sequencing pirfenidone and nintedanib. As mentioned by the ERG, this feature was also incorporated for comparability with the modelling methodology undertaken by Boehringer Ingelheim as part of the submission for nintedanib. Additionally, there is uncertainty regarding the exact mechanism of action for both nintedanib and pirfenidone for the treatment of IPF, and hence there is uncertainty regarding the clinical effectiveness of sequencing these treatments. Consequently, we would propose the following amendment:	The reasoning for not including sequencing of active treatment is important, given the reasons presented in response to question B7 in the clarification response. We feel it is important to recognise that the rationale for not including this within the model was not based on prior work alone, but was based on a combination of prior work and availability of appropriate evidence regarding the safety, clinical effectiveness and mechanisms of action for the use of pirfenidone and nintedanib in sequence.	This is not a matter of factual accuracy. A fuller description of the company's reason for not considering treatment switches within the model is provided in section 2.2 where we discuss the current pathway of care. We do not feel that it is necessary to re-iterate all of this discussion within the health-economics section. No change has been made to the report.

does not state the full reasoning for why sequential therapies were not included in the model.	*"Clinical advisors to the ERG suggested that in practice, people initiating pirfenidone may switch to nintedanib upon discontinuation, and vice versa. This was acknowledged in the company's clarification response (see clarification response, question B7), but the absence of sequences was justified by the company for the following reasons:* - *There is currently no clinical or safety evidence available on the impact of sequencing pirfenidone and nintedanib.* - *A similar approach was used in the nintedanib appraisal.* - *The exact mechanisms of action for both pirfenidone and nintedanib (for the treatment of IPF) are associated with uncertainty, hence the impact of sequencing these treatments is unknown."*	The impact of amending this error will not affect cost-effectiveness results, but may improve understanding regarding the decision not to include sequential active therapies within the model.	

Issue 17 Use of DLco as measure of pulmonary function within the *de novo* economic model

Description of problem	**Description of proposed amendment**	**Justification for amendment**	**ERG response**
On page 34 of the report, the ERG states *"Although there are some data to suggest that DLco is a good prognostic indicator for mortality in IPF, it is not as well accepted as a clinical trial endpoint. Clinical advisors to the ERG agreed that DLco is harder to measure and is more variable than FVC. The variability of DLco has commonly been recognised to be as high as 15%, whereas the minimal clinically important difference for FVC is reported to be between 2% and 6%. The ERG therefore concludes that whilst DLco may provide important relevant information in clinical practice, it is*	To address this missing conclusion, we would propose the following amendment *"The final NICE scope specified that if evidence allows, subgroup analysis by disease severity, defined by percent predicted forced vital capacity (FVC) (such as above and below or 80% FVC) and/or diffusing capacity for carbon monoxide (DLco), should be*	The impact of amending this error will clarify that although DLco is a potentially good prognostic indicator for mortality in IPF, available data does not allow this to be appropriately included in the model and it is reasonable to not include DLco within the model structure.	This is not a matter of factual accuracy. No change has been made to the report.

reasonable for the CS to focus on FVC as the main measure of pulmonary function as it is more accepted as a reliable outcome in a clinical trial setting." We agree that FVC is the main measure of pulmonary function, and was therefore used within the model as the primary measure of clinical effectiveness (i.e. used within the measure of disease progression). However, it is important that this statement should feature prominently within the summary of the ERG's critique of cost effectiveness evidence submitted (Section 1.5). At present, the conclusion that it is reasonable to use percent predicted FVC as the main measure of pulmonary function is not clear in the critique of the cost effectiveness evidence submitted, we would therefore ask for this to be included as current discussion is as follows (on page 10 of the report): *"The final NICE scope specified that if evidence allows, subgroup analysis by disease severity, defined by forced vital capacity (FVC) (such as above and below or 80% FVC) and/or diffusing capacity for carbon monoxide (DLco), should be considered. However, the CS states that available data only allowed subgroups by FVC to be assessed."*	*considered.* *The CS states that available data only allowed subgroups by percent predicted FVC to be assessed, and based on input from clinical advisors to the ERG it is reasonable to focus on percent predicted FVC as the main measure of pulmonary function as it is more accepted as a reliable outcome in a clinical trial setting."*		

Issue 18 Statement on FVC / DLco is misrepresentative

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 23 of the report, the ERG states: "*Clinical advisors to the ERG considered that whilst percent*	To avoid the potential for misinterpretation, we propose the following amendment:	Clarification	This is not a matter of factual inaccuracy as this

predicted FVC had been used to define severity in clinical trials, this measure was not widely used in clinical practice, except to implement the recommendations in TA282. They commented that carbon monoxide diffusing capacity of the lungs (DLco) is clinically more meaningful and that DLco is the primary measure used to determine eligibility for lung transplantation, as some patients can have very low DLco values that suggest lung transplantation would be beneficial whilst maintaining a percent predicted FVC value that, in isolation form other measures, would indicate mild disease." This statement could be misinterpreted to read that FVC is not used in clinical practice, which we do not believe was the intention of the statement.	"*Clinical advisors to the ERG considered that whilst percent predicted FVC had been used to define severity in clinical trials, this was not the only measure used in clinical practice"*		is the advice provided by our clinical experts. The Committee will have the opportunity at the Committee meeting to gain the opinion of other clinical experts.

Issue 19 Choice of curve fit to efficacy data in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 181 of the report, the ERG state: *"Whilst the Weibull and Gompertz distributions provided a relatively similar fit during the observed period, these distributions provided different long-term extrapolations (Table 36)* [of the ERG report]. *Contrary to the company, the ERG considers the Gompertz distribution to provide a more*	To interpret the differences in the reported AIC scores, the following logic should be considered (taken from the following publication: *Burnham, K. P.; Anderson, D. R. (2002), Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed.), Springer-Verlag, ISBN 0-387-95364-7.)* - Suppose there are *R* candidate models. - Denote the AIC values of those models by AIC_1, AIC_2, AIC_3, …, AIC_R. - Let AIC_{min} be the minimum of those values. - Then $\exp((AIC_{min} - AIC_i)/2)$ can be interpreted as the relative probability that the i^{th} model minimises the (estimated) information loss.	The Gompertz is statistically inappropriate for consideration, as presented in the description of the proposed amendment.	This is not a matter of factual accuracy but a matter of opinion. As highlighted on page 183 of the ERG report, *"The ERG considers that both the Weibull and Gompertz distributions provide a similar fit to the observed period and that it is difficult to differentiate between the two. The ERG notes that both curves provided a very similar visual fit to the observed period and had broadly similar BIC values (861.89 for Weibull vs. 869.44 for the Gompertz for the ITT population – see CS, Table 72, page 212). The ERG reiterates that goodness of fit criteria only provide an indication of the goodness of fit during the observed*

realistic long-term extrapolation" It is incorrect to state that the Weibull and Gompertz curve have a similar fit to the observed data.	The AIC scores for OS in the model were given as **Model** / **AIC** **WEIBULL** / **844.15** **GOMPERTZ** / 851.70 Consequently, each distribution has the following relative probability of minimising the (estimated) information loss: **Model** / **AIC** **WEIBULL** / **1.00** **GOMPERTZ** / 0.02 From these results, it can be seen that the Gompertz model is exp((844.15-851.70)/2) = 0.02 times as probable as the Weibull model to minimise the information loss. Additionally, the Gompertz distribution was deemed to be appropriate by demonstrating for the mean age within the trials that the Weibull distribution produces hazards lower than those of the general population. We would not consider this analysis appropriate, as it relies on the mean age of entry into the model, and does not take into account the distribution of age around this mean. Consequently, we would therefore consider the Gompertz distribution to provide a poor fit to the data, and conclude that the Weibull distribution should be used instead. As a result, we would therefore propose that the Weibull should be used and that the Gompertz be removed as the ERG-preferred choice for long-term extrapolation.		*period and do not categorically indicate that one distribution should be preferred over alternative distributions.*" Whilst the Weibull and Gompertz distributions provided a relatively similar fit during the observed period, these distributions provided different long-term extrapolations. In the FACT-Check, the company justifies the use of the Weibull (for both the observed and unobserved period) on the basis of the AIC only. However, in the CS, the company recognises that the distribution should be selected on the basis of (i) visual inspection of the fit during the observed period; (ii) statistical goodness of fit during the observed period (as measured by the AIC and BIC), and; (iii) plausibility of the long-term extrapolation. It is the view of the ERG that the Gompertz distribution provides a more plausible extrapolation of OS than the Weibull distribution after the observed period and therefore was used in the ERG preferred base-case. No changes have been made to the report in response to this issue.

Issue 20 Analyses of trial data for differential treatment effects according to disease severity

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 93-94 of the report, the ERG describe analyses conducted in response to a clarification question investigating a differential treatment effect by disease severity at baseline (Table 23 & 24 of the report). In their description of these analyses of OS and PFS, the ERG state: "*In response to a clarification request from the ERG, the company also provided subgroup analyses according to disease severity for OS and PFS from the ASCEND and CAPACITY trials, although exact numbers within each subgroup in each trial arm were not reported. The findings did not suggest differential treatment effects according to disease severity for either outcome (as judged by the reported HR and 95% CI), however a treatment-by-subgroup interaction test was not reported so it is unclear if the difference between these subgroups was statistically significant*." Similar statements are also made in the accompanying summary statements on pages 17, 150 and 231. The interaction test was performed, but mistakenly not presented in our response to the clarification question. The analyses presented within the clarification response showed that confidence intervals for subgroups in each analysis were overlapping.	Consistent with the overlap of the confidence intervals, the interaction test across all analyses was not significant. These results are presented in Table 1 and Table 2 below, along with the number of patients contributing towards each analysis by treatment arm. Therefore, we would propose the following amendment to the statement: "*In response to a clarification request from the ERG, the company also provided subgroup analyses according to disease severity for OS and PFS from the ASCEND and CAPACITY trials. The findings did not suggest differential treatment effects according to disease severity for either outcome (as judged by the reported HR and 95% CI), with the treatment-by-subgroup interaction test not significant for any trial or outcome at 52 or 72 weeks*."	Whilst the confidence intervals presented in response to clarification question A31 highlighted the lack of a statistically significant difference between subgroups, the tests presented below formally confirm these results.	This is not a factual inaccuracy. We have accurately reported the information available to the ERG at the time the report was prepared. However, we thank the company for providing this additional information for consideration by the Committee. No changes have been made to the report in response to this issue.

Table 1: PFS results by baseline FVC percent predicted subgroup at 52 and 72 weeks

Study & time point	Baseline FVC ≤80% predicted				Baseline FVC >80% predicted				Interaction test: p-value (likelihood ratio)
	n (PFN / pla)	Adjusted HR	95% CI	p-value	n (PFN / pla)	Adjusted HR	95% CI	p-value	
CAPACITY 1									
52 weeks	117 / 125	0.84	0.53-1.32	0.4438	53 / 47	0.63	0.29-1.41	0.2571	0.5798
72 weeks		0.85	0.58-1.26	0.4128		0.56	0.28-1.11	0.0919	0.2962
CAPACITY 2									
52 weeks	124 / 107	0.60	0.40-0.92	0.0159	48 / 66	0.40	0.18-0.89	0.0193	0.3832
72 weeks		0.58	0.39-0.86	0.0590		0.48	0.25-0.92	0.0233	0.6938
ASCEND									
52 weeks	231 / 218	0.56	0.41-0.76	0.0002	45 / 55	0.64	0.30-1.40	0.2584	0.9074
72 weeks	N/A								
Pooled trials									
52 weeks	472 / 450	0.62	0.52-0.78	<0.0001	146 / 168	0.54	0.35-0.75	0.0069	0.4656
72 weeks		0.64	0.52-0.79	<0.0001		0.53	0.35-0.79	0.0017	0.4106

Table 2: OS results by baseline FVC percent predicted subgroup at 52 and 72 weeks

Study & time point	Baseline FVC ≤80% predicted				Baseline FVC >80% predicted				Interaction test: p-value (likelihood ratio)
	n (PFN / pla)	Adjusted HR	95% CI	p-value	n (PFN / pla)	Adjusted HR	95% CI	p-value	
CAPACITY 1									
52 weeks	118 / 125	0.6	0.17-2.04	0.4051	53 / 48	0.77*	0.11-5.59	0.7932	0.752
72 weeks		0.89	0.40-1.99	0.7763		0.77	0.11-5.59	0.7932	0.9881
CAPACITY 2									
52 weeks	126 / 108	0.25	0.08-0.76	0.0080	48 / 66	NE**	**	**	NE
72 weeks		0.29	0.10-0.79	0.0102		4.04***	0.42-38.87***	0.19***	NE
ASCEND									
52 weeks	233 / 221	0.63	0.29-1.34	0.2215	45 / 56	<0.01	0.00-NE	0.1231	0.12
72 weeks	N/A								
Pooled trials									
52 weeks	477 / 454	0.48	0.27-0.83	0.0071	146 / 170	0.59	0.14-2.51	0.4682	0.6452
72 weeks		0.58	0.36-0.94	0.0240		0.90	0.27-2.99	0.8610	0.4728

NE: not estimatable
* Only two deaths occurred in CAPACITY 1 before 52 weeks
** There were no additional deaths observed in either arm of CAPACITY 2 between 52 and 72 weeks in patients with FVC >80% predicted
*** Low number of events

Issue 21 Inaccurate statement on measures available within the clinical trial datasets

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
All of these measures are not available in all of the trials discussed. On page 11 of the report, the ERG state: "*These trials also reported all-cause and IPF-related mortality, PFS (using different definitions), 6-Minute Walking Distance (6MWD), DLco, and patient-reported outcomes, as measured by the University of San Diego Shortness of Breath Questionnaire (UCSD SOBQ) for dyspnoea, and the St George's Respiratory Questionnaire (SGRQ).*" A lack of subgroup information for DLco is repeatedly mentioned elsewhere within the document.	Clarify where mentioning DLco that this endpoint was not collected within the ASCEND trial as mentioned by the ERG on page 178.	It is important to clarify that DLco was not collected within the ASCEND trial, and therefore cannot be used (for example) within a consistent measure of disease progression.	The text on page 11 is simply stating which outcomes were reported across the trials without claiming that each trial reported data on each outcome. The text on page 14 correctly states which trials collected data on DLco. We therefore don't believe this text is misleading. 4.9 Conclusions of the clinical efficacy section excludes ASCEND from the findings on DLco as an outcome. Otherwise, DLco is only mentioned in relation to ASCEND as an eligibility criterion for inclusion No changes have been made to the report in response to this issue.

Issue 22 Inclusion of additional study outside of decision problem

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Inclusion of the Huang study is incorrect as this lies outside of the decision problem.	This study should be removed	PFN is not licensed for combination with NAC.	The study by Huang *et al* (2015) was explicitly included as supporting evidence only, providing information which the Appraisal Committee might find useful, "given the importance of data from randomised controlled trials in this orphan disease population" (EMA SAP for ISE, p.5). No changes have been made to the report in response to this issue.

Issue 23 Inconsistent results of the PSA performed by the ERG

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The cost and QALY results for BSC reported in Tables 64 and 65 do not match. This is also the case for QALY / LYG results for pirfenidone across the tables	These values should match, if a sufficient number of simulations are run in the probabilistic sensitivity analyses.	Inconsistency in analyses across list price vs. with-PAS scenarios	The results for absolute costs and absolute QALYs by treatment arm presented in Table 64 and 65 for values that aren't affected by the pirfenidone PAS are within 1% of each other. We believe that this is an acceptable level of variability between PSA runs and the PSA has been run with sufficient samples to give

			sufficiently accurate results. We also note that the majority of the results in Tables 64 and 65 are taken from the company's response to the clarification request and the company PAS template and were not generated by the ERG.

Issue 24 Incorrect statement on mortality outcomes

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 20 of the report, the ERG states: "*Individual trials do not report any statistically significant treatment effect compared to placebo for mortality outcomes; a statistically significant treatment effect is only observed when pooling or meta-analysing studies*." This is incorrect: CAPACITY-1 identified a statistically significant difference in treatment-emergent IPF-related mortality at 72 weeks.	The statement on page 20 (and associated statements elsewhere in the report) should be amended to reflect this statistically significant outcome from the CAPACITY 1 trial.	Correction	This statement was referring to all-cause mortality. This has been amended to make this clearer. We could not identify any similar statements which needed correcting elsewhere in the report.

Issue 25 Inaccurate reporting of trial outcomes

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
In Table 14 of the report, the ERG incorrectly report trial outcomes from the CAPACITY trials at 72 weeks. We	The outcomes beyond 52 weeks in the CAPACITY trials should be	Correction	52-week data and some separate 72-week data were provided by

believe this is due to use of Noble 2011 as the supporting reference, rather than the data presented within our submission (Tables 23-25). The "†" footnote also incorrectly makes reference to the Noble 2011 publication	based on the results presented within Tables 24 and 25 of our submission, where analyses were run to support the NMA. We propose the "†" footnote be amended to: "*The 72-week data were published in manufacturer submission*"		the CS in Tables 24 and 25, but the stated figures are accurately reproduced from the stated source, Noble 2011, as confirmed by the text of CS, p.96, which reports the same figures: "In the pooled analysis of CAPACITY 1 & 2 at 72 weeks…, there was a 23% reduction in all-cause mortality vs placebo among patients treated with pirfenidone 2403 mg/day (HR=0.77; 95% CI: 0.47-1.28; p=0.315) at Week 72. For TE IPF-related mortality, the HR also favoured pirfenidone at Week 72 (HR=0.48; 95%CI: 0.24-0.95; p=0.03) (Noble 2011)." No change has been made to the report

Issue 26 Missing rows in Table 30. SAEs reported in ≥ 2 patients in CAPACITY 1 & 2 at 72 weeks

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG response</th></tr>
<tr><td>p.105 Table 30. Data are missing from this table that show SAE reported in ≥ 2 patients in CAPACITY 2: myocardial infarction, bronchitis, lobar pneumonia, non-cardiac chest pain.</td><td>Add details on myocardial infarction, bronchitis, lobar pneumonia, non-cardiac chest pain from CAPACITY 2</td><td>Incomplete data set on SAEs</td><td>The following rows have been added to Table 30 of the ERG report:
<table>
<tr><th></th><th colspan="2">CAPACITY 1[36]</th><th colspan="3">CAPACITY 2[36]</th></tr>
<tr><th>Adverse event, n (%)</th><th>PFN 2,403mg/d (n=171)</th><th>PBO (n=173)</th><th>PFN 2,403mg/d (n=174)</th><th>PFN 1,197mg/d (n=87)</th><th>PBO (n=174)</th></tr>
<tr><td>Bronchitis</td><td>0 (0)</td><td>5 (2.9)*</td><td>2 (1.1)</td><td>NR</td><td>2 (1.1)</td></tr>
</table></td></tr>
</table>

Results show similarity between PFN and PBO arms, with the exception of MI, which was higher in PBO arm n=4 [2.3%] vs PFN 2403 mg/day arm n=0.			Lobar pneumonia	-	-	2 (1.1)	NR	2 (1.1)
			Non-cardiac chest pain	-	-	2 (1.1)	NR	2 (1.1)
			Myocardial infarction	-	-	0 (0)	NR	4 (2.3)*
			*p<0.05					

Issue 27 Incorrect interpretation of AEs leading to discontinuation in ASCEND

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
p. 107 ERG report states "*the only other AEs leading to discontinuation of treatment in ≥1% of patients in the PFN group were elevated hepatic enzyme levels, pneumonia, rash and decreased weight, which occurred in 3 patients (1.1%) in each trial arm*" for ASCEND	Change to: "*The only other AEs leading to discontinuation of treatment in ≥1% of patients in the PFN group were elevated hepatic enzyme levels, pneumonia, rash and decreased weight, which occurred in 3 patients (1.1%) in each*."	Factual inaccuracy 3 patients (1.1%) reflects number for each AE category not trial arm.	Change text to: ", which each occurred in 3 patients (1.1%)".

Issue 28 Inaccuracies on reported primary efficacy outcomes for ASCEND

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
p. 56 Table 6. Reported outcome states "*Change in percent predicted FVC **and death** from baseline to week 52*", this is incorrect. ASCEND reported the	p. 56 Reported outcome for ASCEND should be "*change in percent predicted FVC from baseline to Week 52*".	Inaccurate description of reported outcomes for ASCEND	Error in Table 6: Reported outcomes for ASCEND should read: "*change in percent predicted FVC from baseline to Week 52*" and "*Categorical*

primary outcome of change in percent predicted FVC from baseline to Week 52. In row box below, it states for ASCEND, "*Categorical decline of ≥10% in percent predicted FVC" this should include "or death*"	p. 56, for ASCEND, change to "*Categorical decline of ≥10% in percent predicted FVC or death*"		*decline of ≥10% in percent predicted FVC or death"* The same correction has been applied to the columns for the CAPACITY trials and in the text on page 54. We have also amended the discussion of the outcome "percent predicted FVC or death" on page 54 to make it clearer that whilst this doesn't appear to be listed as a trial outcome in the trial protocols, it appears to describe the method used by the company in order to impute a FVC measurement for patients who have died. The report has been amended so that this is longer described as post hoc as this method for handling FVC values in patients who have died was pre-specified.

Issue 29 Inaccurate statement on provision of definition for acute exacerbations in the CAPACITY studies

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
p. 57 Table 7. For CAPACITY 1 & 2, it states that the definition for acute exacerbation is not provided in the protocols and publication. This is incorrect, it is in Appendix	p. 57 Delete "*Definition not provided in protocols or publication*"	Factual inaccuracy	Change text to "*Definition not provided in clinical trial register protocols*"

H in both trial protocols and in the publication as a supplementary web appendix.			

Issue 30 Presentation of scenario results by percent predicted FVC subgroup

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On pages 21-22 of the report, the ERG states: *"For the mild population (percent predicted FVC >80%) the ICERs incorporating the PAS ranged from £31,722 to £186,260. Fort the moderate population (percent predicted FVC 50 – 80%), the ICERs for pirfenidone versus BSC ranged from £31,722 - £186,260 when incorporating the PAS. Results incorporating the PAS for pirfenidone versus nintedanib in the moderate population are presented in the confidential appendix."* This is an error, as the same results have been presented for both percent predicted FVC subgroups.	Given the copy and paste error, we would propose the following amendment: *"For the mild population (percent predicted FVC >80%) the ICERs incorporating the PAS ranged from £31,722 to £186,260. For the moderate population (percent predicted FVC 50 – 80%), the ICERs for pirfenidone versus BSC ranged from £27,432 - £70,234 when incorporating the PAS. Results incorporating the PAS for pirfenidone versus nintedanib in the moderate population are presented in the confidential appendix."*	The impact of amending this error will clarify the differences in model results for the mild and moderate percent predicted FVC subgroups.	The ICERS for the moderate group should have read "£27,432 to £104,915" as the range presented covers both with and without the stopping rule. These have been corrected.

Issue 31 Incorrect naming of nintedanib clinical trial

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Clinical trials assessing the efficacy and safety of nintedanib are incorrectly named as "IMPULSIS" (pages 113, 157-8,	There should be renamed as "INPULSIS".	Correction	Each instance of the use of IMPULSIS has been corrected to INPULSIS

201, 218,223).			

Issue 32 Description of the model reported by Loveman et al

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
ERG describes the Loveman model as the following “which is based upon three main health states” (page 167).	The authors of the model describe their model as follows: "It uses four distinct health states: unprogressed IPF; progressed IPF; lung-transplant; and dead."	Correction	Corrected to say “four main health states (unprogressed IPF, progressed IPF, lung-transplant and dead)”.

Issue 33 Description of the registry reported by Strand et al

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG state: “Fisher al (2015) reports OS from a modelling study whereby the OS in patients initiating BSC is modelled from a log-normal distribution which is fitted to data from the which is fitted to data from the National Jewish Health Interstitial Lung Disease database and not the US Strand registry as suggested by the company.” (page 212)	The data presented by Strand et al is also sourced from the National Jewish Health Interstitial Lung Disease database.	Clarification	This has been corrected to say “Fisher al (2015)[64] reports OS from a modelling study whereby the OS in patients initiating BSC is modelled from a log-normal distribution which is fitted to data from the US strand registry”.

Issue 34 Inaccurate description of CAPACITY 1 inclusion criteria & ASCEND patient number (Table 5)

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The description of the IPF diagnosis used in CAPACITY 1	Match IPF diagnosis description for CAPACITY 1 to that used for CAPACITY 2, i.e.:	Correction	These errors have been

has been incorrectly merged with that for ASCEND, rather than CAPACITY 2 (Table 5, page 51). The patient numbers in each arm of the ASCEND trial have been transposed.	• Confident clinical and radiographic diagnosis of IPF, confirmed locally (diagnosis previous 48 months) • No improvement of IPF in preceding year For ASCEND: adjust patient numbers to: • PFN 2403mg/day (n=278), and ; • Placebo (n=277)		corrected in Table 5: IPF diagnosis criteria has been changed to be the same for both CAPACITY 1 and 2, and different for ASCEND (6-48 months) We have changed numbers in ASCEND arms from PFN=277 to PFN=278, and from PBO=278 to PBO=277.

Issue 35 Misinterpretation of ASCEND manuscript not reporting primary outcome

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG report states that the ASCEND manuscript did not report the change in percent predicted FVC, but this was reported in the CS to inform the NMA (page 64).	For clarity, propose the following: The ASCEND manuscript reported that treatment with pirfenidone resulted in a significant difference vs. placebo in change in % predicted FVC at 52 weeks ($p<0.001$). Mean change in % predicted FVC were reported in the CS to inform the NMA.	Wording of the report could be misinterpreted to imply that the primary endpoint was not reported in the primary manuscript. Whilst the manuscript did not report the actual values for change in % predicted FVC for the PFN and PBO groups, it did report the statistical significance (i.e. p-value) on the difference.	The ERG statement is not factually inaccurate: "The ASCEND manuscript[34] did not report the change in percent predicted FVC, but this was reported in the CS,[4] principally to inform the NMA (see CS,[4] Table 20)."

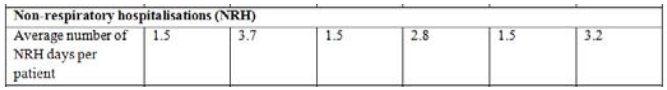
Issue 36 Error on footnote in Table 13 on definition of no decline in FVC

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
p.69 Footnote of Table 13 for * states: "*Change in FVC≥10%; CAPACITY trials data not reported in original publications (Noble 2011).*"	Suggest reword to: *Improvement in predicted FVC >0%; CAPACITY trials data are reported in the respective clinical study report.	The data for 'no decline in FVC' for the CAPACITY trials were taken from the respective CSRs. No decline was the total of mild and moderate improvement in percent predicted FVC (i.e. improvement of >0% to 10%, and ≥10%).	Changed footnote of Table 13 for * to: "*CAPACITY trials data not reported in original publications (Noble 2011), but taken from respective CSRs*"

Issue 37 Incorrect p-values for ACM and TEACM for pooled ASCEND, CAPACITY 1 & 2 studies

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
p.72 top of the page: p-values for pooled ASCEND, CAPACITY 1 & 2 on overall deaths and TE IPF-related deaths is incorrect	Correction: …significantly fewer overall deaths (*p*=0.011) and TE IPF-related deaths (*p*=0.006)	Correction	This is an error: We have changed: "there were significantly fewer overall deaths (*p*=0.047) and TE IPF-related deaths (*p*=0.012) in the pirfenidone groups" to: "significantly fewer overall deaths (*p*=0.011) and TE IPF-related deaths (*p*=0.006) in the pirfenidone groups …"

Issue 38 Incorrect labelling of information on Table 16

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
p. 82 Table 16. There are two rows of "*Average number of NRH days per patients*"' listed under Non-respiratory hospitalisations. The top row under Non-respiratory hospitalisations should be under Respiratory hospitalisations.	Move row to Respiratory hospitalisations, and rename 	Correction	We have made the suggested correction in the ERG report.

Non-respiratory hospitalisations (NRH)						
Average number of NRH days per patient	1.5	3.7	1.5	2.8	1.5	3.2

Issue 39 Incorrect reference to support pooled analysis of ASCEND and CAPACITY 1 & 2 for improvement in 6MWD

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
p. 86 The cited reference is incorrect: this should refer to Noble 2014, rather than Nathan 2014.	Change reference to Noble 2014 Full reference below: Noble PW, Albera C, Bradford WZ et al. Analysis of pooled data from three phase 3 multinational, randomized, double-blind, placebo controlled trials evaluating pirfenidone in patients with idiopathic pulmonary fibrosis [A11]. Poster presented at ATS 2014 annual congress; 16-21 May, San Diega, USA.	Correction.	This was the reference reported in the CS (page 108) and we noted that this appeared to be the wrong reference. Thank you for clarifying where this data came from. No change has been made to the report as it was accurate given the information we had at the time it was prepared.

Issue 40 Incorrect statement on most frequently-reported serious AEs in PFN arm

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
p. 103 ERG report states the following for ASCEND "*The other most frequently-reported serious AEs in the pirfenidone arm were pneumonia, prostate cancer, angina pectoris, nausea, congestive cardiac failure and rib fracture*" Similar statements are also made on pages 18 and 151.	On p. 103 remove pneumonia and prostate cancer from the sentence. "*The other most frequently-reported serious AEs higher in the pirfenidone arm compared to placebo were angina pectoris, nausea, congestive cardiac failure and rib fracture*" Also remove this statement from the summary sections, p. 18 and 151	Inaccurate description of SAEs that were higher in the PFN arm vs placebo arm for ASCEND. This is incorrect as the reported % are higher in the PBO arm for pneumonia and prostate cancer.	This is not a factual error. The statement does not make a comparison with placebo; it does not state that these AEs had a higher frequency in the PFN than in the placebo arm; it merely comments on their frequency in the PFN arm (2 or more patients), and the comparative data are provided in Table 29. No change has been made to the report.