NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Appraisal consultation document

Nintedanib for treating idiopathic pulmonary fibrosis

The Department of Health has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using nintedanib in the NHS in England. The Appraisal Committee has considered the evidence submitted by the company and the views of non-company consultees and commentators, and clinical experts and patient experts.

This document has been prepared for consultation with the consultees. It summarises the evidence and views that have been considered, and sets out the draft recommendations made by the Committee. NICE invites comments from the consultees and commentators for this appraisal (see section 9) and the public. This document should be read along with the evidence base (the Committee papers).

The Appraisal Committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the provisional recommendations sound and a suitable basis for guidance to the NHS?
- Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

Note that this document is not NICE's final guidance on this technology. The recommendations in section 1 may change after consultation.

After consultation:

- The Appraisal Committee will meet again to consider the evidence, this appraisal consultation document and comments from the consultees.
- At that meeting, the Committee will also consider comments made by people who are not consultees.
- After considering these comments, the Committee will prepare the final appraisal determination (FAD).
- Subject to any appeal by consultees, the FAD may be used as the basis for NICE's guidance on using nintedanib in the NHS in England.

For further details, see the Guide to the processes of technology appraisal.

The key dates for this appraisal are:

Closing date for comments: 2 October 2015

Second Appraisal Committee meeting: 15 October 2015

Details of membership of the Appraisal Committee are given in section 8, and a list of the sources of evidence used in the preparation of this document is given in section 9.

Note that this document is not NICE's final guidance on this technology. The recommendations in section 1 may change after consultation.

1 Appraisal Committee's preliminary recommendations

1.1 Nintedanib 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 company provides nintedanib with the discount agreed in the patient access scheme.

1.2 Treatment with nintedanib that is recommended according to 1.1 should be stopped if disease progresses (a confirmed decline in percent predicted FVC of 10% or more) in any 12 month period.

1.3 People whose treatment with nintedanib is not recommended in this NICE guidance, but was started within the NHS before this guidance was published, should be able to continue treatment until they and their NHS clinician consider it appropriate to stop.

2 The technology

2.1 Nintedanib (Ofev, Boehringer Ingelheim) targets 3 growth factor receptors involved in pulmonary fibrosis. Nintedanib is thought to block the signalling pathways involved in fibrotic processes, and may reduce disease progression by slowing the decline of lung function. It is administered orally. Nintedanib has a marketing authorisation in the UK for treating idiopathic pulmonary fibrosis in adults.

2.2 The summary of product characteristics states that the most frequently reported adverse reactions associated with the use of nintedanib are diarrhoea, nausea and vomiting, abdominal pain, decreased appetite, decreased weight and increased hepatic enzyme concentrations in the blood. For full details of adverse reactions and contraindications, see the summary of product characteristics.

2.3 The recommended dosage of nintedanib is 150 mg twice daily. The undiscounted price of nintedanib is £2151.10 for 60 capsules (taken from company submission, confirmed in Monthly Index of Medical Specialities [MIMS] online, accessed June 2015). This equates to a daily cost of £71.70 (2 capsules per day). The company has agreed a patient access scheme with the Department of Health. This scheme provides a simple discount to the list price of nintedanib, with the discount applied at the point of purchase or invoice. The level of the discount is commercial in confidence. The Department of Health considered that this patient access scheme does not constitute an excessive administrative burden on the NHS.

3 The company's submission

The Appraisal Committee (section 8) considered evidence submitted by Boehringer Ingelheim and a review of this submission by the Evidence Review Group (ERG; section 9).

Clinical effectiveness

3.1 The clinical evidence for nintedanib came from 3 multicentre, double-blind, randomised controlled trials comprising 2 phase III trials (INPULSIS 1 [n=513] and INPULSIS 2 [n=548]) and a phase IIb dose-ranging trial (TOMORROW [n=428]). All 3 trials compared nintedanib with placebo for 52 weeks in adults aged

40 years or older with idiopathic pulmonary fibrosis. The primary outcome was the rate of decline (ml per year) in forced vital capacity (FVC). The trials included only people with an FVC of at least 50% of the predicted normal value, and a diffusion capacity of the lung for carbon monoxide of 30–79% of the predicted normal value at baseline. The mean percent predicted FVC at baseline was approximately 80% in all 3 trials.

3.2 The key outcomes from the phase III nintedanib trials are presented in Table 1. The annual rate of decline in FVC with nintedanib (114.7 ml/year) was approximately half that of placebo; this difference was statistically significant ($p<0.001$). Fewer people randomised to nintedanib died compared with placebo, but this difference was not statistically significant. The time to first acute exacerbation was inconsistent across the trials:

- In INPULSIS 1, there was no statistically significant difference between nintedanib and placebo.
- In INPULSIS 2 the difference showed a benefit in favour of nintedanib, and was statistically significant.
- The pooled analysis of the 2 trials showed a trend in favour of nintedanib, which was not statistically significant.

The company noted that the INPULSIS trials were not powered to detect the effect of nintedanib on acute exacerbations.

Table 1 Outcomes from INPULSIS 1, INPULSIS 2, and the TOMORROW trials (shaded cells reflect significant differences)

Study	Annual rate of FVC decline (ml/year)	FVC responders[a]	≥1 acute exacerbation in 52 weeks[b]	Death (all cause)
INPULSIS 1				
Nintedanib 150 mg twice daily	−114.7	218/309 (70.6%)	19/309 (6.1%)	13/309 (4.2%)
Placebo	−239.9	116/206 (56.9%)	11/206 (5.4%)	13/206 (6.4%)
Measure of effect HR/MD/OR (95% CI) p value	MD: 125.3 (77.7, 172.8) p<0.001	OR: 1.91 (1.32, 2.79) p<0.001	HR: 1.15 (0.54, 2.42) p=0.67	HR: 0.63 (0.29,1.36) p=0.29
INPULSIS 2				
Nintedanib 150 mg twice daily	−113.6	229/331 (69.6%)	12/331 (3.6%)	22/331 (6.7%)
Placebo	−207.3	140/220 (63.9%)	21/220 (9.6%)	20/220 (9.1%)
Measure of effect HR/MD/OR (95% CI) p value	MD: 93.7 (44.8, 142.7) p<0.001	OR: 1.29 (0.89, 1.86) p=0.18	HR: 0.38 (0.19, 0.77) p=0.005	HR: 0.74 (0.40, 1.35) p=0.30
Pre-planned pooled analysis (INPULSIS 1 and INPULSIS 2)[c]				
Nintedanib 150 mg twice daily	−113.6	447/638 (70.1%)	31/638 (4.9%)	35/638 (5.5%)
Placebo	−223.5	256/423 (60.5%)	32/423 (7.6%)	33/423 (7.8%)
Measure of effect HR/MD/OR (95% CI) p value	MD: 109.9 (75.9, 144.0) p<0.0001	OR: 1.58 (1.21, 2.05) p=0.0007	HR: 0.64 (0.39, 1.05) p=0.08	HR: 0.70 (0.43, 1.12) p=0.14

[a] People with absolute decline in percent predicted FVC <10% at 52 weeks
[b] Investigator-reported acute exacerbations (according to the criteria described by the trial protocol); hazard ratio is based on analysis of time to first event
[c] Source of pooled results: nintedanib summary of product characteristics (individual trial results were presented in the company submission)
Key: CI, confidence interval; FVC, forced vital capacity; HR, hazard ratio; MD, mean difference; OR, odds ratio

3.3 Subgroup analyses showed that there were no statistically significant differences between the clinical effectiveness of nintedanib in people with a percent predicted FVC of 50–80% compared with people who have a percent predicted FVC more than 80%.

3.4 To compare nintedanib with pirfenidone, the company did a network meta-analysis including the 3 nintedanib trials and 5 placebo-controlled trials of pirfenidone (SP2, SP3, CAPACITY 1, CAPACITY 2 and ASCEND) which informed its economic model. The company chose different trials for different end points in the model:

- it included all evidence for overall survival
- it excluded the 2 pirfenidone trials in Japanese populations (SP2 and SP3) for acute exacerbations, because of heterogeneity (differences between studies)
- it excluded the ASCEND study of pirfenidone for decline in lung function.

The results are presented in tTable 2. The base-case results for overall survival were the same for nintedanib and pirfenidone, and neither drug showed a statistically significant difference in mortality compared with placebo. The base-case analysis of acute exacerbations showed comparable benefits for nintedanib and pirfenidone, but the company reported uncertainty in the results, which it considered to be a result of heterogeneity in the Japanese trials of pirfenidone (SP2 and SP3). After excluding these trials from the network meta-analysis (‘scenario 3’ of the sensitivity analyses for this outcome), the results showed fewer acute exacerbations with nintedanib than pirfenidone. The company’s analysis of loss of lung function (defined by the company as an absolute decline in percent predicted FVC of at least 10%) gave

similar results for nintedanib and pirfenidone using the base-case network meta-analysis. The differences between each drug and placebo were statistically significant. After excluding ASCEND from the network meta-analysis because of heterogeneity ('scenario 2' of the sensitivity analyses for this outcome), the results showed a statistically significant reduction in loss of lung function with nintedanib compared with placebo. The company did not report the results of significance testing for comparing nintedanib with pirfenidone.

3.5 The company evaluated 4 safety outcomes in its network meta-analysis. It reported that, compared with those receiving placebo, people receiving nintedanib were more likely to have severe gastrointestinal events (p=0.055), stop the study drug (p=0.014), and have adverse events that led to stopping the study drug (p=0.007). These differences were statistically significant. Nintedanib was associated with fewer serious cardiac events than placebo and pirfenidone, but the odds ratios were not statistically significant. Nintedanib was associated with more serious gastrointestinal events than pirfenidone (odds ratio 3.96, 95% confidence interval [CI] 1.18 to 14.51, p value not reported). The company reported that nintedanib was associated with lower rates of stopping because of adverse events than pirfenidone (odds ratio 0.88, 95% CI 0.57 to 1.37, p value not reported).

Table 2. Network meta-analysis of efficacy: scenarios used in company cost-effectiveness model (shaded cells reflect significant differences)

Comparison	Median odds ratio (95% CI), fixed effect model		
	Overall survival (NMA base case: all evidence)	**Acute exacerbations** (NMA scenario 3: excluded heterogeneous trials)	**Loss of lung function**[a] (NMA scenario 2: excluded heterogeneous trial)
Nintedanib compared with placebo	0.70 (0.45, 1.10)	0.56 (0.35, 0.89)	0.54 (0.42, 0.69)
Pirfenidone compared with placebo	0.70 (0.46, 1.05)	1.01 (0.22, 4.50)	0.69 (0.47, 1.00)
Nintedanib compared with pirfenidone[b]	1.00 (0.55, 1.85)	0.56 (0.12, 2.68)	0.78 (0.49, 1.22)

[a] Defined as an absolute decline in percent predicted FVC of over 10% by the end of the study follow-up
[b] Results of significance testing not reported
Key: CI, confidence interval; FVC, forced vital capacity; NMA, network meta-analysis

ERG comments

3.6 The ERG highlighted that the 3 nintedanib trials enrolled people with a percent predicted FVC of at least 50% and therefore do not provide evidence for people with more severe disease.

3.7 The ERG was concerned that the company did not fully explain how lung function, physical function or acute exacerbations predict the course and outcome of the disease in patients. Therefore it was unclear which specific outcomes were the most clinically meaningful.

3.8 The ERG's key concerns with the network meta-analysis were that there is the potential for bias in favour of nintedanib because the company excluded studies in some scenarios.

Cost effectiveness

3.9 The company provided a Markov model to assess the cost effectiveness of nintedanib compared with pirfenidone or best supportive care in adults with idiopathic pulmonary fibrosis. The company modelled people with a percent predicted FVC of 50% or more (although the marketing authorisation does not have a restriction related to FVC). The model used a lifetime time horizon, with a cycle length of 3 months.

3.10 The 19 health states in the model used a combination of 2 measures: percent predicted FVC (defined as approximately 10 percentage point increments) and the occurrence of an acute exacerbation. People entered the model at different health states based on percent predicted FVC and without having had an exacerbation. They could remain in the same health state or move through the model to different health states by:

- death
- loss of lung function (representing disease progression, defined as a 10 percentage point decrease in percent predicted FVC)
- exacerbation
- loss of lung function and exacerbation.

Once a person progressed to a health state with a lower percent predicted FVC it was not possible to return to a health state with better lung function. Once an exacerbation occurred, a person could not move back to a health state without exacerbation. Exacerbation health states had different health outcomes and costs than health states without exacerbation. If a person had a second exacerbation they did not move into a different health state. Instead they incurred a short term cost and disutility associated with an exacerbation. Because there was no evidence on the incidence of recurrent exacerbations, the company assumed that a person who

had at least 1 exacerbation had the same risk of another exacerbation as a person who had never had an exacerbation. Death could occur at any point in the model, or when a person's percent predicted FVC reduced to 39.9% or less.

3.11 The company modelled the baseline risks of mortality, disease progression (loss of lung function), and acute exacerbations using the results from the placebo arm of the nintedanib clinical trials (INPULSIS and TOMORROW). It based the efficacy of best supportive care on the placebo arms of the INPULSIS trials. The company applied odds ratios from its network meta-analysis to the baseline risks to estimate the relative effectiveness and safety of nintedanib and pirfenidone compared with best supportive care. To extrapolate data beyond what was available from clinical trials, the company fitted the following parametric models:

- a log logistic model to estimate overall survival
- an exponential model to estimate the probability of exacerbation and stopping medication
- a logistic regression model to predict loss of lung function.

3.12 The company included adverse events in the model if they substantially affected costs and quality-adjusted life years (QALYs), had an incidence of more than 5%, or an incidence 1.5 times greater than in the comparator arm. The company excluded the adverse event diarrhoea, even though it occurred commonly in the INPULSIS trials (reported in over 60% of people receiving nintedanib compared with 19% of people receiving placebo), because the condition was usually mild to moderate in severity and resulted in less than 5% of people stopping treatment.

3.13 The company included the following costs in its model: drug treatments (including concomitant medications), adverse events,

liver function tests, resource use (for drug acquisition, patient monitoring, treating acute exacerbations and adverse events), oxygen use, exacerbations, and end-of-life care. The company assigned utility values to each health state in the model using EQ-5D data collected in the INPULSIS trials. The model also incorporated disutilities from exacerbations and treatment-related adverse events.

3.14 Both nintedanib and pirfenidone had a confidential patient access scheme (price discount) agreed with the Department of Health. At the request of NICE, the company (Boehringer Ingelheim) provided its base-case results and sensitivity analyses using the list prices of nintedanib and pirfenidone. NICE requested that the ERG provide the results of its own exploratory analyses including the list prices, and, separately in a confidential appendix, with both discounts incorporated.

3.15 In the company's deterministic base case, best supportive care was associated with 3.27 QALYs; pirfenidone with 3.62 QALYs and nintedanib with 3.67 QALYs. Using the list prices for nintedanib and pirfenidone, nintedanib dominated pirfenidone (that is, nintedanib was more effective and was cost saving) and produced an incremental cost-effectiveness ratio (ICER) of £149,361 per QALY gained compared with best supportive care. The company did sensitivity and scenario analyses around its base case (using list prices for nintedanib and pirfenidone). The comparison between nintedanib and pirfenidone was sensitive to using stopping rules (when people stop treatment if their percent predicted FVC declines by 10% or more in 1 year).

- When the stopping rule was applied only to people receiving pirfenidone, the ICER for nintedanib was £82,784 per QALY gained compared with pirfenidone.

- When the stopping rule was applied to both the nintedanib and pirfenidone arms, the ICER for nintedanib was £17,096 per QALY gained compared with pirfenidone.

The comparison between nintedanib and best supportive care was very sensitive to estimates of mortality risk associated with treatment. Changing the baseline survival risk (by using an alternative method of extrapolation) increased the ICER by approximately:

- £91,000 per QALY gained when the company used a Weibull parametric model
- £320,000 per QALY gained when it used a Gompertz parametric model.

3.16 When the ERG applied the patient access schemes for nintedanib and pirfenidone to the company base case, pirfenidone was extendedly dominated by nintedanib and best supportive care (meaning that a combination of best supportive care and nintedanib would give more benefit than pirfenidone and would be cost saving). The ICER for nintedanib compared with best supportive care was substantially over £30,000 per QALY gained. In a pairwise comparison, the ICER for nintedanib compared with pirfenidone was between £20,000 and £30,000 per QALY gained. NICE cannot report the exact ICERs due to the confidentiality of the patient access schemes.

ERG comments and alternative base case

3.17 The ERG's clinical adviser considered that people who have had one exacerbation were at higher risk of recurrent exacerbation than those who have not had any.

3.18 The ERG had concerns about the company using a log logistic model to extrapolate overall survival data in the base-case model.

The company explained the validation of the log logistic model in response to clarification, following which the ERG was satisfied with the company's selection.

3.19 The ERG suggested that the population in the company's model may not represent those treated in clinical practice in England because it included people with percent predicted FVC of more than 80% (accounting for approximately 45% of people in the model). The ERG noted that clinical advice during the pirfenidone appraisal suggested that this FVC represents disease that is milder than would typically be treated in current practice.

3.20 The ERG suggested that the results of the company's cost-effectiveness analysis may have been biased, because the company chose a different scenario analysis from its network meta-analysis, to inform the relative effectiveness of nintedanib and pirfenidone for each different outcome.

3.21 The ERG suggested that the company model overestimated disutilities for adverse events because:

- Adverse events in the company's model last for 1 year; the ERG considered that for gastrointestinal and skin disorders the duration would be significantly shorter than this, based on published data.
- The company used a disutility of −0.118 for gastrointestinal perforations, instead of the reported value of −0.025 from published data.
- Data from a long-term open-label extension study of the CAPACITY trials of pirfenidone (the RECAP study) suggest that the incidence of rash is lower than the estimates in the company model.

- The company may have overestimated the incidence of photosensitivity, which the ERG suggested patients can prevent by avoiding sun exposure.

The ERG also noted that the disutility associated with new exacerbations that the company included in its submission (−0.14) did not match the disutility the company used in its model (0.0987).

3.22 The ERG applied its preferred assumptions to the company model to address its methodological concerns. It produced an alternative base case including:

- a restricted population with reduced lung function: percent predicted FVC of 50–79.9%
- all evidence from the company network meta-analysis, using results from the fixed effect model
- a different disutility for new exacerbations (−0.14), because of an error in the company model (0.0987)
- lower disutility and shorter duration for photosensitivity and rash.

In the ERG's alternative base case (including the patient access schemes for both drugs) nintedanib and pirfenidone were associated with similar QALYs (3.437 and 3.444 respectively) but pirfenidone dominated nintedanib (based on an incremental QALY of 0.008). The ICER for nintedanib compared with best supportive care was substantially over £30,000 per QALY gained.

3.23 Full details of all the evidence are in the Committee papers.

4 Consideration of the evidence

The Appraisal Committee reviewed the data available on the clinical and cost effectiveness of nintedanib, having considered evidence on the nature of idiopathic pulmonary fibrosis and the value placed on the benefits of nintedanib by people with the

condition, those who represent them, and clinical experts. It also took into account the effective use of NHS resources.

Clinical management

4.1 The Committee understood that idiopathic pulmonary fibrosis is a distressing illness that limits physical activity because of breathlessness, and can lead to hypoxia, pulmonary hypertension, heart failure and death. The Committee heard from clinical experts that 80% of people with idiopathic pulmonary fibrosis die from the condition or from respiratory failure. The median overall survival in the clinical trials, which excluded people with the most severe disease, was 3 to 4 years. It heard that there is no cure for idiopathic pulmonary fibrosis, although the Committee acknowledged that pirfenidone and lung transplant can be used to manage the condition. It heard however, that pirfenidone and lung transplant are not appropriate for many people due to the severity of their condition, tolerability of pirfenidone, or other comorbidities. Patient experts explained that a key aim of treatment is to slow the progression of disease, and stated that people with idiopathic pulmonary fibrosis would benefit from alternative treatment options. The Committee concluded that idiopathic pulmonary fibrosis is associated with substantial morbidity and mortality, and that there are few treatment options.

4.2 The Committee discussed the effect of acute exacerbations in idiopathic pulmonary fibrosis, which it understood to be more serious than exacerbations in other respiratory diseases. It heard from clinical experts that half of the people with idiopathic pulmonary fibrosis will die within 30 days of an acute exacerbation. People who survive an exacerbation have permanent and substantially reduced lung function (up to a 20% decline in percent predicted forced vital capacity [FVC]). Clinical experts explained

that acute exacerbations can be difficult to define in clinical trials, because they can be confused with respiratory infections, but are clearly recognised in clinical practice by experienced clinicians. The Committee heard from patient experts that preventing or delaying acute exacerbations is an important way to maintain quality of life. The Committee concluded that exacerbations are an important clinical event, but can be difficult to define, particularly in clinical trials.

4.3 The Committee considered how clinicians assess lung function in people with idiopathic pulmonary fibrosis. It understood that clinicians use a number of measures of lung function, and heard that they routinely use percent predicted FVC to guide clinical decision-making. Clinical experts noted some disadvantages with using percent predicted FVC. For example, the equations used to calculate percent predicted FVC (adjusting for age, sex and body weight) extrapolate data from middle-aged white men and may under- or overestimate the expected lung volume for current clinical practice in England. Clinical experts explained that it is difficult to know how a person's lung function would have progressed without treatment. However, the Committee acknowledged that this is not unique to idiopathic pulmonary fibrosis. The Committee heard from clinical experts that other measures of lung function (such as the 6-minute walk test distance and diffusion capacity of the lung for carbon monoxide) are less reliable than FVC. 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.

4.4 The Committee discussed the treatment options for people with idiopathic pulmonary fibrosis.

- It heard that pirfenidone is normally offered to people whose disease meets the criteria in the NICE guidance: that is, people with a percent predicted FVC of 50–80%. It heard that this group represents around half of the idiopathic pulmonary fibrosis population in the UK, but that around 30% of could not tolerate pirfenidone. Clinical experts noted that they follow the stopping rule in the NICE recommendations for pirfenidone, but explained that before withdrawing treatment they retest FVC to confirm that the 10% reduction is not temporary, as might occur with an infection.
- The Committee heard that people with a percent predicted FVC of more than 80% represent around one third of people with idiopathic pulmonary fibrosis. It heard from clinical experts that this group would be offered best supportive care because pirfenidone is not recommended in this population.
- Clinical experts explained that drug treatment might not be appropriate for people with a percent predicted FVC of less than 50% (more severe disease). The focus of treatment in this population is to maintain quality of life, and lung transplant might be explored as an option.

The Committee concluded that in clinical practice nintedanib would be appropriate for treating people with a percent predicted FVC more than 50%.

Clinical effectiveness

4.5 The Committee discussed the clinical trial evidence for nintedanib and heard that the trials reflected current clinical practice. It understood that there were inconsistencies in some of the results across the 2 phase III trials. However, it noted that a pre-planned pooled analysis showed statistically significant differences between nintedanib and placebo for loss of lung function and a

non-significant trend in favour of nintedanib for reducing acute exacerbations. The Committee heard that, based on pre-planned and post-hoc subgroup analyses, nintedanib was effective regardless of baseline FVC. It noted that the mean baseline percent predicted FVC was approximately 80% across all 3 nintedanib trials, indicating that the trials provided evidence for treating idiopathic pulmonary fibrosis in people with a percent predicted FVC above 80%. The Committee concluded that the trials provided an appropriate basis for its decision-making, and showed that nintedanib is more effective than placebo in all subgroups.

4.6 The Committee considered whether the company network meta-analysis was robust. It heard from the ERG that the company had included all relevant trials, but that the company had explored heterogeneity in the results by excluding trials in sensitivity analyses. The Committee understood that the results of the network meta-analysis informed the relative effectiveness of nintedanib and pirfenidone in the company model. It heard that the company used the results of different sensitivity analyses (that is, using data from different sets of trials) for different outcomes in the model (see section 3.4). The Committee agreed that this introduced a potential bias in favour of nintedanib because the analyses chosen by the company showed nintedanib more favourably than the results from analyses including all trials, and concluded that the same trials should be included for all outcomes. It agreed with excluding one of Japanese studies (SP2) as it considered it to be an outlier. However, the Committee concluded that the network meta-analysis provided an appropriate basis for its decision-making.

4.7 The Committee considered the effectiveness of nintedanib and discussed the results of the company network meta-analysis. The Committee heard that nintedanib and pirfenidone provided similar

benefits in survival, and that neither treatment produced a statistically significant difference compared with placebo. For loss of lung function and acute exacerbations, the Committee heard that there was a trend favouring nintedanib over pirfenidone, but was unclear whether this was statistically significant because the company did not provide results of significance testing. The Committee concluded that the clinical effectiveness of nintedanib is similar to, if not slightly better than, pirfenidone.

4.8 The Committee considered the position of nintedanib in the treatment pathway and heard from clinical experts that there was no evidence for using nintedanib and pirfenidone at the same time or sequentially, and no data to identify people who are more or less likely to respond to either drug.

Cost effectiveness

4.9 The Committee considered whether the company model, in which health states were based on percent predicted FVC and occurrence of acute exacerbations, accurately represents the progression of idiopathic pulmonary fibrosis. It heard from patient experts that idiopathic pulmonary fibrosis is a progressive disease which does not improve, and noted that the model reflects this. The Committee heard from the company that exacerbations increased the rate of disease progression (that is, loss of lung function) in the model, which the Committee considered appropriate. 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. It considered that, based on the network meta-analysis, nintedanib was probably more effective than pirfenidone at reducing exacerbations, and that this benefit may not be fully reflected in the

model. The ERG stated that it considered the methods applied in the company's economic analyses to be generally appropriate. The Committee concluded that a model based on percent predicted FVC and exacerbations was appropriate for decision-making.

4.10 The Committee discussed the population included in the economic model. It noted that the company analyses included a population with a percent predicted FVC more than 50%, and that the ERG modelled a restricted population with a percent predicted FVC of 50–79.9%. The Committee recognised that the ERG had assumed that people with a percent predicted FVC of 80% or more do not present in clinical practice. The Committee, however, heard the opposite from the clinical experts who stated that people with a percent predicted FVC of 80% or more represented a third of people with idiopathic pulmonary fibrosis (see section 4.4), and that the relevant comparator for this population is best supportive care. The Committee agreed that to compare nintedanib with best supportive care, it would have preferred to see a model of only people with a percent predicted FVC of more than 80%. But, having not seen this population modelled, it concluded that the company model was appropriate for its decision-making. The Committee concluded that to compare nintedanib and pirfenidone, it was appropriate to model a population with a percent predicted FVC of 50–80%.

4.11 The Committee considered the extrapolation of overall survival in the company model. It noted that the company and the ERG had used the log logistic curve in their base-case analyses. The Committee discussed whether this curve was appropriate given that it has a long tail. It noted that using other methods to extrapolate survival substantially increased the incremental cost-effectiveness ratio (ICER) for nintedanib compared with best supportive care. The Committee heard from the clinical experts that

the median survival estimated with the log logistic curve generally reflected the natural history of treated disease. The Committee agreed that the survival modelling was uncertain, but noted that it had little effect on the ICER when comparing nintedanib with pirfenidone because the company had assumed equal survival with these drugs. Any differences in quality-adjusted life year (QALY) gain between nintedanib and pirfenidone were therefore derived from differences in quality of life. The Committee concluded that the log logistic curve was sufficient for decision-making, but recognised that the ICER for nintedanib compared with best supportive care, when using the log logistic curve, may be an underestimate.

4.12 The Committee discussed the utility values in the company's model. It approved of the company using trial-based EQ-5D data to estimate health-state utility values. The Committee expressed some concern that the company did not include a disutility for diarrhoea in the model, because this is a common adverse event with nintedanib that it considered would worsen quality of life. It heard from the company that including a diarrhoea-related disutility would not affect the model results because the event was mild-to-moderate and led to fewer than 5% of people stopping treatment in the nintedanib clinical trials. However, the Committee did not agree that diarrhoea would have no clinical impact.

4.13 The Committee considered the changes to adverse-event-related disutilities in the ERG's alternative base case. It heard from clinical experts that the ERG's estimate of rash-related disutility (based on data from the RECAP study) was inaccurate because the RECAP study underestimated the incidence of rash by 50%. Clinical experts agreed with the ERG's choice of a lower incidence of photosensitivity with pirfenidone, and stated that it was reasonable to assume people have or tolerate adverse events in the model for approximately 1 month (as in the ERG alternative base case) rather

than for 1 year (as in the company base case). The Committee concluded that that the estimate of rash-related disutility in the company base case was more appropriate than the ERG's estimate, but that the ERG's alternative base case provided more accurate estimates of other disutilities.

4.14 The Committee compared the company model with the model submitted for the NICE technology appraisal 282 of pirfenidone, for external validation. It noted that the nintedanib model produced different results to the pirfenidone model, when comparing pirfenidone with best supportive care. The ERG could not fully compare the 2 models due to confidentiality, but explained some key differences. For example, treatment effect in the pirfenidone model was based on FVC and the 6-minute walk test distance. The Committee, however, understood from clinicians that the 6-minute walk test distance was an unreliable measure. The pirfenidone model did not include acute exacerbations, which the Committee understood to be an important clinical event in idiopathic pulmonary fibrosis. The pirfenidone model used a mapping algorithm to calculate utility values, whereas the nintedanib model included trial-based EQ-5D data. The Committee concluded that it could not compare the models fully, but that the current model for nintedanib was robust, and appropriate for decision-making.

4.15 The Committee discussed the most plausible ICER. It noted that the most plausible scenario would include assumptions from both the ERG and company base cases for:

- the disutilities for adverse events (see section 4.12)
- estimating the overall survival, acute exacerbations, loss of lung function, adverse events, and stopping treatment odds ratios for nintedanib compared with pirfenidone (see section 4.6)

- the population modelled (that is, limiting this to people with a percent predicted FVC of 50–80% for the comparison with pirfenidone).

The Committee considered the population for which pirfenidone and best supportive care were comparators (that is, people with a percent predicted FVC of 50–80%). It heard that when the ERG included the patient access schemes for nintedanib and pirfenidone in the company base case, pirfenidone was extendedly dominated by nintedanib and best supportive care (meaning that a combination of best supportive care and nintedanib would give more benefit than pirfenidone and would cost less). It understood that in the ERG alternative base case (including the patient access schemes for both drugs) pirfenidone dominated nintedanib (that is, pirfenidone was more effective and cost less). The Committee agreed that the difference in the results of the company base case and ERG alternative base case could be because of the small incremental costs and very small incremental QALYs. Nintedanib produced 0.05 more QALYs than pirfenidone in the company base case, and pirfenidone produced 0.008 more QALYs that nintedanib in the ERG alternative base case. The Committee agreed that nintedanib was likely to be more clinically effective than pirfenidone (see section 4.7), and that this may not be fully reflected in the cost-effectiveness results as the model was not sensitive to acute exacerbations (see section 4.9). The Committee understood that current NICE guidance for pirfenidone recommends that pirfenidone treatment should be stopped if there is evidence of disease progression (a decline in percent predicted FVC of 10% or more in any 12-month period). Applying this stopping rule for pirfenidone and not nintedanib substantially reduced the cost effectiveness of nintedanib. The Committee agreed that nintedanib and pirfenidone were associated with comparable costs and

benefits in the model and therefore nintedanib could be considered cost effect compared with pirfenidone in people with a percent predicted FVC of 50–80% when the stopping rule was applied.

4.16 The Committee considered the population with a percent predicted FVC of more than 80%, for whom the comparator is best supportive care. The Committee would have preferred to see a model of only people with a percent predicted FVC of more than 80%, but in the absence of this the Committee considered the company model of people with a percent predicted FVC of more than 50%, and the ERG model of people with a percent predicted FVC of 50–79.9%. In both the company's base case and the ERG's alternative base case (both including the discount for nintedanib), the ICER for nintedanib was substantially over £30,000 per QALY gained. NICE cannot report the ICERs because of the confidentiality of the patient access schemes. The Committee concluded that the ICERs for nintedanib as a replacement for best supportive care in people with a percent predicted FVC of more than 80% were not within the range considered to be a cost-effective use of NHS resources.

4.17 The Committee heard from patient experts and Committee members that nintedanib was innovative in its potential to make a significant and substantial impact on health-related benefits. After looking at the evidence and hearing a patient expert compare his experience with pirfenidone and nintedanib, the Committee agreed that nintedanib may be considered innovative because the benefits of its improved tolerability profile compared with pirfenidone are not fully captured in the model. The patient expert valued his opportunity to pursue outdoor activities while receiving nintedanib, which he had been unable to do when receiving pirfenidone because of the associated photosensitivity. The Committee heard differing views about the value of the reduced dosing frequency with nintedanib compared with pirfenidone, and considered this a

small advantage. The Committee concluded that there may be some additional gains in health-related quality of life over those already included in the QALY calculations, further supporting its recommendation that nintedanib should be offered as an alternative to pirfenidone.

4.18 The Committee considered whether it should take into account the consequences of the Pharmaceutical Price Regulation Scheme (PPRS) 2014, and in particular the PPRS payment mechanism, when appraising nintedanib. The Appraisal Committee noted NICE's position statement in this regard, and accepted the conclusion 'that the 2014 PPRS payment mechanism should not, as a matter of course, be regarded as a relevant consideration in its assessment of the cost effectiveness of branded medicines'. The Committee heard nothing to suggest that there is any basis for taking a different view with regard to the relevance of the PPRS to this appraisal of nintedanib. It therefore concluded that the PPRS payment mechanism was not applicable for considering the cost effectiveness of nintedanib.

Summary of Appraisal Committee's key conclusions

TAXXX	**Appraisal title: Nintedanib for treating idiopathic pulmonary fibrosis**	**Section**
Key conclusion		
Nintedanib 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 company provides nintedanib with the discount agreed in the patient access scheme		1.1, 1.2, 4.15, 4.16

Treatment with nintedanib that is recommended according to 1.1 should be is stopped if disease progresses (a confirmed decline in percent predicted FVC of 10% or more) in any 12 month period. The incremental cost-effectiveness ratios (ICERs) for nintedanib as a replacement for best supportive care in people with a percent predicted FVC of more than 80% were not within the range considered to be a cost-effective use of NHS resources.		
Current practice		
Clinical need of patients, including the availability of alternative treatments	There is no cure for idiopathic pulmonary fibrosis; median overall survival is 3–4 years and 80% die from the disease or respiratory failure. Lung transplant is effective in improving survival but few people are eligible. Pirfenidone is normally offered to people whose disease meets the criteria in the NICE guidance: that is, people with a percent predicted FVC of 50–80%. But 30% of people with idiopathic pulmonary fibrosis in the UK cannot tolerate pirfenidone. People with a percent predicted FVC of more than 80% (who represent around one third of people with the disease) are offered best supportive care. Drug treatment might not be appropriate for people with more severe disease (a percent predicted FVC of less than 50%).	4.1, 4.4
The technology		
Proposed benefits of	Nintedanib is likely to be more clinically effective than pirfenidone, for example at	3.5, 4.7, 4.9,

the technology How innovative is the technology in its potential to make a significant and substantial impact on health-related benefits?	preventing acute exacerbations. Nintedanib may be considered innovative, because the benefits of its improved tolerability profile and reduced dosing frequency compared with pirfenidone are not fully captured in the model.	4.17
What is the position of the treatment in the pathway of care for the condition?	Nintedanib is an option for people with a percent predicted FVC of 50% or more.	4.7
Adverse reactions	The most frequently reported adverse reaction associated with the use of nintedanib is diarrhoea. Clinical and patient experts indicated that people may tolerate nintedanib better than pirfenidone.	2.2, 4.17
Evidence for clinical effectiveness		
Availability, nature and quality of evidence	The clinical evidence for nintedanib came from 3 multicentre, double-blind, randomised controlled trials comprising 2 phase III trials (INPULSIS 1 and INPULSIS 2) and a phase IIb dose-ranging trial (TOMORROW). In the absence of head-to-head trials of nintedanib and pirfenidone, the company submitted a network meta-analysis, which generally provided an appropriate basis for decision-making.	3.1, 4.5, 4.6

Relevance to general clinical practice in the NHS	The 3 nintedanib trials enrolled people with a percent predicted FVC of at least 50% and therefore provided an appropriate basis for decision-making because clinical experts stated that people with severe disease would not likely be treated with nintedanib.	3.6, 4.4, 4.5
Uncertainties generated by the evidence	The company used the results of different sensitivity analyses from its network meta-analysis (using data from different sets of trials) for different outcomes in the model. The trials selected by the company potentially biased the results in favour of nintedanib. The Committee would have preferred the company to have used the same trials for all end points in the model, excluding only the SP2 study of pirfenidone (to reduce heterogeneity).	3.4, 4.6
Are there any clinically relevant subgroups for which there is evidence of differential effectiveness?	There are no subgroups for which there is evidence of differential effectiveness.	3.3, 4.5
Estimate of the size of the clinical effectiveness including strength of supporting evidence	Clinical trials showed that nintedanib is more effective than placebo in all subgroups. There were inconsistencies in some of the results across the 2 phase III trials. However, a pre-planned pooled analysis showed statistically significant differences between nintedanib and placebo for loss of lung function and a non-significant trend in favour	4.5, 4.7

	of nintedanib for reducing acute exacerbations. The clinical effectiveness of nintedanib is similar to, if not slightly better than, pirfenidone based on the results of the network meta-analysis.	
Evidence for cost effectiveness		
Availability and nature of evidence	The company model (based on percent predicted FVC and occurrence of acute exacerbations) was appropriate for decision-making. It was appropriate not to include people with a percent predicted FVC of less than 50% in the model.	4.4, 4.9, 4.10
Uncertainties around and plausibility of assumptions and inputs in the economic model	The Committee was concerned that: • the results of the model were not sensitive to changes in the rate of exacerbations (it considered that nintedanib was probably more effective than pirfenidone at reducing exacerbations) • using the log logistic curve to extrapolate survival data could underestimate the true ICER for nintedanib compared with best supportive care • the company did not include a disutility for diarrhoea in the model • the company had overestimated some adverse event-related disutilities (this was amended in the Evidence Review Group's (ERG) alternative base case) • the company's use of different network	4.6, 4.9, 4.10, 4.11, 4.12, 4.13, 4.15

	meta-analysis scenarios for different outcomes in the model could bias the results in favour of nintedanib (this was amended, in part, in the ERG's alternative base case) • neither the company nor the ERG provided a model of only people with percent predicted FVC over 80%, for the comparison of nintedanib with best supportive care in people for whom pirfenidone is not recommended.	
Incorporation of health-related quality-of-life benefits and utility values Have any potential significant and substantial health-related benefits been identified that were not included in the economic model, and how have they been considered?	The Committee were concerned that the company did not include a disutility for diarrhoea in the model, because this is a common adverse event with nintedanib that the Committee considered would worsen quality of life. The company overestimated some adverse-event-related disutilities, but these were amended in the ERG's alternative base case. The Committee concluded that there were some additional gains in health-related quality of life over those already included in the quality-adjusted life year (QALY) calculations (its innovative mechanism of action, tolerability and reduced dosing frequency). These further supported the Committee's recommendation that nintedanib should be offered as an alternative to pirfenidone.	4.12, 4.17
Are there specific	Nintedanib had different comparators (either	4.10,

groups of people for whom the technology is particularly cost effective?	pirfenidone or best supportive care) for different subgroups according to percent predicted FVC. Nintedanib was cost effective compared with pirfenidone, but not when compared with best supportive care. As pirfenidone is a comparator for a subgroup (those with FVC 50–80%) nintedanib was only cost effective for this group.	4.15, 4.16
What are the key drivers of cost effectiveness?	Any differences in QALY gain between nintedanib and pirfenidone were derived from differences in quality of life (because the survival gain with each drug was the same). Nintedanib and pirfenidone were associated with very similar QALYs (3.437 and 3.444 respectively). The cost effectiveness of nintedanib compared with best supportive care was sensitive to survival rates.	3.22, 4.11, 4.15
Most likely cost-effectiveness estimate (given as an ICER)	For people with a percent predicted FVC of 50–80%: • Nintedanib is likely to be more clinically effective than pirfenidone, which may not be fully reflected in the cost-effectiveness results. The Committee agreed that nintedanib and pirfenidone were associated with comparable costs and benefits in the model and therefore nintedanib could be considered cost effective compared with pirfenidone in people with a percent	4.15, 4.16

	predicted FVC of 50–80% when the stopping rule was applied. For people with a percent predicted FVC of more than 80%: • The ICER for nintedanib compared with best supportive care (with patient access schemes applied) was substantially over £30,000 per QALY gained. NICE cannot report the exact ICERs because of the confidentiality of the patient access schemes.	
Additional factors taken into account		
Patient access schemes (PPRS)	Not applicable	4.18
End-of-life considerations	Not applicable	–
Equalities considerations and social value judgements	Not applicable	–

5 Implementation

5.1 Section 7(6) of the National Institute for Health and Care Excellence (Constitution and Functions) and the Health and Social Care Information Centre (Functions) Regulations 2013 requires clinical commissioning groups, NHS England and, with respect to their public health functions, local authorities to comply with the

recommendations in this appraisal within 3 months of its date of publication.

5.2 The Welsh Assembly Minister for Health and Social Services has issued directions to the NHS in Wales on implementing NICE technology appraisal guidance. When a NICE technology appraisal recommends the use of a drug or treatment, or other technology, the NHS in Wales must usually provide funding and resources for it within 3 months of the guidance being published.

5.3 When NICE recommends a treatment 'as an option', the NHS must make sure it is available within the period set out in the paragraphs above. This means that, if a patient has idiopathic pulmonary fibrosis and the doctor responsible for their care thinks that nintedanib is the right treatment, it should be available for use, in line with NICE's recommendations.

5.4 The Department of Health and Boehringer Ingelheim have agreed that nintedanib will be available to the NHS with a patient access scheme which makes it available with a discount. The size of the discount is commercial in confidence. It is the responsibility of the company to communicate details of the discount to the relevant NHS organisations. Any enquiries from NHS organisations about the patient access scheme should be directed to [NICE to add details at time of publication]

5.5 NICE has developed tools [link to www.nice.org.uk/guidance/TAXXX] to help organisations put this guidance into practice (listed below). [NICE to amend list as needed at time of publication]

- Slides highlighting key messages for local discussion.
- Costing template and report to estimate the national and local savings and costs associated with implementation.

- Implementation advice on how to put the guidance into practice and national initiatives that support this locally.
- A costing statement explaining the resource impact of this guidance.
- Audit support for monitoring local practice.

6 Related NICE guidance

Details are correct at the time of consultation and will be removed when the final guidance is published. Further information is available on the NICE website.

Published

- Idiopathic pulmonary fibrosis: The diagnosis and management of suspected idiopathic pulmonary fibrosis. NICE guideline 163 (2013).
- Pirfenidone for treating idiopathic pulmonary fibrosis. NICE technology appraisal guidance 282 (2013). Guidance under review, publication expected May 2016.

7 Proposed date for review of guidance

7.1 The recommendations provided in this guidance are directly linked to the recommendations in TA282 pirfenidone for treating idiopathic pulmonary fibrosis. Therefore, NICE is considering options for reviews of both this guidance and TA282. NICE welcomes comment in consultation responses.

Amanda Adler
Chair, Appraisal Committee
September 2015

8 Appraisal Committee members, guideline representatives and NICE project team

Appraisal Committee members

The Appraisal Committees are standing advisory committees of NICE. Members are appointed for a 3-year term. A list of the Committee members who took part in the discussions for this appraisal appears below. There are 4 Appraisal Committees, each with a chair and vice chair. Each Appraisal Committee meets once a month, except in December when there are no meetings. Each Committee considers its own list of technologies, and ongoing topics are not moved between Committees.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each Appraisal Committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Dr Amanda Adler (Chair)
Consultant Physician, Addenbrooke's Hospital

Professor Ken Stein (Vice Chair)
Professor of Public Health, University of Exeter Medical School

Dr Ray Armstrong
Consultant Rheumatologist, Southampton General Hospital

Dr Jeff Aronson
Reader in Clinical Pharmacology, University Department of Primary Health Care, University of Oxford

Professor John Cairns

Professor of Health Economics Public Health and Policy, London School of Hygiene and Tropical Medicine

Mr Matthew Campbell-Hill

Lay member

Mr David Chandler

Lay member

Mr Mark Chapman

Health Economics and Market Access Manager, Medtronic UK

Professor Imran Chaudhry

Lead Consultant Psychiatrist and Deputy Associate Medical Director, Lancashire Care NHS Foundation Trust

Professor Daniel Hochhauser

Consultant in Medical Oncology, UCL Cancer Institute

Dr Neil Iosson

Locum General Practitioner

Mrs Anne Joshua

NHS 111 Pharmacy Lead, Patients and Information, NHS England

Dr Sanjay Kinra

Reader in Clinical Epidemiology and Honorary Consultant in Paediatrics, London School of Hygiene and Tropical Medicine and University College London NHS Hospitals Trust

Dr Miriam McCarthy

Consultant, Public Health, Public Health Agency, Northern Ireland

Mr Christopher O'Regan

Head of Health Technology Assessment & Outcomes Research, Merck Sharp & Dohme

Dr John Pounsford

Consultant Physician, Frenchay Hospital, Bristol

Dr Danielle Preedy

Lay Member

Ms Marta Soares

Research Fellow, Centre for Health Economics, University of York

Dr Nicky Welton

Senior Lecturer in Biostatistics/Health Technology Assessment, University of Bristol

NICE project team

Each technology appraisal is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the appraisal) and a project manager.

Sophie Laurenson

Technical Lead(s)

Jeremy Powell

Project Manager

9 Sources of evidence considered by the Committee

A. The Evidence Review Group (ERG) report for this appraisal was prepared by Southampton Health Technology Assessments Centre (SHTAC):

- Cooper K, Kalita N, Rose M, et al. Nintedanib for treating idiopathic pulmonary fibrosis: A Single Technology Appraisal, July 2015

B. The following organisations accepted the invitation to participate in this appraisal as consultees and commentators. They were invited to comment on the draft scope, the ERG report and the appraisal consultation document (ACD). Organisations listed in I were also invited to make written submissions. Organisations listed in II and III had the opportunity to make written submissions. Organisations listed in I, II and III also have the opportunity to appeal against the final appraisal determination.

I. Company:

- Boehringer Ingelheim

II. Professional/expert and patient/carer groups:

- Action for Pulmonary Fibrosis
- British Thoracic Society
- Royal College of Nursing
- Royal College of Pathologists
- Royal College of Physicians
- United Kingdom Clinical Pharmacy Association

III. Other consultees:

- Department of Health
- NHS England
- NHS Nottingham City CCG
- Welsh Government

IV. Commentator organisations (did not provide written evidence and without the right of appeal):

- Department of Health,Social Services and Public Safety for Northern Ireland
- Healthcare Improvement Scotland
- Roche Products

C. The following individuals were selected from clinical expert and patient expert nominations from the consultees and commentators. They gave their expert personal view on nintedanib by attending the initial Committee discussion and providing a written statement to the Committee. They are invited to comment on the ACD.

- Dr Toby Maher, Consultant Respiratory Physician, nominated by NHS England – clinical expert
- Phillip Lloyd Mayers, Specialist Respiratory and ILD Pharmacist, nominated by United Kingdom Clinical Pharmacy Association – clinical expert
- Michael Bray, Chair of trustees for Action for Pulmonary Fibrosis, nominated by Action for Pulmonary Fibrosis – patient expert
- Peter Burns, Secretary of Papworth IPF patient support group, nominated by Pulmonary Fibrosis Trust – patient expert

D. Representatives from the following company attended Committee meetings. They contributed only when asked by the Committee chair to clarify specific issues and comment on factual accuracy.

- Boehringer Ingelheim