NICE National Institute for Health and Care Excellence

For public – no confidential information included

Ustekinumab for treating moderately to severely active ulcerative colitis [ID1511]
Lead team presentation – part 1

Mark Upton, Pam Rees and Mohit Sharma
ERG: Southampton Health Technology Assessments Centre (SHTAC)
Technical team: Brian Shine, Juliet Kenny, Joanna Richardson, Janet Robertson
Company: Janssen
December 2019

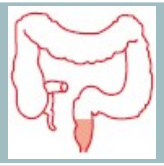
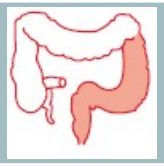
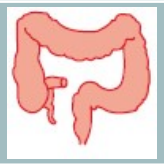
© NICE 2018. All rights reserved. Subject to notice of rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

Ulcerative colitis (UC)

- Lifelong, progressive disease characterised by relapsing and remitting episodes of inflammation of the rectal and colonic mucosa; tiny ulcers develop on the surface of the lining of the colon and these may bleed and produce pus
- Around 146,000 people in England have UC (about 52% have moderate to severe)
- Cause unknown - hereditary, infectious and immunological factors proposed as possible causes
- Can develop at any age; peak incidence between 15 and 25 years; second, smaller peak between 55 and 65 years
- Symptoms: bloody diarrhoea, colicky abdominal pain, urgency and tenesmus (recurrent feeling of needing to empty the bowel). Some patients may have extra-intestinal manifestations involving joints, eyes, skin and liver
  - can recur or the disease can go into remission for months or even years: around 50% of people will have at least one relapse per year; about 80% of these are mild to moderate and about 20% are severe
- Complications: include haemorrhage, perforation, stricture formation, abscess formation and anorectal disease. People with long-standing disease have an increased risk of bowel cancer

Classification

UC is classified according to its maximal extent seen on colonoscopy

Distribution	Description	% presenting with this distribution
	Proctitis: Involvement limited to rectum	30-60%
	Left-sided Colitis: Involvement limited to left portion of colon; extends from rectum up colon and stops at splenic flexure (point where the colon bends)	16-45%
	Extensive pancolitis: Involves inflammation of entire colon	14-47%

Source: CS, section B.1.3.1, table 3, figure 6 and p16

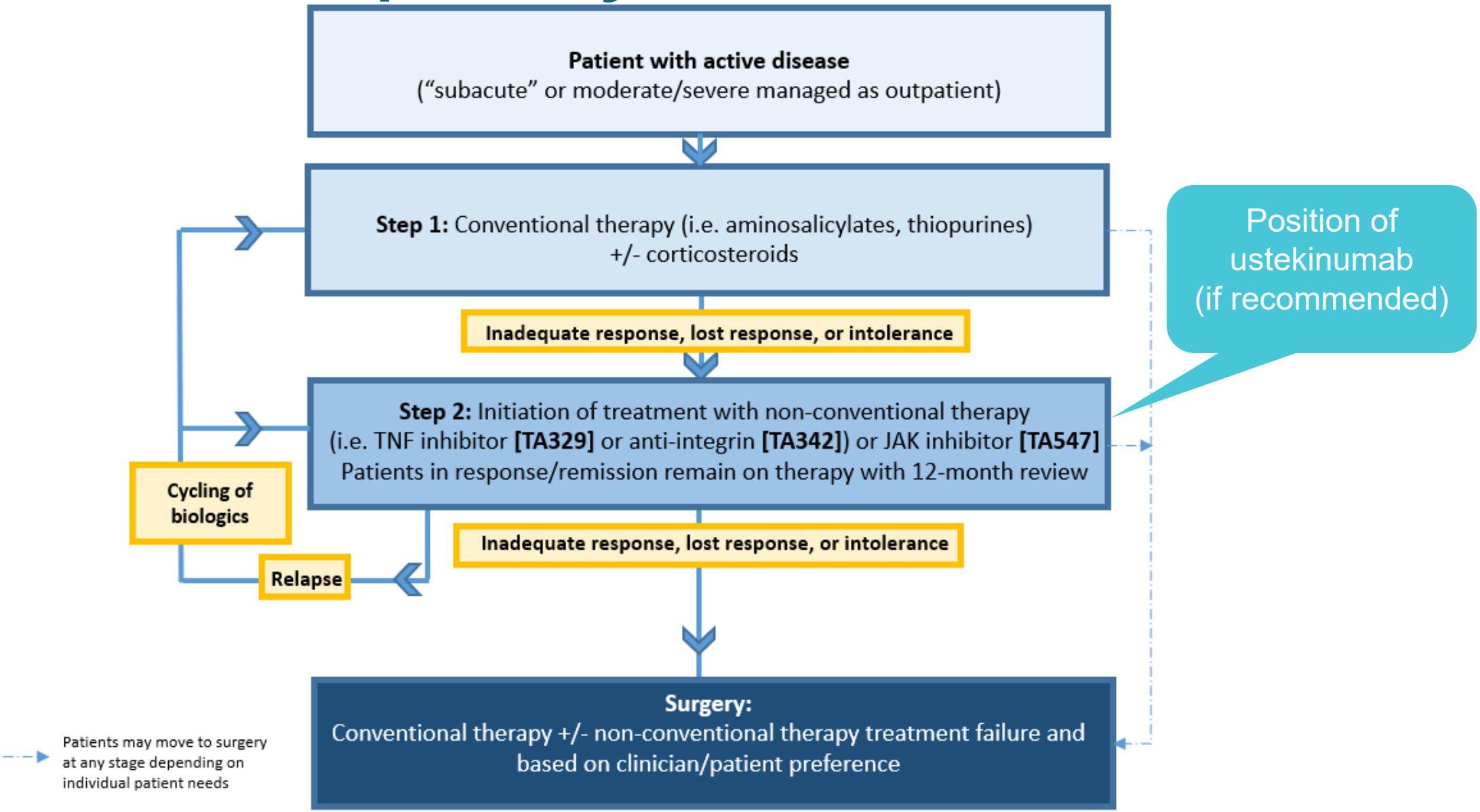
No consensus or validated definitions for severity stages; Mayo score is typically used to measure disease activity in clinical trials (0-12 points, consisting of four sub-scores)

Mayo Index	0	1	2	3
Stool frequency	Normal	1-2/day – normal	3-4/day – normal	≥5/day – normal
Rectal bleeding	None	Streaks	Obvious	Mostly blood
Mucosa (i.e. endoscopic findings)	Normal	Mild friability	Moderate friability	Spontaneous bleeding
PGA	Normal	Mild	Moderate	Severe

Source: CS, section B.1.3.1, table 4 | Abbreviations: PGA, Physician global assessment

[redacted]

Treatment pathway



Source: CS, section B.1.3.3, figure 9

Abbreviations: JAK = janus kinase; TA = technology appraisal; TNF = tumor necrosis factor

Existing guidance

Number	Class	Intervention	Population covered by recommendations
TA329 (Feb 2015)	Tumour necrosis factor [TNF] alpha inhibitor	Infliximab Adalimumab Golimumab (MTA)	Adults with moderately to severely active UC whose disease has responded inadequately to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who cannot tolerate, or have medical contraindications for, such therapies
TA342 (Jun 2015)	Anti-integrin agent	Vedolizumab	Adults with moderately to severely active ulcerative colitis who have had an inadequate response with, or lost response to, or were intolerant to either conventional therapy or a tumour necrosis factor-alpha antagonist
TA547 (Nov 2018)	Janus kinase (JAK) inhibitor	Tofacitinib	Adults with moderately to severely active ulcerative colitis when conventional therapy or a biological agent cannot be tolerated, or the disease has responded inadequately or lost response to treatment

The technology: Ustekinumab

Marketing authorisation September 2019	Treatment of adult patients with moderately to severely active ulcerative colitis who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a biologic or have medical contraindications to such therapies.
Method of administration and dosage	Induction: intravenous weight-based dose (aligns to a dose of approximately 6mg/kg) Maintenance: subcutaneous injection; fixed dose of 90mg • first dose given at week 8 following induction. After this, dosing every 12 weeks is recommended • Patients who have not shown adequate response 8 weeks after the first subcutaneous dose (week 16), may receive a second subcutaneous dose at this time to allow for delayed response • Patients who lose response on dosing every 12 weeks may benefit from an increase in dosing frequency to every 8 weeks • Patients may subsequently be dosed every 8 weeks or every 12 weeks according to clinical judgment
Additional tests or investigations	No additional tests or investigations are expected (as compared to other currently available biologic therapies)
List price and average cost of a course of treatment	130mg vial concentrate for solution for infusion: £2,147; 90mg vial solution for injection: £2,147 (Annual treatment costs: induction year: £14,482; maintenance Year 2 and onwards: £9,304)
Commercial arrangements	Commercial Medicines Unit (CMU) arrangement

[redacted]

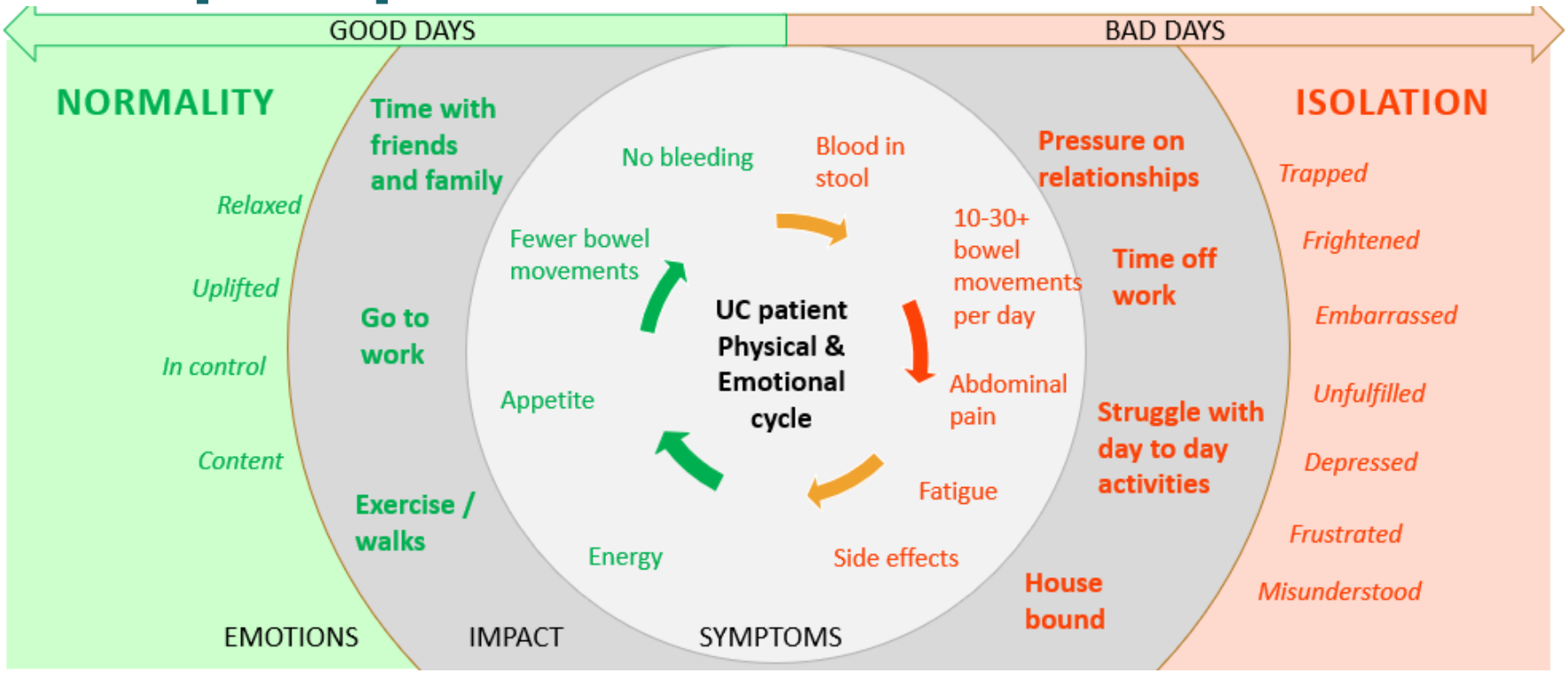
Source: CS, section B.1.2, table 2

Decision problem

- The company submission reflects the NICE scope with the following exceptions:
  - The NICE scope is not specific to any age group but the company have only submitted evidence for adults in line with the anticipated marketing authorisation
  - No evidence has been submitted that included populations with prior exposure to tofacitinib
- The NICE scope population includes all patients with moderately to severely active UC and suggests subgroup analyses be considered based on prior exposure to biologic agents. No estimates for the cost effectiveness of ustekinumab in the overall population have been provided by the company. ICERs are provided for subgroups defined by biologic failure status

[redacted]

Patient perspective



Source: CS, section B.1.3.2, figure 8

I was constantly ill over a period of years, I had my relationship break down

I can't take my children to the park, for a walk or play date […] it is simply not possible for me to go out when I may need to open my bowels with no warning

I have suffered with UC for 13 years. It's always been moderate to severe. I have tried all drugs including all biologics. All failed after a while

When I am unwell, I struggle with extreme tiredness and extended periods in the bathroom which makes my working life very difficult. I work in construction so spend a lot of time away from toilets

[redacted]

Key clinical issues

- UNIFI trial (technical report [TR] issue 1)
  - Does the committee agree that issues raised at technical engagement are resolved?
  - Do the results for non-ITT groups impact interpretation of ITT results?
- Exclusion of trials conducted in Asia from the network meta-analyses (NMA) (TR issue 2)
  - Is the committee willing to accept the current analyses that exclude the trials conducted in Asia for decision making?
- Dose regimen pooling in the maintenance phase (TR issue 7)
  - Does the committee have a preference between the company's or ERG's approach to dose pooling?
- Synthesising maintenance phase trial data - which method is most appropriate? (TR issues 4 & 5)
  - Is the ERG's maintenance only NMA scenario analyses relevant?
  - Does the committee have a preference between the company's or ERG's choice of clinical inputs?

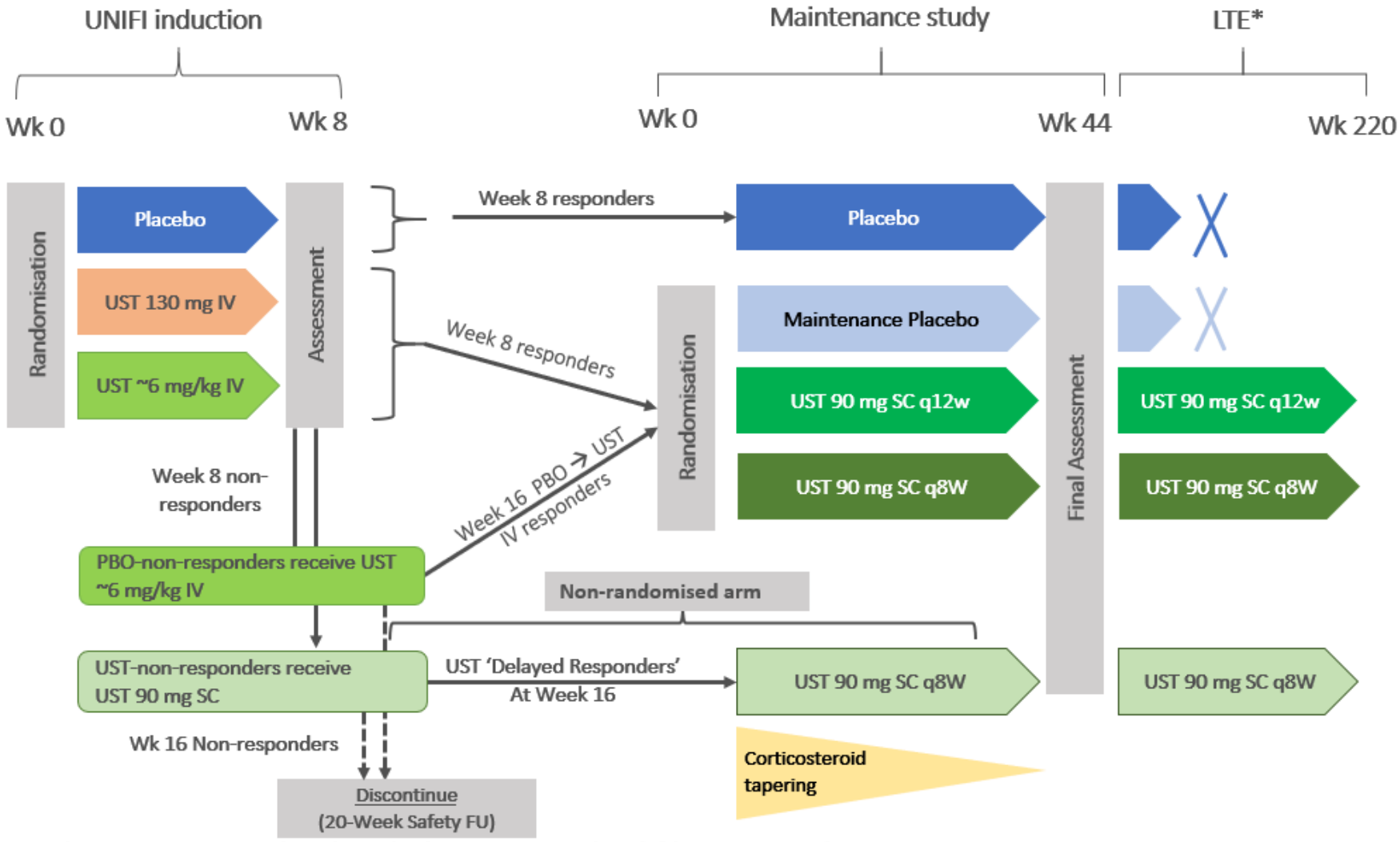
[redacted]

Clinical evidence

- The company's pivotal trial, UNIFI (NCT02407236), compared ustekinumab against placebo for treating patients with moderately to severely active UC
- UNIFI included both an induction study (outcomes measured at week 8 and week 16) and a maintenance study (outcomes measured at week 44)
- The population was stratified by various factors including biologic failure status subgroups
- Like most other recent UC trials UNIFI had a re-randomised design, meaning that patients who had a response to induction therapy were subject to re-randomisation prior to the maintenance phase. The design differs from that of earlier trials of UC therapies where a 'treat through' approach was taken i.e. treatments are randomly allocated at the start of the induction phase and participants continue to receive these treatments throughout the maintenance phase. The company notes that the newer trial designs based on response aim to reduce patients' exposure to placebo treatments that are ineffective and are considered more ethical than treat-through designs
- Clinical remission was defined as Mayo score of ≤2 points, with no individual scores >1 was the primary outcome for both the induction and maintenance studies
- Health-related quality of life was measured in the UNIFI trial primarily using the IBDQ, SF-36, EQ-5D (5L version) and the Work Productivity and Activity Impairment Questionnaire-General Health (WPAI-GH)

[redacted]

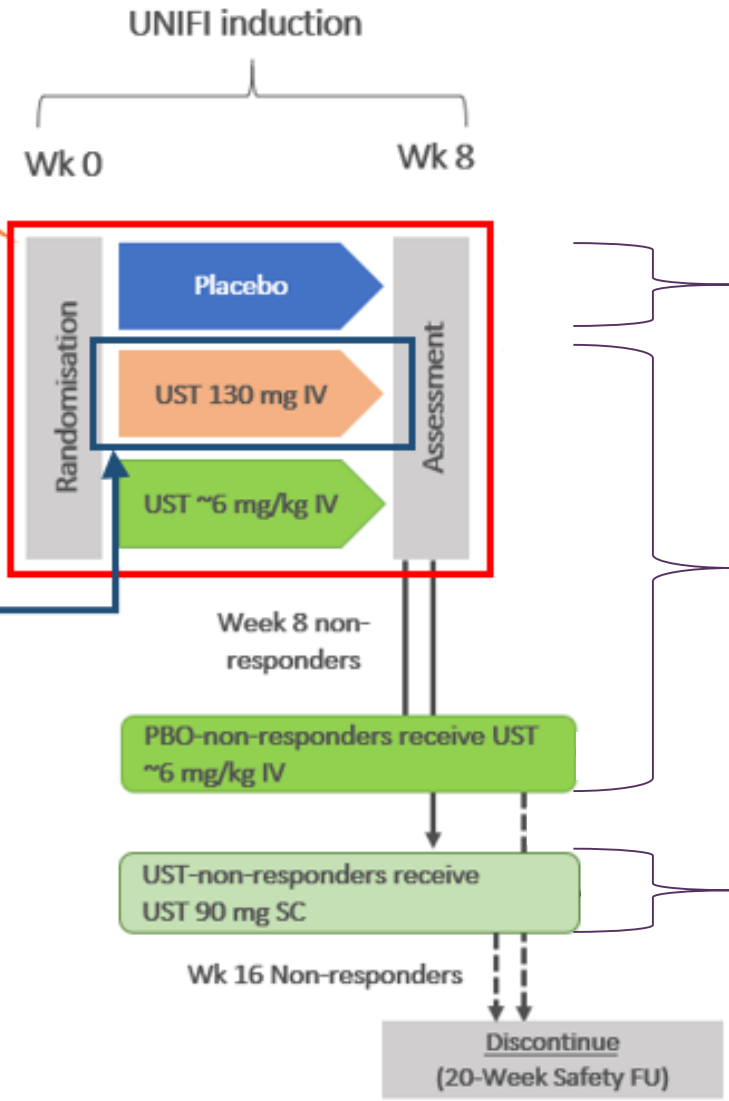
UNIFI trial design



*Patients will continue to receive the same treatment regimen during the LTE that they were receiving at the end of the maintenance study
Note: Conventional therapy is the background treatment for patients on placebo and ustekinumab.

UNIFI trial design (induction)

Induction study ITT population



Dose not covered by draft SPC but patients who responded to it were included in maintenance ITT

Responders progress to maintenance as a non-randomized, non-ITT population treated with placebo

Responders progress to maintenance and become ITT population; patients are re-randomized to either UST 90 mg q8w, UST 90 mg q12w or placebo

Responders progress to maintenance as a further, non-randomized non-ITT population of 'delayed responders'; all patients receive 90 mg q8w

Key trial results (induction ITT)

Included in company model via induction NMA

	Overall population (induction ITT)			Non-biologic failure population			Biologic failure population		
End point	**PBO N=319**	**6mg/kg (p-value)[a] N=322**	**130mg (p-value) N=320**	**PBO N=158**	**6mg/kg (p-value)[a] N=156**	**130mg (p-value) N=156**	**PBO N=161**	**6mg/kg (p-value)[a] N=166**	**130mg (p-value) N=164**
Clinical remission	5.3%	15.5% (<0.001)	15.6% (<0.001)	9.5%	18.6% (0.022)	19.9% (0.009)	1.2%	12.7% (<0.001)	11.6% (<0.001)
Clinical response[b]	31.3%	61.8% (<0.001)	51.3% (<0.001)	35.4%	66.7% (<0.001)	57.7% (<0.001)	27.3%	57.2 (<0.001)	45.1% (<0.001)

Source: CS, section B.2.6.1.1 figure 12, table 12, section B.2.7.1, table 17 | Abbreviations: PBO, Placebo | a Weight-range based UST doses approximating 6 mg/kg: 260 mg (weight ≤ 55 kg), 390 mg (weight > 55 kg and ≤ 85 kg), and 520 mg (weight > 85 kg), b Patients who had a prohibited change in concomitant UC medication or an ostomy or colectomy prior to the Week 8 visit were considered not to be in clinical remission; patients who had all 4 Mayo subscores missing at Week 8 were considered not to be in clinical remission or response

[redacted]

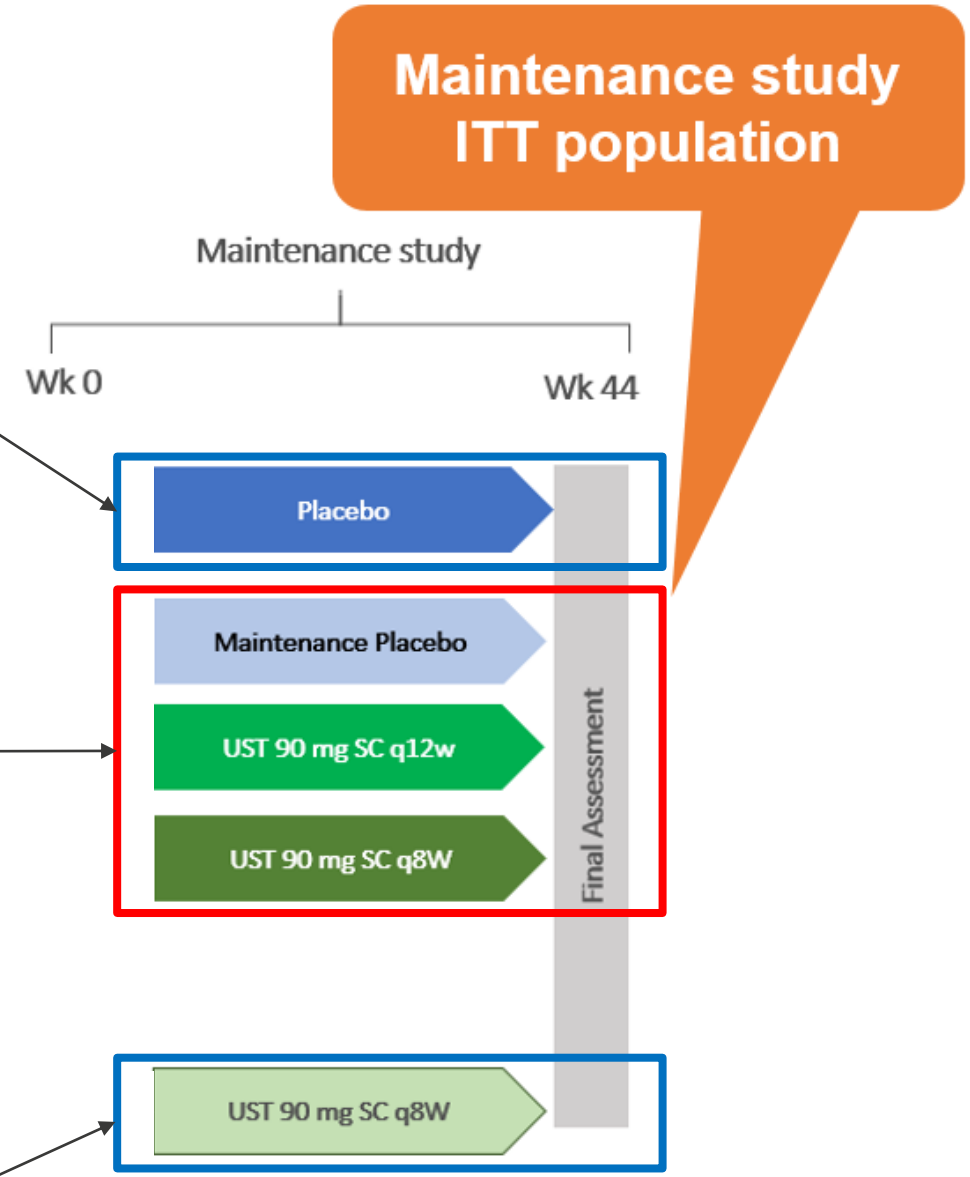
UNIFI trial design (maintenance)

Non-randomised group consisting of patients who were randomised to placebo group at induction and were in response at week 8

Re-randomised population consisting of:

- patients who had been randomised to UST 130 mg IV or UST ~6mg/kg IV during induction and were in response at week 8 PLUS
- patients who were randomised to placebo group at induction and did not respond, then were given ~6 mg/kg UST IV at week 8 and were in response at week 16

Non-randomised group consisting of 'delayed responders' i.e. patients who were in response at week 16 having received an additional 90 mg UST SC at week 8 following non-response to active treatment during weeks 0-8 of induction phase



Key trial results (maintenance ITT)

Pooled results included in company model directly (pooling = simple mean of two regimens with 30% assumed to have escalated regimen)

Un-pooled results included in company model directly

	Overall population (maintenance ITT)			Non-biologic failure population			Biologic failure population		
End point	**PBO[a] N=175**	**90mg SC q8w (p-value) N=176**	**90mg SC q12w (p-value) N=172**	**PBO[a] N=87**	**90mg SC q8w (p-value) N=85**	**90mg SC q12w (p-value) N=102**	**PBO[a] N=88**	**90mg SC q8w (p-value) N=91**	**90mg SC q12w (p-value) N=70**
Clinical remission	24%	43.8% (<0.001)	38.4% (0.002)	31.0%	48.2% (0.024)	49.0% (0.020)	17.0%	39.6% (<0.001)	22.9% (0.044)
Clinical response[b]	44.6%	71% (<0.001)	68% (0.001)	50.6%	77.6% (<0 .001)	76.5% (<0 .001)	38.6%	64.8% (<0 .001)	55.7% (<0 .001)

Source: CS, section B.2.6.2.1 figure 14, section B.2.7.2, table 18, figures 19 and 20 | Abbreviations: PBO, Placebo; UST, ustekinumab; q12w, every 12 weeks; q8w, every 8 weeks | a Patients who were in clinical response to ustekinumab IV induction dosing and were randomised to maintenance placebo SC on entry into this maintenance phase, b Maintenance of clinical response through end of maintenance

Key trial results (adverse events)

	Induction phase			Maintenance phase		
	UST ~6 mg/kg N=322	UST 130 mg/kg N=320	PBO N=319	UST 90 mg q8w n=176	UST 90 mg q12w n=172	PBO n=175
Infection	49 (15.3%)	51 (15.9%)	48 (15%)	86 (48.9%)	58 (33.7%)	81 (46.3%)
Serious infection	1 (0.3%)	2 (0.6%)	5 (1.6%)	3 (1.7%)	6 (3.5%)	4 (2.3%)
Injection-site reactions [a]	-	-	-	5 (2.8%)	1 (0.6%)	4 (2.3%)
Opportunistic infections	0%	0%	0%[b]	1 (0.6%)	2 (1.2%)	0%
Malignancies	0 %	0%[c]	0%	1 (0.6% [2 patients in non-randomized])	1 (0.6%)	0 (0% [1 patient in non-randomized])
Cardiovascular events	0%	0%	1 (0.3%)	1 (0.6% [1 patient in non-randomized])	1 (0.6% [1 patient in non-randomized])	1 (0.6% [1 patient in non-randomized])
Deaths	1[a] (0.3%)[b]	0	0	1c (0.6%)b	0	0

Source: CS, section B.2.10.5-6, table 33

Abbreviations: PBO, Placebo; UST, ustekinumab; q12w, every 12 weeks; q8w, every 8 weeks

a patient experienced sudden death on Study Day 42 attributed to a SAE of oesophageal varices haemorrhage. The event was not considered to be related to the study agent by the investigator

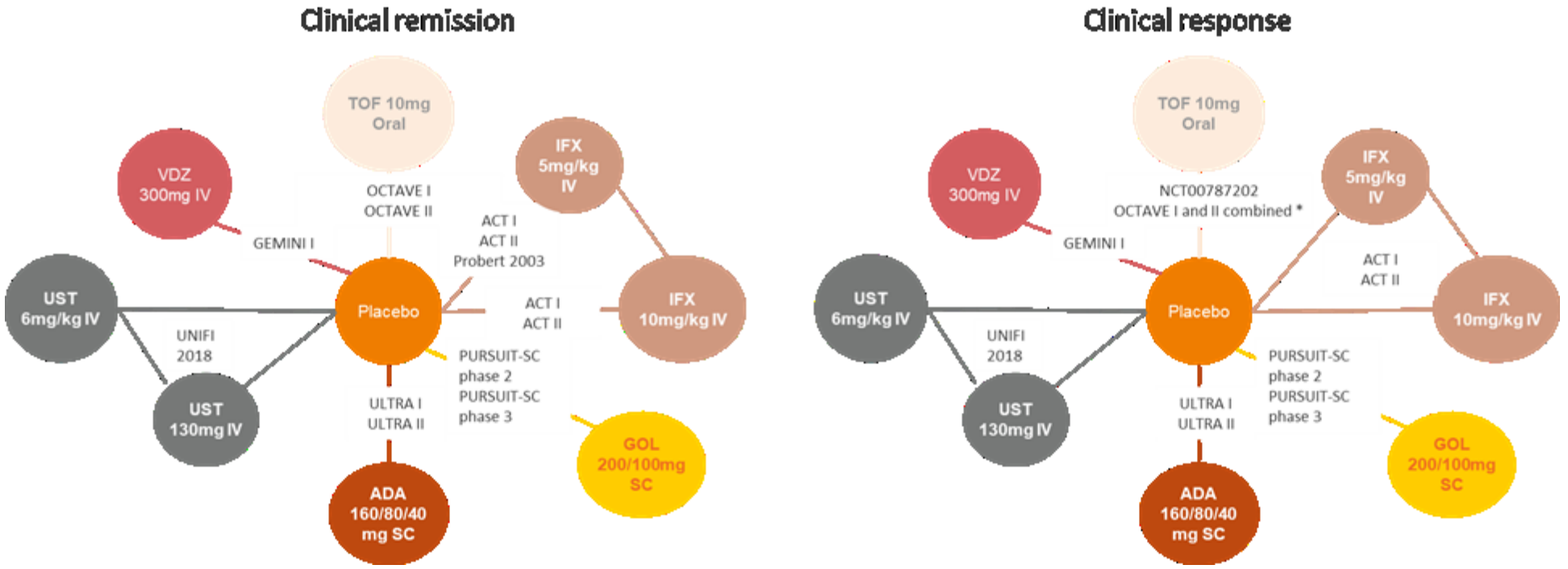
b calculated by technical team

c patient experienced death on maintenance Day 85 attributed to acute respiratory failure that occurred during thyroid surgery for a multinodular goiter. The event was not considered to be related to the study agent by the investigator

Indirect treatment comparisons – induction phase

- Limited evidence: no head-to-head trials of active treatments reporting outcomes for the relevant subgroups identified
- Company conducted adjusted indirect comparisons (network meta-analyses [NMAs]) using placebo as a common comparator ('star-shaped' evidence networks)
- Separate NMAs conducted for the non-biologic failure and biologic failures subgroups, but not for the overall (ITT) trial population (consistent with economic modelling approach)
- NMAs conducted for 2 outcomes: clinical remission and clinical response
- Trials conducted exclusively in Japan and China were excluded from the base case networks but included in a sensitivity analyses
- Results of the NMAs informed the company's economic base case (the transition probabilities in the induction phase of model were based on the NMA results)
- Company conducted both fixed and random effects models but preferred the fixed effects model

Indirect treatment comparisons – induction cont.



Summary of results:

- **Non-biologic failure patients:** ustekinumab superior to placebo for both clinical remission and response (clinical remission (median OR [CrI]) 2.19 [1.14; 4.39]; clinical response 3.66 [2.31 ; 5.88]); but for all active comparators, the credible intervals surrounding the point estimates were wide, overlapping and crossed 1
- **Biologic failure patients:** ustekinumab superior to both placebo and adalimumab for both clinical remission and response (vs. placebo clinical remission 13.41 [3.62; 94.58]; clinical response 3.58 [2.27; 5.74]; vs. adalimumab clinical remission 9.97 [1.77; 88.37]; clinical response 2.48 [1.17; 5.31]), but for all other active comparators, the credible intervals surrounding the point estimates were wide, overlapping and crossed 1

Indirect treatment comparisons – maintenance

Differences in trial designs (treat-through versus re-randomised) meant NMAs that included maintenance trial data could not be carried out using standard methods. Consequently different ITC methods were explored by the company and ERG

Name	Company 1 year NMA	Company 1-year NMA conditional on response	Company direct trial loss of response analyses	ERG maintenance only NMA
Description	Data from re-randomised trials recalculated to correspond to treat-through designs. Data for induction non-responders included	Data from re-randomised trials recalculated to correspond to treat-through designs. Data for induction non-responders excluded	Absolute data on clinical remission and response from individual trial arms included in economic model directly (data effectively become observational in nature)	Data from treat through trials re-calculated to correspond to re-randomised design (assumes number of induction responders is a proxy for entering maintenance)
Company base case	**No**	**No**	**Yes**	**No**
Company scenario	**No**	**Yes**	**No**	**No**
ERG base case	**No**	**Yes**	**No**	**No**
ERG scenario	**No**	**No**	**No**	**Yes**

As for induction separate analyses conducted for each subgroup and Asian trials excluded from both company's and ERG's base cases

UNIFI trial (TR issue 1) (1)

- Clarification style questions asked at engagement
- Prompted by the following concerns:
  - clinical response results in the CS did not appear to match those in the New England Journal of Medicine (NEJM) trial report published in September 2019 (after clarification had been completed)
  - it was not clear from the information in the CS that blinding had been maintained between induction week 8 and maintenance, or that baseline characteristics of patients in the re-randomized groups were well balanced, and therefore if the study was at high risk of bias
  - results amongst placebo non-responders who received ~6 mg/kg UST IV at week 8 and assessed at week 16 were not reported in the CS

Summary of technical engagement responses:

- Company
  - Response results not inconsistent between CS and NEJM – these are different data points reflecting different ways of calculating clinical response
  - Blinding maintained until week 44 – more detailed description of the interactive web response system (IWRS) used
  - Provided additional baseline and results data – see next slides

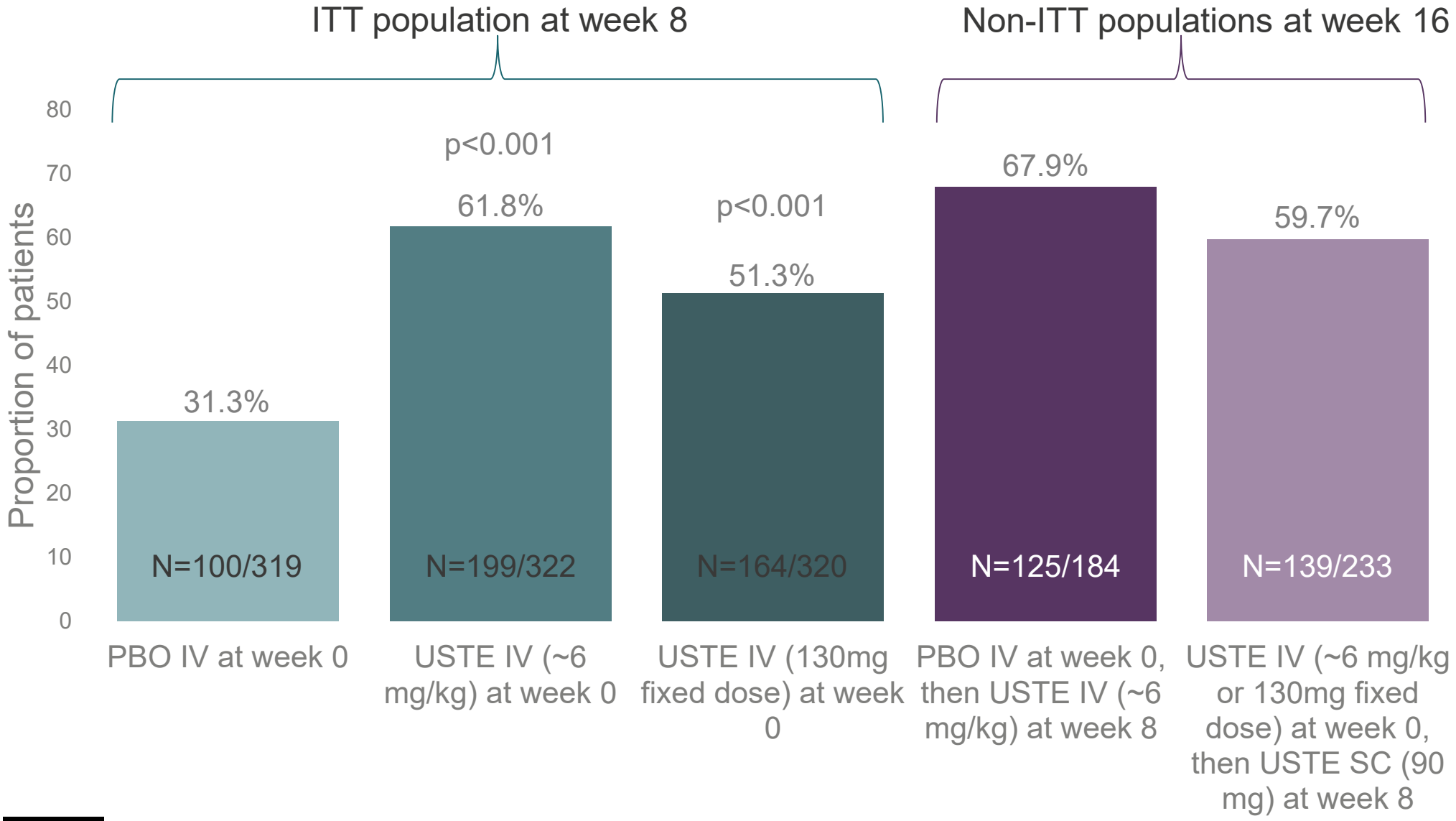
ERG comments on company TE response

- All discrepancies resolved/explained
- Clarifications/new data do not change ERG's original conclusions about UNIFI or affect ERG analyses

Key questions: (1) Does the committee agree that issues raised at technical engagement are resolved? (2) Do the results for non-randomised groups impact interpretation of ITT results?

UNIFI trial (TR issue 1) (2)

Results including non-randomised groups: Clinical response at the end of induction



Exclusion of trials conducted in Asia from the network meta-analyses (NMA) (TR issue 2)

- Company excluded five trials conducted in Asian countries from its NMAs (Hibi 2017 [PURSUIT-J], Jiang 2015, Kobayashi 2016 [Japic CTI-060298], Motoya 2019 [NCT02039505], Suzuki 2014) - inconsistent with previous appraisals
- Sensitivity analysis including Asian trials was conducted by the company to determine impact of excluding Asian trials - ERG agreed with approach in principle but found methodological problems with the analyses
- At engagement we asked if there is a clinical rationale as to why trials including only patients recruited in China or Japan should not be included in the analyses of the clinical effectiveness of ustekinumab

Summary of technical engagement responses:

- Company:
  - said there was no clinical rationale to exclude these trials
  - defended the validity of original sensitivity analyses with explanations regarding interpretation of trials in question

ERG comments on company TE response

- ERG preference for inclusion of Asian trials unchanged
- Company's explanations resolved one of ERG's original concerns
- Further inconsistencies mean results still uncertain but not of significant concern in terms of estimating cost effectiveness

Key question: is committee willing to accept the current analyses that exclude the trials conducted in Asia for decision making?

[redacted]

Dose regimen pooling in the maintenance phase (TR issue 7)

- Company pooled effectiveness rates for the standard and escalated regimens in the non-biologic failure subgroup but not in biologic failure group on the basis that there is not evidence of an exposure-response relationship in this subgroup
- This means, when it comes to economic modelling, the % of patients receiving the escalated dose impacts the cost of maintenance therapy in both subgroups, but it only affects *effectiveness* in biologic failure group
- ERG argued evidence for company approach was weak – preferred using pooled estimates in both subgroups because of high uncertainty over the exposure-response relationships
- At engagement we asked stakeholders their views on whether sufficient evidence had been presented by the company to support different approaches across subgroups, and what the benefits were of taking the same approach in both subgroups

Summary of technical engagement responses:

- Company
  - did not address the question for engagement - re-iterated that evidence was provided in relation to clarification question A22
  - made no changes to its base case in light to this issue
  - Argued it made little difference to cost effectiveness results

ERG comments on company TE response

- ERG preference for pooling in both groups unchanged by company response

Key question: does the committee have a preference preference between the company's or ERG's approach to dose pooling?

Synthesising maintenance phase trial data - which method is most appropriate? [TR issues 4 & 5] (1)

- Important driver of cost effectiveness –company's approach and ERG's preferences differ
- All approaches have big limitations so committee must decide which imperfect option they want to use to inform decision making – there are 3 main options
  - Company direct trial loss of response analyses i.e. 'naïve' comparison (company base case)
  - Company 1-year NMA conditional on response (company scenario preferred by ERG [ERG base case])
  - ERG maintenance only NMA (additional ERG scenario described by company as 'extremely implausible scenario that directly contradicts the evidence')
- Company and ERG preferences influenced by observed heterogeneity in results from placebo arms of re-randomised studies (concern is these are not 'true placebos') but
  - company argues (1) heterogeneity mainly due to effects of induction treatment being carried over into the maintenance phase (2) particular issue for ustekinumab due to half-life
  - ERG argues could be due to carry over effects or other factors
- Because of known limitations with each approach stakeholders not asked to pick between them. Instead feedback sought on
  - Evidence for carry over (to determine relevance of ERG maintenance only NMA (TR issue 4)
  - whether there were any other population-adjusted approaches to data synthesis not yet explored by either the company or ERG that could help reduce uncertainty (TR issue 5)

[redacted]

Synthesising maintenance phase trial data - which method is most appropriate? [TR issues 4 & 5] (2)

Summary of technical engagement responses:

- Company:
  - same comments as in original CS, plus data from UNIFI long-term extension (LTE)
  - population-adjusted approaches not appropriate but 1 additional DSA and 2 additional PSAs provided
- Comparator company: there is no correlation between the half-life of the various treatment options and the placebo response rates at week 52

Trial	Non biologic failure placebo response at end of maintenance	Half-life (taken from SPC)
Octave Sustain (tofacitinib)	24.8%	3 hours
PURSUIT-M (golimumab)	31.2%	12 days
UNIFI (ustekinumab)	50.6%	21 days
GEMINI I (vedolizumab)	26.6%	25 days

ERG comments on company TE response

- No new data have been provided to support the assertion that heterogeneity in placebo effects during maintenance is mainly due to carry-over effects
- Key limitations cited by company about NMA preferred by ERG are common to the company's preferred direct trial approach
- Agree with company that population-adjusted approaches not appropriate
- UNIFI LTE study data are supportive of a continued clinical benefit of ustekinumab for up to 2 years but no inferences can be drawn from these data about cost effectiveness due to lack of equivalent comparator data
- Company's new DSA has similar limitations to their base-case approach
- Company's 2 new PSAs do not 'improve confidence in' direct trial approach

Synthesising maintenance phase trial data - which method is most appropriate? [TR issues 4 & 5] (3)

Pros and cons of different clinical input options have been debated at length throughout the appraisal. Lead team have summarized key considerations in table below to help committee choose. Some apply to more than one model.

Company direct trial loss of response analyses (preferred by company)	**Company 1-year NMA conditional on response (preferred by ERG)**	**ERG maintenance only NMA**
Breaks within trial randomisation - data are of observational standard	Within trial randomisation preserved	Within trial randomisation preserved
Assumes that the placebo-placebo arms are similar (there are some differences in terms of response rates and these rates are low leading to a weak evidence base)		Post-re randomisation placebo arm data included (meaning more observations contribute to final estimates) BUT assumes re-randomised placebo arms are similar (not supported by evidence)
Does not use the post-re randomisation placebo arm data and is therefore not prone to carry-over effects BUT relative treatment effects are based on data from a small subset of placebo arms which may not be representative		
No imputation required	Imputation required – imputation method not used in previous UC appraisals	Imputation required – imputation method accepted (despite limitations) in previous UC appraisals
Results very uncertain – uncertainty not quantified by credible intervals	Results very uncertain – see ERG report table 36. Some of these uncertainties are reflected in the very credible intervals around point estimates. Impact of other uncertainties cannot be estimated statistically	

Key questions: (1) Is the ERG's maintenance only NMA scenario analyses relevant? (2) Does the committee have a preference between the company's or ERG's choice of clinical inputs?

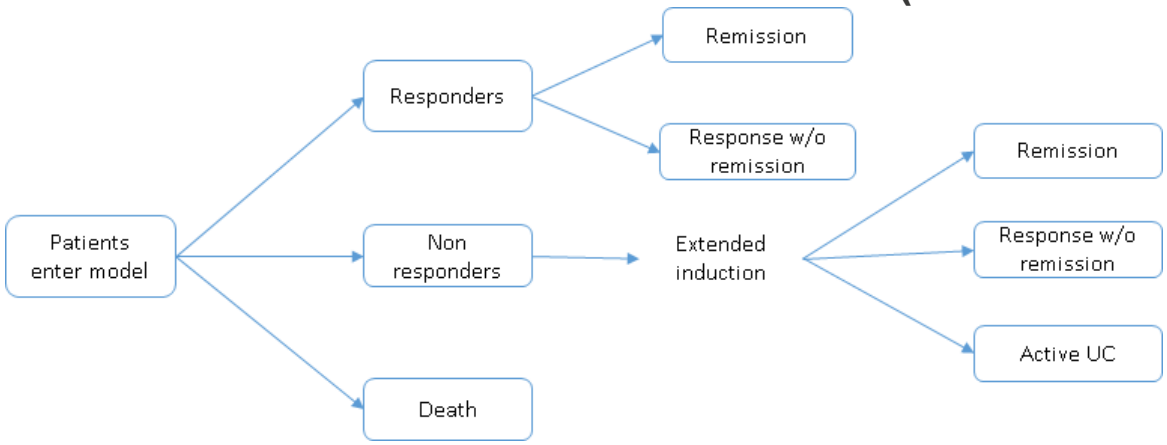
Key clinical issues (re-cap)

- UNIFI trial (technical report [TR] issue 1)
  - Does the committee agree that issues raised at technical engagement are resolved?
  - Do the results for non-ITT groups impact interpretation of ITT results?
- Exclusion of trials conducted in Asia from the network meta-analyses (NMA) (TR issue 2)
  - Is committee willing to accept the current analyses that exclude the trials conducted in Asia for decision making?
- Dose regimen pooling in the maintenance phase (TR issue 7)
  - Does the committee have a preference between the company's or ERG's approach to dose pooling?
- Synthesising maintenance phase trial data - which method is most appropriate? (TR issues 4 & 5)
  - Is the ERG's maintenance only NMA scenario analyses relevant?
  - Does the committee have a preference between the company's or ERG's choice of clinical inputs?

[redacted]

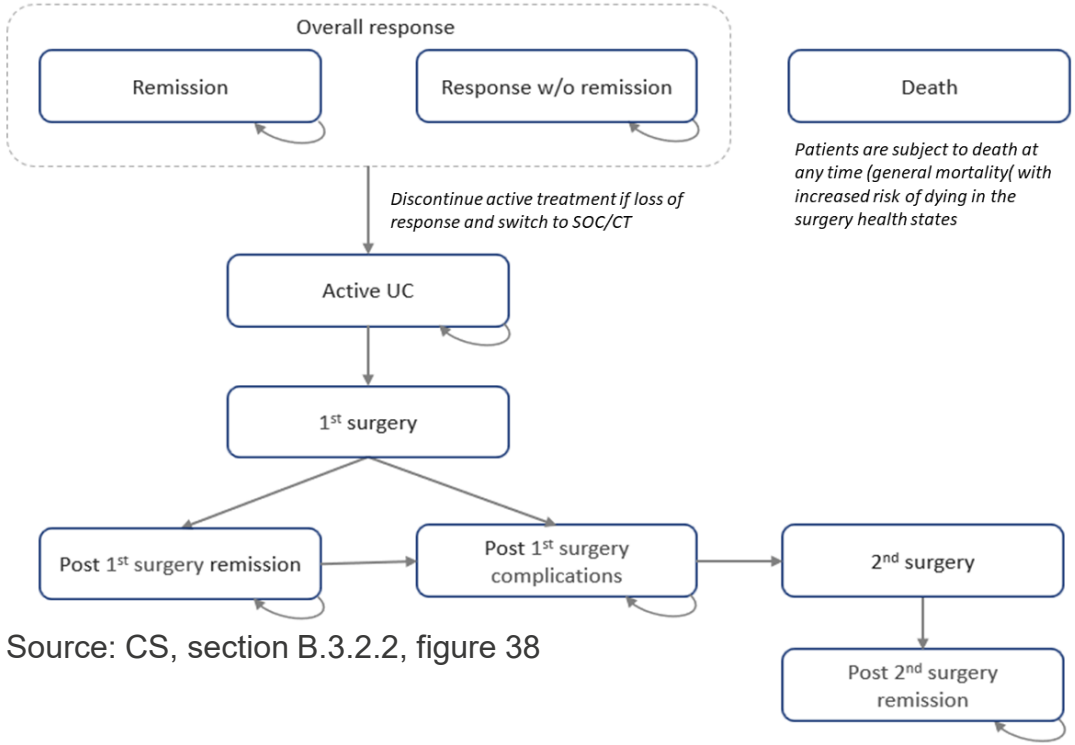
Company model structure – updated base case

Decision Tree for the Induction Phase (ERG's illustration)



Source: ERG report, section 4.3.3, figure 13

Markov model for the Maintenance Phase



Source: CS, section B.3.2.2, figure 38

- Conventional design for UC, but with some changes to previous TA models
- Hybrid - decision tree (for the induction phase) / Markov model (for maintenance and ongoing care)
- Markov has a cycle length = 2 weeks, designed to accommodate induction periods of different lengths
- 50-year time horizon (effectively lifetime from a starting age of 41 years), with a half-cycle correction
- Costs and QALYs are discounted at an annual rate of 3.5%

Description of model health states

Health State	Definition
Remission	Total Mayo score ≤2 with no individual subscore >1
Response without remission	Decrease from baseline in total Mayo score of at least 3 points and at least 30%, with accompanying decrease in subscore for rectal bleeding of at least 1 point, or an absolute subscore for rectal bleeding of 0 or 1, but not meeting remission definition
Active UC	Mayo score between 6 and 12 points (remission or response without remission not achieved)
1st surgery	First surgical intervention to resolve UC (with assumed duration of six months); could include acute complications
Post-1st surgery remission	No chronic complications from first surgery
Post-1st surgery complications	Chronic complications from first surgery such as wound infection, bowel obstruction, intra-abdominal abscess, or anastomotic leak
2nd surgery	Second surgical intervention due to pouch failure (with assumed duration of six months); could include acute complications
Post-2nd surgery remission	No chronic complications from second surgery
Death	Absorbing state

Source: CS, section B.3.2.2, table 35

[redacted]

Patients who do not achieve response after extended induction + and those who lose response to maintenance treatment enter Markov model in the Active UC health state on conventional therapy alone. Subsequently, patients can continue with Active UC, have surgery or die. Differs from models in previous NICE TAs (TA547 and TA342), which also included transitions from Active UC to Remission and Response without Remission after switching to conventional treatment alone

Key model assumptions (company's updated base case)

Population characteristics	Reflect characteristics of the patients in equivalent subgroups in the UNIFI trial
Clinical inputs	Identical clinical efficacy rates were used for infliximab/infliximab biosimilar and adalimumab/adalimumab biosimilar, for all efficacy outcomes in the model
• Standard induction:	Company's induction NMA (fixed effects model; excludes trials in Asian-only populations)
• Extended induction:	Direct trial data for people who did not respond during standard induction
• Maintenance phase:	
- Active arms	Proportion of induction responders who lost response by end of maintenance taken directly from individual trials active treatment arms. Standard and escalated dose results pooled (by taking simple means for the two regimens) in non-biologic failure subgroup but not in biologic failure subgroup (with 30% of patients assumed to have the escalated regimen in the base case). In both subgroups, escalated regimens for adalimumab were excluded due to lack of data
- Conventional therapy	Loss of response rates taken as a weighted mean for induction responders who received placebo during both induction and maintenance (PBO-PBO). This restricted the data source to UNIFI, ACT1, PURSUIT-M and ULTRA for the non-biological failure subgroup, and UNIFI and ULTRA 2 for the biological-failure subgroup
• Risk of loss of response	Remains constant during maintenance treatment, loss of response for delayed responders is the same as for those who responded to the first induction

Key model assumptions (company's updated base case, cont.)

Clinical inputs (cont.)

• Surgery:	Incidence of surgery and surgery related complications same across subgroups
• Adverse events:	Only serious infections included (treated as a one-time event); rates in the model based on multinational registry for systemic treatment of psoriasis: the PSOLAR study
• Mortality:	General population all-cause mortality rates adjusted for age and gender from UK Life tables, the only excess mortality for UC was a relative risk of 1.3 for surgery (applied during the six-month first and second surgery health states)
Utilities	Based on published literature not UNIFI trial data, adjusted by age and gender to account for the natural decline in quality of life associated with age
Costs	Costs for drug acquisition and administration, monitoring and follow-up care and treatment of serious infections included • Ustekimumab: CMU price used • All comparators: list prices used • Costs for concomitant treatment with conventional drugs alongside biologics not included

Key cost issues

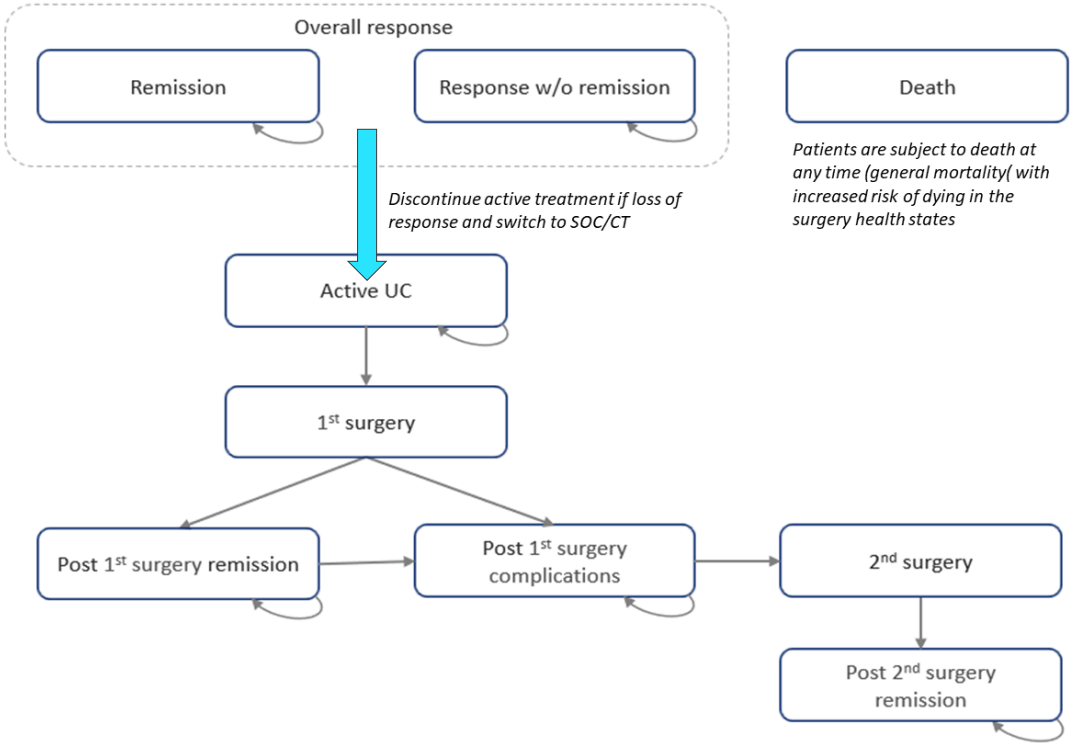
- Model structure – response and remission rates after failing therapy (TR issue 3)
  - What does committee think is a plausible percentage of patients that might experience spontaneous response after initial treatment failure?
- Infliximab dose escalation (TR issue 6)
  - Is committee happy to accept the ERG/company approach?
- What is the most appropriate source of utility data?
- Is it appropriate to use the same utilities across both sub-groups?

[redacted]

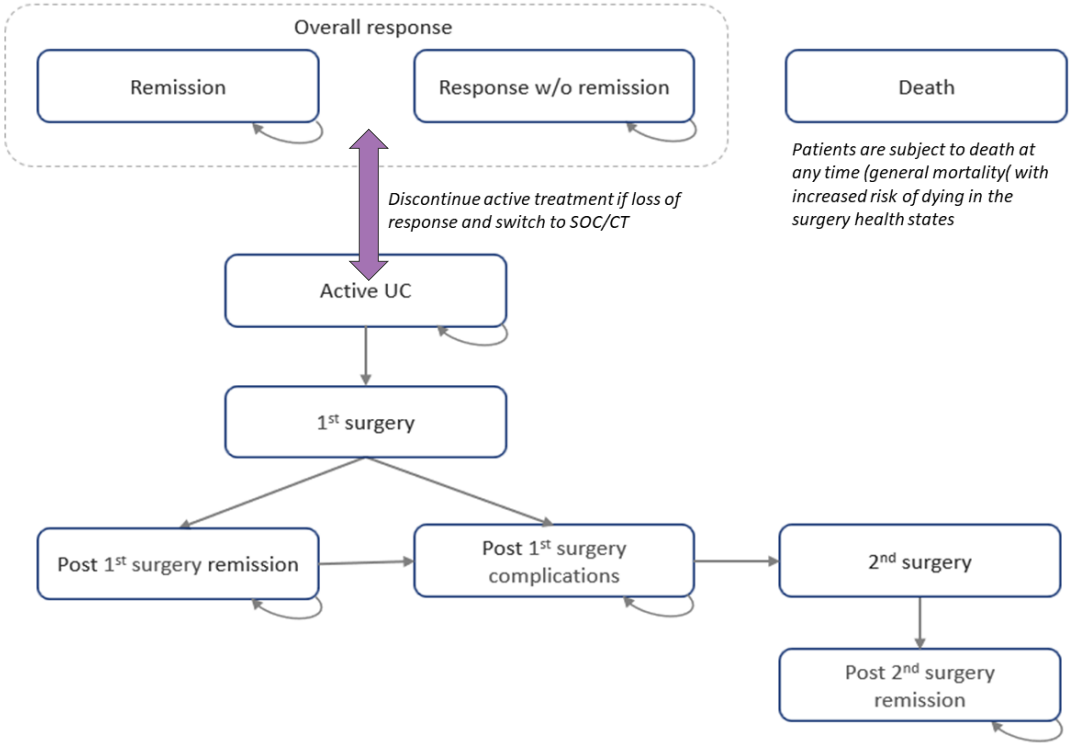
Response and remission rates for patients that do not respond or who lose response to therapy [TR issue 3] (1)

- Biggest driver of cost effectiveness *where company and ERG disagree*
- No hard evidence so committee must decide which model best represents reality

Company model



ERG model



- Company: assume that no further remission or response would be achieved on CT alone
- ERG: assume 5.5% will respond per 8 weeks (of whom 4.0% have response without remission) – this response is then lost at same rate for maintenance conventional therapy
- At engagement we asked for advice about how realistic these assumptions were

[redacted]

Response and remission rates for patients that do not respond or who lose response to therapy [TR issue 3] (2)

Summary of technical engagement responses:

- Company:
  - a survey of 10 clinicians proves plausibility of company approach
  - ERG approach results in unrealistic proportions of people experiencing response and remission over long term (same comments as made at factual accuracy stage)
  - 2 additional scenario analyses provided to explore uncertainty around rates of response and remission in this group
- Patient experts: the majority of patients […] will continue to experience active disease whilst on conventional therapy until surgery, entering a clinical trial or death

ERG comments on company TE response

- survey is unreliable – asks the wrong question and is at high risk of bias
- 1st of company's additional scenarios is helpful, but second has similar limitations to company base case as it assumes that no patient can experience more than two periods of reduced disease activity when treated with conventional therapy
- lack of evidence and outstanding uncertainty in ERG base case is acknowledged and a range of alternative distributions provided for committee to consider
- other criticisms not addressed (because were addressed previously at factual accuracy)

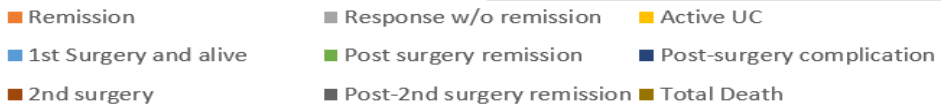
Key question: What does Committee think is a plausible percentage of patients that might experience spontaneous response after initial treatment failure?

[redacted]

Visual comparison of company and ERG assumptions

Company approach: overall response after initial treatment failure 0% - non-biologic failure

ERG approach: overall response after initial treatment failure 5.5% - non-biologic failure



Visual comparison of company and ERG assumptions

Company approach: overall response after initial treatment failure 0% - biologic failure

[redacted]

[redacted]

ERG approach: overall response after initial treatment failure 5.5% - biologic failure

[redacted]

[redacted]

[redacted]

Remission
Response w/o remission
Active UC
1st Surgery and alive
Post surgery remission
Post-surgery complication
2nd surgery
Post-2nd surgery remission
Total Death

Infliximab dose escalation (TR issue 6) (1)

- Company included evidence relating to two unlicensed higher (10mg/kg) doses of infliximab in their NMAs to strengthen the network
  - Infliximab 10 mg/kg IV at weeks 0, week 2 and week 6
  - Infliximab 10mg/kg IV every 8 weeks in maintenance
- did not include these doses in its economic analysis – not recommended in the SmPC
- On clinical advice, ERG did include escalated maintenance dose (30% patients receive infliximab 10mg/kg IV every 8 weeks)
- At engagement we asked if infliximab maintenance dose escalation was standard NHS practice

[redacted]

Infliximab dose escalation (TR issue 6) (2)

Summary of technical engagement responses:

- Company:
  - accepted ERG approach in updated base case
- Clinical experts: many, but not all, centres have the option of infliximab maintenance dose escalation available so it cannot be considered standard NHS practice
- Patient experts: doses are escalated based an individual circumstances/monitoring
- Comparator company:
  - dose escalation common for all biologic treatments (including infliximab)
  - escalation rates vary considerably across treatments (e.g. vedolizumab generally below 5% vs. ustekinumab generally above 75%) – should be considered when calculating cost-effectiveness

ERG comments on company TE response

- Agree with company conclusions

Key question: ERG and company analyses now both assume use of escalated maintenance doses for all biologics in 30% of patients. We have scenario analyses varying the proportion who received the escalated dose (10% or 50%) but we don't have any analyses with different escalation rates across treatments. Are the committee happy to accept the ERG/company approach?

[redacted]

Choice of utility values for response and remission health states [no feedback sought at TE] (1)

- Important driver of cost effectiveness; using UNIFI data instead of Woehl 2008 increases company base case ICER vs. CT by £55,344 and £61,651 per QALY for the non-biologic failure and biologic failure groups respectively
- Company's approach and ERG's approach did not differ (hence why this was not raised at TE)
- In both company and ERG base case, utilities for the 'Remission', 'Response without remission' and 'Active UC' health states are all derived from Woehl et al. (2008) UK EQ-5D-3L study of 180 UC patients
  - Only available as abstract - limited methodological information available about how study was conducted
  - same data were considered in all of the previous technology appraisals of the comparator treatments (TA329, TA342 and TA547)
- Utilities from EQ-5D data collected in the UNIFI trial also presented by the company and used in a scenario analyses

	Woehl et al. (2008) values	Values estimated from the UNIFI trial using EQ-5D-3L			
Health state	Based on total sample size of N=180	Average (sample size)	Standard deviation	Minimum	Maximum
Remission	0.87	[redacted]	[redacted]	[redacted]	[redacted]
Response without remission	0.76	[redacted]	[redacted]	[redacted]	[redacted]
Active UC	0.41	[redacted]	[redacted]	[redacted]	[redacted]

Source: ERG report section 4.3.5 tables 45 and 46

Choice of utility values for response and remission health states [no feedback sought at TE] (2)

The percentages in these tables show the proportions of total QALYs per intervention (CT or UST) and the proportions of incremental QALYs (UST vs. CT) that accrue in each health state category depending on the choice of utility data

UNIFI vs. Woehl 2008 utilities: discounted QALYs gained disaggregated by health state (all other assumptions as per updated company base case) – non-biologic failure

Drug	Remission	Response w/o remission	Active UC	Surgery, post-surgery & AE	Total QALYs
Company revised TE base case with Woehl et al. utilities					
CT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2%	1%	89%	8%	100%
UST	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	23%	5%	66%	6%	100%
Increment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	168%	29%	-87%	-10%	100%
Company revised TE base case with UNIFI utilities					
CT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	1%	1%	93%	5%	100%
UST	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	15%	3%	77%	4%	100%
Increment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	532%	100%	-498%	-34%	100%

Source: ERG Comparator PAS addendum 3 – 2nd December 2019, table 39

[redacted]

Choice of utility values for response and remission health states [no feedback sought at TE] (3)

The percentages in these tables show the proportions of total QALYs per intervention (CT or UST) and the proportions of Incremental QALYs (UST vs. CT) that accrue in each health state category depending on the choice of utility data

UNIFI vs. Woehl 2008 utilities: discounted QALYs gained disaggregated by health state (all other assumptions as per updated company base case) – biologic failure

Drug	Remission	Response w/o remission	Active UC	Surgery, post-surgery & AE	Total QALYs
Company revised TE base case with Woehl et al. utilities					
CT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	1%	1%	90%	8%	100%
UST	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	10%	6%	77%	7%	100%
Increment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	135%	67%	-91%	-11%	100%
Company revised TE base case with UNIFI utilities					
CT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	0%	1%	94%	5%	100%
UST	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	6%	4%	85%	4%	100%
Increment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	425%	233%	-521%	-37%	100%

Source: ERG Comparator PAS addendum 3 – 2nd December 2019, table 40

Choice of utility values for response and remission health states [no feedback sought at TE] (4)

- Company's reasons for not using the UNIFI-derived utility data (ERG summary)
  - Patients in the trial continue to receive ustekinumab, unlike in the model where they are assumed to switch to CT on loss of response
  - Inconsistency in the summary results from the UNIFI trial and published literature
  - Insufficient duration of trial follow up to assess the change of utility over time
- ERG comments
  - EQ-5D-5L scores from UNIFI cross-walked to the 3L scale using a published algorithm (van Hout et al. (2012) - recommended by NICE)
  - Number of observations in three severity health states large, analysis appears well-conducted
  - Agreed with company's decision not to use UNIFI EQ-5D data because it is inconsistent with values used in previous UC TAs but consider scenario analysis with UNIFI utilities important
- Lead team comments
  - Utilities source of controversy in previous appraisals
  - Need to be aware of methodological uncertainties in Woehl et al. – in particular, we know little about disposition of patients included (less or more sick than trial patients)
  - Woehl et al. attributes lower utility value to active disease compared with UNIFI, this results in higher QALY gains for effective treatments when Woehl et al. is used

Key question: (1) What is the most appropriate source of utility data? (2) Is it appropriate to use the same utilities across both sub-groups?

Updated company base case with no comparator discounts – non-biologic failure

Technologies	Total Discounted costs (£)	Total Discounted QALYs	Incremental costs (£) vs CT	Incremental QALYs vs CT	ICER (£/QALY) Fully incremental	ICER (£/QALY) ustekinumab vs comparator
CT	[redacted]	[redacted]	[redacted]	[redacted]	-	£23,712
Adalimumab biosimilar	[redacted]	[redacted]	[redacted]	[redacted]	Extended Dominated	£19,350
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£18,251
Biosimilar - Inflectra	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£13,423
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£11,067
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	Extended Dominated	£12,243
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	Extended Dominated	£13,653
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£1,968
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	£23,712	-

Source: Company appendices to technical engagement response, table 1

[redacted]

Updated company base case with no comparator discounts – biologic failure

Technologies	Total Discounted costs (£)	Total Discounted QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY) Fully incremental	ICER (£/QALY) ustekinumab vs comparator
CT	[redacted]	[redacted]	[redacted]	[redacted]	-	£26,593
Adalimumab biosimilar	[redacted]	[redacted]	[redacted]	[redacted]	Extended Dominated	£19,938
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£18,480
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	Extended Dominated	£5,718
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	£26,593	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominant

Source: Company appendices to technical engagement response, table 2

[redacted]

Key cost issues (re-cap)

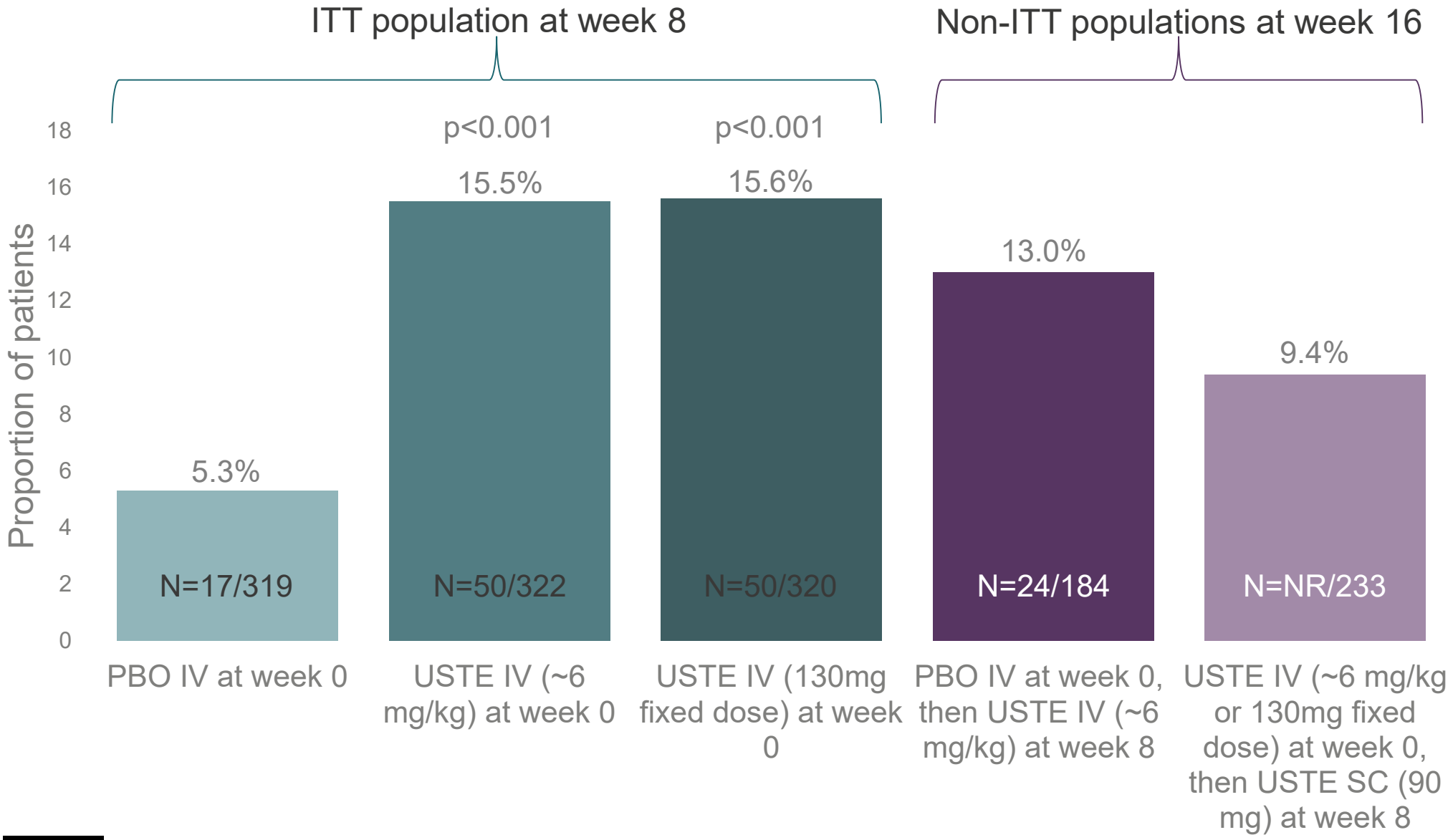
- Model structure – response and remission rates after failing therapy (TR issue 3)
  - What does committee think is a plausible percentage of patients that might experience spontaneous response after initial treatment failure?
- Infliximab dose escalation (TR issue 6)
  - Is committee happy to accept the ERG/company approach?
- What is the most appropriate source of utility data?
- Is it appropriate to use the same utilities across both sub-groups?

[redacted]

Back up slides

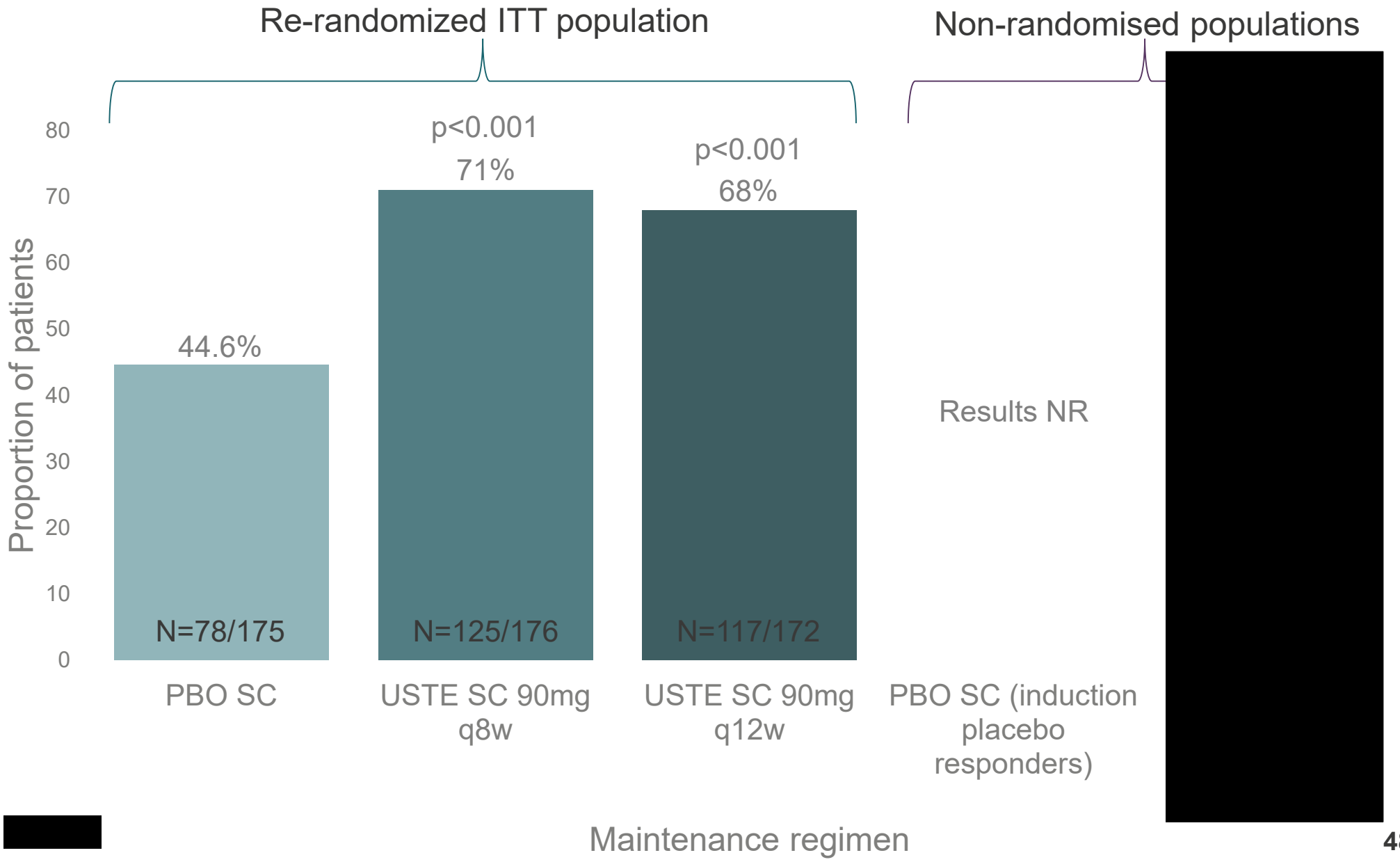
UNIFI trial (TR issue 1) (3)

Results including non-randomised groups: Clinical remission at the end of induction



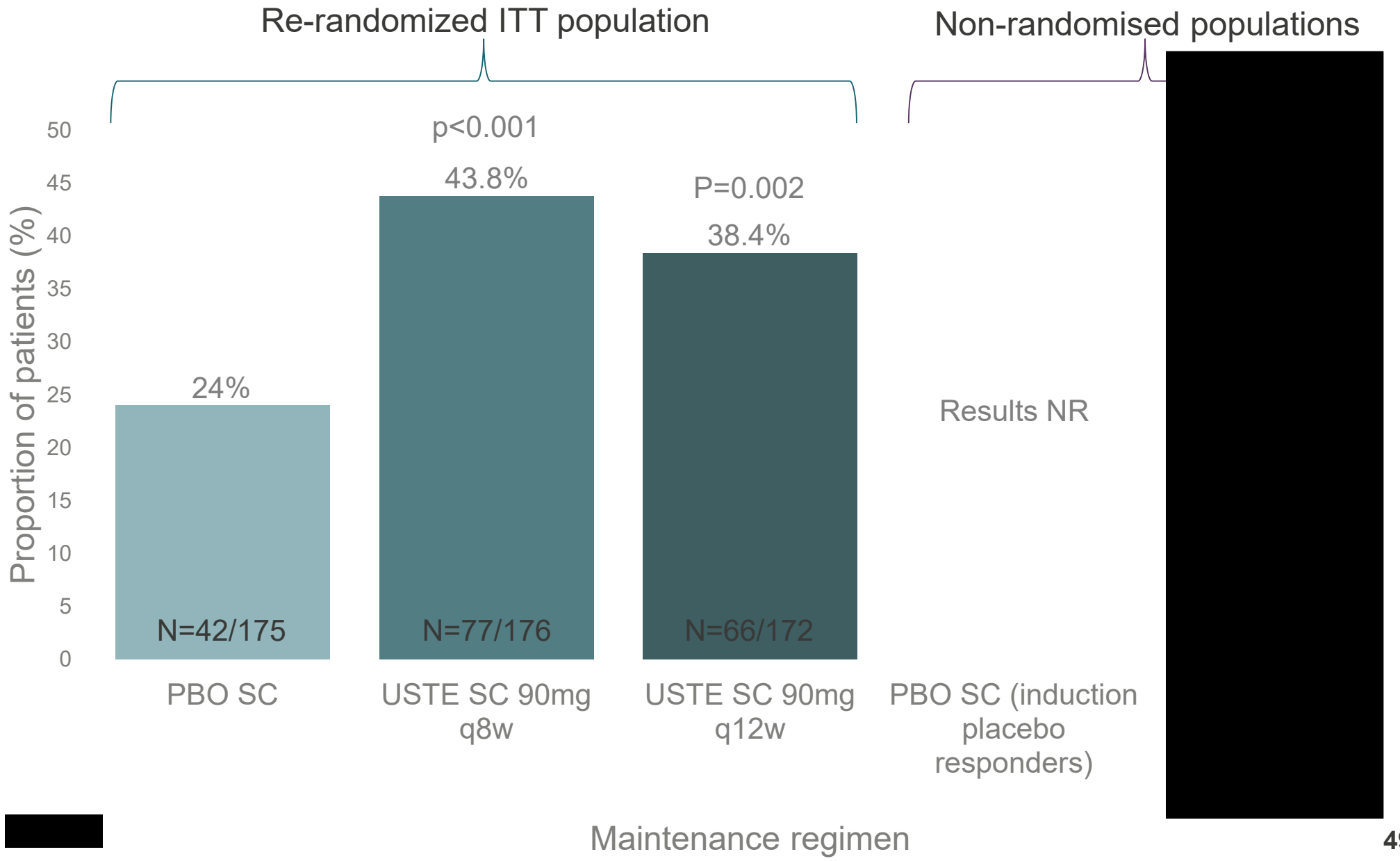
UNIFI trial (TR issue 1) (4)

Results including non-randomised groups: Clinical response at the end of maintenance (wk 44)



UNIFI trial (TR issue 1) (5)

Results including non-randomised groups: Clinical remission at the end of maintenance (wk 44)



Indirect treatment comparisons – maintenance cont.

Comparison of NMA results: Non-biologic failure patients

Comparator		Median OR [CrI], comparator vs. PBO			
		Clinical remission		Clinical response	
Induction	Maintenance	1 year NMA conditional on response	ERG maintenance only NMA	1 year NMA conditional on response	ERG maintenance only NMA
VED 300mg	VED 300mg pooled	4.76 [1.82; 15.24]	3.86 [1.57; 9.64]	4.18 [1.82; 10.68]	4.34 [1.83; 10.43]
INF pooled	INF pooled	3.18 [1.76; 6.12]	1.80 [0.67; 5.07]	3.82 [2.18; 7.06]	2.29 [0.91; 5.85]
GOL 200/100mg	GOL pooled	1.63 [1.03; 2.59]	1.79 [0.83; 3.89]	2.47 [1.58; 3.85]	2.08 [0.98; 4.40]
ADA 160/80/40mg	ADA 40mg EOW	2.65 [1.31; 5.57]	1.47 [0.55; 3.97]	2.11 [1.21; 3.74]	1.31 [0.52; 3.31]
TOF 10mg	TOF pooled	3.51 [1.83; 7.34]	6.25 [2.56; 15.94]	3.46 [2.00; 6.31]	4.67 [2.08; 10.58]
UST 6mg/kg	UST 90mg pooled	5.59 [2.92; 11.21]	2.13 [0.93; 4.89]	6.21 [3.59; 11.05]	3.30 [1.44; 7.59]

- Neither NMA provides consistently more or less conservative results
- In 1 year conditional on response analyses all treatments out-performed placebo
- In ERG maintenance-only NMA
  - no evidence of treatment effect for infliximab, golimumab, adalimumab (clinical remission or response) or ustekinumab (clinical remission)
  - lower point estimates for vendolizumb (clinical remission) and ustekinumab (clinical response) but higher point estimates for vendolizumb (clinical response) and tofacitinib (clinical remission and response) - difficult to interpret due to corresponding changes in credible intervals

Indirect treatment comparisons – maintenance cont.

Comparison of NMA results: Biologic failure patients

Comparator		Median OR [CrI], comparator vs. PBO			
		Clinical remission		Clinical response	
Induction	Maintenance	1 year NMA conditional on response	ERG maintenance only NMA	1 year NMA conditional on response	ERG maintenance only NMA
VED 300mg	VED 300mg q8w	8.88 [1.32; 144.60]	NR	2.99 [0.75; 12.24]	NR
VED 300mg	VED 300mg q4w	8.28 [1.15; 135.37]	NR	2.64 [0.61; 11.43]	NR
VED 300mg	VED 300mg pooled	NR	12.16 [2.72; 96.06]	NR	4.53 [1.46; 15.58]
ADA 160/80/40mg	ADA 40mg EOW	6.77 [1.50; 58.44]	3.17 [0.70; 18.38]	2.98 [1.13; 9.01]	2.85 [0.80; 10.98]
TOF 10mg	TOF 5mg	6.17 [1.94; 27.94]	NR	3.43 [1.68; 7.77]	NR
TOF 10mg	TOF 10mg	10.25 [3.40; 45.06]	NR	5.07 [2.57; 11.26]	NR
TOF 10mg	TOF pooled	NR	3.61 [1.39; 9.85]	NR	6.59 [2.69; 16.83]
UST 6mg/kg	UST 90mg q12w	7.89 [2.52; 26.60]	NR	5.21 [2.33; 11.65]	NR
UST 6mg/kg	UST 90mg q8w	10.33 [3.87; 31.22]	NR	5.24 [2.64; 10.54]	NR
UST 6mg/kg	UST 90mg pooled	NR	2.37 [0.97; 5.93]	NR	2.50 [1.10; 5.71]

Again neither NMA provides consistently more or less conservative results, plus even wider credible intervals