NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Tofacitinib for previously treated active ulcerative colitis [ID1218]

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Tofacitinib for previously treated active ulcerative colitis [ID1218]

Contents:

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

NICE National Institute for Health and Care Excellence

Tofacitinib for ulcerative colitis

Pre-meeting briefing

PART 1

This slide set is the pre-meeting briefing for this appraisal. It has been prepared by the technical team with input from the committee lead team and the committee chair. It is sent to the appraisal committee before the committee meeting as part of the committee papers. It summarises:

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

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

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

The lead team may use, or amend, some of these slides for their presentation at the Committee meeting

NICE

2

Key abbreviations (shaded rows contain comparator technologies)			
ADA	Adalimumab	INF	Infliximab
AE	Adverse event	mFAS	Modified full analysis set
BID	Twice daily	NMA	Network meta-analysis
CHMP	Committee for Medicinal Products for Human Use	NRI	Non-responder imputation
CI	Confidence interval	PAS	Patient access scheme
CrI	Credible interval	QALY	Quality-adjusted life year
CSR	Clinical study report	SAE	Serious adverse event
EMA	European Medicines Agency	SF-36	36-Item Short Form survey
ERG	Evidence review group	SUCRA	Surface under cumulative ranking curve
EQ-5D	5-dimensions EuroQol questionnaire	TA	Technology appraisal
FAS	Full analysis set	TNFi	Tumour necrosis factor inhibitor
GOL	Golimumab	TOF	Tofacitinib
HRQoL	Health-related quality of life	UC	Ulcerative colitis
IBDQ	Inflammatory Bowel Disease Questionnaire	VED	Vedolizumab
ICER	Incremental cost-effectiveness ratio		

NICE

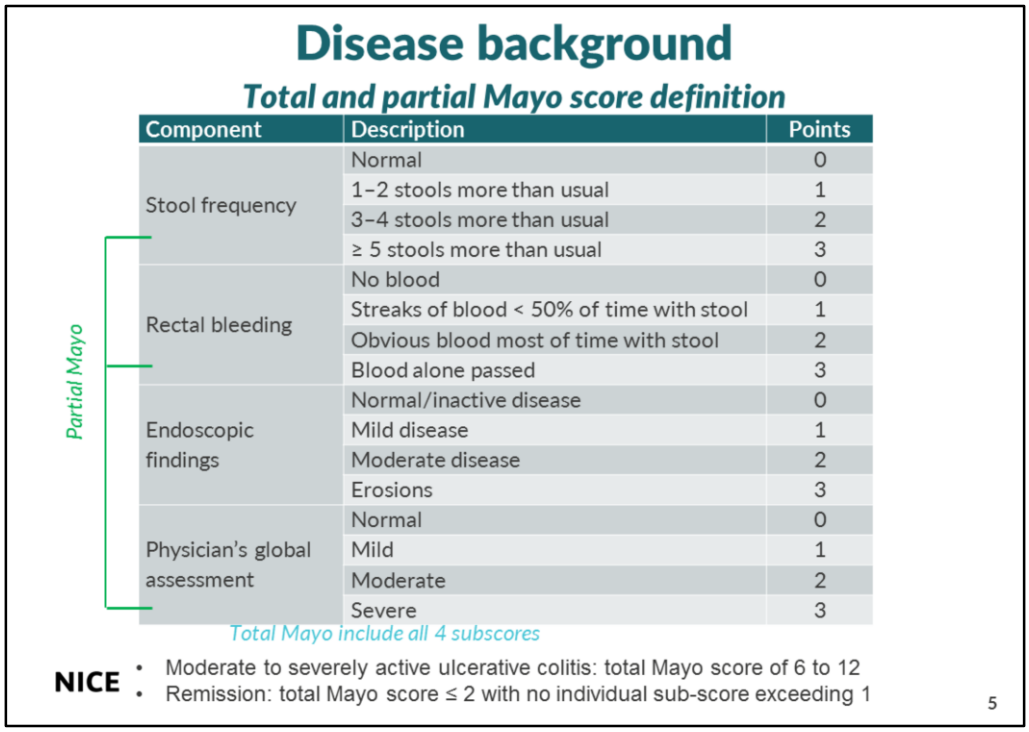
3

Disease background

Ulcerative colitis (UC)

- Most common inflammatory bowel disease
- Unknown cause; possible hereditary, infectious, immunological factors
- Approximately 146,000 people have UC in England, of whom about 52% have moderate to severe active disease
  - defined as Mayo score = 6 to 12
- Symptoms are bloody diarrhoea, colicky abdominal pain, urgency and tenesmus; extra-intestinal manifestations (joints, eyes, skin and liver)
- Onset of symptoms and diagnosis usually occurs between 15 and 25 years, and second peak of incidence between 55 and 65 years
- Symptoms can relapse and go into remission for months or even years:
  - 50% of people will have at least 1 relapse per year
- Complications of ulcerative colitis may include haemorrhage, perforation, stricture formation, abscess formation and anorectal disease
- High risk of surgery
- No increased mortality (only in more severe disease); increased risk of bowel cancer

NICE



Mayo score: Scores on the Mayo scale range from 0 to 12, and scores on each of the four subscores range from 0 to 3, with higher scores indicating more severe disease

Partial Mayo score: is calculated based on the following Mayo subscores: Physician's Global Assessment, stool frequency and rectal bleeding, and ranges from 0 to 9, with higher scores indicating more severe disease.

Relevant NICE guidance

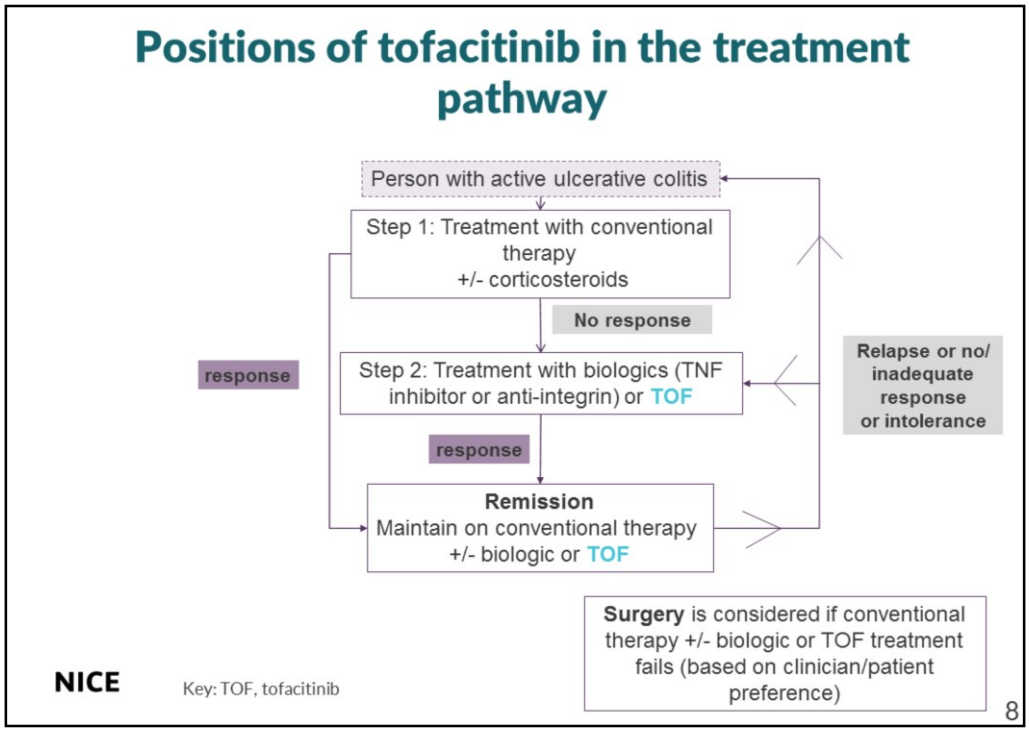
Technology appraisal (TA)		
TA	Intervention	Population
TA329 (Feb. 2015)	Infliximab Adalimumab Golimumab (MTA)	Adults with moderately to severely active UC whose disease has responded inadequately to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who cannot tolerate, or have medical contraindications for, such therapies
TA342 (Jun. 2015)	Vedolizumab	Adults with moderately to severely active ulcerative colitis
NICE clinical guideline (CG)		
CG166: Ulcerative colitis: management (2013, partially updated in 2017)		

CONFIDENTIAL

Tofacitinib citrate (Xeljanz)

Pfizer

Marketing authorisation	Treatment of adults with moderately to severely active ulcerative colitis (UC) who have had an inadequate response, lost response, or were intolerant to either conventional therapy or a biologic agent (MA granted on 1 August)
Mechanism of action	Intracellular janus kinase inhibitor that transmit signals arising from cytokine or growth factor-receptor interactions on the cellular membrane to influence cellular processes of creating new blood cells in the body (hematopoiesis) and immune cell function
Administration & dose	Oral; recommended dose: • induction: 10 mg twice daily for 8 weeks • maintenance: 5 mg twice daily *Patients who do not achieve adequate therapeutic benefit by week 8: extension of induction for 8 weeks, followed by maintenance. Patients who have failed prior TNF antagonist or those who experience a decrease in response on 5 mg can receive 10 mg for maintenance. If therapy is interrupted, restarting treatment can be considered. If there has been a loss of response, reinduction with 10 mg may be considered.*
Stopping rules	Induction should be discontinued if no evidence of benefit by week 16
List price and PAS discount	• List price: 5 mg x 56 tab: £690.03; 10 mg x 56 tab: £1,380.06 (average yearly treatment: £10,350.42 per patient; subsequent annual cost: £8,970.39 per patient) • Simple discount PAS approved



Step 1 therapy: left-sided and extensive ulcerative colitis

Adults

To induce remission in adults with a mild to moderate first presentation or inflammatory exacerbation of left-sided or extensive ulcerative colitis:

offer a high induction dose of an oral aminosalicylate

consider adding a topical aminosalicylate or oral beclometasone dipropionate[1], taking into account the person's preferences.

Step 2 therapy: all extents of disease

Prednisolone and tacrolimus

Consider adding oral prednisolone[1] to aminosalicylate therapy to induce remission in people with mild to moderate ulcerative colitis if there is no improvement within 4 weeks of starting step 1 aminosalicylate therapy or if symptoms worsen despite treatment. Stop beclometasone dipropionate if adding oral prednisolone.

Consider adding oral tacrolimus[2] to oral prednisolone to induce remission in people with mild to moderate ulcerative colitis if there is

an inadequate response to oral prednisolone after 2–4 weeks.

Infliximab for subacute manifestations of ulcerative colitis

This guidance relates only to the use of infliximab for subacute manifestations of moderately to severely active ulcerative colitis. It does not cover the use of infliximab for acute manifestations of moderately to severely active ulcerative colitis.

A subacute manifestation of moderately to severely active ulcerative colitis is defined as disease that would normally be managed in an outpatient setting and that does not require hospitalisation or the consideration of urgent surgical intervention.

Infliximab **is not recommended** for the treatment of subacute manifestations of moderately to severely active ulcerative colitis.

These recommendations are from infliximab for subacute manifestations of ulcerative colitis (NICE technology appraisal guidance 140).

NICE has written information for the public explaining its guidance on infliximab.

Adalimumab for the treatment of moderate to severe ulcerative colitis

The appraisal of adalimumab for the treatment of moderate to severe ulcerative colitis (NICE technology appraisal 262) was terminated because no evidence submission was received from the manufacturer or sponsor of the technology. Therefore NICE is unable to recommend the use in the NHS of adalimumab for the treatment of moderate to severe ulcerative colitis.

Decision problem

	Final NICE scope	Company submission
Population	People with moderately to severely active UC who are intolerant of, or whose disease has had an inadequate response or loss of response to conventional therapy (oral corticosteroids and/or immunosuppressant) or a TNF-alpha inhibitor	
Comparator	• TNF-alpha inhibitors (infliximab, adalimumab, golimumab) • Vedolizumab • Conventional therapies	Same as final scope with addition of placebo
Outcome	• measures of disease activity, including rates and duration of response, relapse and remission • achieving mucosal healing • health-related quality of life • rates of surgical intervention • time to surgical intervention • rates of hospitalisation • adverse effects of treatment • mortality	Absence of 'time to surgical intervention'; company explained that it was not assessed in the OCTAVE trials

NICE

9

Clinician perspective

Tofacitinib

- First drug of this class offering an alternative treatment to patients whose disease has not responded to current treatment options (treatment-refractory or corticosteroid-dependent)
- Step-change in the management of the ulcerative colitis (UC)
- Oral medication, does not require infusion facilities
- Increase chance of avoiding surgical intervention (e.g. colectomy, which can impact on education, relationships and pregnancy)
- Small molecule so reduced chance of immunogenicity and loss of response over time compared to monoclonal antibody therapies (biologics)
- Good safety profile
- OCTAVE trials reflect UK clinical practice (although excluded people with protitis*)
- Variability of access in England due to commissioners interpreting of NICE guidance differently
- Locally defined treatment pathways (commissioners /secondary care)

NICE

*Protitis: disease extent less than 15cm from anal verge

Patient perspective

Tofacitinib and current UC treatment

- Tofacitinib offers an additional treatment option with a different mechanism of action; reduced likelihood of loss of response
- Convenience of oral therapy
- Concerns with current available treatments
  - far from optimal due to lack of response and safety concern
  - surgery associated with considerable anxiety and potential complications; can interfere with religious and cultural belief
  - injections and infusions (at hospital or home) can impact significantly on patient's lives and work (e.g., travel/parking cost; cannot travel due to storage requirement)
- Profound and devastating impact of UC symptoms on all aspects of life: study, socialise, participate in leisure activities, have intimate relationships.
- Burden on carer as UC is (to some degree) an invisible condition, unpredictable symptoms, extremely uncomfortable to talk about

"Tofacitinib has completely changed my life…. I am now in my 4th year of taking tofacitinib and it is like I am a new person"

NICE

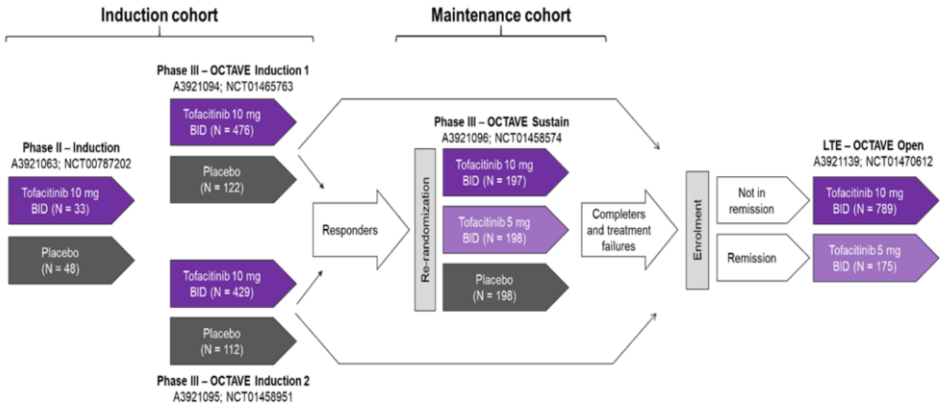
11

CLINICAL EFFECTIVENESS

NICE

CONFIDENTIAL

Overview of company tofacitinib trial programme



- Phase II was a small dose-finding study (n=194 patients, of whom only 33 received TOF10) therefore company submission only focuses on Phase III OCTAVE trials (although included in NMA because it met inclusion criteria)
- ERG comments: reasonable; each study included UK patients although number was low ([redacted]

NICE

Source: Figure 4 p.39 of CS, company response to clarification A9

13

Summary of OCTAVE trials

	Induction cohort		Maintenance cohort
	OCTAVE Induction 1	OCTAVE Induction 2	OCTAVE Sustain
N (n prior TNF exposed)	598 (319)	541 (299)	593 (283)
Design	Phase III	Phase III	Phase III
Duration	9 weeks	9 weeks	52 weeks
Population	Patients with moderate to severe active UC **(results are presented by prior TNFi exposure)**		Patients who achieved clinical response in OCTAVE Induction 1 or 2 **(results are presented by prior TNFi exposure)**
Intervention	TOF 10mg BID	TOF 10mg BID	TOF 5mg & 10mg BID
Comparator	placebo	placebo	placebo
Endpoints	1° remission (central & local read)* 2° mucosal healing; clinical response; clinical remission; IBDQ; SF-36; EQ-5D (all outcomes measured at 8 wk)		1° remission (central & local read)* 2° mucosal healing; sustained steroid-free remission** (24 wk); clinical response; clinical remission (all outcomes measured at 52 wk)

NICE

BID, twice daily; EQ-5D, 5-dimension EuroQol questionnaire; IBDQ, Inflammatory Bowel Disease Questionnaire; SF-36, 36-Item Short Form Survey

*Remission is measured based on centrally & locally assessment of endoscopic subscores; only locally read included in NMA; **although corticosteroids used for induction of remission, because of their side-effect profile they are not typically used for long-term management of UC, making corticosteroid-free remission an important goal

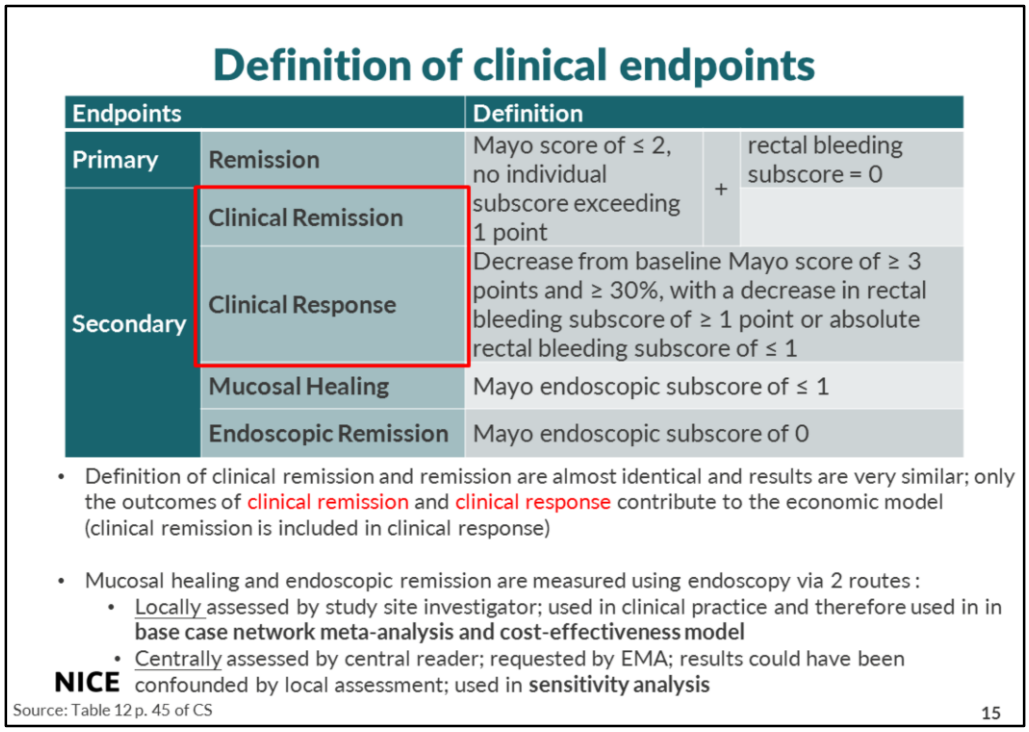
14

Note:

- OCTAVE Induction 1 and 2 are identical studies
- a third 15 mg BID tofacitinib arm was discontinued prior to full recruitment based on feedback from regulatory authorities (Randomisation to the 15-mg dose was discontinued after 38 patients in the OCTAVE Induction 1 trial, and 18 in the OCTAVE Induction 2 trial, had undergone randomisation across three treatment groups)

Mayo score: Scores on the Mayo scale range from 0 to 12, and scores on each of the four subscores range from 0 to 3, with higher scores indicating more severe disease

Partial Mayo score: is calculated based on the following Mayo subscores: Physician's Global Assessment, stool frequency and rectal bleeding, and ranges from 0 to 9, with higher scores indicating more severe disease.



Clinical remission is an outcome with an almost identical definition to the primary outcome of remission. The difference being that the rectal bleeding sub-score of the Mayo score does not have to be zero to achieve clinical remission. The outcomes of clinical remission and clinical response contribute data to the economic model.

CONFIDENTIAL

Results by subgroup: Proportion of patients in clinical remission & response (1)

OCTAVE 1 & 2 pooled data at week 8 (locally read scores)

Subgroup: prior-TNFi treatment	TOF 10 mg n/N (%)	PBO n/N (%)	Difference (95% CI); p-value
Clinical remission			
TNFi-naive	[redacted]	[redacted]	[redacted]
TNFi-exposed	[redacted]	[redacted]	[redacted]
Clinical response			
TNFi-naive	[redacted]	[redacted]	[redacted]
TNFi-exposed	[redacted]	[redacted]	[redacted]

- [redacted]

NICE

16

CONFIDENTIAL

Results by subgroup: Proportion of patients in clinical remission & response (2)

OCTAVE Sustain at week 52 (locally read scores)

Subgroup: prior-TNFi treatment	TOF 5 mg n/N (%)	PBO n/N (%)	Difference vs PBO (95% CI)	TOF 10 mg n/N (%)	Difference vs PBO (95% CI)
Clinical Remission					
TNFi-naïve	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clinical Response					
TNFi-naïve	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

- [redacted]

NICE

17

Adverse events (AEs)
OCTAVE trials

	OCTAVE trial programme (Phase II, OCTAVE Induction, OCTAVE Sustain and OCTAVE Open)	
	Placebo (n not reported)	Tofacitinib (n=1157)
Deaths	0	5* (0.4%)
Serious AEs	<10%	<10%
Serious infections (included in the economic model)	2 (0%)	38 (3.3%)**

*1 death was related to treatment **n not reported in OCTAVE Open

- **ERG comments**: Overall, and in comparison with evidence from the use of tofacitinib in patients with rheumatoid arthritis, no new safety signals were identified.

NICE

18

Company's network meta-analysis (NMA)

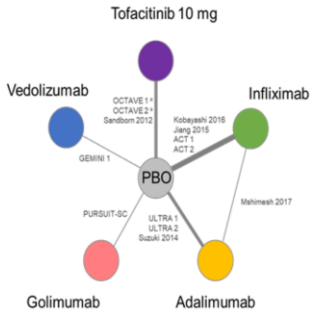
Description

- Company performed a Bayesian network meta-analysis (NMA) to estimate the relative efficacy and safety between tofacitinib [TOF] (5mg and 10mg) and
  - TNF-alpha inhibitors: adalimuab [ADA] (40/80/160mg); golimumab [GOL] (200/100mg and 100 mg); infliximab [INF] (5mg/kg)
  - vedolizumab [VED] (300mg Q4W and Q8W)
  - conventional therapies (placebo)
- 2 evidence networks for each induction and maintenance cohort, to match OCTAVE trials
  - ***TNFi naïve/ TNFi experienced***
- used a **multinomial probit model** for clinical response and clinical remission-this modelled clinical response and remission jointly to avoid impossible predictions such as more patients experience clinical remission than experience clinical response.
- **Efficacy endpoints**: clinical response, clinical remission, mucosal healing *(only clinical response, clinical remission were included in the economic model)*
- **Safety endpoints**: discontinuations due to adverse events, serious adverse events, and serious infections *(only serious infections were included in the economic model).*

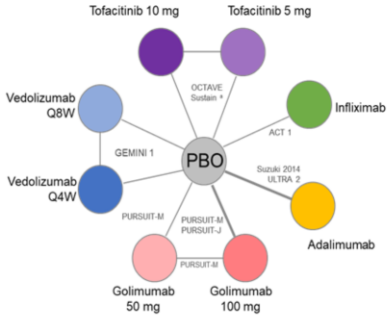
NICE 19

Company's network for TNFi Naïve

Induction Phase



Maintenance Phase



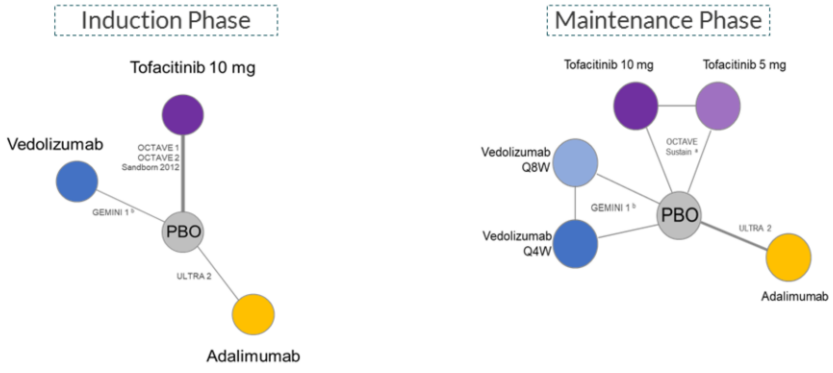
[a] Local read

Abbreviations: PBO, placebo; Q4W, every 4 weeks; G8W, every 8 weeks.

NICE

20

Company's network for TNFi Exposed



[b] TNFi failures

Abbreviations: PBO, placebo; Q4W, every 4 weeks; G8W, every 8 weeks.

NICE

21

CONFIDENTIAL

Company's NMA Results: TNFi Naïve (1)

- In the Induction phase for TNFi naïve [redacted]

Random effects model used	Treatment effect (probit scale*) Median (95% CrI)	Proportion of patients in		SUCRA**
		Clinical response	Clinical remission	
Induction Phase				
Placebo	[redacted]	[redacted]	[redacted]	[redacted]
TOF	[redacted]	[redacted]	[redacted]	[redacted]
INF	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	[redacted]
VED	[redacted]	[redacted]	[redacted]	[redacted]

*On the probit scale a negative coefficient indicates treatment is more effective than placebo

**The surface under cumulative ranking curve (SUCRA) value is used to rank treatments based on their probability of ranking first through to last among the treatment options. If the SUCRA probability is 0% the treatment always ranks last and if it is 100% the treatment always ranks first

NICE

22

CONFIDENTIAL

Company's NMA Results: TNFi Naïve (2)

- In the maintenance phase for TNFi naïve [redacted]

Fixed effects model used	Treatment effect (probit scale) Median (95% CrI)	Proportion of patients in		SUCRA
		Clinical response	Clinical remission	
Maintenance Phase				
PBO		[redacted]	[redacted]	[redacted]
TOF 5 mg	[redacted]	[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]	[redacted]
INF	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]
GOL 50 mg	[redacted]	[redacted]	[redacted]	[redacted]
VED Q8W	[redacted]	[redacted]	[redacted]	[redacted]
VED Q4W	[redacted]	[redacted]	[redacted]	[redacted]

ERG comments: Preferred Random effects model. Results of this analysis showed wider credible intervals *=> ERG used this model in its preferred base case*

NICE

23

CONFIDENTIAL

Company's NMA results: TNFi Exposed

- In both the induction and maintenance phase for TNFi exposed [redacted]

Fixed effects model used	Treatment effect (probit scale) Median (95% CrI)	Proportion of patients in		SUCRA
		Clinical response	Clinical remission	
Induction Phase				
Placebo	[redacted]	[redacted]	[redacted]	[redacted]
TOF	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]
VED	[redacted]	[redacted]	[redacted]	[redacted]
Maintenance Phase				
PBO	[redacted]	[redacted]	[redacted]	[redacted]
TOF 5 mg	[redacted]	[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]	[redacted]
ADA**	[redacted]	[redacted]	[redacted]	[redacted]
VED Q8W	[redacted]	[redacted]	[redacted]	[redacted]
VED Q4W	[redacted]	[redacted]	[redacted]	[redacted]

NICE **ERG comments:** Preferred Random effects model for induction phase. Results of this analysis showed wider credible intervals *=> ERG used this model in its preferred base case*

CONFIDENTIAL

Company's Serious Infections NMA *(induction phase only)*

Comparator	Treatment effect vs placebo, median (95% CrI), logit scale	
	Company base-case (random effects)	ERG alternative model selection (fixed effects)
TOF	[redacted]	[redacted]
INF	[redacted]	[redacted]
ADA	[redacted]	[redacted]
GOL	[redacted]	[redacted]
VED	[redacted]	[redacted]
AZA	[redacted]	[redacted]

ERG comments:

- ERG replication of company fixed effect model showed high level of uncertainty with very wide credible intervals which persisted in a fixed effect model
- ERG noted this is probably caused by the lack of any serious infections across placebo arms in the 3 TOF studies (in the other studies included in the NMA only 1 trial had 0 events)
- ERG ran a frequentist NMA to adjust for this *(added 0.5 to zero events)*. Results showed a **non-significant** increased risk of serious infection with TOF [redacted]

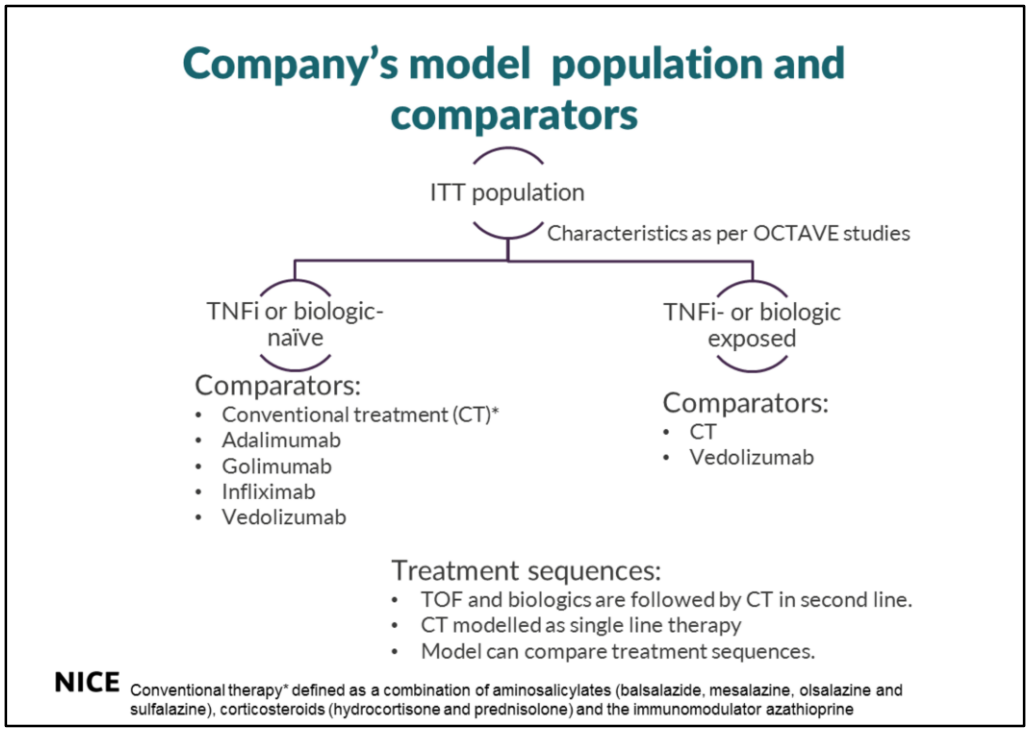
=> ERG used this analysis in its preferred base case

NICE

25

Key Clinical Issues

- Where is tofacitinib used in the treatment pathway?
- Are the results of the OCTAVE trials generalisable to NHS clinical practice?
  - Is it appropriate to subgroup the results based on prior treatment with TNF alpha inhibitors?
- Is the ERG's use of a frequentist approach for the NMA of serious infections appropriate?
- Is tofacitinib associated with an increased risk of serious infections?

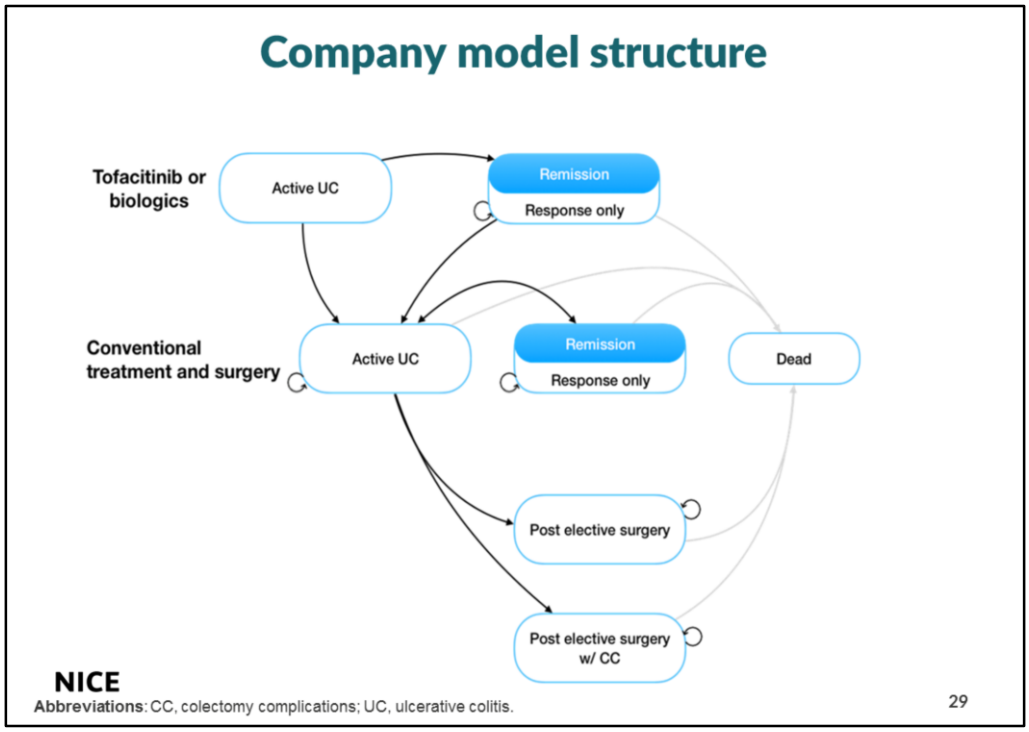

Company's model population and comparators
ITT population
Characteristics as per OCTAVE studies
TNFi or biologic-naïve
TNFi- or biologic exposed
Comparators:
Conventional treatment (CT)*
Adalimumab
Golimumab
Infliximab
Vedolizumab
Comparators:
CT
Vedolizumab
Treatment sequences:
TOF and biologics are followed by CT in second line.
CT modelled as single line therapy
Model can compare treatment sequences.
NICE
Conventional therapy* defined as a combination of aminosalicylates (balsalazide, mesalazine, olsalazine and sulfalazine), corticosteroids (hydrocortisone and prednisolone) and the immunomodulator azathioprine

Company's model population and comparators - ERG critique

- **Subgroups by TNFi or biologics exposure:**
  - Company: labelling by biologics exposure because
    - prior exposure to biologics is an important treatment effect modifier
    - patients' treatment history is a deciding factor in the treatment pathway
  - ERG agree but note that labelling is misleading, as NMA results are defined by prior exposure to TNFi alone (and not by prior biologic exposure)
- **Characteristics of the population**
  - Company: subgroups as per the OCTAVE trials
  - ERG: same gender, age and weight mix regardless of prior TNFi exposure
    => ERG explore impact of age and body weight in scenario analysis
- **Comparators**
  - Company did not include ADA
  - ERG considers ADA is a relevant comparator
    => ERG include ADA in their base case
- **Sequences:**
  => ERG explore effect of switching within or between classes and compare 'step-up' and 'step-down' strategies

NICE

28



- Induction phase: patients start with TOF or biologic comparator, with CT
- Maintenance: if respond continue with the same drug, with CT
- Patients who do not respond to induction treatment and those who lose response during maintenance treatment continue to receive CT alone.
- Subsequent treatment with CT is assumed to continue regardless of the disease state, except when the patient undergoes surgery. Similarly, for the CT comparison arm, CT is not assumed to stop when the patient has active disease.

Model health state description

	Health states	Description
Tofacitinib / biologic	1. Active UC	Patients enter model, start 8-week induction of TOF or biologic
	2.Remission	Patients in remission (after induction) continue to receive maintenance with TOF or biologic as they remain in response. Patients achieving response and patients responders in remission in each 8-week cycle are estimated from NMA
	3.Response only	
Conventional treatment	4.Active UC	Patients transition to the Active UC state on conventional treatment if: • Non response to tofacitinib/biologic induction • Loss of response in tofacitinib/biologic maintenance • Loss of response during conventional therapy (CT).
	5.Remisson	Patients who respond with or without remission while on CT. Transitions between active UC, remission and response only health states continue to occur while patients receive ongoing CT
	6.Response only	
Surgery (transient event)	Emergency surgery	Patients not in remission require emergency surgery due to acute exacerbation events in each model cycle.
	Elective surgery	Patients in Active UC assumed to undergo elective surgery in each cycle.
	7.Post surgery without CC	Surgery is associated with perioperative risks of complications and mortality. Patients who survive surgery transition to 1 of 2 health states: with- or without long-term complications.
	8.Post surgery with CC	
	9.Dead	

ERG critique on the model structure

- Economic model of good quality
- Appropriate reflection of clinical practice, in line with previous UC models.
- Includes risk of relapse and immediate cessation of response at each cycle.
- Assumes a fixed duration of induction of 8 weeks, followed by cessation of treatment for patients whose disease does not show a response in this time
  - TOF SPC recommends assessment 8-16 weeks after initiation and annual reassessment.
  - NICE MTA329 and NICE TA342 recommend assessment of response at 12 months. ERG's clinical experts agree that benefit is assessed annually.
  - NICE MTA329 and NICE TA342 recommend consideration of treatment withdrawal. ERG's clinical experts consider that withdrawal is unlikely in clinical practice.
- Adverse drug reactions only include serious infection, which in the model do not cause treatment discontinuation (although clinical advice is that TOF would be temporarily withheld).

NICE

31

Clinical parameters in the model

	Efficacy	Safety	Complications
Parameters and rationale	Locally read clinical response/ remission; choice of NMA models based on DIC statistics, with preference for FE if no difference	Serious infections only included as model already accounts for UC related conditions since model health states are defined based on clinical response and clinical remission corresponding to Mayo scores	Incidence and complication/mortality rates for surgery (perioperative complication and mortality, incidence of emergency and elective surgery
Source	NMA (clinical) and assumption	NMA (safety) for serious infections	Literature and assumptions
ERG comments	• Prefer NMA results using RE models to better reflect uncertainty related to heterogeneity in efficacy outcomes => ERG test alternative NMA in scenario analysis • Safety: in clinical practice, patients would be temporarily withheld following serious infection so assuming no discontinuation due to serious infections or other AEs is unrealistic and likely to introduce bias		

NICE

CONFIDENTIAL

Effectiveness in the model: TNFi-naïve

	Distribution by health state at end of induction			Response and remission given response (over 8 weeks)		
	Active UC	Response only	Remission		Probability of maintaining response	Percentage of responders in remission
ADA	[redacted]	[redacted]	[redacted]	ADA	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	GOL 50mg	[redacted]	[redacted]
				GOL 100mg	[redacted]	[redacted]
INF	[redacted]	[redacted]	[redacted]	INF	[redacted]	[redacted]
TOF	[redacted]	[redacted]	[redacted]	TOF 5mg	[redacted]	[redacted]
				TOF 10mg	[redacted]	[redacted]
VED	[redacted]	[redacted]	[redacted]	VED Q8W	[redacted]	[redacted]
				VED Q4W	[redacted]	[redacted]
CT	[redacted]	[redacted]	[redacted]	CT	[redacted]	[redacted]

NICE
Source: Adapted from Table 57 p. 149 and Table 58 p. 150 of ERG report

- NMA results of the clinical response/remission were transformed to transition probabilities
- Proportion of patients in 'Active UC' is estimated for proportion of patients in response only and in remission (= 1 – (patients in response + patients in remission))
- The Company used assumptions to calculate 8-week transition probabilities from the 52-week NMA response/remission rates
- Company could not apply method seen in NICE TA329 and TA342 due to a lack of mid maintenance period results for some comparators; and a failure to accurately predict the target data with calibration.
- Therefore company assumed constant risks within and beyond the 52-week trial data
    - **probability of loss of response** is calculated from the probability of no response over 52 weeks from the NMA (1 - probability of response), adjusted to 8-week model cycle
    - patients in cohort who maintain a response in each cycle are then split between remission and response only health states using a fixed proportion (ratio of 52-week

probabilities of response with and without remission)

Method seen in NICE TA329 and TA342

In the TA329 MTA (adalimumab, infliximab and golimumab), the assessment group had access to mid-point response and remission data for the maintenance period.[25] They used these data to estimate transition probabilities for two phases of maintenance - week 8 to 32 and week 32 to 52. The results are generally more favourable for the TNFi drugs in the second period than in the first.

In the TA342 STA (vedolizumab), the company used a calibration approach to fit transition probabilities to the 52 week NMA results. This involved applying certain constraints, such as that no more than 20% of people with mild disease would enter remission. This approach was criticised by the TA342 ERG for using arbitrary constraints and assumptions.

CONFIDENTIAL

Effectiveness in the model: TNFi-exposed

	Distribution by health state at end of induction			Response and remission given response (over 8 weeks)		
	Active UC*	Response only	Remission		Probability of maintaining response	Percentage of responders in remission
ADA	[redacted]	[redacted]	[redacted]	ADA	[redacted]	[redacted]
TOF	[redacted]	[redacted]	[redacted]	TOF 5mg	[redacted]	[redacted]
				TOF 10mg	[redacted]	[redacted]
VED	[redacted]	[redacted]	[redacted]	VED Q8W	[redacted]	[redacted]
				VED Q4W	[redacted]	[redacted]
CT	[redacted]	[redacted]	[redacted]	CT	[redacted]	[redacted]

ERG comments:

- Assumption does not reflect clinical experience; clinical experience shows risk is greatest in first 6-12 months; and falls thereafter
- Likely to underestimate the duration of treatment and hence costs and QALYs of active treatments; unknown direction of bias in ICERs

NICE

Source: Adapted from Table 57 p. 149 and Table 58 p. 150 of ERG report

CONFIDENTIAL

Safety outcomes in the model: Serious infections

- probabilities of serious infections used in the company base case, with ranges for sensitivity analysis vs ERG preferred frequentist approach

Treatment	Company (Bayesian NMA RE)			ERG (Frequentist NMA RE)		
	Base case	Lower limit	Upper limit	Base case	Lower limit	Upper limit
Placebo	[redacted]	[redacted]	[redacted]	[redacted]		
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib *	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG comments

- ERG frequentist estimates, give more plausible ranges of uncertainty
- Uncertainty associated with serious infections due to the rarity of events.

NICE

Source table 59, page 153 of ERG report

Surgical complication parameter
Sources

	Value	Source
Colectomy rates	Elective colectomy: 0.058% per cycle; emergency colectomy: 0.021% per cycle	Misra et al. (2016), HES analysis; ERG scenario analysis: Chhaya et al. (2015)
Perioperative complications and mortality	2.8% mortality risk per operation	UK IBD audit 2008-2014
Post-surgery complications	1.5% per cycle	Ferrante et al. (2007); ERG scenario analysis: Japanese study by Arai et al. (2010)
All-cause mortality	Same as general population, adjusted for age and gender-mix	

NICE

CONFIDENTIAL

Health-related quality of life

- Company used utilities for pre and post-surgical states from Woehl *et al.* 2008; and the background utility ('no disease') is based on EQ-5D by age and gender in general population (Ara *et al.* 2010):

Health state	Woehl et al. 2008 (company base case)	OCTAVE trials		Swinburn et al. 2012
		8 weeks	52 weeks	
Active UC	0.47	[redacted]	[redacted]	0.6317
Response	0.87	[redacted]	[redacted]	0.8944
Remission	1.00	[redacted]	[redacted]	1.0000
Post-surgery	0.82	NA	NA	0.6596

ERG comments:

- ERG believes that utilities from OCTAVE trials are problematic because of the re-randomisation design and lack of intermediate assessments between week 8 and 52. Therefore, ERG agrees that utilities by Woehl *et al.* provide a more appropriate source for base case parameters and also, are consistent with previous NICE TAs for UC.
- ERG use these estimates in ERG preferred analyses, and test scenarios based on the company's OCTAVE analyses and published sources (Swinburn et al.)

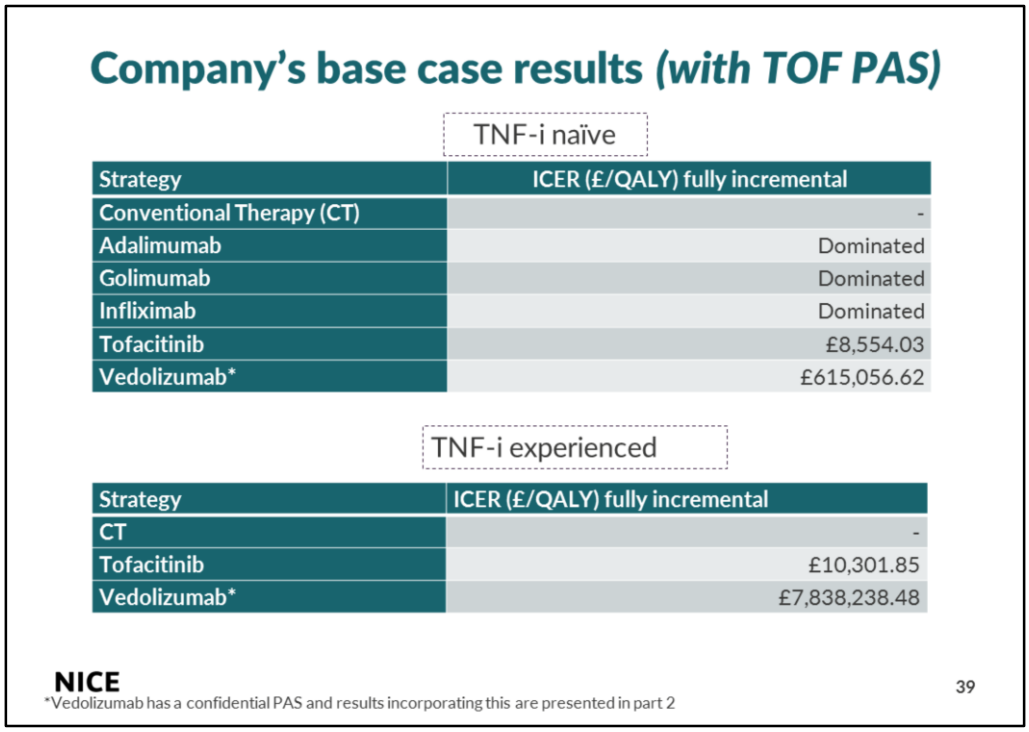
NICE

37

Resource use and costs

Items	Company assumption	ERG comments
Drug acquisition	• TOF: confidential patient access scheme (PAS) discount • GOL: PAS discount assume 50 and 100 mg dose at same cost • INF: biosimilar cost included	• ERG analysis also include VED confidential PAS (results in part 2) • INF: biosimilar cost included
Conventional therapy	Assumed equal usage for balsalazide, mesalazine, olsalazine and sulfalazine	• Does not reflect UK practice; mesalazine is prescribed more • Minor change of NHS price
Outpatient visit	Assumed 2 outpatient visits for patients in remission on maintenance treatment and 4.5 visits/ year for patients with a response but no remission	Monitoring and follow-up costs might not reflect clinical practice whereby treatment can be withdrawn within 8 weeks of a relapse => ERG explore scenario with additional costs for outpatient visits to enable treatment cessation within 8 weeks of a relapse (6.5 visits/year)
Drug administration	Assumed no administration cost for self-administered sub-cutaneous injections (golimumab, adalimumab)	=> ERG explore impact of assuming an initiation of self-administration
Stoma care	Company model omits ongoing costs of stoma care for post-colectomy health states (£426.36 per person in post-surgery assuming 40% have a stoma)	ERG include these costs in their base case and explore variation in scenario analysis

38



Company's base case results *(with TOF PAS)*

TNF-i naïve

Strategy	ICER (£/QALY) fully incremental
Conventional Therapy (CT)	-
Adalimumab	Dominated
Golimumab	Dominated
Infliximab	Dominated
Tofacitinib	£8,554.03
Vedolizumab*	£615,056.62

TNF-i experienced

Strategy	ICER (£/QALY) fully incremental
CT	-
Tofacitinib	£10,301.85
Vedolizumab*	£7,838,238.48

NICE

*Vedolizumab has a confidential PAS and results incorporating this are presented in part 2

39

The company present their base case results in CS section B.3.7, page 155. These incorporate the confidential PAS discount for tofacitinib but not the PAS discount for vedolizumab. The base case assume use of biosimilar drugs for infliximab

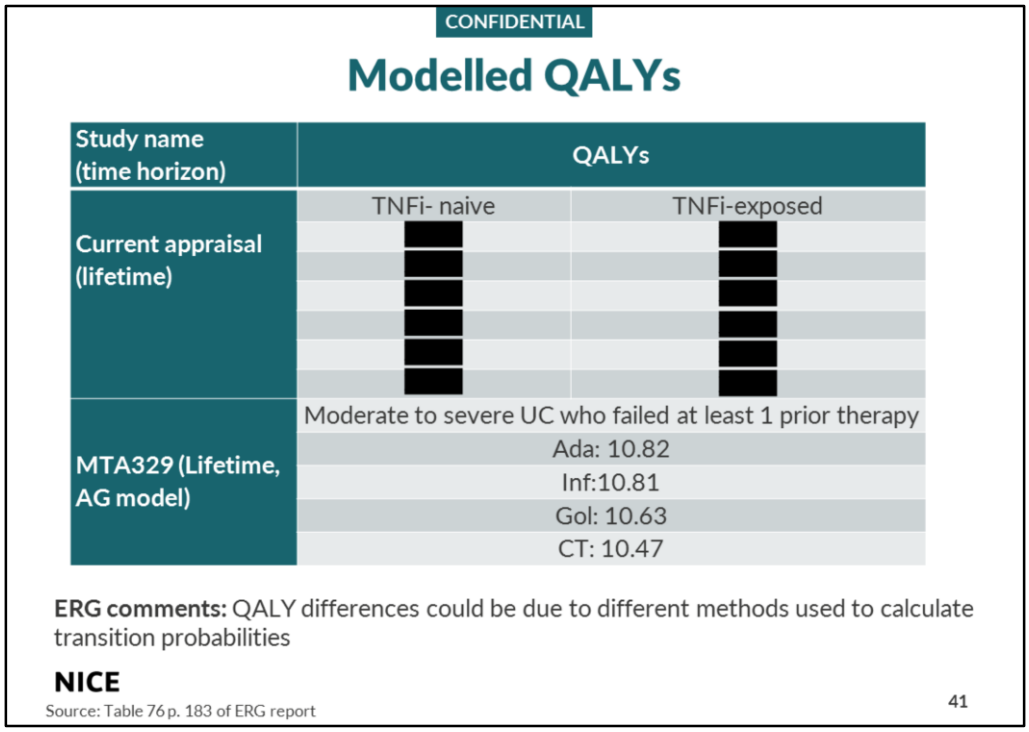
Company's scenario analysis

Company scenarios	Brief rationale/assumption	ICERs for Tofacitinib vs CT (£/QALY)	
		TNFi-naïve	TNFi-exposed
Company base case		£8,554	£10,302
Tofacitinib maintenance dose mix	[redacted] of patients receiving 5mg; [redacted] of patients receiving 10mg	£12,628	£13,947
OCTAVE trial utilities	EQ-5D data were collected in Tofacitinib Phase III clinical trials	£15,508	£18,276

ERG comments: company do not explore impact of key assumptions such as inclusion of costs associated with stoma care, cost-effectiveness results from alternative NMA models. ERG extend the range of scenario analyses in ERG additional analyses.

NICE

Source table 74, page 180 of ERG report



CONFIDENTIAL

Modelled QALYs

Study name (time horizon)	QALYs	
	TNFi- naive	TNFi-exposed
Current appraisal (lifetime)	[redacted]	[redacted]
MTA329 (Lifetime, AG model)	Moderate to severe UC who failed at least 1 prior therapy	
	Ada: 10.82	
	Inf:10.81	
	Gol: 10.63	
	CT: 10.47	

ERG comments: QALY differences could be due to different methods used to calculate transition probabilities

NICE

Source: Table 76 p. 183 of ERG report

41

The most relevant analysis for the current appraisal is the final version from the NICE TA of vedolizumab (TA342). This appraisal relates to same patient population as the current appraisal and comparators overlap, except Tofacitinib and surgery.

ERG additional analyses (*with PAS for TOF*)

- ERG made some corrections* to company base case and developed a ERG preferred base case. ERG also ran several scenario analyses which results are presented in part 2 (as these include PAS for VED)

TNF-i naïve

ICER TOF vs conventional	ICER TOF vs ADA	ICER TOF vs GOL	ICER TOF vs INF	ICER TOF vs VED
Reminder: Company base case corrected by ERG				
£8,564	TOF dominant	TOF dominant	TOF dominant	£615,077 (SW)
Average age: 41 years				
£8,562	TOF dominant	TOF dominant	TOF dominant	£614,916 (SW)
+ ERG preferred NMAs for remission and response				
£8,584	TOF dominant	TOF dominant	TOF dominant	£590,046 (SW)
+ Frequentist NMA for serious infections				
£7,886	TOF dominant	TOF dominant	TOF dominant	£607,571 (SW)
*+ Cost of stoma-care = **ERG base case***				
£7,815	**TOF dominant**	**TOF dominant**	**TOF dominant**	**£607,571 (SW)**

42

*ERG corrected 3 main errors: Error in cost calculation for elective surgery and conventional therapy, Error in estimation of weight – wastage, Error in incremental cost & QALY **Vedolizumab has a confidential PAS and results incorporating this are presented in part 2

ERG additional analyses (*with PAS for TOF*)

TNF-i experienced

ICER TOF vs CT	ICER TOF vs ADA	ICER TOF vs VED
Reminder: Company base case corrected by ERG		
£10,311	TOF dominant	£7,838,381 (SW)
Average age: 41 years		
£10,304	TOF dominant	£7,798,892 (SW)
+ ERG preferred NMAs for remission and response		
£10,148	TOF dominant	TOF dominant
+ Frequentist NMA for serious infections		
£9,458	TOF dominant	TOF dominant
*+ Cost of stoma-care = **ERG preferred***		
£9,389	**TOF dominant**	**TOF dominant**

SW: south-west

NICE

43

Equality issues

- No potential equality issues raised during scoping or by the company
- Patient perspective: Potential equality issues that should be considered are:
  - women who have not yet completed their family
  - people who consider surgery to be unacceptable due to cultural or religious factors
  - cost may also be a factor associated with lower income.

Innovation (Company)

- First therapy in its class; offers a new mechanism of action in ulcerative colitis
- Oral therapy given as monotherapy; alternative to current parenteral treatments
- Small molecule that should not be associated with issues relating to immunogenicity
- Opportunity to stop treatment and restart with similar efficacy
- Rapid improvements in ulcerative colitis symptoms

Key cost effectiveness issues

- Are the comparators appropriate for each sub-group?
  - Company base case excludes ADA in TNFi-exposed group
- What is the committee's view on:
  - The most appropriate source of health-related quality of life data?
  - Patient characteristics (e.g. age) being different depending on TNFi exposure status?
  - Importance of stoma care costs and surgery costs ?
  - Application of stopping rules in the model vs. clinical practice
- What is the committee preferred scenario?

NICE

46

Authors

- **Aminata Thiam**
  Technical Lead
- **Victoria Kelly**
  Technical Adviser
- with input from the Lead Team (Mohit Sharma and Rita Faria)

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Tofacitinib for moderately to severely active ulcerative colitis [ID 1218]

Document B

Company evidence submission

Pfizer confirm that all information in the submission summary is an accurate summary or replication of evidence in the main submission and accompanying appendices and that wherever possible a cross reference to the original source is provided

May 2018

File name	Version	Contains confidential information	Date
ID1218 UC_tofacitinib Document B	1.0	Yes	15th of May 2018

Contents

Submission length excluding template and references: 150 pages

Tables

Figures

Abbreviations

ADA	adalimumab
AE	adverse event
AG	assessment group
ANCOVA	analysis of covariance
ASA	aminosalicylates
AZA	azathioprine
BID	twice daily
CC	colectomy complications
CHMP	Committee for Medicinal Products for Human Use
CI	confidence interval
CrI	credible interval
CPRD	Clinical Practice Research Database
CRC	colorectal cancer
CRP	C-reactive protein
CSR	clinical study report
DIC	Deviance Information Criterion
ECCO	European Crohn's and Colitis Organisation
EMA	European Medicines Agency
EPAR	European public assessment report
EQ-5D	5-dimension EuroQol questionnaire
FAS	full analysis set
GI	gastrointestinal
GOL	golimumab
HDL	high-density lipoprotein
HRQoL	health-related quality of life
IBD	inflammatory bowel disease
IBDQ	Inflammatory Bowel Disease Questionnaire
ICER	incremental cost-effectiveness ratio
IM	immunomodulatory agent
INF	infliximab

INMB	incremental net monetary benefit
IPAA	ileal pouch-anal anastomosis
IR	incidence rate
JAK	Janus kinase
LDL	low-density lipoprotein
LTE	long-term extension
LYG	life years gained
MACE	major cardiovascular events
MCID	minimal clinically important difference
MCS	mental health component summary
MedDRA	Medical Dictionary for Regulatory Activities
MP	mercaptopurine
NICE	National Institute for Health and Care Excellence
NMA	network meta-analysis
NMB	net monetary benefit
NMSC	non-melanoma skin cancer
NRI	non-responder imputation
OP	outpatient
PAS	patient access scheme
PBO	placebo
PbR	payment-by-results
PCS	physical component summary
PPAS	per-protocol analysis set
PY	patient-year
QALY	quality-adjusted life year
QW	every week
Q2W	every 2 weeks
Q4W	every 4 weeks
Q8W	every 8 weeks
SAE	severe adverse event
SAS	safety analysis set

SD	standard deviation
SE	standard error
SF-36	36-Item Short Form survey
SmPC	summary of product characteristics
SoC	Standard of Care
SOC	system organ class
STA	single technology appraisal
STAT	signal transducer and activator of transcription
SUCRA	surface under cumulative ranking curve
TEAE	treatment-emergent adverse event
TNF	tumour necrosis factor
TNFi	tumour necrosis factor inhibitor
TOF	tofacitinib
UC	ulcerative colitis
UCSS	ulcerative colitis symptom score
UI	utility index
ULN	upper limit of normal
VAS	visual analogue scale
VED	vedolizumab
WPAI	Work Productivity and Activity Impairment-Ulcerative Colitis

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 Decision problem

The submission covers the technology's full (anticipated) marketing authorisation for this indication.

Table 1 The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	People with moderately to severely active ulcerative colitis	People with moderately to severely active ulcerative colitis	N/A
Intervention	Tofacitinib	Tofacitinib	N/A
Comparator(s)	• TNF-alpha inhibitors (infliximab, adalimumab and golimumab) • Vedolizumab • Conventional therapies, which may include a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclometasone, budesonide, hydrocortisone or prednisolone), thiopurines (mercaptopurine or azathioprine)	• TNF-alpha inhibitors (infliximab, adalimumab and golimumab) • Vedolizumab • Conventional therapies, which may include a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclometasone, budesonide, hydrocortisone or prednisolone), thiopurines (mercaptopurine or azathioprine)	N/A

Outcomes	The outcome measures to be considered include: **Efficacy** • measures of disease activity, including rates and duration of response, relapse and remission • achieving mucosal healing **Health outcomes** • health-related quality of life • rates of surgical intervention • time to surgical intervention • rates of hospitalisation **Safety** • adverse effects of treatment • mortality	The outcomes considered are: **Efficacy** • measures of disease activity, including rates and duration of response, relapse and remission • achieving mucosal healing **Health outcomes** • health-related quality of life (IBDQ, EQ-5D, SF-36) • rates of surgical intervention • rates of hospitalisation **Safety** • adverse effects of treatment • mortality	Time to surgical intervention was not assessed in the OCTAVE trials.
Subgroups to be considered	If evidence allows the following subgroups will be considered: • people who have been previously treated with one or more TNF-alpha inhibitors and people who have not received prior TNF-alpha inhibitor therapy.	Due to evidence limitations following prior treatment subgroups have been considered as a decision tool for people with moderately to severely active ulcerative colitis: • people who are biologic naïve • people with prior exposure to biologics	

Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; IBDQ, Inflammatory Bowel Disease Questionnaire; SF-36, 36-Item Short Form Survey; TNF, tumour necrosis factor.

B.1.2 Description of the technology being appraised

A draft version of the summary of product characteristics (SmPC) has been included in Appendix C. This document is subject to being updated until publication of the European public assessment report.

Table 2 Technology being appraised

UK approved name and brand name	• UK approved name: Tofacitinib citrate • Brand name: XELJANZ.
Mechanism of action	• Janus kinase (JAK) inhibitor.
Marketing authorisation/CE mark status	• Regulatory submission to EMA: The application was submitted on 27th July 2017. • [redacted] • [redacted] • [redacted]
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	XELJANZ is indicated for the treatment of adult patients with moderately to severely active ulcerative colitis (UC) who have had an inadequate response, lost response, or were intolerant to either conventional therapy or a biologic agent.
Method of administration and dosage	The recommended dose is 10 mg given orally twice daily for induction for 8 weeks and 5 mg given twice daily for maintenance [a]
Additional tests or investigations	No additional tests or investigations are required beyond those that are already part of current clinical practice for NICE recommended biologic treatments in ulcerative colitis.
List price and average cost of a course of treatment	Acquisition costs: List price: • Tofacitinib 5 mg: £690.03 per pack of 56 tablets • Tofacitinib 10 mg: [redacted] per pack of 56 tablets Discounted Price: • Tofacitinib 5 mg: [redacted] per pack of 56 tablets • Tofacitinib 10 mg: [redacted] per pack of 56 tablets Average cost per course of treatment: List price: • Year one: £10,350.45 • Subsequent annual cost: £8,970.39 Discounted price: • Year one: [redacted] • Subsequent annual cost: [redacted]
Patient access scheme (if applicable)	Tofacitinib (Xeljanz®) is currently recommended by NICE (TA480) for the indication in rheumatoid arthritis, which includes a confidential patient access scheme (PAS). [redacted]

[a] draft SmPC

[redacted]

B.1.3 Health condition and position of the technology in the treatment pathway

Summary

- Ulcerative colitis is a lifelong inflammatory disorder of the colon. Its incidence is highest in developed countries and is increasing. In the UK, the prevalence is an estimated 243–260 per 100,000 individuals.
- Ulcerative colitis is characterised by alternating periods of relapse and remission. Its physical symptoms, characterised by urgent bloody diarrhoea, are disabling. It has a significant negative impact on patient quality of life, social and psychological well-being, as well as daily functioning. Acute disease-related complications, such as severe bleeding or toxic megacolon, are associated with high mortality; chronic complications include an increased risk of colorectal cancer.
- Ulcerative colitis varies in severity with moderate-to-severe disease associated with worsening gastrointestinal symptoms and the development of systemic signs, such as fever and tachycardia.
- The primary goals of treatment are:
    - to rapidly induce remission
    - to maintain remission once achieved
    - to improve quality of life
    - to prevent complications.
- The treatment chosen depends on severity and extent of disease, prior therapies and patient preference. Mild-to-moderate disease is managed with conventional therapies (such as aminosalicylates, corticosteroids and immunomodulators). Biologic agents may be used in moderate-to-severe ulcerative colitis not responding to conventional therapies. Surgery may be considered for severe or refractory disease, or in the case of complications.
- Although the management of moderate-to-severe disease has improved, patients still have limited therapeutic choice and current options have a number of limitations. Current biologic agents induce remission in a minority of patients (18-39%, depending on the definition and type of analysis). Patients continue to live with a considerable symptom burden and a high risk of disability, and the rate of surgery remains at 20–30% within 25 years of diagnosis.
- Tofacitinib (Xeljanz, Pfizer) is a small molecule Janus kinase (JAK) inhibitor. It offers a new mechanism of action and is administered orally. It acts intracellularly to inhibit the JAK/STAT pathway, preferentially inhibiting JAK1 and JAK3, thereby interrupting the abnormal interactions between the gut and immune system.
- Tofacitinib is indicated for the treatment of adult patients with moderately to severely active ulcerative colitis who have had an inadequate response, lost response, or were intolerant to either conventional therapy or a biologic agent.
- Tofacitinib provides an additional treatment choice with clear benefits over current biological treatments in ulcerative colitis, including the absence of immunogenicity.

B.1.3.1 Disease overview

Ulcerative colitis is a lifelong inflammatory condition of the colon (1) and is the most common form of inflammatory bowel disease (IBD) (2). Its incidence has increased worldwide over the last 50 years and continues to rise (2, 3). The highest prevalence and incidence are seen in the developed world, particularly in Northern Europe, the United Kingdom and North America (4, 5). The incidence of ulcerative colitis is reported to be as high as 24.3 per 100,000 persons per year in Europe with prevalence as high as 505 per 100,000 (3). In the UK, the annual incidence of ulcerative colitis has been reported as ranging from 9 to 15 cases per 100,000 individuals (6), with prevalence ranging from 243 to 260 per 100,000 (6, 7). Based on these figures it was estimated that in 2011 146 000 people in the UK population of 60 million suffered from this condition (8), but this is probably a substantial underestimate if the common age of onset and lifelong duration are considered. Ulcerative colitis may present at any age, but most commonly affects adults in the second to fourth decades of life (2, 4, 9), resulting in disability that impacts patients in their most economically productive years.

Ulcerative colitis develops though a complex interaction of factors. The precise aetiology is unknown, so curative medical therapy is not yet available (10, 11). Current evidence suggests that innate and adaptive cellular immunity is key to disease pathogenesis in conjunction with the gut microbiota in genetically susceptible individuals, though epithelial barrier defects, and other environmental factors all play a role (9).

Clinically, ulcerative colitis is characterised by intermittent flares of symptoms interposed between variable periods of remission (4, 12). Flares can range in severity from minor to life-threatening (4, 13, 14) and are unpredictable both in severity and timing. About 50% of patients have a relapse in any year, with an appreciable minority having frequently relapsing or chronic, continuous disease (8).

The diagnosis of ulcerative colitis is based on the history of symptoms, endoscopic findings on colonoscopy, histology, and the exclusion of other causes of colonic inflammation such as infection (10). Patients with ulcerative colitis typically present with bloody diarrhoea (5). Other symptoms include faecal urgency and even incontinence, fatigue, increased frequency of bowel movements, tenesmus (a feeling of incomplete defaecation despite evacuation of the bowels), nocturnal defaecations, and abdominal pain (5, 8, 9, 15). The symptoms of ulcerative colitis vary according to the severity and extent of disease activity with greater severity and extent associated with worsening bloody diarrhoea and the development of systemic signs (9, 11).

The key endoscopic feature of ulcerative colitis is diffuse continuous mucosal inflammation affecting the rectum and a variable extent of the colon. The classic findings include erythema, loss of normal vascular pattern, bleeding, erosions and ulcerations (9). The extent of inflammation observed at colonoscopy is related to the risk of disease complications (8-10), but any extent of colitis can be associated with constitutional symptoms, including fatigue and fever (9, 11).

This classification by disease extent results in three main types:

- Proctitis (40% of patients), where only the rectum or the rectum and the sigmoid colon are inflamed. These patients primarily have rectal bleeding, urgency and tenesmus. They may not have diarrhoea even if they open their bowels many times a day.
- Left-sided colitis (20–45% of patients), involving the rectum, sigmoid colon and the descending colon. This is referred to as 'distal colitis' in the standard (Montreal) classification (16). These patients have symptoms of urgent bloody diarrhoea and abdominal cramping.
- Extensive or pancolitis (15–35% of patients), involves the left colon as well as some or all of the colon proximal to the splenic flexure. These patients usually present with bloody diarrhoea, more often become anaemic and are more likely to suffer from complications.

The British Society of Gastroenterology and the European Crohn's and Colitis Organisation define severity of disease activity based on clinical presentation. Patients with mild ulcerative colitis have fewer than 4 bowel movements per day with minimal blood. With moderate-to-severe disease patients have more than 4 bowel movements per day with increasing blood in the stool, increasing systemic symptoms and signs such as fever, tachycardia or anaemia (8, 11). Patients with severe disease have potentially life-threatening attacks (10). Moderate-to-severe disease, which is referred to as "subacute" in the NICE guidance (17) can be managed as an outpatient, but acute severe colitis is a medical emergency requiring hospitalisation (17, 18).

Numerous indices and scoring systems of disease activity have been developed; most are primarily used in clinical trials (5, 19). These may be based on clinical scores, clinical endoscopic scores or combined scoring systems.

In the clinical trial setting the Mayo Score has been the most widely used (20). It is clinically relevant, correlates well with both disease-specific Inflammatory Bowel Disease Questionnaire (IBDQ) and generic quality of life scores, such as 36-item Short Form survey (SF-36), and short-term (8-week) response has been used to predict steroid withdrawal or colectomy at 26-52 weeks (19, 21, 22). Disease severity is based on stool frequency, rectal bleeding, endoscopic findings and physician's global assessment (see Table 3). Full Mayo Scores range from 0 to 12 points, with higher scores indicating more severe disease. The 9-point non-invasive partial Mayo Score (i.e. without endoscopy) has also been shown to perform as well as the full Mayo Score to identify patient perceived clinical response (23).

Table 3 Mayo Score

Component	Description	Points
Stool frequency	Normal	0
	1–2 stools more than usual	1
	3–4 stools more than usual	2
	≥ 5 stools more than usual	3
Rectal bleeding	No blood	0
	Streaks of blood < 50% of time with stool	1
	Obvious blood most of time with stool	2
	Blood alone passed	3
Endoscopic findings	Normal/inactive disease	0
	Mild disease [a]	1
	Moderate disease [b]	2
	Erosions	3
Physician's global assessment	Normal	0
	Mild	1
	Moderate	2
	Severe	3

[a] Erythema, decreased vascular pattern and mild friability; [b] Marked erythema, lack of vascular pattern, friability and erosions. Walsh *et al.* 2016 (18).

B.1.3.2 Burden of disease

Impact on patient quality of life

Ulcerative colitis has wide-ranging effects on psychological and emotional health, education and employment, family life and social interactions, and fertility and pregnancy (3, 24). This is exacerbated by the chronic nature of the disease, its unpredictable course, the young age of onset and current therapies (24). This has a marked negative effect on patients' daily functioning: a European survey has reported that ulcerative colitis symptoms affect the ability to enjoy leisure activities in 73% of patients, and the ability to perform job functions in 66% of patients (25).

The physical symptoms of ulcerative colitis, including frequent diarrhoea and abdominal pain, have a negative impact on patients' health-related quality of life (HRQoL): a UK cross-sectional study has found moderately or severely active ulcerative colitis to be associated with significantly worse quality of life and significantly more work impairment compared to those whose disease was in remission (26). Coping with the unpredictable symptoms of ulcerative colitis can cause patients to experience anxiety and depression (27, 28). The condition has an impact not only on the patient, but the whole family, especially when it starts at a young age (29). The embarrassment and social stigma associated with faecal incontinence cause patients to fear this, which is extremely limiting for them (30-32). Some patients with ulcerative colitis are unable to sleep adequately, leading to daytime somnolence or fatigue (28). The presence of chronic fatigue has been found to be associated with worse HRQoL as measured by both the physical and emotional components of the generic 36-item Short Form survey (SF-36) and the disease-specific Inflammatory Bowel Disease Questionnaire (IBDQ) (33).

The profound impact on patient quality of life is underscored by the fact that patients with moderately to severely active ulcerative colitis have similar or worse SF-36 mental

component summary scores than patients with other chronic diseases such as breast cancer, chronic obstructive pulmonary disease, or patients on dialysis (34-37).

If patients undergo surgery, procedure-related complications can further affect their health-related Quality of Life (38, 39). Colectomy may remove the colon, but it does not return bowel function or quality of life to normal and extra-intestinal manifestations commonly persist after colectomy (40).

Extra-Intestinal Manifestations

The clinical burden of ulcerative colitis is not limited to gastrointestinal manifestations. Ulcerative colitis is associated with extra-intestinal manifestations, most commonly affecting the joints (peripheral arthritis or ankylosing spondylitis), skin (erythema nodosum, pyoderma gangrenosum) eyes (uveitis), or the hepatobiliary tract (primary sclerosing cholangitis) (41). 30 years after diagnosis, 50% of patients with ulcerative colitis have at least one extra-intestinal manifestation (41), and up to one-quarter of those will suffer from more than one (42). As some extra-intestinal manifestations occur with disease flares, they are expected to improve with treatment of bowel inflammation (41). Given the commonness and diversity of these manifestations, they represent a considerable source of morbidity and add to the overall disease burden of ulcerative colitis.

Complications

Patients with ulcerative colitis can suffer from acute or chronic disease-related complications. Acute complications include severe bleeding, toxic megacolon (a potentially life-threatening dilatation of the colon with systemic toxicity), peritonitis and bowel perforation (5). While the latter is rare and often associated with colonoscopy or toxic dilatation, it carries a mortality of up to 50% (10, 43).

Colonic dysplasia and cancer are well-recognised chronic complications of ulcerative colitis. The reported cumulative risk of colorectal cancer for all patients ranges broadly and historically this has been estimated up to 43% after 35 years of disease (44). More recent studies suggest that over time this risk has decreased and might be approaching the general population; however, the risk remains elevated in certain populations, such as those with extensive or long duration of disease, primary sclerosing cholangitis (an extra-intestinal manifestation), and uncontrolled inflammation (9). As underlying bowel inflammation is the driver for this colorectal cancer risk, medical therapies that reduce inflammation appear to reduce this risk (45-48).

The rate of surgery has decreased over past decades with improved management, but is still substantial, with approximately 20–30% of patients undergoing colectomy within 25 years of diagnosis. Rates can be as high as 40% if admitted once or more often to hospital with acute severe colitis (UK data) (49), or 30% within 10 years in cases of extensive disease (4, 14, 50, 51). If patients undergo surgery there is a risk of further complications including ileostomy or pouch dysfunction, adhesions causing obstruction, infertility (especially in women), and changes in sexual function (8, 38).

Despite the risk of disease-related complications, overall mortality is not increased in European patients with ulcerative colitis compared to the general population (6, 52). However, mortality is increased among newly diagnosed patients, in patients with extensive disease, or those with acute severe colitis, especially in older patients. Mortality is higher than the general population for three years after hospital admission with severe colitis, especially for patients who avoid emergency colectomy, presumably because of persistent, poorly controlled disease or delayed decision making (53). Over the long-term, there is an increased risk of dying from ulcerative colitis-related complications (primary sclerosing cholangitis, cholangiocarcinoma, colorectal cancer) (4).

Impact on healthcare system

The burden of ulcerative colitis on healthcare resource utilisation is substantial due to its early age of onset, chronic relapsing and remitting course, the likelihood of hospitalisation or surgery, and the association with extra-intestinal manifestations(6). When combined with the indirect costs related to lost work productivity and daily activity impairment in an economically active patient group, the overall costs of ulcerative colitis pose a significant economic burden to society (54). This overall burden in Europe has been estimated to be in the range of €12.5 to €29.1 billion (2008 values) (54). Previously the cost of hospitalisation was the largest component of direct total medical costs. (54). More recent studies have shown that with the introduction of newer therapies this has shifted. Medication costs now predominate, with a corresponding reduction in those for hospitalisation and surgery, so that overall direct costs have remained stable (55). Compared with patients with quiescent disease, symptom flares have been found to lead to a 2–3-fold increase in healthcare costs for patients managed in an ambulatory setting. With hospitalisation this increases to a more than 20-fold increase in costs. This reinforces the view that novel therapies capable of maintaining remission or reducing the need for inpatient care may prove cost effective despite their high acquisition costs when compared with other drug therapies (56). Indirect costs associated with ulcerative colitis account for between 54% and 68% of the overall burden (54), with loss-of-productivity costs accounting for 31% in some studies (57). In addition, the management of complications associated with ulcerative colitis surgery represents a considerable expense to healthcare systems, with pouchitis, pouch failure and small bowel obstruction carrying the greatest burden (58). Direct costs, hospitalisations and surgeries all increase with increasing severity of disease (54).

B.1.3.3 Treatment overview

Objective of treatment

The primary goals of treatment are:

- to rapidly induce remission (11)
- to maintain remission once achieved (11)
- to improve quality of life (17)
- to prevent complications (9).

The European Crohn's and Colitis Organisation agreed that the best definition of remission was a combination of clinical parameters (stool frequency ≤ 3/day with no bleeding) and no mucosal lesions at endoscopy, which in clinical trials is frequently defined as an endoscopic Mayo Score of zero or one (1, 11). For patients, what matters most is that clinical remission is steroid-free (59), as approximately 50% of patients receiving steroids experience side effects (60), of which patients are fearful (61, 62).

Mucosal healing has been associated with long-term clinical remission, decreased risk of surgery, and corticosteroid-free clinical remission (9, 63). This is probably why mucosal healing is also associated with patients achieving a good HRQoL (6). Furthermore, active ulcerative colitis, particularly when extensive and associated with moderate or severe mucosal inflammation, is a risk factor for developing colorectal cancer which makes assessing its resolution important (64).

Clinical pathway of care

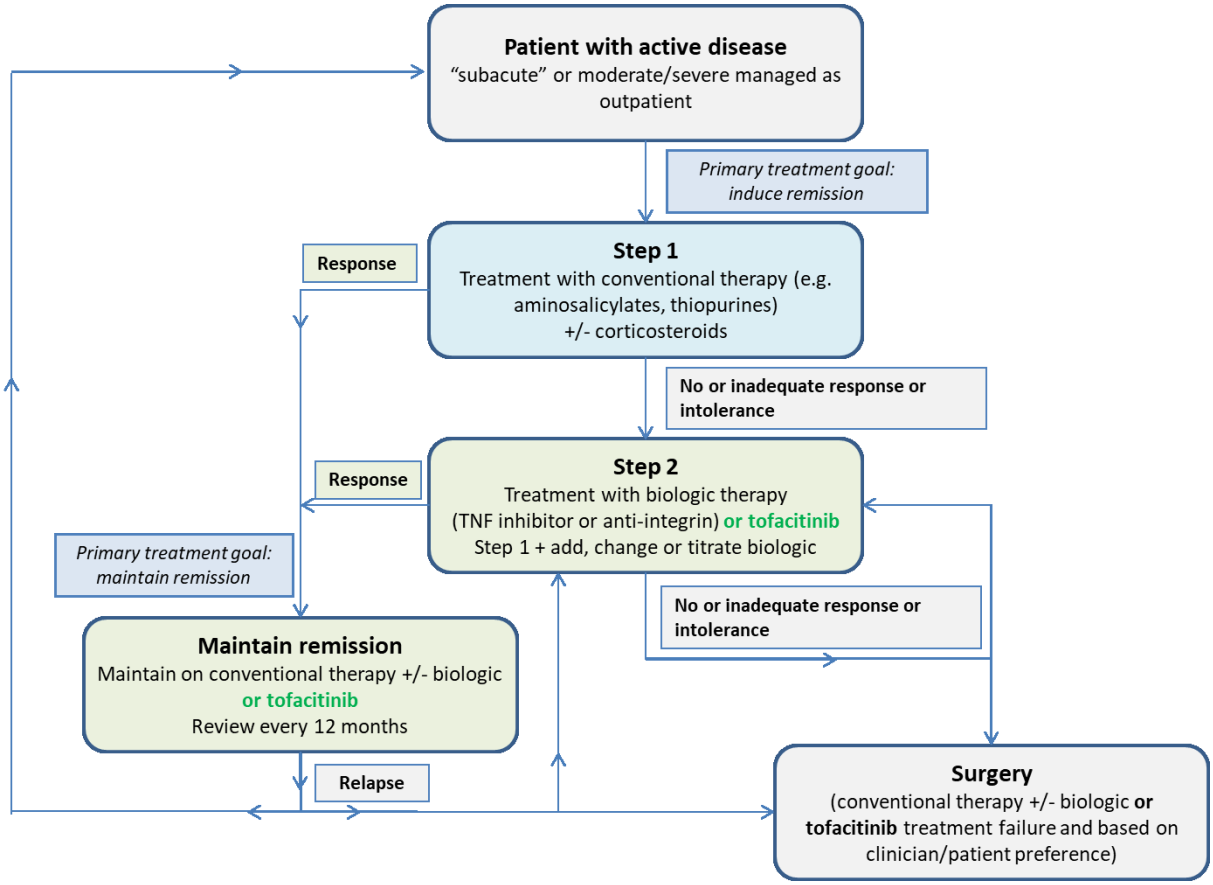
The treatment of patients with ulcerative colitis depends on severity and extent of disease, prior therapies and patient preference (5, 17). These treatments may be medical or surgical, with all patients managed medically, before surgery in some cases. Patients with mild disease are offered oral or topical aminosalicylates conventional therapies, oral immunomodulators (usually azathioprine or mercaptopurine) and corticosteroids. These therapies are generally adequate for managing disease of lesser severity (5).

Management of patients with moderate to severely active ulcerative colitis can be more challenging and typically involves a step-up approach based on patient history, treatment response, and tolerance of individual therapies (59, 65). If there is inadequate response to conventional therapies then a biological therapy (either a tumour necrosis factor alpha inhibitor [TNFi] or the anti-integrin agent vedolizumab) may be considered (5, 65, 66). If symptoms cannot be adequately controlled with medical therapy, if patients feel that medication does not give them adequate quality of life, or if there are other grounds for doing so (for example evidence of dysplasia or recurrent flares), surgery may be considered (8) (see Figure 1). Decision making is an iterative process since there are as yet no biomarkers to determine response to therapy. The median time to colectomy in those who undergo surgery is around 7–11 years (49).

Unmet need in treatment

The therapies currently available for the management of ulcerative colitis have resulted in improvements in treatment and outcomes, but there is substantial unmet need. Patients continue to live with a considerable symptom burden and high risk of disability with few treatment options (25, 67). This is especially true in patients in need of biologics or surgery; these patients have been shown to have significantly lower HRQoL at follow-up than patients with Crohn's disease treated with biological therapy (6). This may be particularly true in younger patients with ulcerative colitis (below 40 years) who tend to have more aggressive disease and require more intensive medical and surgical management compared to those with later-onset disease (11).

Figure 1 Proposed position of tofacitinib within the treatment pathway for patients with moderately to severely active ulcerative colitis, in accordance with NICE recommendations and clinical practice



Currently available therapies, both conventional and biologic have several limitations (see Table 4).

Table 4 Key limitations of current pharmacotherapies for moderately to severely active ulcerative colitis

Class/Drug/Route of Administration	Key Limitations
Corticosteroids	
Oral and IV	• Not suitable for maintenance use due to side effects associated with long-term use (68, 69) • Significant side effects, including extensive endocrine, metabolic, musculoskeletal, neurological, and infectious complications (68, 69) • Corticosteroid dependence or refractoriness in approximately half of the patients over 1 year after receiving first course of corticosteroid (70)
Immunomodulators	
AZA/MP (Oral)	• Cochrane meta-analysis shows efficacy in maintenance but not in induction (71) • Slow therapeutic response that may take several months and therefore not suitable as an induction agent (72) • Safety concerns including pancreatitis, serious infections, myelosuppression, hepatotoxicity, lymphoma, non-melanoma skin cancer, possibly other malignancies (72)
TNFi agents	
Adalimumab (SC) **Golimumab (SC)** **Infliximab (IV)**	• Failure to respond to induction therapy (i.e., primary non-response) in approximately one-third to half of patients (73-75) • Substantial rate of loss of response over time (i.e., secondary non-response) in up to 50% of initial responders (76, 77) • No controlled data on efficacy in patients with prior TNFi failure, except data that showed lack of efficacy with adalimumab treatment in patients with prior secondary non-response to TNFi (78, 79) • Added burden of therapeutic drug monitoring for optimisation in both induction and maintenance treatment (80) • Safety concerns including serious infections (e.g., bacterial, TB, fungal, other opportunistic infections); malignancy (e.g., lymphoma (81)) • Need for concomitant immunosuppressant therapy, especially with infliximab, to optimise efficacy and/or reduce immunogenicity (78, 82) • No current oral options; regular visits to infusion centre settings for IV route of administration (infliximab) or need for refrigeration for SC route of administration (adalimumab/golimumab); potential for infusion or injection site reactions
Anti-integrin agent	
Vedolizumab (IV)	• Onset of action not viewed as sufficiently rapid in some patients with moderately to severely active disease (83, 84) • Bridging therapy commonplace (often with steroids or ciclosporin) until vedolizumab takes effect • Lower induction efficacy (placebo adjusted) for TNFi failure patients: 6.6% and 18.4% remission and response respectively versus 16.5% and 26.8% for TNFi-naïve patients (85) • No oral options; regular visits to infusion centre settings for IV administration and need for refrigeration; potential for infusion site reactions

Abbreviations: MP, mercaptopurine; AZA, azathioprine; IV, intravenous; SC, subcutaneous; TB, tuberculosis; TNFi, tumour necrosis factor inhibitor.

In summary, the key limitations of current therapeutics for moderately to severely active disease are:

- primary non-response to induction (TNFi, anti-integrin agent)
- secondary non-response (TNFi, anti-integrin agent)
- slow onset on action (anti-integrin agent, immunomodulators)
- need for therapeutic drug monitoring (immunomodulators, TNFi)
- lack of suitability as a long-term maintenance therapy (corticosteroids)
- lack of oral options (all biologics)
- lower efficacy/lack of data with previous TNFi therapy (TNFi, anti-integrin agent)
- side-effect and safety concerns (all)

This unmet need underscores the necessity for novel treatments with new mechanisms of action to increase therapeutic choice for patients.

B.1.3.4 Technology

Tofacitinib citrate

Tofacitinib citrate (Xeljanz®) is an innovative, orally administered small molecule with a novel mode of action for the treatment of ulcerative colitis: inhibition of the JAK family of kinases. The tofacitinib molecule is similar in structure to adenosine triphosphate (ATP) without the phosphate group, thereby competing with ATP at target sites, and is a potent, selective JAK inhibitor (86).

In kinase assays, tofacitinib inhibits JAK1, JAK2, JAK3, and to a lesser extent tyrosine kinase 2 (TYK2). In cellular settings where JAK kinases signal in pairs, tofacitinib preferentially inhibits signalling by heterodimeric receptors associated with JAK3 and/or JAK1 with functional selectivity over receptors that signal via pairs of JAK2 (87).

The JAK-STAT pathway in ulcerative colitis

The pathogenesis of IBD is complex and multifactorial, but the role of the immune system and inflammatory cascade is key to understanding the disease and the role of current and future treatment options.

The pathogenesis of IBD is partly related to altered barrier function in the form of structural changes to intestinal epithelial cells (88). Substantial changes in the microbiota of the gut can elicit an inflammatory response from both the innate and adaptive immune systems. In a chronic disease such as ulcerative colitis, an exaggerated immune response causing infiltration of the lamina propria by immune cells (e.g. macrophages, T and B cells) leads to over production of pro-inflammatory cytokines such as TNF-α, interferon-γ, IL-1β, interleukins 2, 5, 6, 7, 9, 12, 21, 22, and 23, all of which may result in damage to the mucosal barrier (9, 88, 89).

Many pro-inflammatory cytokines implicated in the pathogenesis of ulcerative colitis utilise the JAK-STAT pathway to induce intracellular signalling. JAKs are tyrosine kinases that consist of 4 members in the JAK family: JAK1, JAK2, JAK3, and TYK2. JAKs are activated by external stimulus of the cell receptor by pro-inflammatory cytokines, and initiate a cascade of phosphorylation and dimerisation of signal transducer and activation of

transcription (STAT) molecules, that translocate to the cell nucleus, triggering gene transcription (Figure 2).

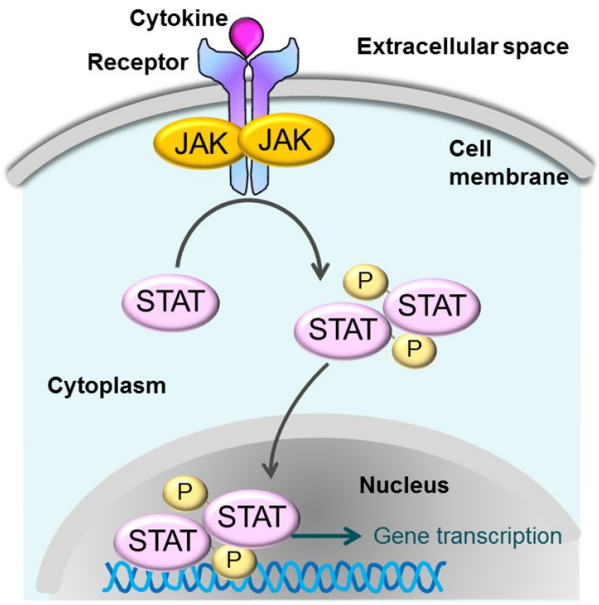


Figure 2 The JAK-STAT signalling pathway

Cytokine binding to its cell surface receptor leads to receptor polymerisation and autophosphorylation of associated JAKs

Activated JAKs phosphorylate the receptors that dock STATs

Activated JAKs phosphorylate STATs, which dimerise and move to the nucleus to activate new gene transcription

Abbreviations: JAK, Janus Kinase; P, phosphate group; STAT, signal transducer and activation of transcription

The way that JAKs pair will determine their role in cytokine signalling, for example, JAK1 pairs with JAK3 to control signalling of the common γ-chain cytokines such as IL-9, which in ulcerative colitis is expressed in the lamina propria where it correlates with disease severity and negatively impacts the wound healing of the intestinal epithelium (Figure 3) (88, 90).

Figure 3 Key Cytokines in IBD and JAK Combinations for Cytokines That Depend on JAK Pathways for Signalling (88, 91)

Abbreviations: IL, interleukin; JAK, Janus Kinase; TKY, tyrosine kinase

The current biologic treatment options infliximab, adalimumab and golimumab have a mode of action that targets one cytokine, TNF-α. The intracellular mode of action of tofacitinib targeting the JAK/STAT pathway means it can modulate the response to multiple cytokines implicated in the pathogenesis of ulcerative colitis. Inhibition of JAK1/3 by tofacitinib is expected to block signalling through the gamma common chain containing receptors for several cytokines, including IL-2, -4, -7, -9, -15, and -21. These cytokines are integral to lymphocyte activation, proliferation, and function and inhibition of their signalling may result in modulation of multiple aspects of the immune response. In addition, crossover to JAK1 may result in some attenuation of signalling through additional cytokines such as IL-6 and Type 1 and Type 2 interferon-γ (Figure 3).

Tofacitinib therefore offers a new therapeutic approach to treating ulcerative colitis. Unlike current biologic therapies, it is a small molecule, taken orally rather than by injection or infusion, providing a further treatment option for patients. As a small molecule tofacitinib is likely to avoid issues related to immunogenicity, for example the production of anti-drug antibodies that can reduce efficacy over time, as seen with large proteins such as the TNFi and anti-integrin monoclonal antibodies (92). Development of anti-drug antibodies to current biologic treatment has been shown to be associated with secondary loss of response in patients with ulcerative colitis, resulting in a requirement to dose-escalate in many patients (93-95). Azathioprine is often co-administered with biologic therapy to prevent the development of anti-drug antibodies and improve efficacy (96). By contrast, tofacitinib is a monotherapy and does not require concomitant administration of immunomodulators.

Therapeutic drug monitoring (TDM) with biologics is common in gastroenterology clinical practice in order to determine plasma levels of active drug in patients who have developed secondary loss of response. TDM can support the clinical decision to dose escalate with biologics in these patients. In the OCTAVE trials, pharmacokinetic data were collected for patients on both 5 mg and 10 mg doses of tofacitinib for up to 52 weeks (97). The data show that tofacitinib plasma concentration in an individual patient reaches steady state within 24 hours of the start of therapy and remains stable (that is, no significant change between visits) over the course of maintenance treatment. Within the 5 mg and 10 mg dosing groups, small variations in plasma concentration of tofacitinib at 52 weeks of treatment did not correlate with changes in remission status. Taken together, these data support the understanding that tofacitinib, as a synthetic small molecule, does not provoke the immunogenicity response associated with protein-based therapeutics, and that TDM is unnecessary in patients treated with tofacitinib.

B.1.4 Equality considerations

It is not anticipated that this appraisal will exclude from consideration any people protected by the equality legislation, lead to a recommendation that has a different impact on people protected by equality legislation than on the wider population, or lead to recommendations that have any adverse impact on people with a particular disability or disabilities.

B.2 Clinical effectiveness

Summary

The efficacy and safety of tofacitinib for the treatment of moderately to severely active ulcerative colitis was investigated in three Phase III trials: OCTAVE Induction 1 and 2, and the OCTAVE Sustain maintenance trial. Tofacitinib demonstrated rapid and sustained improvements of symptoms and effective control of disease activity, with a safety profile generally similar to that of biologics and consistent with that of tofacitinib in rheumatoid arthritis.

All three OCTAVE trials met their primary endpoints:

- In OCTAVE Induction 1 and 2, the proportion of patients achieving remission at week 8 was significantly higher with tofacitinib 10 mg than with placebo (OCTAVE Induction 1, 18.5% vs 8.2%; $p = 0.007$; OCTAVE Induction 2, 16.6% vs 3.6%; $p = 0.0005$).
- In OCTAVE Sustain, 40.6% of patients treated with tofacitinib 10 mg and 34.3% receiving tofacitinib 5 mg group achieved remission at week 52, compared with 11.1% in the placebo group (both $p < 0.0001$).

Results for all key clinical and quality of life secondary endpoints showed significantly greater efficacy with tofacitinib than with placebo across Induction and Maintenance (section B.2.6):

- Clinical response and clinical remission (used in the NMA and economic model):
  - The proportion of patients achieving clinical response at week 8 and week 52 was significantly higher with tofacitinib (TOF) than with placebo (OCTAVE pooled Induction, 57.6% vs 30.8%, $p < 0.0001$ and OCTAVE Sustain, 61.9% (TOF 10 mg), 51.5% (TOF 5 mg) vs 20.2% (both $p < 0.0001$).
  - The proportion of patients achieving clinical remission at week 8 and week 52 was significantly higher with tofacitinib than with placebo (OCTAVE pooled Induction, 17.7% vs 6.0%, $p < 0.0001$ and OCTAVE Sustain, 41.1% (TOF 10 mg), 34.3% (TOF 5 mg) vs 11.1% (both $p < 0.0001$).
- Mucosal healing:
  - The proportion of patients achieving mucosal healing at week 8 and week 52 was significantly higher with tofacitinib than with placebo (OCTAVE pooled Induction, 29.9% vs 13.7%, p < 0.0001 and OCTAVE Sustain, 45.7% (TOF 10 mg), 37.4% (TOF 5 mg) vs 13.1% (both $p < 0.0001$).
- Sustained corticosteroid-free remission:
  - Among patients in remission at OCTAVE Sustain baseline, patients treated with tofacitinib were more likely to achieve sustained corticosteroid-free remission than those receiving placebo (tofacitinib 10 mg, 47.3%; tofacitinib 5 mg, 35.4%; placebo, 5.1%; both $p < 0.0001$).
- Health Related Quality of Life (HRQoL):
  - Compared to placebo, patients receiving tofacitinib had significant and meaningful improvements in HRQoL measures assessed in the OCTAVE studies. Significant improvements were seen as early as week 2 in the Induction studies and HRQoL differences between patients treated with tofacitinib and those receiving placebo persisted over 52 weeks of maintenance therapy in OCTAVE Sustain.

A consistent treatment effect was observed with tofactinib between TNFi experienced and TNFi-naive patients across Induction and maintenance studies

- The results of subgroup analyses of the OCTAVE trials demonstrated that tofacitinib was significantly more efficacious than placebo regardless of prior use of TNFi therapies or corticosteroids (section B.2.7).

Network meta-analysis (section B.2.9)

- An NMA comparing the effects of tofacitinib and comparators (adalimumab, golimumab, infliximab and vedolizumab) on clinical response and clinical remission showed tofacitinib to be an efficacious induction and maintenance treatment in both patients with and without prior TNFi exposure.
    - The NMA suggests that across induction and maintenance, tofacitinib presents comparable or numerically better efficacy than vedolizumab; comparable or sinificantly better efficacy than most TNFi therapies.

Adverse event rates in the OCTAVE studies were similar with tofacitinib and placebo, and the tofacitinib safety profile in ulcerative colitis is generally similar to that of TNFi and consistent with that of tofacitinib in rheumatoid arthritis (section B.2.10)

- The rates of adverse events was similar in the tofacitinib and placebo groups during both induction and maintenance therapy, and in OCTAVE Sustain, adverse event rates were similar in the two tofacitinib groups, and were generally mild and manageable.
- Serious adverse events rates weren't significantly different between treatment groups; with event rates being numerically higher for placebo compared to both tofacitinib groups.

Key OCTAVE Open (open-label extension study) results (section B.2.6.3)

- An additional 8 weeks of induction treatment with tofacitinib 10 mg, for a total of 16 weeks of induction therapy, resulted in [redacted] and [redacted] of the OCTAVE Induction non-responders gaining clinical response or clinical remission, respectively
- Sustained remission was demonstrated across doses, with [redacted] and [redacted] of patients in remission in OCTAVE Sustain on tofacitinib 5 mg and 10 mg, respectively, maintaining remission up to 12 months in OCTAVE Open.
- Dose-increase and re-treatment was demonstrated, with remission achieved at 8 weeks of treatment with tofacitinib 10 mg in [redacted] and [redacted] of patients who had treatment failure during OCTAVE Sustain on tofacitinib 5 mg or placebo, respectively.

Conclusion:

Tofacitinib is an innovative therapy for moderately to severely active ulcerative colitis, with a novel mechanism of action providing rapid improvement in symptoms and long-term treatment response, even after interruption to treatment. It is an oral therapy that is effective in both TNF-naïve and TNF-experienced patients, and as a small molecule is not expected to be associated with the production of anti-drug antibodies.

To further aid the interpretation of the clinical outcomes from the Phase III clinical trials of tofacitinib in relation to the decision problem set out by NICE, top line results are presented in Table 5 and Table 6. For definitions of central and local reads, please refer to Section B.2.3.1.2.4 and B.2.3.1.3.4.

Table 5 Summary of statistical significance of OCTAVE Induction 1 and 2 outcomes relevant to the decision problem

Clinical Impact	Outcome assessed	Used in CEA?	Time points (weeks)	Endoscopic read	OCTAVE Induction 1	OCTAVE Induction 2
					Tofacitinib 10 mg (n=476)	Tofacitinib 10 mg (n=429)
Disease Activity	Remission (primary endpoint)	No	8	Central	Sig	Sig
				Local	Sig	Sig
	Mucosal healing (key secondary endpoint)	No	8	Central	Sig	Sig
				Local	Sig	Sig
	Clinical Remission	Yes	8	Central	Sig	Sig
				Local	Sig	Sig
	Clinical Response	Yes	8	Central	Sig	Sig
				Local	Sig	Sig
	Endoscopic Remission	No	8	Central	Sig	Sig
				Local	Sig	Sig
	Symptomatic Remission	No	8	Central	NS	Sig
				Local	Sig	Sig
	Deep Remission	No	8	Central	Sig	NS
				Local	Sig	Sig
	Total Mayo Score	No	8	Central	Sig	Sig
				Local	NR	NR
	Partial Mayo Score	No	2	NA	Sig	Sig
			4	NA	Sig	Sig
			8	NA	Sig	Sig
HRQoL	IBDQ treatment response and remission	No	4	NA	Sig	Sig
			8	NA	Sig	Sig
	Change from baseline in SF-36 MCS and PCS	No	8	NA	Sig	Sig
	Change from baseline in EQ-5D utility index and EQ-5D VAS score	Yes	2	NA	Sig	Sig
			8	NA	Sig	NS (UI) / Sig (VAS)

Statistical significance = $p < 0.05$.
Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; HRQoL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; MCS, Mental Component Summary; NA, not applicable; NR, not reported; NS, not significant; PCS, Physical Component Summary; SF-36, 36-Item Short Form Survey; Sig, significant difference versus placebo; UI, utility index; VAS, visual analogue scale

Table 6 Summary of statistical significance of OCTAVE Sustain outcomes relevant to the decision problem

Clinical impact	Outcome assessed	Used in CEA?	Time points (weeks)	Endoscopic read	Tofacitinib 5 mg (n = 198)	Tofacitinib 10 mg (n = 197)
Disease activity	Remission (primary endpoint)	No	52	Central	Sig	Sig
				Local	Sig	Sig
	Mucosal healing (key secondary endpoint)	No	52	Central	Sig	Sig
				Local	Sig	Sig
	• Sustained remission • Sustained mucosal healing • Sustained corticosteroid-free remission (key secondary endpoint) • Sustained clinical response	No	24 and 52	Central	Sig	Sig
				Local	Sig	Sig
	Clinical remission	Yes	52	Central	Sig	Sig
				Local	Sig	Sig
	Clinical response	Yes	52	Central	Sig	Sig
				Local	Sig	Sig
	Endoscopic, symptomatic and deep remission	No	52	Central	Sig	Sig
				Local	Sig	Sig
HRQoL	IBDQ treatment response and remission	No	8	NA	Sig	Sig
			24	NA	Sig	Sig
			52	NA	Sig	Sig
	Change from baseline in SF-36 MCS and PCS	No	24	NA	Sig	Sig
			52	NA	Sig	Sig
	Change from baseline in EQ-5D utility index and EQ-5D VAS score	Yes	8	NA	Sig	Sig
			24	NA	Sig	Sig
			52	NA	Sig	Sig

Statistical significance = $p < 0.05$.
Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; HRQoL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; MCS, Mental Component Summary; NA, not applicable; NR, not reported; NS, not significant; PCS, Physical Component Summary; SF-36, 36-Item Short Form Survey; Sig, significant difference versus placebo; VAS, visual analogue scale.

B.2.1 Identification and selection of relevant studies

Pfizer conducted a systematic review to identify all relevant clinical data from the published literature regarding the clinical effectiveness and safety of treatments in ulcerative colitis. Full details of the methodology used to identify and select the RCT and non-RCT clinical evidence relevant to the technology being appraised are reported in Appendix D along with a PRISMA flow diagram, full summary of the included and excluded studies and reasons for study exclusion, where applicable.

B.2.2 List of relevant clinical effectiveness evidence

The systematic review of clinical evidence identified five randomised controlled trials (RCT) with tofacitinib in ulcerative colitis. All five trials provide evidence supporting the application for marketing authorisation and are in populations relevant to the decision problem.

Tofacitinib has been investigated for the treatment of moderately to severely active ulcerative colitis in three Phase III RCTs: the OCTAVE Induction 1 and 2 trials and OCTAVE Sustain (NCT01465763, NCT01458951 and NCT01458574, respectively). Tofacitinib has also been compared with placebo in a Phase II trial (NCT00787202), which is not described in detail in this submission, but is included in the NMA (section B.2.9).

Patients in the OCTAVE Phase III trial programme were eligible to enter an ongoing open-label extension study, OCTAVE Open (NCT01470612). OCTAVE Open was not used to inform the NMA or the economic model because of its open-label, uncontrolled design but interim results, which provide additional evidence for the long-term efficacy and safety of tofacitinib, are included in section B.2.6.3.1.

Table 7 Clinical effectiveness evidence with most relevance to the decision problem

Study	**OCTAVE Induction 1 and 2 (NCT01465763, NCT01458951)**		
Study design	Two identical, multicentre, randomised, double-blind, placebo-controlled trials with an 8-week induction phase.		
Population	Patients aged 18 years or older who had a confirmed diagnosis of ulcerative colitis for at least 4 months. Patients had moderately to severely active disease, which was defined as a Mayo score of 6 to 12, with a rectal bleeding subscore of 1 to 3 and an endoscopic subscore of 2 or 3.		
Intervention(s)	Tofacitinib 10 mg twice daily		
Comparator(s)	Placebo		
Indicate if trial supports application for marketing authorisation	**Yes**	Indicate if trial used in the economic model	**Yes**
Rationale for use/non-use in the model	Studies provide evidence of the efficacy of tofacitinib and were included in the network meta-analysis used in the economic model.		
Reported outcomes specified in the decision problem	• Measures of disease activity: Mayo score and partial Mayo score • Rates of and duration of response and remission: Mayo score • Achieving mucosal healing (endoscopic findings) • Adverse effects of treatment		

	• HRQoL: IBDQ, SF-36, EQ-5D-3L and EQ-5D VAS • Hospitalisation and surgery due to ulcerative colitis.		
Study	**OCTAVE Sustain (NCT01458574)**		
Study design	Multicentre, randomised, double-blind, placebo-controlled trial with a 52-week maintenance phase.		
Population	Patients who completed the OCTAVE Induction 1 or 2 trial and had a clinical response during the Induction trials		
Intervention(s)	Tofacitinib 5 mg twice daily or 10 mg twice daily		
Comparator(s)	Placebo		
Indicate if trial supports application for marketing authorisation	**Yes**	Indicate if trial used in the economic model	**Yes**
Rationale for use/non-use in the model	Study provides evidence of the efficacy of tofacitinib and was included in the network meta-analysis used in the economic model.		
Reported outcomes specified in the decision problem	• Measures of disease activity: Mayo score and partial Mayo score • Rates of and duration of response and remission: Mayo score • Achieving mucosal healing (endoscopic findings) • Adverse effects of treatment • HRQoL: IBDQ, SF-36, EQ-5D-3L and EQ-5D VAS • Hospitalisation and surgery due to ulcerative colitis • Corticosteroid-free remission		
Study	**OCTAVE Open (NCT01470612)**		
Study design	Open-label extension study		
Population	Patients who completed 52 weeks of maintenance therapy in OCTAVE Sustain, and patients who did not have a clinical response in OCTAVE Induction 1 or 2, or who withdrew from OCTAVE Sustain due to treatment failure		
Intervention(s)	Tofacitinib 5 mg twice daily or 10 mg twice daily		
Comparator(s)	None		
Indicate if trial supports application for marketing authorisation	**Yes**	Indicate if trial used in the economic model	**No**
Rationale for use/non-use in the model	Study was not included in the economic model because of its open-label, uncontrolled design and it had not completed at time of submission (interim data available).		
Reported outcomes specified in the decision problem	• Measures of disease activity: Mayo score and partial Mayo score • Rates of and duration of response and remission: Mayo score • Achieving mucosal healing (endoscopic findings) • Adverse effects of treatment		
Study	**Phase II trial (NCT00787202)**		
Study design	Multicentre, randomised, double-blind, placebo-controlled trial with an 8-week induction phase.		
Population	Adults with a confirmed diagnosis of ulcerative colitis, total Mayo scores of 6 to 12 and endoscopic Mayo scores of 2 to 3		
Intervention(s)	Tofacitinib 0.5 mg, 3 mg, 10 mg or 15 mg twice daily		
Comparator(s)	Placebo		

Indicate if trial supports application for marketing authorisation	**Yes**	Indicate if trial used in the economic model	**Yes**
Rationale for use/non-use in the model	Study provides evidence of the efficacy of tofacitinib and was included in the network meta-analysis used in the economic model		
Reported outcomes specified in the decision problem	• Measures of disease activity: Mayo score and partial Mayo score • Rates of and duration of response and remission: Mayo score a • Achieving mucosal healing (endoscopic findings) • Adverse effects of treatment		

[a] Data from the Phase II trial are included in the NMA and economic model.
Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; HRQoL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; SF-36, 36-Item Short Form Survey; VAS, visual analogue scale.

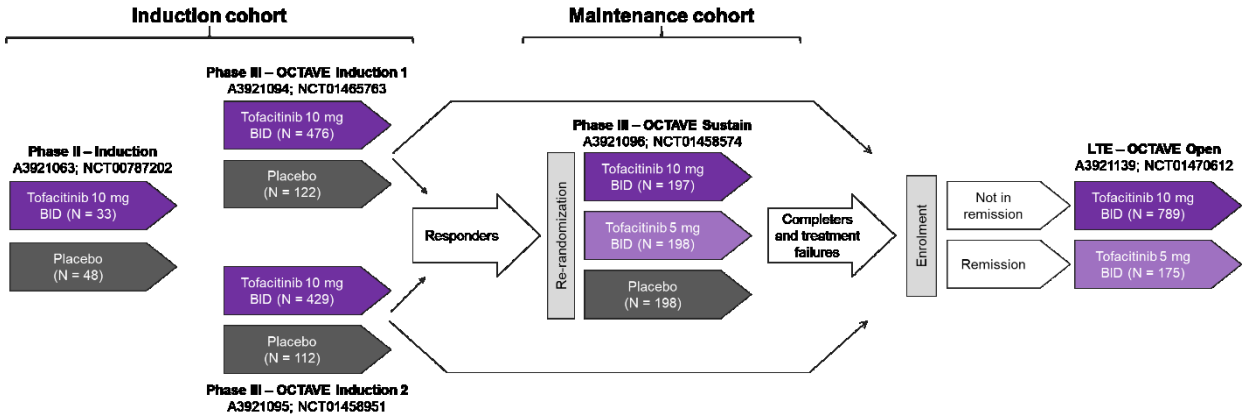
B.2.3 Summary of methodology of the relevant clinical effectiveness evidence

B.2.3.1 Methodology

B.2.3.1.1 Overall summary of OCTAVE clinical trial programme

The clinical trial programme for tofacitinib includes one Phase II and three Phase III placebo-controlled studies, and one long-term open-label extension study (98-103), and is summarised in Figure 4.

Figure 4 OCTAVE clinical trial programme overview



For the phase II study, only the tofacitinib 10 mg arm is shown.
Abbreviations: BID, twice daily; LTE, long-term extension.

The pivotal evidence used to support the decision problem is predominantly based on the Phase III studies. These included:

- OCTAVE Induction 1 and 2: two identical studies of patients with moderately to severely active disease to assess the efficacy of tofacitinib in inducing remission
- OCTAVE Sustain: a study of patients who had achieved clinical response in the two OCTAVE Induction studies to assess the efficacy of tofacitinib in maintaining remission (98).

A brief summary of the study details for the five studies in the clinical trial programme is presented in Table 8 (98, 99, 103).

Table 8 Summary of studies in tofacitinib clinical trial programme

Study	Population	Design	Dose Regimen	Key Selected Endpoints
Phase II study (A3921063) (NCT00787202)	Adult patients with moderately to severely active ulcerative colitis (N = 194)	Double-blind Placebo-controlled 8 weeks	• tofacitinib 0.5 mg BID (n = 31) • tofacitinib 3 mg BID (n = 33) • tofacitinib 10 mg BID (n = 33) • tofacitinib 15 mg BID (n = 49) • placebo (n = 48)	**Primary** Clinical response at week 8 **Secondary** Clinical remission at week 8 IBDQ
OCTAVE Induction 1 (A3921094) (NCT01465763)	Adult patients with moderately to severely active ulcerative colitis (N = 598)	Double-blind Placebo-controlled 9 weeks (primary efficacy endpoint at 8 weeks)	• tofacitinib 10 mg BID (n = 476) • placebo (n = 122)	**Primary** Remission at week 8 **Secondary** Mucosal healing at week 8 Clinical response at week 8 Clinical remission at week 8 IBDQ SF-36 EQ-5D
OCTAVE Induction 2 (A3921095) (NCT01458951)	Adult patients with moderately to severely active ulcerative colitis (N = 541)	Double-blind Placebo-controlled 9 weeks (primary efficacy endpoint at 8 weeks)	• tofacitinib 10 mg BID (n = 429) • placebo (n = 112)	**Primary** Remission at week 8 **Secondary** Mucosal healing at week 8 Clinical response at week 8 Clinical remission at week 8 IBDQ SF-36 EQ-5D
OCTAVE Sustain (A3921096) (NCT01458574)	Patients who had achieved clinical response in OCTAVE Induction 1 or 2 (N = 593)	Double-blind Placebo-controlled 53 weeks (primary efficacy endpoint at 52 weeks)	• tofacitinib 10 mg BID (n = 197) • tofacitinib 5 mg BID (n = 198) • placebo (n = 198)	**Primary** Remission at week 52 **Secondary** Mucosal healing at week 52 Sustained steroid-free remission at week 24 and week 52 Clinical response at week 52 Clinical remission at week 52
OCTAVE Open (A3921139) (NCT01470612)	Patients who completed OCTAVE Induction 1 or 2 without clinical response, OR patients who completed or had early withdrawal due to treatment failure in OCTAVE Sustain (N = 886)	Open-label up to 6 years duration	• tofacitinib 10 mg BID (n = 742) • tofacitinib 5 mg BID (n = 144)	**Primary** Safety and tolerability of long-term tofacitinib therapy in patients with UC **Secondary** Remission at month 12 Mucosal healing at month 12

Abbreviations: BID, twice daily; UC, ulcerative colitis.

B.2.3.1.2 OCTAVE Induction 1 and 2 methodology

B.2.3.1.2.1 Summary of trial methodology

The methodology of the OCTAVE Induction 1 and 2 trials is summarised in Table 9.

Table 9 Summary of OCTAVE Induction 1 and 2 methodology

Trial no. (acronym)	**NCT01465763** **OCTAVE Induction 1 (A3921094)**	**NCT01458951** **OCTAVE Induction 2 (A3921095)**
Study objective	To demonstrate the efficacy of tofacitinib in inducing remission in patients with moderately to severely active ulcerative colitis	
Trial design	Multicentre, randomised, double-blind, placebo-controlled trials	
Duration of study	9 weeks (primary efficacy endpoint was assessed at 8 weeks)	
Method of randomisation	Randomisation was performed centrally with the use of a tele-randomisation system; stratified according to previous treatment with TNFi therapies, glucocorticoid use at baseline, and geographic region	
Method of blinding	Trials were patient-, investigator- and sponsor-blinded	
Eligibility criteria for participants	Adult patients aged 18 years or older with moderately to severely active ulcerative colitis. Details of inclusion and exclusion criteria are provided in section B.2.3.1.2.2	
Settings and locations where the data were collected	OCTAVE Induction 1 was conducted at 144 sites worldwide (two in the UK)	OCTAVE Induction 2 was conducted at 169 sites worldwide (three in the UK)
Trial drugs	4:1 ratio of oral tofacitinib 10 mg twice daily and placebo; trials initially included tofacitinib 15 mg twice daily (see below)	
Permitted and disallowed concomitant medications	Permitted concomitant medications for ulcerative colitis were oral aminosalicylates and oral glucocorticoids (at a maximum dose of 25 mg per day of prednisone or a prednisone equivalent), provided that the medications were administered at a stable dose throughout the induction trials. Chronic treatment with antibiotics was permitted providing that the dose was stable for 2 weeks prior to baseline. Prohibited medications include azathioprine, ciclosporin, TNFi therapy within 8 weeks of baseline, intravenous corticosteroids.	
Primary outcomes (see Table 12for definitions)	Primary endpoint: remission at week 8, based on centrally read endoscopic subscores	
Secondary/tertiary outcomes (including scoring methods and timings of assessments) (see Table 11 and Table 12 for definitions)	Key secondary endpoint: mucosal healing at week 8 Secondary endpoints: Clinical response at week 8 Clinical remission at week 8 Endoscopic remission at week 8 Symptomatic remission at week 8 Deep remission at week 8 Partial Mayo score at week 8 and change from baseline over time Change from baseline in total Mayo score at week 8 Key PRO and resource use endpoints: IBDQ remission at week 4 and week 8	

	IBDQ treatment response at week 4 and week 8 Score and change from baseline in EQ-5D/VAS over time Score and change from baseline in SF-36 PCS and MCS at week 8 Score and change from baseline in WPAI domains at week 8 Incidence and duration of ulcerative colitis-related hospitalisations Number of patients undergoing colectomy for ulcerative colitis or ulcerative colitis-related complications
Pre-planned subgroups	Prior TNFi exposure (yes vs no) Prior TNFi failure (yes vs no) Baseline corticosteroid use (yes vs no) Geographic Region (3 groups: North America, Europe, other).
Protocol amendments	Trials initially included tofacitinib 15 mg twice daily, but exploration of this dose was discontinued based on feedback from regulatory authorities. Randomisation to the 15-mg dose was discontinued after 38 patients in the OCTAVE Induction 1 trial, and 18 in the OCTAVE Induction 2 trial, had undergone randomisation across three treatment groups.

Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; IBDQ, Inflammatory Bowel Disease Questionnaire; MCS, Mental Health Component Summary; PCS, Physical Health Component Summary; PRO, patient-reported outcome; SF-36, 36-Item Short Form Survey; TNFi, tumour necrosis factor inhibitor; VAS, visual analogue scale.
Source: Sandborn *et al.* 2017 (98).

B.2.3.1.2.2 Eligibility criteria

Full inclusion and exclusion criteria for OCTAVE Induction 1 and 2 are listed in Table 10 (98).

Prohibited concomitant therapies included TNFi therapies, azathioprine, methotrexate, and mercaptopurine. Permitted concomitant medications for ulcerative colitis during the Induction trials were oral aminosalicylates at a stable dose for at least 4 weeks prior to baseline and during the study period, and oral glucocorticoids (at a maximum dose of 25 mg per day of prednisone or a prednisone equivalent) at a stable dose for at least 2 weeks prior to baseline and during the study period. Patients currently receiving chronic treatment for ulcerative colitis with antibiotics (e.g., metronidazole and rifaximin) were also eligible provided the dose was stable for at least 2 weeks prior to baseline and during the study period.

Table 10 Inclusion and exclusion criteria in the OCTAVE Induction 1 and 2 studies

Inclusion criteria

- Patient has provided informed consent
- Patient is aged 18 years or older
- Patient has had a confirmed diagnosis of ulcerative colitis for at least 4 months
- Patient has moderately to severely active disease (defined as a Mayo score of 6 to 12, with a rectal bleeding subscore of 1 to 3 and an endoscopic subscore of 2 or 3; see section B.2.3.1.2.2).
- Patients were required to have had treatment failure with or to have had unacceptable side effects from treatment with at least one of the following agents:
    - oral or intravenous glucocorticoids
    - azathioprine
    - mercaptopurine
    - infliximab
    - adalimumab.

Exclusion criteria

- Presence of clinical findings suggestive of Crohn's disease, ulcerative colitis limited to the distal 15 cm of colon, clinical signs of fulminant colitis, toxic megacolon, or indeterminate, microscopic, ischaemic, or infectious colitis.
- Inadequate washout for the following medications prior to baseline:
    - azathioprine, 6- mercaptopurine, or methotrexate within 2 weeks
    - TNFi therapies or interferon therapy within 8 weeks
    - cyclosporine, mycophenolate mofetil/mycophenolic acid, or tacrolimus within 4 weeks
    - intravenous corticosteroids within 2 weeks
    - rectally administered corticosteroid or 5-aminosalicylic acid within 2 weeks
    - anti-adhesion molecule therapy within one year
    - lymphocyte-depleting agents/therapies within one year
    - other marketed immunosuppressants or biologics with immunomodulatory properties within 3 months
    - leukocyte apheresis within 6 months.
- Patients were also excluded if, at screening, they had:
    - haemoglobin levels < 9.0 g/dL
    - absolute white blood cell count of < 3.0 x 10^9/L (< 3000/mm^3) or absolute neutrophil count of < 1.2 x 10^9/L (< 1200/mm^3) or absolute lymphocyte count of < 0.5 x 10^9/L (< 500/mm^3) (< 0.75 x 10^9/L [< 750/mm^3] in the UK)
    - thrombocytopenia, as defined by a platelet count < 100 x 10^9/L (< 100,000/mm^3)
    - estimated glomerular filtration rate < 40 mL/min based on Cockcroft-Gault calculation
    - total bilirubin, aspartate transaminase or alanine transaminase more than 1.5 times the upper limit of normal.

Source: Sandborn *et al.* 2017 (98).

B.2.3.1.2.3 Outcomes

Outcomes were measured for disease activity, health-related quality of life (HRQoL) and health utility.

Different instruments and scoring systems were used to measure these (see Table 11) and to define efficacy endpoints (see Table 12). Adverse events were also recorded as safety endpoints.

Table 11 Outcome measures used in the OCTAVE Induction trials

Outcome	Definition
Efficacy	
Mayo score	Scores on the Mayo scale range from 0 to 12, and scores on each of the four subscores range from 0 to 3, with higher scores indicating more severe disease (see Table 3).
Partial Mayo score	The partial Mayo score is calculated based on the following Mayo subscores: Physician's Global Assessment, stool frequency and rectal bleeding, and ranges from 0 to 9, with higher scores indicating more severe disease.
Patient-reported outcomes	
IBDQ	IBDQ scores range from 32 to 224, with higher scores indicating better HRQoL.
SF-36 v2, acute	The SF-36 version 2, acute assesses eight domains of functional health and well-being. Physical Health Component Summary (PCS) and Mental Health Component Summary (MCS) scores were calculated from the domain scores. The acute form uses a recall period of 1 week. Higher scores indicate a better HRQoL.
EQ-5D utility index score	The EQ-5D is a standardised instrument developed by the EuroQoL Group for use as a generic, preference-based measure of health outcome. The EQ-5D questionnaire is used to calculate a utility score based on a descriptive profile, or 'health state'. Data in this submission are based on the 3-level version (EQ-5D-3L), with UK preference weights.
EQ-5D VAS	In the EQ-5D VAS, patients indicate their overall health on a vertical scale, ranging from "worst possible" to "best possible" health.
WPAI-UC v2	The 6-item WPAI-UC version 2 questionnaire is a validated instrument designed to measure the ability to work and perform regular activities, specifically as a result of the target health problem (ulcerative colitis). The WPAI-UC yields four scores: absenteeism (work time missed); presenteeism (impairment at work/reduced on-the-job effectiveness); work productivity loss (overall work impairment/absenteeism plus presenteeism) and non-work activity Impairment.

Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; HRQoL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; SF-36, 36-Item Short Form Survey; VAS, visual analogue scale; WPAI-UC, Work Productivity and Activity Impairment-Ulcerative Colitis.
Sources: Sandborn *et al.* 2017 (98); OCTAVE CSRs (100-102).

Adverse events (which were classified with the use of the Medical Dictionary for Regulatory Activities) were recorded as safety endpoint. Laboratory test results and concomitant medications were recorded throughout the trials (98). Opportunistic infections, cancers, and cardiovascular events were assessed by external adjudication committees (Appendix L.1.1). The incidence and duration of ulcerative colitis-related hospitalisation and surgery was also recorded. Details of adverse events in the OCTAVE trials are presented in section B.2.9.

Endpoint definitions used in the OCTAVE Induction 1 and 2 trials are shown in Table 12.

Table 12 Summary of OCTAVE Induction 1 and 2 endpoints

Definition		Endpoint	Type	Assessed at:
Efficacy endpoints using the Mayo score				
Mayo score of ≤ 2, no individual subscore exceeding 1 point +	rectal bleeding subscore = 0	Remission	**Primary**	Week 8
		Clinical remission	Secondary	Week 8
	rectal bleeding and stool frequency subscore = 0	Symptomatic Remission	Secondary	Week 8
	rectal bleeding subscore and endoscopic subscore = 0	Deep Remission	Secondary	Week 8
Mayo endoscopic subscore of ≤ 1		Mucosal Healing	**Key Secondary**	Week 8
Mayo endoscopic subscore of 0		Endoscopic Remission	Secondary	Week 8
Decrease from baseline Mayo score of ≥ 3 points and ≥ 30%, with a decrease in rectal bleeding subscore of ≥ 1 point or absolute rectal bleeding subscore of ≤ 1)		Clinical Response	Secondary	Week 8
Total of Physician's Global Assessment, stool frequency and rectal bleeding subscores		Partial Mayo	Secondary	Weeks 2, 4, and 8
HRQoL endpoints				
IBDQ score of ≥ 170		IBDQ remission	Secondary	Weeks 4 and 8
Increase in IBDQ score of ≥ 16 points from induction trial baseline		IBDQ treatment response	Secondary	Weeks 4 and 8
Change from baseline in EQ-5D utility index score		EQ-5D utility	Secondary	Weeks 2 and 8
Change from baseline in EQ-5D VAS score		EQ-5D VAS	Secondary	Weeks 2 and 8
Change from baseline in SF-36 MCS score		SF-36 MCS	Secondary	Week 8
Change from baseline in SF-36 PCS score		SF-36 PCS	Secondary	Week 8

Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; HRQoL, health-related quality of life; IBDQ, Inflammatory Bowel Disease Questionnaire; SF-36, 36-Item Short Form Survey; VAS, visual analogue scale.
Source: Sandborn *et al.* 2017 (98).

B.2.3.1.2.4 Central and local assessment of endoscopic subscores

In the OCTAVE Induction 1 and 2 trials, the Mayo endoscopic subscore, based on mucosal appearance during endoscopy, was assessed by both the study site investigator (local assessment) and by a central reader using video recorded during the procedure (central assessment). Centrally read endoscopic subscores were used for both eligibility and efficacy analyses. The tofacitinib OCTAVE development programme was the first in ulcerative colitis to use central reads to assess primary efficacy endpoints, as requested by the regulatory authority. The benefit of using central reads has not yet been established. In addition, it is not known whether the knowledge that endoscopic appearances were also being scored by central readers may have consciously or subconsciously influenced the results of the local reading. In clinical practice, physicians use their own assessment of endoscopic findings, and not that of central readers, to complement other data to make clinical decisions. Therefore, results based on local reading may be closer to real-world data than central reading and remain relevant for prescribers.

In this submission, results based on locally read endoscopic subscores are presented in addition to those based on central reads. Locally read endoscopic subscores are also used

in the base case of the NMA and economic model (see sections B.2.9 and B.3); centrally read data are used in sensitivity analysis.

B.2.3.1.3 OCTAVE Sustain – methodology

B.2.3.1.3.1 Summary of trial methodology

The methodology of the OCTAVE Sustain trial is summarised in Table 13.

Table 13 Summary of OCTAVE Sustain methodology

Trial no. (acronym)	**NCT01458574** **OCTAVE Sustain** **(A3921096)**
Study objective	To demonstrate the efficacy of tofacitinib as maintenance therapy in patients with moderately to severely active ulcerative colitis
Trial design	Multicentre, randomised, double-blind, placebo-controlled trials
Duration of study	52 weeks
Method of randomisation	Randomisation was performed centrally with the use of a tele-randomisation system; stratified according to previous treatment with TNFi therapies, glucocorticoid use at baseline, and geographic region
Method of blinding	Trial was patient-, investigator- and sponsor-blinded
Eligibility criteria for participants	Patients were eligible to enter OCTAVE Sustain if they met the entry criteria for the Induction trials (see Section B 2.3.1.2.2. and had completed 8 weeks of induction therapy. They also had to have achieved the criteria for clinical response in OCTAVE Induction 1 and 2.
Settings and locations where the data were collected	OCTAVE Sustain was conducted at 297 sites worldwide (five in the UK)
Trial drugs	1:1:1 ratio of oral tofacitinib 5 mg twice daily, tofacitinib 10 mg twice daily and placebo
Permitted and disallowed concomitant medications	Permitted concomitant medications for ulcerative colitis were oral aminosalicylates (at a stable dose) and chronic treatment for ulcerative colitis with antibiotics (e.g., metronidazole, rifaximin). If patients were using oral glucocorticoids at study entry, tapering was mandatory starting the first week of the study at a specified rate depending on starting dose: the daily dose of prednisone or equivalent was decreased at a rate of 5 mg per week until the dose reached 20 mg/day, then 2.5 to 5.0 mg per week until the dose reached 10 mg/day, then by 2.5 mg per week until the dose was 0 mg.
Primary outcomes (see Table 12 for definitions)	Primary endpoint: • remission at week 52, based on centrally read endoscopic subscores
Secondary/tertiary outcomes (including scoring methods and timings of assessments) (see Table 11, Table 12 and section B.2.3.1.3.3 for definitions)	Key secondary endpoints: • mucosal healing at week 52 • sustained corticosteroid-free remission among patients in remission at baseline Secondary endpoints: • Clinical response at week 52 • Clinical remission at week 52 • Endoscopic remission at week 52 • Symptomatic remission at week 52

	• Deep remission at week 52 • Key health outcome endpoints: • IBDQ remission over time • IBDQ treatment response over time • Score and change from baseline in EQ-5D/VAS over time • Score and change from baseline in SF-36 PCS and MCS at weeks 24 and 52 • Score and change from baseline in WPAI domains at week 52 • incidence and duration of ulcerative colitis-related hospitalisations • Number of patients undergoing colectomy for ulcerative colitis or ulcerative colitis-related complications
Pre-planned subgroups[a]	• Duration of disease at induction study baseline (< 6 years vs ≥ 6 years) • Prior TNFi exposure at induction study baseline (yes vs no) • Prior TNFi failure at induction study baseline (yes vs no) • Prior corticosteroid failure at induction study baseline (yes vs no) • Induction study treatment assignment (tofacitinib 10 mg vs tofacitinib 10 mg or 15 mg vs placebo) • Remission at maintenance study baseline (yes vs no) • Mucosal healing at maintenance study baseline (yes vs no) • Corticosteroid use at maintenance study baseline (yes vs no)

Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; IBDQ, Inflammatory Bowel Disease Questionnaire; MCS, Mental Health Component Summary; PCS, Physical Health Component Summary SF-36, 36-Item Short Form Survey; TNFi, tumour necrosis factor inhibitor; VAS, visual analogue scale.
Source: Sandborn *et al.* 2017 (98).
[a] further subgroups were planned and are listed in the CSR in detail

Patients in OCTAVE Sustain who met treatment failure criteria were required to withdraw from the study; treatment failure was defined as an increase in Mayo score ≥ 3 points from baseline, increase in rectal bleeding ≥ 1 point, increase of endoscopic subscore ≥ 1 point, yielding an absolute endoscopic subscore of ≥ 2 after a minimum treatment of 8 weeks in the study.

B.2.3.1.3.2 Eligibility criteria

Patients were eligible to enter OCTAVE Sustain if they met the entry criteria for the two Induction studies and had completed 8 weeks of induction therapy. In addition, they must have had a clinical response to therapy in the OCTAVE Induction 1 and 2 trials. Oral glucocorticoids were permitted in the OCTAVE Induction 1 and 2 trials (Table 10), provided that the dose was stable throughout the trials, however, tapering of glucocorticoids was mandatory in OCTAVE Sustain starting from entry into the maintenance trial. A fixed schedule of tapering was applied according to glucocorticoid and starting dose, but the protocol required eventual withdrawal from corticosteroids. Patients could receive an increase in their steroid dosing once during the maintenance study to treat flare if necessary, however, tapering must then be restarted and no further dose increases were permitted. (98).

B.2.3.1.3.3 Outcomes

Outcome definitions in OCTAVE Sustain were identical those in the OCTAVE Induction 1 and 2 trials, as described in section B.2.3.1.2.3, Table 11 (98). OCTAVE Sustain endpoints based on Mayo scores and IBDQ scores were defined in the same way as those in the OCTAVE Induction 1 and 2 trials (see Table 12). In addition, the proportion of patients who achieved Mayo score endpoints at both week 24 and week 52 was calculated (e.g. sustained remission) (98). Data for the subset of patients in remission at OCTAVE Sustain baseline (i.e. at week 8 in OCTAVE Induction 1 and 2) were used to assess the key secondary endpoint of sustained corticosteroid-free remission. Sustained corticosteroid-free remission comprised of remission (as defined in Table 12) plus no treatment with steroids for ≥ 4 weeks before the 24-and 52-week visits (98).

B.2.3.1.3.4 Central and local assessment of endoscopic subscores

As in the OCTAVE Induction 1 and 2 trials, the Mayo endoscopic subscore, based on mucosal appearance during endoscopy, was assessed by both the study site investigator (local assessment) and by a central reader using video recorded during the procedure (central assessment). This submission includes results based on both centrally- and locally read endoscopic subscores. Locally read endoscopic subscores are also used in the base case of the NMA and economic model (see sections B.2.9 and B.3); centrally read data are used in sensitivity analysis.

B.2.3.1.4 OCTAVE Open – methodology

B.2.3.1.4.1 Summary of trial methodology

Patients without a clinical response in OCTAVE Induction 1 and 2 and those who completed OCTAVE Sustain or had early withdrawal due to treatment failure were eligible to enter an open-label extension phase, OCTAVE Open (see section B.2.6.3).

The methodology of the ongoing OCTAVE Open trial is summarised in Table 14.

Table 14 Summary of OCTAVE Open methodology

Trial no. (acronym)	**NCT01470612 OCTAVE Open (A3921139)**
Study objective	To assess the safety and tolerability of long-term tofacitinib therapy
Trial design	Open-label extension study
Duration of study	Up to 6 years (12-month results are reported in this submission)
Method of randomisation	None
Method of blinding	None
Trial no. (acronym)	**NCT01470612 OCTAVE Open (A3921139)**

Eligibility criteria for participants	Patients who have completed or demonstrated treatment failure in the maintenance study, or who were non-responders after completing 8 weeks of treatment in the induction studies were eligible to enter OCTAVE Open.
Settings and locations where the data were collected	OCTAVE Open was conducted at 215 sites worldwide (five in the UK)
Trial drugs	Oral tofacitinib 5 mg twice daily or 10 mg twice daily, depending on response in OCTAVE Induction 1 and 2 and OCTAVE Sustain (Figure 5)
Permitted and disallowed concomitant medications	Permitted concomitant medications for ulcerative colitis were oral aminosalicylates (at a stable dose) and chronic treatment for ulcerative colitis with antibiotics (e.g., metronidazole, rifaximin); if patients were using oral glucocorticoids at study entry, tapering was mandatory as per the OCTAVE Sustain schedule.
Primary outcomes	Primary objective: • To assess the safety and tolerability of long-term tofacitinib therapy in patients with ulcerative colitis. There were no primary efficacy endpoints
Secondary/tertiary outcomes (including scoring methods and timings of assessments) (see Table 12 for definitions)	Secondary objectives: • To evaluate the efficacy of long-term tofacitinib therapy in patients with ulcerative colitis To evaluate the effect of long-term tofacitinib therapy on HRQoL in patients with ulcerative colitis. Secondary efficacy endpoints based on Mayo scores • Remission • Clinical remission • Mucosal healing • Partial Mayo score remission (total score ≤ 2, no individual subscore > 1)
Pre-planned subgroups	Subgroups according to response in OCTAVE Induction 1 and 2 and OCTAVE Sustain (see Figure 5): 1. Patients who completed maintenance study in remission: 'maintenance remission' 2. Patients who completed maintenance not in remission: 'other remission completers' 3. Patients who withdrew from maintenance study due to treatment failure: 'maintenance treatment failure' Patients who did not have a clinical response to induction therapy: 'induction non-response'

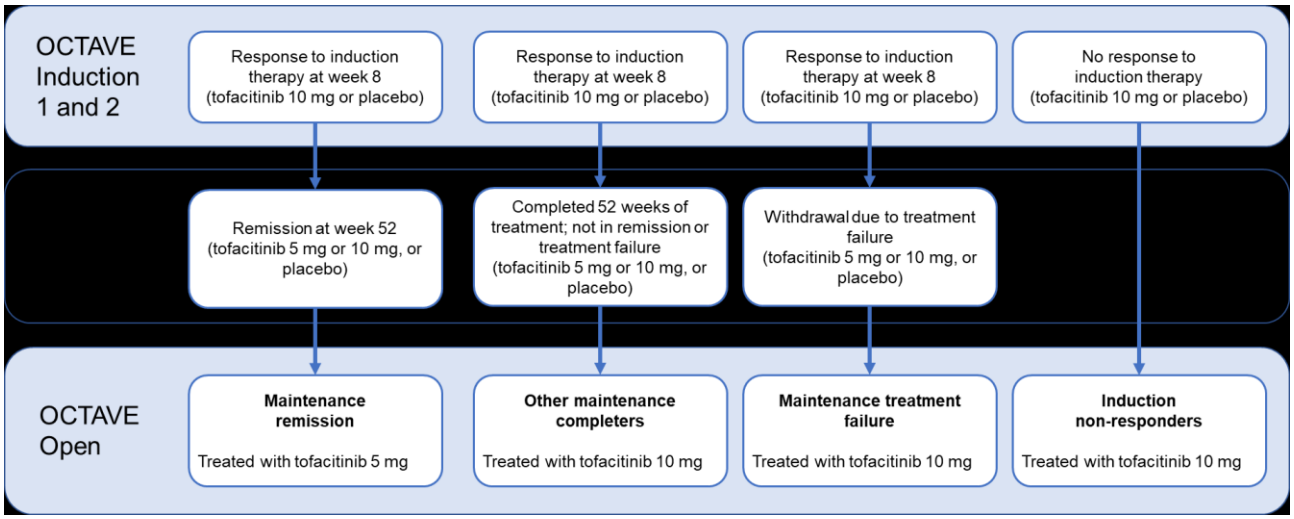
Abbreviations: HRQoL, health-related quality of life.
Source: OCTAVE Open CSR (103).

Treatment assignment in OCTAVE Open was determined according to patients' response in OCTAVE Induction 1 and 2 and OCTAVE Sustain and is summarised in Figure 5.

B.2.3.1.4.2 Outcomes

Outcome definitions in OCTAVE Open were consistent those in the OCTAVE Induction and Sustain trials (see section B.2.3.1.2.3).

Figure 5 Treatment assignment in OCTAVE Open



B.2.3.1.4.3 Central and local assessment of endoscopic subscores

Endoscopic subscores based on central reading were used to determine remission status at entry to OCTAVE Open.

B.2.3.1.5 Relevance of endpoints to decision problem

Remission is a stringent endpoint that requires both symptomatic improvement and endoscopic evidence of mucosal healing. In addition, the secondary endpoint of mucosal healing is regarded as an important therapeutic endpoint in clinical practice; achieving mucosal healing is associated with sustained clinical remission, a reduced need for corticosteroids and a decreased risk of surgery being required (63, 104).

Sustained corticosteroid-free remission is regarded as an important clinical endpoint: although corticosteroids may be used for induction of remission, because of their side-effect profile they are not typically used for long-term management of ulcerative colitis, making corticosteroid-free remission an important goal (105).

The endpoints used in the NMA and economic model (see sections B.2.9 and B.3) are clinical response and clinical remission (see Table 12 for definitions). These endpoints are used instead of the primary OCTAVE trial endpoint, remission, to ensure comparability with trials of biological therapies for ulcerative colitis, which have typically used clinical remission or clinical response as a primary endpoint. In OCTAVE Induction 1 and 2 and OCTAVE Sustain the results for clinical remission are very similar to those for the primary endpoint of remission (see sections B.2.6.1.1.3 and B.2.6.2.1.4). Both clinical response and clinical remission are considered to be clinically meaningful endpoints (106).

B.2.3.2 Baseline characteristics

Demographics and baseline characteristics of patients included in the OCTAVE studies are shown in Table 15. The baseline characteristics of the patients were similar across treatment groups in all the trials, except for sex in the OCTAVE Induction 2 trial and smoking status in the OCTAVE Sustain trial (98).

The OCTAVE trial populations are relevant to the NICE decision problem. Approximately half of OCTAVE Induction trial participants had extensive disease or pancolitis (49–54% across groups), with mean total Mayo scores of 8.9–9.1 (98). More than half of participants had previously received TNFi therapy (53–58% across groups); of these TNFi-experienced patients, the majority (over 95%) had experienced failure of at least one TNFi therapy. In addition, of this TNFi-experienced group, 33% in the tofacitinib group had received more than 2 TNFi agents. More than two-thirds of patients in the Induction trials had had treatment failure with an immunosuppressant (such as azathioprine or mercaptopurine; 67–76%), and around three-quarters had treatment failure with a glucocorticoid (71–80%) (98).

B.2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

B.2.4.1 Analysis sets

The main analysis sets in the OCTAVE RCTs are defined below.

Full Analysis Set (FAS): The primary analysis population for efficacy endpoints was the FAS defined as all subjects randomly assigned to either placebo, tofacitinib 10 mg twice daily, or tofacitinib 5 mg twice daily (OCTAVE Sustain only).

OCTAVE Induction modified Full Analysis Set (mFAS): In OCTAVE Induction 1 and 2 the mFAS was a subset of the FAS with 3 patients excluded from a site in Japan due to potential unblinding during the study.

OCTAVE Sustain mFAS: in OCTAVE Sustain the mFAS was a subset of the FAS that included only patients who received tofacitinib in the induction studies (excluding patients from the OCTAVE Induction 1 and 2 placebo groups).

Per-Protocol Analysis Set (PPAS): A subset of the FAS population who had no major protocol violations that could have potentially had a significant impact on efficacy analyses, as determined by the sponsor prior to database lock.

Safety Analysis Set (SAS): The safety analysis set consisted of all randomised subjects who received at least 1 dose of study medication.

For the purpose of this submission, mFAS and PPAS results are not described in full detail within this document; however, results for the primary endpoints are summarised in Appendix L, Table 206, Table 207 and Table 208.

In addition, both the efficacy and the safety analyses in the induction trials excluded data from patients who were assigned to receive tofacitinib at a dose of 15 mg; those data were analysed separately and are summarised in Appendix L, Table 209.

B.2.4.2 Statistical information

A summary of the statistical methods used in the OCTAVE RCTs is presented in Table 16.

Table 15 Baseline Demographic and Disease Characteristics of the Patients in the OCTAVE Induction 1, OCTAVE Induction 2, and OCTAVE Sustain Trials

Characteristic	OCTAVE Induction 1		OCTAVE Induction 2		OCTAVE Sustain		
	Placebo (N = 122)	Tofacitinib 10 mg (N = 476)	Placebo (N = 112)	Tofacitinib 10 mg (N = 429)	Placebo (N = 198)	Tofacitinib 5 mg (N = 198)	Tofacitinib 10 mg (N = 197)
Male sex, n (%) [a]	77 (63.1)	277 (58.2)	55 (49.1)	259 (60.4)	116 (58.6)	103 (52.0)	110 (55.8)
Age, years [b]	41.8±15.3	41.3±14.1	40.4±13.2	41.1±13.5	43.4±14.0	41.9±13.7	42.9±14.4
Induction trial group assignment, n (%)							
Placebo	—	—	—	—	24 (12.1)	22 (11.1)	24 (12.2)
Tofacitinib, 10 mg twice daily	—	—	—	—	167 (84.3)	170 (85.9)	167 (84.8)
Tofacitinib, 15 mg twice daily	—	—	—	—	7 (3.5)	6 (3.0)	6 (3.0)
Remission at maintenance trial entry, n (%)	—	—	—	—	59 (29.8)	65 (32.8)	55 (27.9)
Duration of disease — years [b]							
Median	6.0	6.5	6.2	6.0	7.2	6.5	6.8
Range	0.5–36.2	0.3–42.5	0.4–27.9	0.4–39.4	0.6–42.7	0.6–40.3	0.6–35.7
Extent of disease, n/total n (%) [c,d]							
Proctosigmoiditis	19/122 (15.6)	65/475 (13.7)	16/111 (14.4)	67/428 (15.7)	21/198 (10.6)	28/196 (14.3)	33/196 (16.8)
Left-sided colitis	37/122 (30.3)	158/475 (33.3)	39/111 (35.1)	149/428 (34.8)	68/198 (34.3)	66/196 (33.7)	60/196 (30.6)
Extensive colitis or pancolitis	66/122 (54.1)	252/475 (53.1)	56/111 (50.5)	211/428 (49.3)	108/198 (54.5)	102/196 (52.0)	103/196 (52.6)
Total Mayo score [b,e]	9.1±1.4	9.0±1.4	8.9±1.5	9.0±1.5	3.3±1.8	3.3±1.8	3.4±1.8
Partial Mayo score [b,e]	6.5±1.2	6.3±1.2	6.4±1.2	6.4±1.3	1.8±1.4	1.8±1.3	1.8±1.3
C-reactive protein, mg/litre [b]							
Median	4.7	4.4	5.0	4.6	1.0	0.7	0.9
Range	0.1–82.5	0.1–208.4	0.2–205.1	0.2–156.0	0.1–45.0	0.1–33.7	0.1–74.3
Oral glucocorticoid use at baseline — no. (%) [b]	58 (47.5)	214 (45.0)	55 (49.1)	198 (46.2)	100 (50.5)	101 (51.0)	87 (44.2)
Previous treatment with TNFi, n (%) [c]	65 (53.3)	254 (53.4)	65 (58.0)	234 (54.5)	92 (46.5)	90 (45.5)	101 (51.3)

Characteristic	OCTAVE Induction 1		OCTAVE Induction 2		OCTAVE Sustain		
	Placebo (N = 122)	Tofacitinib 10 mg (N = 476)	Placebo (N = 112)	Tofacitinib 10 mg (N = 429)	Placebo (N = 198)	Tofacitinib 5 mg (N = 198)	Tofacitinib 10 mg (N = 197)
Previous treatment failure, n (%) [c,f]							
TNF antagonist	64 (52.5)	243 (51.1)	60 (53.6)	222 (51.7)	89 (44.9)	83 (41.9)	93 (47.2)
Glucocorticoid	98 (80.3)	350 (73.5)	83 (74.1)	303 (70.6)	151 (76.3)	145 (73.2)	149 (75.6)
Immunosuppressant [g]	83 (68.0)	360 (75.6)	75 (67.0)	301 (70.2)	129 (65.2)	143 (72.2)	141 (71.6)
White race, n (%) [h]	98 (83.1)	395 (84.6)	88 (83.0)	331 (80.3)	155 (80.3)	164 (84.5)	153 (81.8)
Weight, kg	72.7 (16.7)	72.9 (16.8)	73.2 (16.2)	74.4 (16.8)	76.2 (16.7)	73.4 (17.8)	74.6 (15.1)
Smoking status, n (%) [c,i]							
Never smoked	80 (65.6)	301 (63.2)	81 (72.3)	268 (62.5)	113 (57.1)	142 (71.7)	128 (65.0)
Current smoker	4 (3.3)	22 (4.6)	5 (4.5)	25 (5.8)	12 (6.1)	7 (3.5)	6 (3.0)
Former smoker	38 (31.1)	153 (32.1)	26 (23.2)	136 (31.7)	73 (36.9)	49 (24.7)	63 (32.0)

Plus–minus values are means ±SD. There were no significant differences between groups within each trial unless otherwise noted.
[a] In the OCTAVE Induction 2 trial, there was a significant difference between groups in the proportion of male patients (p = 0.03).
[b] For the OCTAVE Sustain trial, the baseline values were obtained at the time of entry in the OCTAVE Sustain trial.
[c] For the OCTAVE Sustain trial, the baseline values were obtained at the time of entry into one of the induction trials (OCTAVE Induction 1 or 2).
[d] Data on extent of disease are missing for three patients.
[e] The total Mayo score ranges from 0 to 12 and the partial Mayo score (i.e., the total Mayo score excluding the endoscopic subscore) ranges from 0 to 9, with higher scores indicating more severe disease.
[f] Previous treatment failure was determined by the investigator.
[g] Immunosuppressants included agents such as azathioprine and mercaptopurine and did not include biologic agents (e.g., TNF antagonists) or glucocorticoids.
[h] Unspecified race was treated as missing data.
[i] In OCTAVE Sustain, there was a significant difference for smoking status among placebo and tofacitinib groups (p = 0.03).
Abbreviations: SD, standard deviation; TNFi, tumour necrosis factor inhibitor.
Source: Sandborn *et al.* 2017 (98).

Table 16 Summary of statistical methods in the OCTAVE trials

Trial no. (acronym)	**OCTAVE Induction 1 (A3921094) (NCT01465763)**	**OCTAVE Induction 2 (A3921095) (NCT01458951)**	**OCTAVE Sustain (A3921096) (NCT01458574)**	**OCTAVE Open (A3921139) (NCT01470612)**
Hypothesis objective	To demonstrate a difference between tofacitinib 10 mg and placebo in the proportion of patients in remission at week 8		To demonstrate a difference between tofacitinib and placebo in the proportion of patients in remission at week 52	To assess the safety and tolerability of long-term tofacitinib therapy
Multiple comparisons and multiplicity	The family-wise type 1 error rate was controlled at 0.05 for the primary and key secondary endpoints using a fixed-sequence testing procedure All other efficacy endpoints were evaluated at the 0.05 level of significance, without adjustments for multiple comparisons		The family-wise type 1 error rate was controlled at 0.05 for the primary and key secondary endpoints using a sequentially rejective, Bonferroni-based, iterative multiple test procedure All other efficacy endpoints were evaluated at the 0.05 level of significance, without adjustments for multiple comparisons	Only summary statistics were generated
Statistical analysis of primary endpoint	The primary endpoint was compared between treatment groups by the CMH Chi-square test stratified by prior treatment with TNFi therapy, corticosteroid use at baseline, and geographic region		The primary endpoint was compared between treatment groups by the CMH Chi-square test stratified by prior treatment with TNFi therapy, corticosteroid use at baseline, and geographic region	Only summary statistics were generated
Statistical analysis of secondary efficacy endpoints	Binary endpoints were compared between treatment groups with the use of a stratified CMH Chi-square test Change from baseline in the total Mayo score was analysed with the use of an analysis of covariance model with observed case data For continuous end points, change from baseline was analysed with the use of a linear mixed-effects model (EQ-5D data) or an ANCOVA model (SF-36, WPAI data)		Binary endpoints were compared between treatment groups with the use of a stratified CMH Chi-square test For other continuous end points, change from baseline was analysed with the use of a linear mixed-effects model	Only summary statistics were generated

Trial no. (acronym)	**OCTAVE Induction 1 (A3921094) (NCT01465763)**	**OCTAVE Induction 2 (A3921095) (NCT01458951)**	**OCTAVE Sustain (A3921096) (NCT01458574)**	**OCTAVE Open (A3921139) (NCT01470612)**
Data management, patient withdrawals and the advancement of patients from placebo to active treatment	For binary endpoints, patients with missing data were considered as not having had a response (NRI). [a] For EQ-5D continuous endpoints, analyses were performed using a linear mixed-effects model with repeated measures, where the missing values were assumed to be missing at random Missing SF-36 and WPAI values were not imputed, and analyses were based on observed data		For binary endpoints, patients with missing data were considered as not having had a response (NRI). [a] For EQ-5D and SF-36 continuous endpoints, analyses were performed using a linear mixed-effects model with repeated measures, where the missing values were assumed to be missing at random Missing WPAI values were not imputed, and analyses were based on observed data	For binary endpoints, patients with missing data were considered as not having had a response (NRI). [a]
Sample size, power calculation	A sample size of approximately 545 patients in each trial (436 patients assigned to receive tofacitinib and 109 patients assigned to receive placebo) was calculated to provide the trials with 90% power to detect a difference of 17.5 percentage points between the tofacitinib groups and the placebo groups in the rates of the primary and key secondary endpoints, assuming rates in the placebo groups of 15% for the primary endpoint and 35% for the mucosal healing secondary endpoint.		A sample size of 654 patients (218 in each of the three treatment groups) was calculated to provide the trial with 90% power to detect a difference of 17.5 percentage points between the tofacitinib groups and the placebo group in the rate of the primary endpoint, assuming a rate in the placebo group of 30%.	A sample size of approximately 900 patients combined from OCTAVE Induction 1 and 2 and OCTAVE Sustain was expected

[a] NRI is considered to be a more conservative approach for managing missing data than LOCF (107).
Abbreviations: ANCOVA, analysis of covariance; CMH, Cochran-Mantel-Haenszel; EQ-5D, 5-dimension EuroQol questionnaire; LOCF, last observation carried forward; NRI, non-responder imputation; SF-36, 36-Item Short Form Survey; TNFi, tumour necrosis factor inhibitor; WPAI, Work Productivity and Activity Impairment.
Source: Sandborn *et al.*, 2017 (98); OCTAVE Induction 1 and 2 CSRs (100, 101); OCTAVE Sustain CSR (102); OCTAVE Open CSR (103).

B.2.4.3 Participant flow

Details of patient disposition in OCTAVE Induction 1 and 2 and OCTAVE Sustain are summarised in Table 17 and Table 18 and shown in full in Appendix D, Figure 46 and Figure 47.

Table 17 Summary of patient disposition in OCTAVE Induction 1 and 2

	OCTAVE Induction 1		OCTAVE Induction 2	
	Tofacitinib 10 mg	Placebo	Tofacitinib 10 mg	Placebo
Total patients randomised	476	122	429	112
Patients completing treatment phase	445 (93.5%)	118 (96.7%)	397 (92.5%)	97 (86.6%)
Total discontinuations	31 (6.5%)	4 (3.3%)	32 (7.5%)	15 (13.4%)
Insufficient clinical response	11	1	17	11
Adverse events	9	1	7	2
Protocol violation	4	1	5	0
Withdrawal of consent	4	1	2	2
Death	1	0	0	0
Other	2	0	1	0

Source: Sandborn *et al*., 2017 (98).

In OCTAVE Sustain, discontinuation due to insufficient clinical response was more common in the placebo group (132 of 198; 66.7%) than in the tofacitinib 5 mg (70 of 198; 35.4%) or 10 mg (53 of 197; 27.0%) groups.

Table 18 Summary of patient disposition in OCTAVE Sustain

	Tofacitinib 5 mg	Tofacitinib 10 mg	**Placebo**
Total patients randomised	198	196	198
Patients completing treatment phase	111 (56.1%)	126 (64.3%)	53 (26.8%)
Total discontinuations	87 (43.9%)	70 (35.7%)	145 (73.2%)
Insufficient clinical response	70	53	132
Adverse events	5	9	7
Withdrawal of consent	6	3	5
Protocol violation	0	1	0
Other	1	4	1

Source: Sandborn *et al*., 2017 (98).

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

A summary of the quality assessment for the OCTAVE trials is shown in Table 19, with a detailed description of the quality assessment presented in Appendix D, Table 86.

Table 19 Quality assessment results for OCTAVE trials

Study Question	OCTAVE Induction 1	OCTAVE Induction 2	OCTAVE Sustain
Was randomisation carried out appropriately?	Yes (see Table 9)		Yes (see Table 13)
Was the concealment of treatment allocation adequate?	Yes (see Table 9)		Yes (see Table 13)
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes (see Table 15)		Yes (see Table 15)
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes (see Table 9)		Yes (see Table 13)
Were there any unexpected imbalances in drop-outs between groups?	No (see Table 17)		No (see Table 17)
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No		No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes (see Table 16)		Yes (see Table 16)

B.2.6 Clinical effectiveness results of the relevant trials

B.2.6.1 OCTAVE Induction 1 and 2

The data presented in this submission correspond to the FAS results for the tofacitinib 10 mg twice daily group in the OCTAVE Induction 1 and 2 trials. Pooled results from the two Induction trials are also provided, as are endpoint results based on both central and local endoscopic reads. Data for the small tofacitinib 15 mg twice daily groups are summarised in Appendix L, Table 209. Primary endpoint results for the mFAS and PPAS populations (see section B.2.4.1) are shown in Appendix L, Table 206 and Table 208. Results according to prior treatment with TNFi therapies are summarised in this section, and described in detail in section B.2.7.3.

B.2.6.1.1 Clinical outcomes in OCTAVE Induction 1 and 2

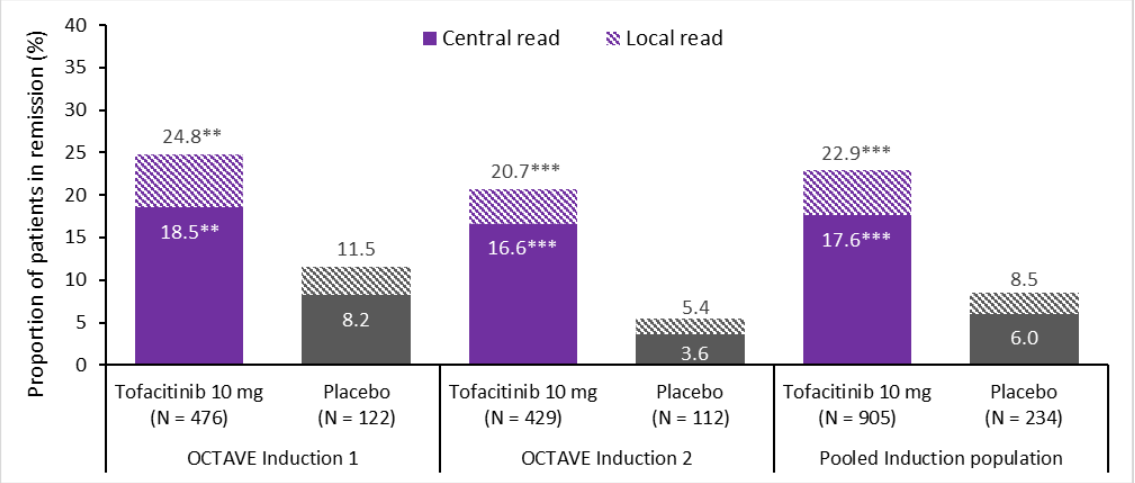
B.2.6.1.1.1 Primary endpoint: remission at week 8

In OCTAVE Induction 1, 18.5% of patients receiving tofacitinib 10 mg achieved remission at week 8, compared with 8.2% in the placebo group (difference, 10.3%; $p = 0.007$) (Figure 6 and Appendix L, Table 213). In OCTAVE Induction 2, the corresponding rates of remission were 16.6% and 3.6% (difference, 13.0%; $p < 0.001$) (98).

In an analysis based on locally read endoscopic subscores (Figure 6 and Appendix L, Table 213), the proportion of patients in remission at week 8 in OCTAVE Induction 1 was 24.8% in the tofacitinib group, compared with 11.5% in the placebo group (difference, 13.3%; $p = 0.0017$) (100). In OCTAVE Induction 2, the corresponding rates of remission were 20.7% and 5.4% (difference, 15.4%; $p = 0.0002$) (101).

Similar results were seen in an analysis of the pooled OCTAVE Induction population (Figure 6 and Appendix L, Table 213). The proportion of patients in remission at week 8 was 17.6% in the tofacitinib group, compared with 6.0% in the placebo group (difference, 11.6%; $p < 0.0001$); in the analysis using locally read endoscopic subscores, the corresponding rates of remission were 22.9% and 8.5% (difference, 14.3%; $p < 0.0001$).

Figure 6 Proportion of patients in remission at week 8 in OCTAVE Induction 1 and 2 (FAS, NRI, central and local reads)



* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98); OCTAVE Induction 1 and 2 CSRs (100, 101).

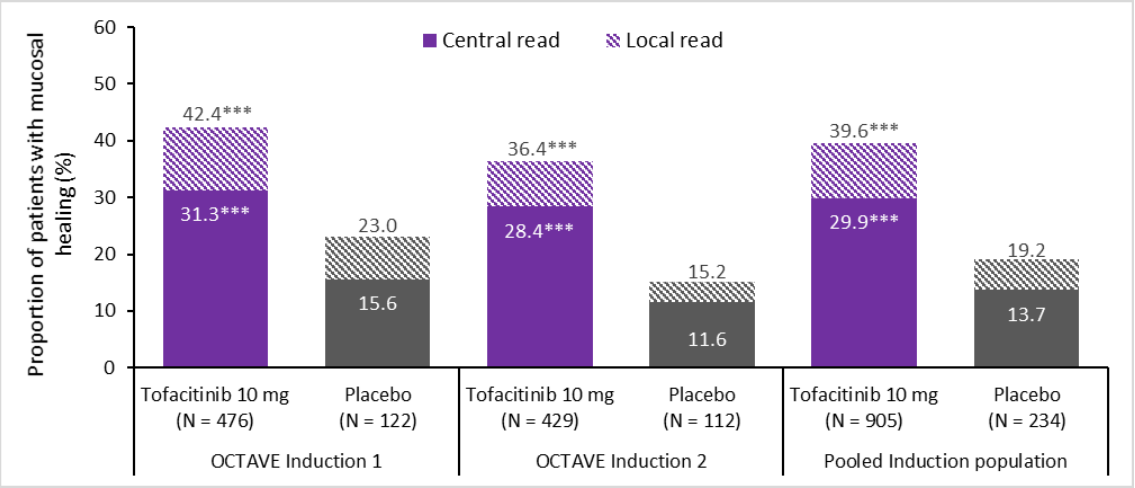
B.2.6.1.1.2 Key secondary endpoint: mucosal healing at week 8

In OCTAVE Induction 1, 31.3% of patients receiving tofacitinib 10 mg had mucosal healing at week 8, compared with 15.6% in the placebo group (difference, 15.7%; $p < 0.001$) (Figure 7 and Appendix L, Table 214). In OCTAVE Induction 2, the corresponding rates of mucosal healing were 28.4% and 11.6% (difference, 16.8%; $p < 0.001$) (98).

In an analysis based on locally read endoscopic subscores (Figure 7 and Appendix L, Table 214), the proportion of patients with mucosal healing at week 8 in OCTAVE Induction 1 was 42.4% in the tofacitinib group, compared with 23.0% in the placebo group (difference, 19.5%; $p < 0.0001$) (100). In OCTAVE Induction 2, the corresponding rates of mucosal healing were 36.4% and 15.2% (difference, 21.2%; $p < 0.0001$) (101).

Similar results were seen in an analysis of the pooled OCTAVE Induction population (Figure 7 and Appendix L, Table 214). The proportion of patients with mucosal healing at week 8 was 29.9% in the tofacitinib group, compared with 13.7% in the placebo group (difference, 16.3%; $p < 0.0001$); in the analysis using locally read endoscopic subscores, the corresponding rates of remission were 39.6% and 19.2% (difference, 20.3%; $p < 0.0001$).

Figure 7 Proportion of patients with mucosal healing at week 8 in OCTAVE Induction 1 and 2 (FAS, NRI, central and local read)



* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98); OCTAVE Induction 1 and 2 CSRs (100, 101).

B.2.6.1.1.3 Week 8 endpoints used in the economic analysis

Clinical Remission

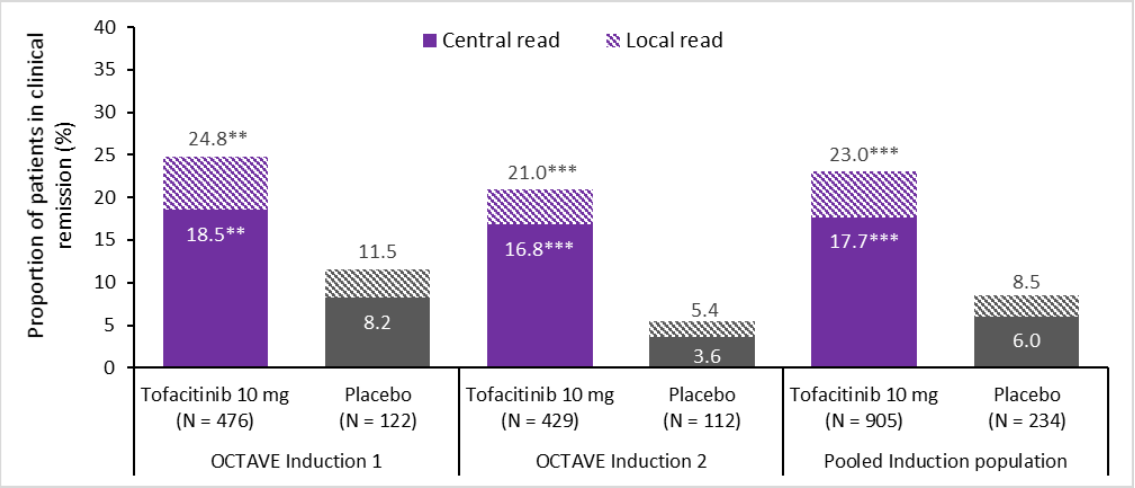
In OCTAVE Induction 1, 18.5% of patients receiving tofacitinib 10 mg achieved clinical remission at week 8, compared with 8.2% in the placebo group (difference, 10.3%; $p = 0.007$) (Figure 8 and Appendix L, Table 215). In OCTAVE Induction 2, the corresponding rates of clinical remission were 16.8% and 3.6% (difference, 13.2%; $p < 0.001$) (98).

In an analysis based on locally read endoscopic subscores (Figure 8 and Appendix L, Table 215), the proportion of patients in clinical remission at week 8 in OCTAVE Induction 1 was 24.8%, compared with 11.5% in the placebo group (difference, 13.3%; $p = 0.0017$) (100). In OCTAVE Induction 2, the corresponding rates of clinical remission were 21.0% and 5.4% (difference, 15.6%; $p = 0.0002$) (101).

The results for clinical remission in the OCTAVE Induction 1 and 2 trials are very similar to those for the primary endpoint of remission; the difference between the two endpoints corresponds to a single patient in OCTAVE Induction 2.

Similar results were seen in an analysis of the pooled OCTAVE Induction population (Figure 8 and Appendix L, Table 215). The proportion of patients in clinical remission at week 8 was 17.7% in the tofacitinib group, compared with 6.0% in the placebo group (difference, 11.7%; $p < 0.0001$); in the analysis using locally read endoscopic subscores, the corresponding rates of remission were 23.0% and 8.5% (difference, 14.4%; $p < 0.0001$).

Figure 8 Clinical remission at week 8 in OCTAVE Induction 1 and 2 (FAS, NRI, central and local reads)



* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98); OCTAVE Induction 1 and 2 CSRs (100, 101).

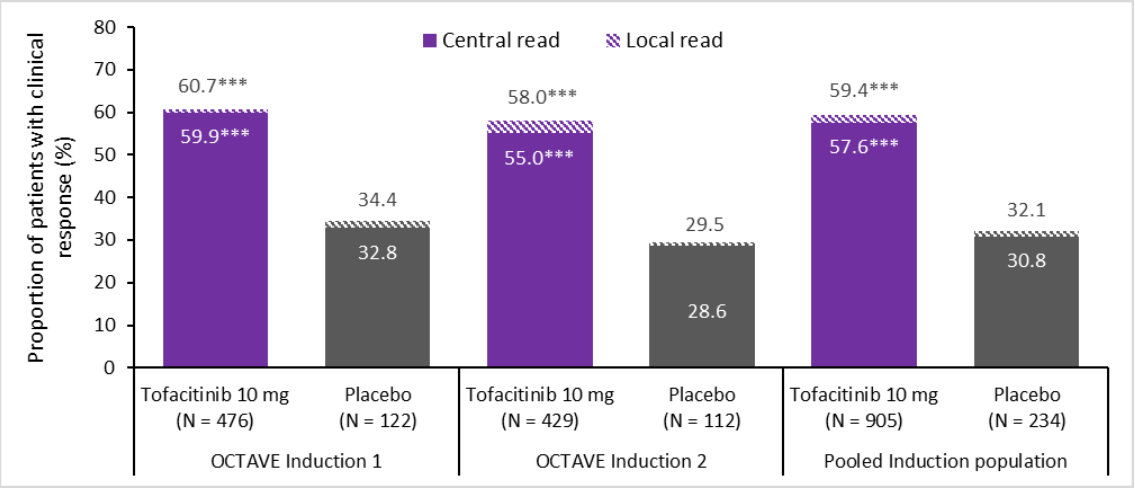
Clinical Response

OCTAVE Induction 1, 59.9% of patients receiving tofacitinib 10 mg had a clinical response at week 8, compared with 32.8% in the placebo group (difference, 27.1%; $p < 0.001$) (Figure 9 and Appendix L, Table 216). In OCTAVE Induction 2, the corresponding rates of clinical response were 55.0% and 28.6% (difference, 26.4%; $p < 0.001$) (98).

In an analysis based on locally read endoscopic subscores (Figure 9 and Appendix L, Table 216), the proportion of patients with a clinical response at week 8 in OCTAVE Induction 1 was 60.7%, compared with 34.4% in the placebo group (difference, 26.3%; $p < 0.0001$) (100). In OCTAVE Induction 2, the corresponding rates of clinical response were 58.0% and 29.5% (difference, 28.6%; $p < 0.0001$) (101).

Similar results were seen in an analysis of the pooled OCTAVE Induction population (Figure 9 and Appendix L, Table 216). The proportion of patients with a clinical response at week 8 was 57.6% in the tofacitinib group, compared with 30.8% in the placebo group (difference, 26.8%; $p < 0.0001$); in the analysis using locally read endoscopic subscores, the corresponding rates of remission were 59.4% and 32.1% (difference, 27.4%; $p < 0.0001$).

Figure 9 Clinical response at week 8 in OCTAVE Induction 1 and 2 (FAS, NRI, central and local reads)



* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98); OCTAVE Induction 1 and 2 CSRs (100, 101).

B.2.6.1.1.4 Other clinical endpoints at week 8

The following secondary outcomes were also assessed in OCTAVE Induction 1 and 2. Detailed results are presented in Appendix L.

- Endoscopic Remission at week 8 (Appendix L, Table 210)
- Symptomatic remission at week 8 (Appendix L, Table 210)
- Deep remission at week 8 (Appendix L, Table 210)
- Change in partial Mayo Score from baseline to week 8 (Appendix L, Figure 52)
- Change in total Mayo Score from baseline to week 8 (Appendix L, Table 211 and Figure 53).

B.2.6.1.2 Patient-reported outcomes in OCTAVE Induction 1 and 2

IBDQ

At week 4, 35.1% of patients receiving tofacitinib 10 mg in OCTAVE Induction 1 were in IBDQ remission (IBDQ score of ≥ 170), compared with 22.1% in the placebo group (difference, 13.0%; $p = 0.008$) (Appendix L, Table 217). In OCTAVE Induction 2, the corresponding rates of IBDQ remission were 28.9% and 8.0% (difference, 20.9%; $p < 0.001$) (98). At week 8, the proportion of patients in IBDQ remission had increased in all groups, to 43.3% and 40.3% in the two tofacitinib groups, compared with 26.2% and 17.9% in the respective placebo groups (differences, 17.0% and 22.5%; both $p < 0.001$) (98).

At week 4, 62.8% of patients receiving tofacitinib 10 mg in OCTAVE Induction 1 had an IBDQ treatment response (increase in IBDQ score of ≥ 16 points from baseline), compared with 45.1% in the placebo group (difference, 17.7%; $p < 0.001$) (Appendix L, Table 217). In OCTAVE Induction 2, the corresponding rates of IBDQ treatment response were 62.0% and 39.3% (difference, 22.7%; $p < 0.001$) (98).

At week 8, the proportion of OCTAVE Induction 1 patients with an IBDQ treatment response was 64.5% in the tofacitinib group and 45.9% in the placebo group (difference, 18.6; $p < 0.001$); in OCTAVE Induction 2, 67.1% of patients treated with tofacitinib had an IBDQ treatment response, compared with 48.2% in the placebo arm (difference 18.9%; $p < 0.001$) (98).

EQ-5D

EQ-5D results from OCTAVE Induction 1 and 2 are shown in Appendix L, Table 218. From baseline to week 8, EQ-5D utility index scores improved by a mean of +0.15 in the OCTAVE Induction 1 tofacitinib 10 mg group, compared with +0.08 in the placebo group (difference, 0.08; $p < 0.0001$). In OCTAVE Induction 2, the corresponding changes were +0.14 and +0.11 (difference, 0.03; $p = 0.22$) (100, 101).

Similar results were seen with the EQ-5D VAS. In OCTAVE Induction 1 and 2, the change from baseline to week 8 in the tofacitinib group was 17.67 and 16.52, respectively, compared with 9.49 and 8.29 in the corresponding placebo groups (differences, 8.19 and 8.23; both $p < 0.0001$; Appendix L, Table 218) (100, 101).

After 2 weeks of treatment, changes in EQ-5D utility index and VAS scores were significantly greater with tofacitinib than with placebo in both Induction trials (Appendix L, Table 218) (100, 101). For both instruments, the improvements with tofacitinib at week 2 exceeded the estimated minimal clinically important difference (MCID) for patients with IBD (MCIDs: EQ-5D utility index, 0.076; EQ-5D VAS, 10.9) (108).

SF-36 component scores (MCS and PCS)

The change from baseline to week 8 in SF-36 MCS and PCS scores is summarised in Appendix L, Table 219. In OCTAVE Induction 1, SF-36 PCS scores improved by a mean of +6.8 in the tofacitinib group and +2.5 in the placebo group (difference, 4.2; $p < 0.0001$). In OCTAVE Induction 2, the corresponding improvements were +6.8 and +4.6 (difference, 2.2; $p = 0.0035$) (100, 101).

Similar improvements were seen in SF-36 MCS scores: in OCTAVE Induction 1 the change from baseline in the tofacitinib and placebo groups was +6.8 and +3.5, respectively (difference, 3.4; $p = 0.0005$); in OCTAVE Induction 2 the corresponding values were +7.6 and +4.4 (difference, 3.2; $p = 0.0037$) (100, 101).

WPAI-UC

The change from baseline in WPAI-UC scores is shown in Appendix L, Table 212. In both OCTAVE Induction 1 and 2, patients treated with tofacitinib had significant improvements in non-work activity impairment, compared with the placebo group (100, 101).

B.2.6.1.3 Healthcare resource use

Details of ulcerative colitis-related hospitalisation and surgery in OCTAVE Induction 1 and 2 are shown in Appendix L, Table 220. The proportion of patients with ulcerative colitis-related hospitalisation was similar in the placebo and tofacitinib groups in OCTAVE Induction 1, but was numerically higher in the placebo group in OCTAVE Induction 2. Few patients had

surgery during the 8-week study period: one patient treated with tofacitinib in OCTAVE Induction 1, and two patients receiving placebo in OCTAVE Induction 2 (100, 101).

B.2.6.2 OCTAVE Sustain

All patients initiating maintenance therapy in OCTAVE Sustain had completed the OCTAVE Induction 1 or 2 trial and had a clinical response to 8 weeks of induction therapy (98).

The data presented in this submission correspond to the OCTAVE Sustain FAS. Primary endpoint results for the PPAS population (see section B.2.4.1) are shown in Appendix L, Table 208. Results according to prior treatment with TNFi therapies are summarised in this section, and described in detail in section B.2.7.3.

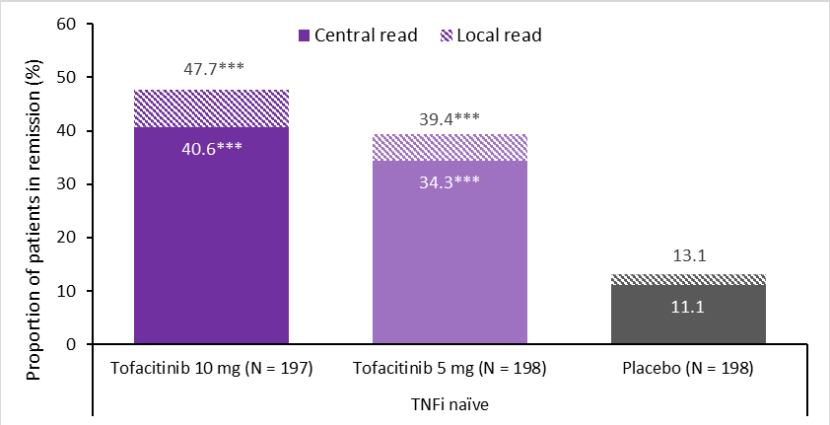
B.2.6.2.1 Clinical outcomes in OCTAVE Sustain

B.2.6.2.1.1 Primary endpoint: remission at week 52

In OCTAVE Sustain, 40.6% of patients receiving tofacitinib 10 mg achieved remission at week 52, compared with 11.1% in the placebo group (difference, 29.5%; $p < 0.001$) (Figure 10 and Appendix L, Table 221). Similarly, 34.3% of patients in the tofacitinib 5 mg group achieved remission at week 52 (difference from placebo, 23.2%; $p < 0.001$) (98).

In an analysis based on locally read endoscopic subscores (Figure 10 and Appendix L, Table 221), the proportion of patients in remission at week 52 was 13.1% in the placebo group, compared with 47.7% of those receiving tofacitinib 10 mg (difference, 34.6%; $p < 0.0001$) and 39.4% in the tofacitinib 5 mg group (difference, 26.3%; $p < 0.0001$) (102).

Figure 10 Proportion of patients in remission at week 52 in OCTAVE Sustain (FAS, NRI, central and local read)



*** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98), OCTAVE Sustain CSR (102).

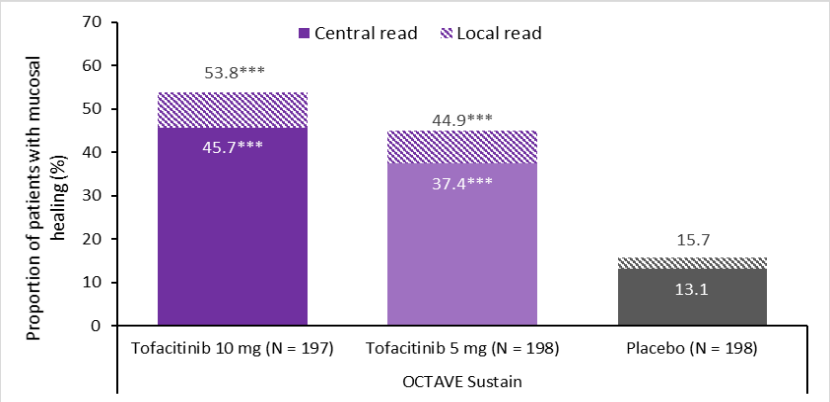
B.2.6.2.1.2 Key secondary endpoint: mucosal healing at week 52

In OCTAVE Sustain, 45.7% of patients receiving tofacitinib 10 mg had mucosal healing at week 52, compared with 13.1% in the placebo group (difference, 32.6%; p < 0.001) (Figure 11 and Appendix L, Table 222). In the tofacitinib 5 mg group, 37.4% of patients had mucosal healing at week 52 (difference from placebo, 24.2%; $p < 0.001$) (98).

In an analysis based on locally read endoscopic subscores (Figure 10 and Appendix L, Table 222), the proportion of patients in remission at week 52 was 15.7% in the placebo

group, compared with 53.8% of those receiving tofacitinib 10 mg (difference, 38.2%; $p < 0.0001$) and 44.9% in the tofacitinib 5 mg group (difference, 29.3%; $p < 0.0001$) (102).

Figure 11 Proportion of patients with mucosal healing at week 52 in OCTAVE Sustain (FAS, NRI, central and local read)



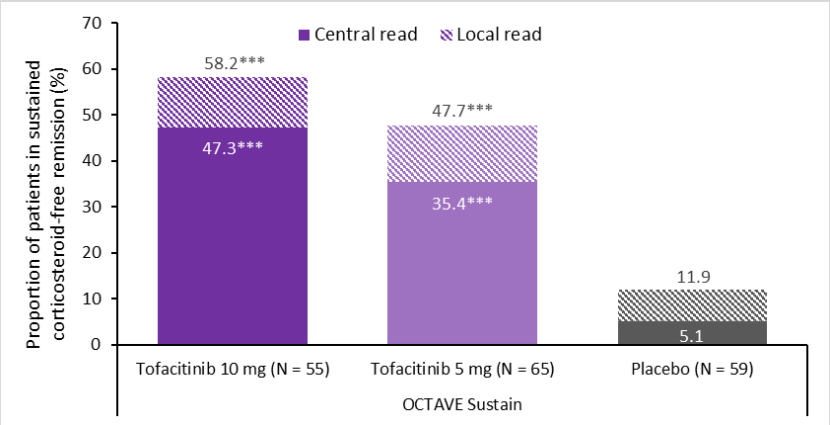
*** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98), OCTAVE Sustain CSR (102).

B.2.6.2.1.3 Key secondary endpoint: sustained corticosteroid-free remission at weeks 24 and 52

Among patients in remission at OCTAVE Sustain baseline, 47.3% (26/55) of those receiving tofacitinib 10 mg had sustained corticosteroid-free remission, compared with 5.1% (3/59) in the placebo group (difference, 42.2%; $p < 0.000$) (Figure 12 and Appendix L, Table 225). In the tofacitinib 5 mg group, 35.4% (23/65) of patients had sustained corticosteroid-free remission (difference from placebo, 30.3%; $p < 0.001$) (98).

Similar results were seen with locally read endoscopic subscores (Figure 12 and Appendix L, Table 225): 11.9% of patients in the placebo group, compared with 58.2% of those receiving tofacitinib 10 mg (difference, 46.3%; $p < 0.0001$) and 47.7% in the tofacitinib 5 mg group (difference, 35.8%; $p < 0.0001$) (102).

Figure 12 Proportion of patients in remission at baseline who had sustained steroid-free remission at weeks 24 and 52 in OCTAVE Sustain (FAS, NRI, central and local read)



*** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98), OCTAVE Sustain CSR (102).

B.2.6.2.1.4 Week 52 endpoints used in the economic analysis

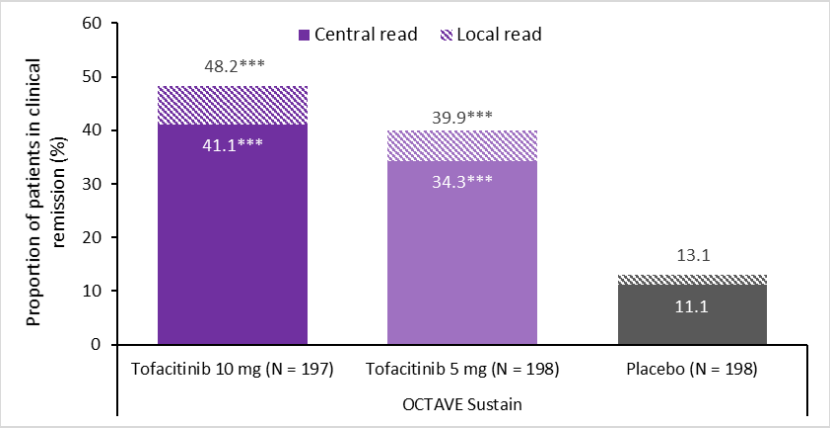
Clinical Remission at week 52

In OCTAVE Sustain, 41.1% of patients receiving tofacitinib 10 mg achieved clinical remission at week 52, compared with 11.1% in the placebo group (difference, 30.0%; $p < 0.001$) (Figure 13 and Appendix L, Table 226). In the tofacitinib 5 mg group, 34.3% of patients achieved clinical remission at week 52 (difference from placebo, 23.2%; $p < 0.001$) (98). In addition, the rate of sustained clinical remission was significantly higher in both tofacitinib groups than in the placebo group (Appendix L, Table 223).

The results for clinical remission in OCTAVE Sustain are very similar to those for the primary endpoint of remission; the difference between the two endpoints corresponds to a single patient in the tofacitinib 10 mg group (98).

In an analysis based on locally read endoscopic subscores (Figure 13 and Appendix L, Table 226), the proportion of patients in clinical remission at week 52 was 13.1% in the placebo group, compared with 48.2% of those receiving tofacitinib 10 mg (difference, 35.1%; $p < 0.0001$) and 39.9% in the tofacitinib 5 mg group (difference, 26.8%; $p < 0.0001$) (102).

Figure 13 Proportion of patients in clinical remission at week 52 in OCTAVE Sustain (FAS, NRI, central and local read)



*** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98), OCTAVE Sustain CSR (102).

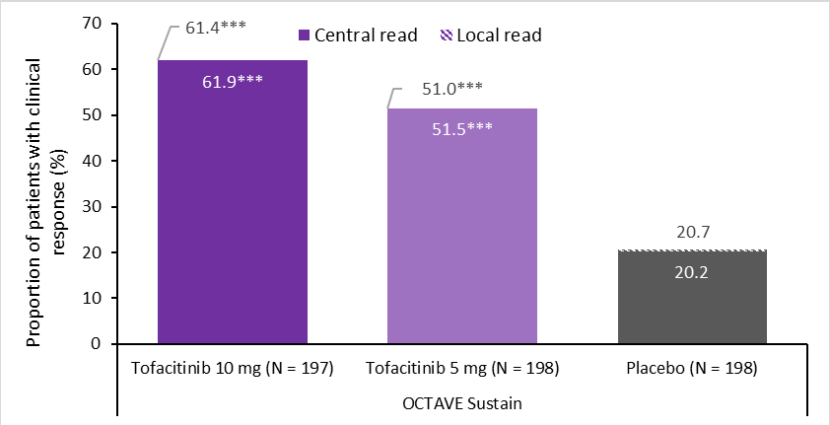
Clinical response at week 52

In OCTAVE Sustain, 61.9% of patients receiving tofacitinib 10 mg had a clinical response at week 52, compared with 20.2% in the placebo group (difference, 41.7%; $p < 0.001$) (Figure 14 and Appendix L, Table 227). Similarly, 51.5% of patients in the tofacitinib 5 mg group had a clinical response at week 52 (difference from placebo, 31.3%; $p < 0.001$) (98). In addition, the rate of sustained clinical response was significantly higher in both tofacitinib groups than in the placebo group (Appendix L, Table 227) (98).

The results for clinical response in OCTAVE Sustain are consistent with the high discontinuation rate observed in the placebo arm. In total, 132 participants randomised to placebo (66.7%) discontinued treatment due to insufficient clinical response, compared with 53 (27.0%) and 70 (35.4%) in the tofacitinib 10 mg and 5 mg arms, respectively (see section B.2.4.3 and Appendix D.1, Figure 47) (98).

In an analysis based on locally read endoscopic subscores (Figure 14 and Appendix L, Table 227), the proportion of patients with a clinical response at week 52 was 20.7% in the placebo group, compared with 61.4% of those receiving tofacitinib 10 mg (difference, 40.7%; $p < 0.0001$) and 51.0% in the tofacitinib 5 mg group (difference, 30.3%; $p < 0.0001$) (102).

Figure 14 Proportion of patients with clinical response at week 52 in OCTAVE Sustain (FAS, NRI, central and local read)



*** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98), OCTAVE Sustain CSR (102).

B.2.6.2.1.5 Other efficacy endpoints at week 52

Sustained remission with tofacitinib

In OCTAVE Sustain, the proportion of patients who achieved sustained remission (remission at both week 24 and week 52) was significantly higher with tofacitinib 10 mg or 5 mg than with placebo (see Appendix L, Table 221).

Sustained mucosal healing with tofacitinib

In OCTAVE Sustain, the proportion of patients who achieved sustained remission mucosal healing (mucosal healing at both week 24 and week 52) was significantly higher with tofacitinib 10 mg or 5 mg than with placebo (see Appendix L, Table 222).

Sustained clinical response and clinical remission

The proportion of patients with clinical response and clinical remission sustained at week 24 and week 52 is shown in Appendix L, Table 223. For both endpoints, rates were significantly higher with both tofacitinib doses than with placebo (98).

Endoscopic remission, symptomatic remission and deep remission

Results for additional binary endpoints based on Mayo scores are shown in Appendix L, Table 224. Significantly more patients achieved endoscopic remission, symptomatic remission and deep remission with tofacitinib than with placebo.

B.2.6.2.2 Patient-reported outcomes in OCTAVE Sustain

IBDQ remission

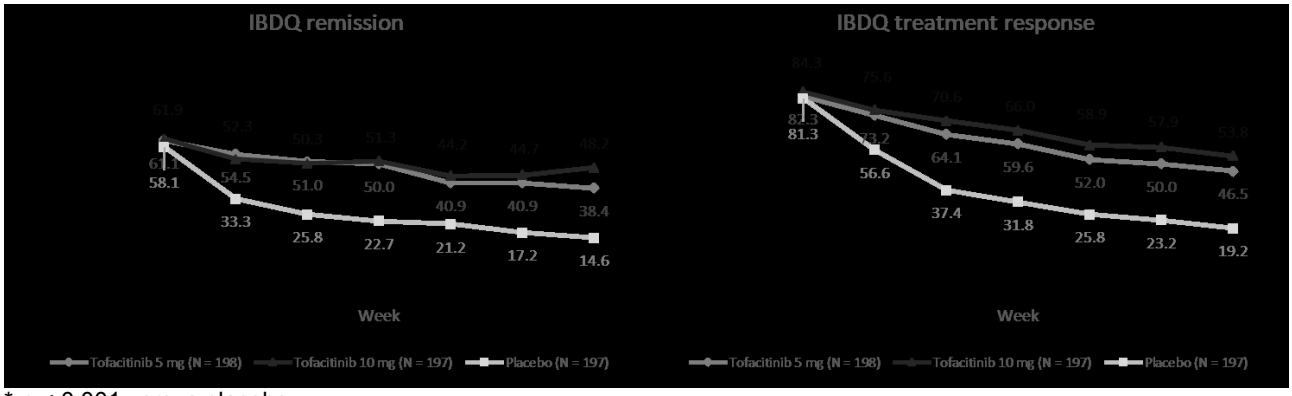
At week 52, 48.2% of patients treated with tofacitinib 10 mg in OCTAVE Sustain, and 38.4% of those receiving tofacitinib 5 mg, were in IBDQ remission, compared with 14.6% in the

placebo group ($p < 0.001$ for both tofacitinib groups vs placebo) (98).The proportion of patients in IBDQ remission was significantly different between the tofacitinib and placebo groups at all timepoints from week 8 to week 52 (Figure 15) (102).

IBDQ treatment response

The proportion of patients with an IBDQ response was significantly higher in the tofacitinib groups than the placebo group as early as week 8 (Figure 15). At week 52, 53.8% of patients treated with tofacitinib 10 mg, and 46.5% of those receiving tofacitinib 5 mg, had an IBDQ response, compared with 19.2% in the placebo group ($p < 0.001$ for both tofacitinib groups vs placebo) (98).

Figure 15 Proportion of patients with IBDQ remission and IBDQ treatment response in OCTAVE Sustain (FAS, NRI)



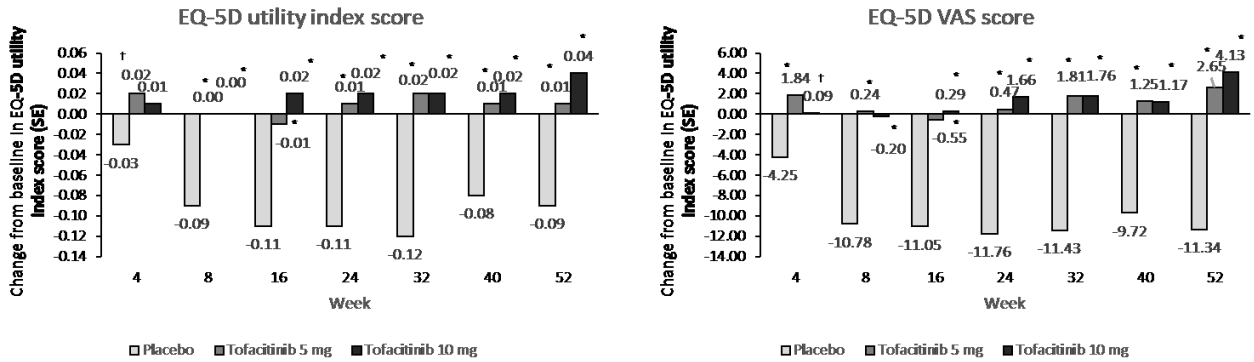
* $p < 0.001$ versus placebo.
p values were calculated using a CMH Chi-squared test stratified by treatment assignment in the induction study and remission at maintenance study baseline.
Abbreviations: CMH, Cochran-Mantel-Haenszel; FAS, full analysis set; IBDQ, Inflammatory Bowel Disease Questionnaire; NRI, non-responder imputation.
Source: OCTAVE Sustain CSR (102).

EQ-5D

EQ-5D results from OCTAVE Sustain are shown in Figure 16. From baseline to week 52, mean EQ-5D utility scores increased slightly in the two tofacitinib groups (5 mg, +0.01; 10 mg, +0.04), but decreased in the placebo group (−0.09; $p < 0.001$ vs both tofacitinib groups). The difference between both tofacitinib groups and placebo was statistically significant as early as week 8 (102).

Similar results were seen with the EQ-5D VAS. With both tofacitinib doses, the mean EQ-5D VAS score increased from baseline to week 52 (5 mg, +2.65; 10 mg, +4.13), compared with a reduction of −11.34 in the placebo group ($p < 0.0001$ vs both tofacitinib doses) (102). In addition, the decrease in EQ-5D VAS score in the placebo group compared with the tofacitinib groups was statistically significant from week 4 (102).

Figure 16 Change from baseline to week 52 in EQ-5D utility index and VAS scores in OCTAVE Sustain (FAS, as observed)



† $p < 0.05$ vs placebo; * $p < 0.001$ vs placebo.
Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; VAS, visual analogue scale.
Source: OCTAVE Sustain CSR (102).

SF-36 component scores (MCS and PCS)

The change from baseline in SF-36 MCS and PCS scores is summarised in Appendix L, Table 228. From baseline to week 24, SF-36 PCS scores decreased by a mean of −5.0 in the placebo group, compared with −0.3 in the tofacitinib 5 mg group and +0.4 in the tofacitinib 10 mg group (differences, 4.8 and 5.4, respectively; both $p < 0.0001$ vs placebo). At week 52, PCS scores were stable in the two tofacitinib groups (5 mg, −0.0; 10 mg, +0.3), compared with a mean reduction of −5.2 in the placebo group (differences, 5.1 and 5.5; $p < 0.0001$ for both tofacitinib groups vs placebo) (102).

Similar results were seen for the SF-36 MCS, with significant decreases in the placebo group compared with both tofacitinib doses at both week 24 and week 52 (Appendix L, Table 228). The difference from placebo was 6.3 and 6.9 in the tofacitinib 5 mg and 10 mg groups, respectively, at week 24; at week 52 the corresponding differences were 5.8 and 6.1 (all $p < 0.0001$) (102).

WPAI-UC

The change from baseline in WPAI-UC scores is shown in Appendix L, Table 229. Patients in both tofacitinib groups had significant improvements in presenteeism and non-work activity impairment compared with the placebo group.

B.2.6.2.3 Healthcare resource use

Details of ulcerative colitis-related hospitalisation and surgery in OCTAVE Sustain are shown in Appendix L, Table 230. The proportion of patients with ulcerative colitis-related hospitalisation was generally low, with six patients in the placebo group, five in the tofacitinib 5 mg group and two in the tofacitinib 10 mg group requiring hospitalisation. The mean duration of ulcerative colitis-related hospitalisation was numerically higher in the placebo group than in the tofacitinib groups. Overall, few patients had surgery during the 52-week study period, but two patients in the placebo arm had ulcerative colitis-related colectomy (Appendix L, Table 230) (102).

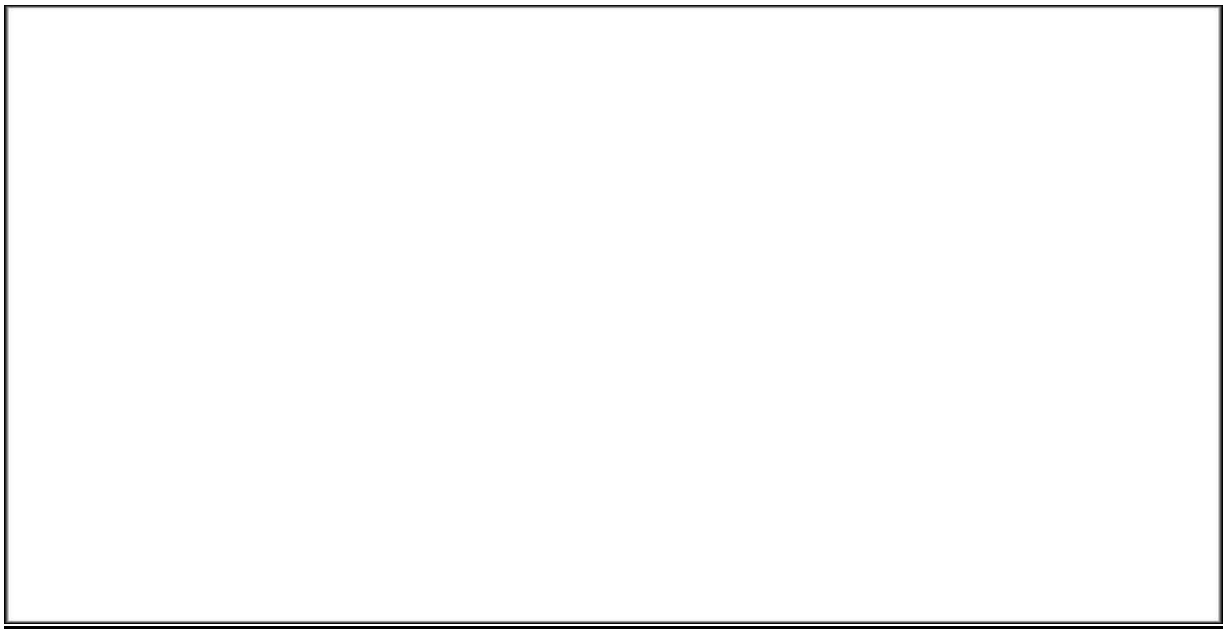
B.2.6.3 OCTAVE Open

B.2.6.3.1 Overview

The OCTAVE Open study included patients who completed 52 weeks of maintenance therapy in OCTAVE Sustain, and patients did not have a response in OCTAVE Induction 1 or 2, or who withdrew from OCTAVE Sustain due to treatment failure (see section B.2.3.1.1) (103). Therefore, the patients in OCTAVE Open comprise four distinct populations (Figure 17) (103):

- The 'maintenance remission' population ([redacted]), composed of patients with a response to induction therapy in OCTAVE Induction 1 and 2 who were in remission at week 52 in OCTAVE Sustain; this group received tofacitinib 5 mg twice daily in OCTAVE Open.
- The 'other maintenance completers' population ([redacted]), comprising patients who at the end of 52 weeks of maintenance therapy in OCTAVE Sustain were not in remission but did not meet the definition of treatment failure.
- The 'maintenance treatment failure' population ([redacted]), comprising patients with a response in OCTAVE Induction 1 and 2 who withdrew from OCTAVE Sustain due to treatment failure on tofacitinib (5 mg, [redacted]; 10 mg, [redacted]) or placebo ([redacted]).
- The 'induction non-responders' population ([redacted]), composed of patients who did not have a response to induction therapy and did not enter OCTAVE Sustain (103).

Figure 17 Summary of OCTAVE Open patient populations



Numbers of patients with data for 12 months of treatment in OCTAVE Open are based on NRI analysis.
Abbreviations: NRI, non-responder imputation. Source: OCTAVE Open CSR (103).
Data in this figure are [redacted].

Demographics, patient baseline characteristics and disease characteristics at entry to the OCTAVE Open study are shown in Appendix L, Table 231 and Table 232. In the maintenance remission group, [redacted] of patients with treatment failure in OCTAVE Sustain were using corticosteroids at entry to OCTAVE Open.

Patient disposition in the OCTAVE Open study is shown in Appendix D, Table 119. As of the cut-off date for available data (8 July 2016), [redacted]

Full tables of OCTAVE Open endpoint results are shown in Appendix L, Table 233, Table 234, Table 235 and Table 236. In this section, results are summarised for the proportion of patients in remission at month 12; although the OCTAVE Open study is ongoing, data are available for only a small number of patients at month 24. Unless otherwise stated OCTAVE Open results described in this submission are based on locally read endoscopic subscores and an NRI analysis (i.e. patients who discontinued treatment within 12 months were imputed as non-responders).

B.2.6.3.2 Maintenance remission

Of the [redacted] OCTAVE Open patients with 12-month data who were in remission at week 52 in OCTAVE Sustain, most had been treated with tofacitinib maintenance therapy (5 mg, [redacted] patients; 10 mg, [redacted] patients). After 12 months of maintenance therapy with tofacitinib 5 mg, the majority were still in remission ([redacted]; [redacted] in the OCTAVE Sustain tofacitinib 5 mg group; [redacted] in the OCTAVE Sustain tofacitinib 10 mg group; NRI analysis, see Appendix L, Table 233) (103).

B.2.6.3.3 Other maintenance completers

Of the [redacted] OCTAVE Open patients who completed OCTAVE Sustain but were not in remission at week 52, [redacted] were classified as being in remission at OCTAVE Open baseline, based on locally read endoscopic subscores. Among the [redacted] patients in this group who had data for 12 months of tofacitinib 10 mg treatment in OCTAVE Open, the proportion in remission had increased to [redacted] ([redacted] patients), including [redacted] patients ([redacted]) previously treated with placebo maintenance therapy (NRI analysis, see Appendix L, Table 234) (103).

B.2.6.3.4 Maintenance treatment failure

The 'maintenance treatment failure' population ([redacted]) comprises patients with a response in OCTAVE Induction 1 and 2 who withdrew from OCTAVE Sustain due to treatment failure on tofacitinib (5 mg, [redacted]; 10 mg, [redacted]) or placebo ([redacted]). Of these, [redacted] achieved

remission after 8 weeks of treatment with tofacitinib 10 mg, including [redacted] and [redacted] of those who had treatment failure during OCTAVE Sustain on tofacitinib 5 mg or placebo, respectively (NRI analysis, see Appendix L, Table 235) (103).

Among the [redacted] patients in the maintenance treatment failure group who had data for 12 months of tofacitinib 10 mg treatment in OCTAVE Open, [redacted] achieved remission. Patients in remission at month 12 in OCTAVE Open included [redacted] of those with an initial response to tofacitinib 10 mg who were randomised to placebo maintenance therapy, and [redacted] of those with an initial response to tofacitinib 10 mg who were randomised to tofacitinib 5 mg maintenance therapy (NRI analysis, see Appendix L, Table 235) (103).

B.2.6.3.5 Induction non-responders

In total, [redacted] patients without a response to induction therapy in OCTAVE Induction 1 and 2 received tofacitinib 10 mg in OCTAVE Open; [redacted] patients in this group had received tofacitinib 10 mg in the Induction trials. Of these, [redacted] were in remission after an additional 8 weeks of induction treatment with tofacitinib 10 mg, for a total of 16 weeks of induction therapy, and [redacted] had a clinical response (NRI analysis; see Appendix L, Table 236) (103).

Among the [redacted] patients in the induction non-responders group who had data for 12 months of tofacitinib 10 mg treatment in OCTAVE Open, [redacted] were in remission at month 12. Among the [redacted] patients who received tofacitinib 10 mg in OCTAVE Induction 1 and 2, [redacted] achieved remission after a further 12 months of treatment, as did [redacted] of those initially randomised to placebo (NRI analysis; see Appendix L, Table 236) (103).

There was a high rate of discontinuation in this analysis population (see Appendix D, Table 119) – when only patients who continued therapy for 12 months (as observed analysis) were considered, the proportion in remission at month 12 was [redacted] and [redacted] in the initial tofacitinib 10 mg and placebo groups, respectively (see Appendix L, Table 236) (103).

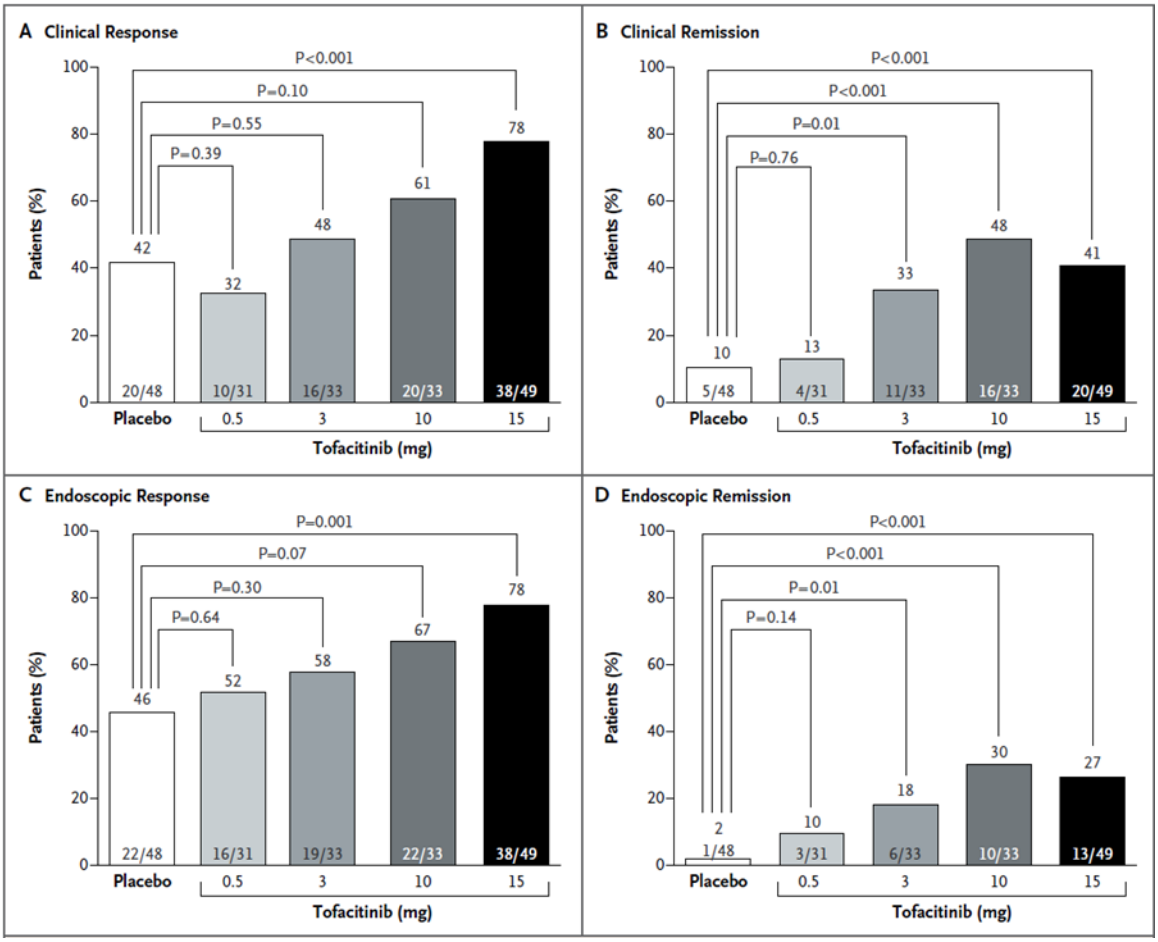
B.2.6.4 Phase II study

Tofacitinib has also been compared with placebo in a Phase II trial (NCT00787202) (99). In the 8-week, double-blind study, adults with a confirmed diagnosis of ulcerative colitis, total Mayo scores of 6 to 12 and endoscopic Mayo scores of 2 to 3 (n = 194) were randomised to receive placebo or 1 of 4 doses of tofacitinib (0.5 mg, 3 mg, 10 mg or 15 mg), administered twice daily (99).

Baseline characteristics for patients in the Phase II study are shown in Appendix L, Table 237. The clinical efficacy results in the Phase II study tofacitinib 10 mg group were consistent with those in the Phase III OCTAVE Induction 1 and 2 trials described in section B.2.6.1.1.3 (Figure 18). The primary endpoint was clinical response at 8 weeks, which was achieved by 61% of patients (20/33) receiving tofacitinib 10 mg, similar to the results in OCTAVE Induction 1 (59.9%) and 2 (60.7%) (see section B.2.6.1.1.3) (98, 99).

Clinical response and clinical remission data from the Phase II study are included in the NMA (section B.2.9)

Figure 18 Phase II study – summary of clinical efficacy results at week 8



Source: Sandborn *et al.* 2012 (99).

B.2.7 Subgroup analysis

B.2.7.1 Subgroup analyses conducted

Efficacy outcomes were analysed according to the following key prespecified subgroups, based on duration of disease, treatment history, and for OCTAVE Sustain, status at the maintenance study baseline. Analysis of heterogeneity was conducted using the Breslow-Day test.

Subgroup analyses of OCTAVE Induction 1 and 2:

- prior TNFi exposure (yes vs no)
- prior TNFi failure (yes vs no)
- baseline corticosteroid use (yes vs no)

Subgroup analyses of OCTAVE Sustain:

- duration of disease (< 6 years vs ≥ 6 years)
- prior TNFi exposure (yes vs no)
- prior TNFi failure (yes vs no)
- prior corticosteroid failure (yes vs no)
- Induction study treatment assignment (tofacitinib 10 mg vs tofacitinib 10 mg or 15 mg vs placebo)
- remission at maintenance study baseline (yes vs no)
- mucosal healing at maintenance study baseline (yes vs no)
- corticosteroid use at maintenance study baseline (yes vs no)

B.2.7.2 Subgroup analysis results

Detailed results for all subgroup analyses are shown in Appendix E, Table 130 to Table 153. In addition, results for the key prior TNFi treatment subgroup are discussed in this section.

In both OCTAVE Induction trials, more than half of participants had previously received a TNFi therapy (53–58% across groups); of these, most had had treatment failure with a TNFi therapy (51–54% of the population; see section B.2.3.2, Table 15) (98).

Overall, subgroup analyses showed higher efficacy with tofacitinib than placebo in all subgroups investigated. There was no evidence of a systematic difference in treatment effect according to prior TNFi treatment. In most analyses the difference between tofacitinib and placebo was statistically significant; however, the OCTAVE trials were not powered to test the statistical significance of subgroup analyses due to the limited patient numbers in the subgroups. Therefore, *p* values from subgroup analyses of the individual OCTAVE Induction trials should be treated with caution. To increase the statistical power, subgroup analyses were also conducted for the pooled population from OCTAVE Induction 1 and 2.

B.2.7.3 Overview of trial results according to prior TNFi treatment

Subgroup analysis results according to prior TNFi failure are shown in full in Appendix E and are generally similar to the results according to prior TNFi treatment. In subgroup analyses of the pooled Induction trials, the efficacy results were highly significant.

Subgroup analysis results according to prior TNFi treatment are summarised in Table 20 and Table 21.

Table 20 Summary of statistical significance of OCTAVE Induction 1 and 2 outcomes according to prior TNFi treatment

Clinical impact	Outcome assessed	Used in CEA?	Time points (weeks)	Endoscopic read	OCTAVE Induction 1	OCTAVE Induction 2	Pooled Induction trials
Disease activity	TNFi-naïve				n = 222	n = 195	n = 417
	Remission (primary endpoint)	No	8	Central	NS	Sig	Sig
				Local	NS	Sig	Sig
	Mucosal healing (key secondary endpoint)	No	8	Central	NS	Sig	Sig
				Local	NS	Sig	Sig
	Clinical remission	Yes	8	Central	NS	Sig	Sig
				Local	NS	Sig	Sig
	Clinical response	Yes	8	Central	Sig	Sig	Sig
				Local	NS	Sig	Sig
	TNFi-experienced				n = 254	n = 234	n = 488
	Remission (primary endpoint)	No	8	Central	Sig	Sig	Sig
				Local	Sig	Sig	Sig
	Mucosal healing (key secondary endpoint)	No	8	Central	Sig	Sig	Sig
				Local	Sig	Sig	Sig
	Clinical remission	Yes	8	Central	Sig	Sig	Sig
				Local	Sig	Sig	Sig
	Clinical response	Yes	8	Central	Sig	Sig	Sig
				Local	Sig	Sig	Sig

Statistical significance = $p < 0.05$.
Abbreviations: NS, not significant; Sig, significant difference versus placebo; TNFi, tumour necrosis factor inhibitor.

Table 21 Summary of statistical significance of OCTAVE Sustain outcomes according to prior TNFi treatment

Clinical impact	Outcome assessed	Used in CEA?	Time points (weeks)	Endoscopic read	Tofacitinib 5 mg	Tofacitinib 10 mg
Disease activity	TNFi-naïve				n = 108	n = 96
	Remission (primary endpoint)	No	52	Central	Sig	Sig
				Local	Sig	Sig
	Mucosal healing (key secondary endpoint)	No	52	Central	Sig	Sig
				Local	Sig	Sig
	Sustained corticosteroid-free remission [a]	No	24 and 52	Central	Sig	Sig
				Local	Sig	Sig
	Clinical remission	Yes	52	Central	Sig	Sig
				Local	Sig	Sig
	Clinical response	Yes	52	Central	Sig	Sig
				Local	Sig	Sig
	TNFi-experienced				n = 90	n = 101
	Remission (primary endpoint)	No	52	Central	Sig	Sig
				Local	Sig	Sig
	Mucosal healing (key secondary endpoint)	No	52	Central	Sig	Sig
				Local	Sig	Sig
	Sustained corticosteroid-free remission [a]	No	24 and 52	Central	NS	Sig
				Local	Sig	Sig
	Clinical remission	Yes	52	Central	Sig	Sig
				Local	Sig	Sig
	Clinical response	Yes	52	Central	Sig	Sig
				Local	Sig	Sig

Statistical significance = $p < 0.05$.
[a] Tofacitinib 5 mg: TNFi-naïve, n = 43; TNF experienced, n = 22. Tofacitinib 10 mg: TNFi-naïve, n = 34; TNF experienced, n = 21.
Abbreviations: NS, not significant; Sig, significant difference versus placebo; TNFi, tumour necrosis factor inhibitor.

B.2.7.4 OCTAVE Induction trial results according to prior TNFi treatment

Remission

In OCTAVE Induction 1 and 2, the difference between remission rates in the tofacitinib 10 mg and placebo groups was generally similar between the prior TNFi 'Yes' and 'No' subgroups, and analysis of heterogeneity did not suggest any significant difference in treatment effect between subgroups (Figure 19 and Appendix E, Table 121).

In OCTAVE Induction 1, 25.2% of patients without prior TNFi exposure who were treated with tofacitinib and 15.8% of those in the placebo group were in remission at week 8; among those with prior TNFi exposure, the corresponding values were 12.6% and 1.5%; the differences between tofacitinib and placebo were 9.4% ($p = 0.13$) and 11.1% ($p = 0.0090$) in the two subgroups, respectively.

In OCTAVE Induction 2, 22.1% of patients without prior TNFi exposure who were treated with tofacitinib and 8.5% of those in the placebo group were in remission at week 8; among those with prior TNFi exposure, the corresponding values were 12.0% and 0.0%; the differences between tofacitinib and placebo were 13.5% ($p = 0.035$) and 12.0% ($p = 0.0060$) in the two subgroups, respectively.

In the OCTAVE 1 and 2 pooled analyses the results were highly significant across both prior-TNFi-exposure subgroups. In the tofacitinib group, [redacted] of patients without prior TNFi-exposure achieved remission at week 8, compared with [redacted] in the placebo group. Among those with prior TNFi exposure, the corresponding values were [redacted] and [redacted]. The differences between tofacitinib and placebo were [redacted] ($p = 0.0122$) and [redacted] ($p < 0.0001$) in the two subgroups, respectively.

Figure 19 Proportion of patients in remission in OCTAVE Induction 1 and 2 according to prior TNFi treatment (FAS, NRI, central and local reads)

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98); OCTAVE Induction 1 and 2 CSRs (100, 101).
Local read data and results from the pooled Induction population in this figure are [redacted].

Mucosal healing

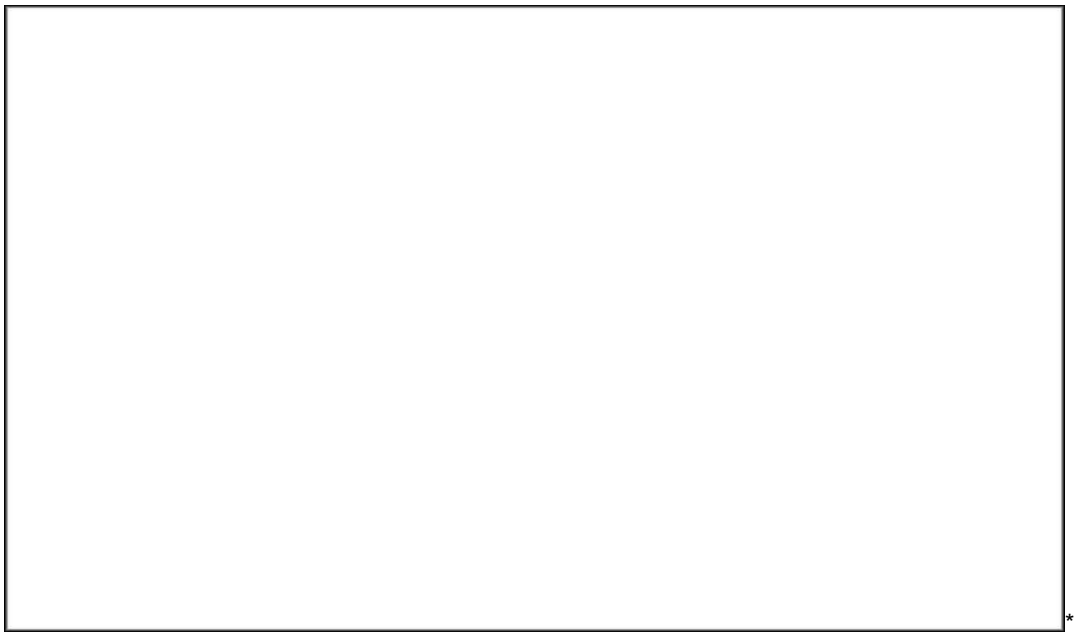
In OCTAVE Induction 1 and 2, the difference between rates of mucosal healing in the tofacitinib 10 mg and placebo groups was numerically higher in the TNFi 'Yes' group than in the 'No' group, but analysis of heterogeneity did not suggest any significant difference in treatment effect between subgroups (Figure 20 and Appendix E, Table 122).

In OCTAVE Induction 1, 39.6% of patients without prior TNFi exposure who were treated with tofacitinib and 26.3% of those in the placebo group had mucosal healing at week 8; among those with prior TNFi exposure, the corresponding values were 24.0% and 6.2%; the differences between tofacitinib and placebo were 13.3% ($p = 0.063$) and 17.9% ($p = 0.0014$) in the two subgroups, respectively.

In OCTAVE Induction 2, 36.4% of patients without prior TNFi exposure who were treated with tofacitinib and 19.1% of those in the placebo group had mucosal healing at week 8; among those with prior TNFi exposure, the corresponding values were 21.8% and 6.2%; the differences between tofacitinib and placebo were 17.3% ($p = 0.024$) and 15.6% ($p = 0.0040$) in the two subgroups, respectively.

In the OCTAVE 1 and 2 pooled analyses the results were highly significant across both prior-TNFi-exposure subgroups. In the tofacitinib group, [redacted] of patients without prior TNFi-exposure had mucosal healing at week 8, compared with [redacted] in the placebo group. Among those with prior TNFi exposure, the corresponding values were [redacted] and [redacted]. The differences between tofacitinib and placebo were [redacted] ($p = 0.0039$) and [redacted] ($p < 0.0001$) in the two subgroups, respectively.

Figure 20 Proportion of patients with mucosal healing in OCTAVE Induction 1 and 2 according to prior TNFi treatment (FAS, NRI, central and local reads)



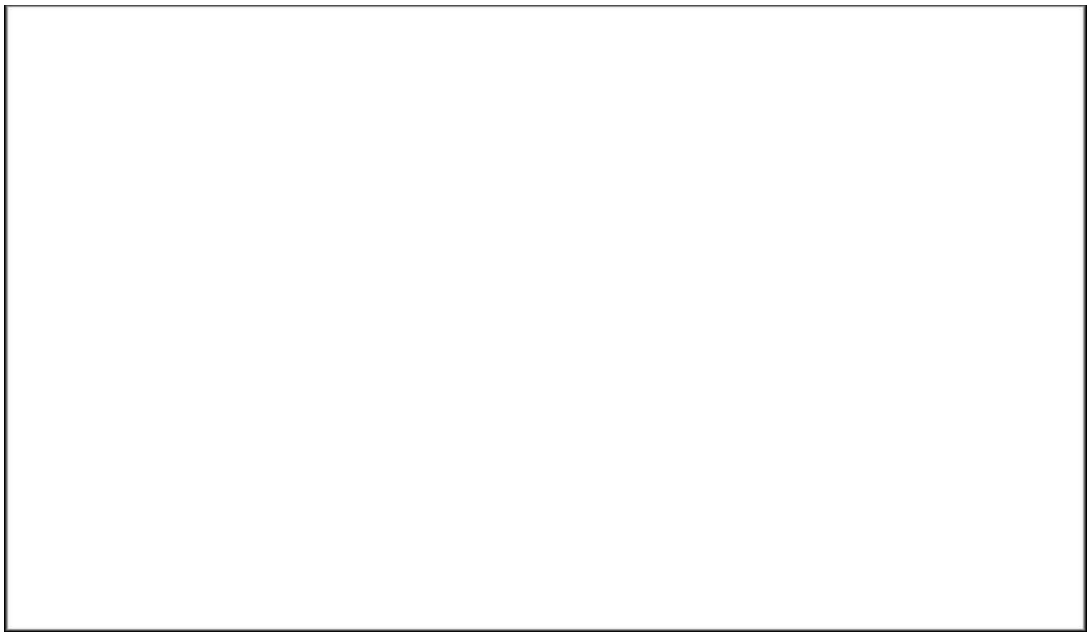
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98); OCTAVE Induction 1 and 2 CSRs (100, 101).
Local read data and results from the pooled Induction population in this figure are [redacted].

Clinical remission

In OCTAVE Induction 1 and 2, the difference between clinical remission rates in the tofacitinib 10 mg and placebo groups was generally similar between the prior TNFi 'Yes' and 'No' subgroups, and analysis of heterogeneity did not suggest any significant difference in treatment effect between subgroups (Figure 21 and Appendix E, Table 123).

Summary results for the OCTAVE 1 and 2 pooled analyses per prior-TNFi-subgroup are as follows: in the tofacitinib group, [redacted] of patients without prior TNFi-exposure achieved clinical remission at week 8, compared with [redacted] in the placebo group. Among those with prior TNFi exposure, the corresponding values were [redacted] and [redacted]. The differences between tofacitinib and placebo were [redacted] ($p < 0.0108$) and [redacted] ($p < 0.0001$) in the two subgroups, respectively.

Figure 21 Proportion of patients in clinical remission in OCTAVE Induction 1 and 2 according to prior TNFi treatment (FAS, NRI, central and local reads)



$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98); OCTAVE Induction 1 and 2 CSRs (100, 101).
Data in this figure are [redacted].

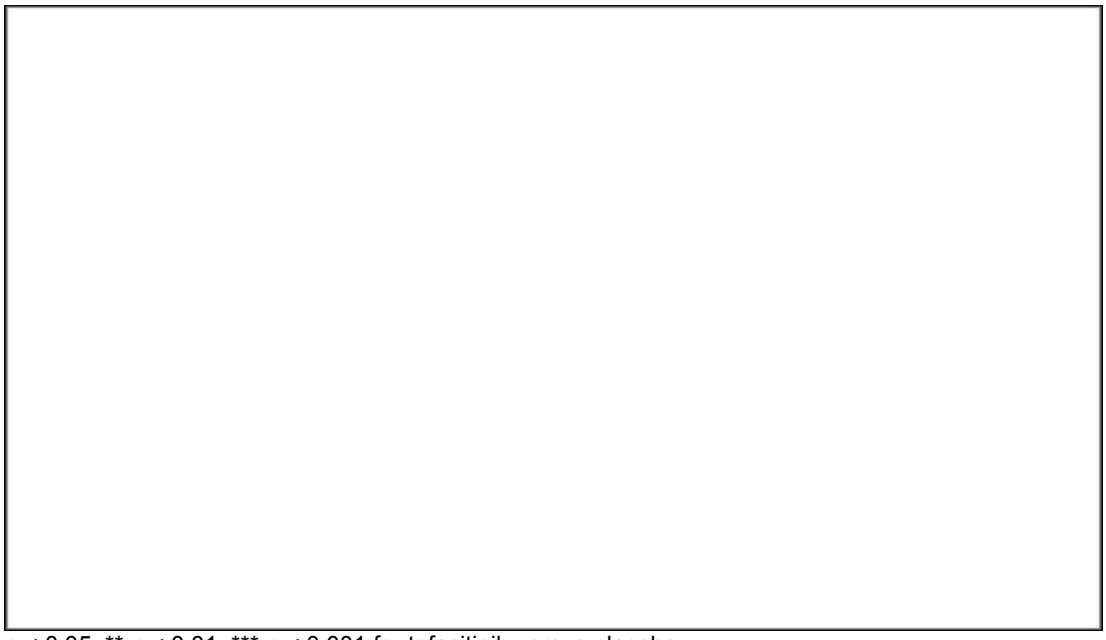
Clinical response

In OCTAVE Induction 1 the difference between clinical response rates in the tofacitinib 10 mg and placebo groups was larger in the prior TNFi 'Yes' subgroup than in the 'No' subgroup, with a higher clinical response rate among placebo arm participants who had not received prior TNFi treatment than among those who had. Analysis of heterogeneity suggested a significant difference in treatment effect between subgroups (Figure 22 and Appendix E, Table 124). By contrast, no significant differences between subgroups were seen in OCTAVE Induction 2.

Summary results for the OCTAVE 1 and 2 pooled analyses per prior-TNFi-subgroup are as follows: in the tofacitinib group, [redacted] of patients without prior TNFi-exposure achieved clinical remission at week 8, compared with [redacted] in the placebo group. Among those with

prior TNFi exposure, the corresponding values were [redacted] and [redacted]. The differences between tofacitinib and placebo were [redacted] ($p < 0.0001$) and [redacted] ($p < 0.0001$) in the two subgroups, respectively.

Figure 22 Proportion of patients with clinical response in OCTAVE Induction 1 and 2 according to prior TNFi treatment (FAS, NRI, central and local reads)



$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: Sandborn *et al.* 2017 (98); OCTAVE Induction 1 and 2 CSRs (100, 101).
Data in this figure are [redacted].

B.2.7.5 OCTAVE Sustain trial results according to prior TNFi treatment

Remission

In OCTAVE Sustain, remission rates with tofacitinib 10 mg or 5 mg were significantly higher than those with placebo in both the prior TNFi 'Yes' and 'No' subgroups. The difference between placebo and tofacitinib remission rates was numerically higher in the prior TNFi 'No' subgroup than in the 'Yes' subgroup – this difference was more apparent in the tofacitinib 5-mg arm than in the 10-mg arm (Figure 23 and Appendix E, Table 125).

Among patients without prior TNFi exposure, the remission rate at week 52 was [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] and [redacted] (both $p < 0.0001$). In the prior-TNFi subgroup, remission rates were [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] ($p < 0.0001$) and [redacted] ($p = 0.0118$).

Figure 23 Proportion of patients in remission in OCTAVE Sustain according to prior TNFi treatment (FAS, NRI, central and local reads)

$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: OCTAVE Sustain CSR (102). Data in this figure are [redacted].

Mucosal healing

In OCTAVE Sustain, the proportion of patients who had mucosal healing was significantly higher with tofacitinib 10 mg or 5 mg than with placebo in both subgroups. The difference between placebo and tofacitinib mucosal healing rates was numerically higher in the prior TNFi 'No' subgroup than in the 'Yes' subgroup (Figure 24 and Appendix E, Table 126).

Among patients without prior TNFi exposure, the mucosal healing rate at week 52 was [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] and [redacted] (both $p < 0.0001$). In the prior-TNFi subgroup, mucosal healing rates were [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] ($p < 0.0001$) and [redacted] ($p = 0.0020$).

Figure 24 Proportion of patients with mucosal healing in OCTAVE Sustain according to prior TNFi treatment (FAS, NRI, central and local reads)

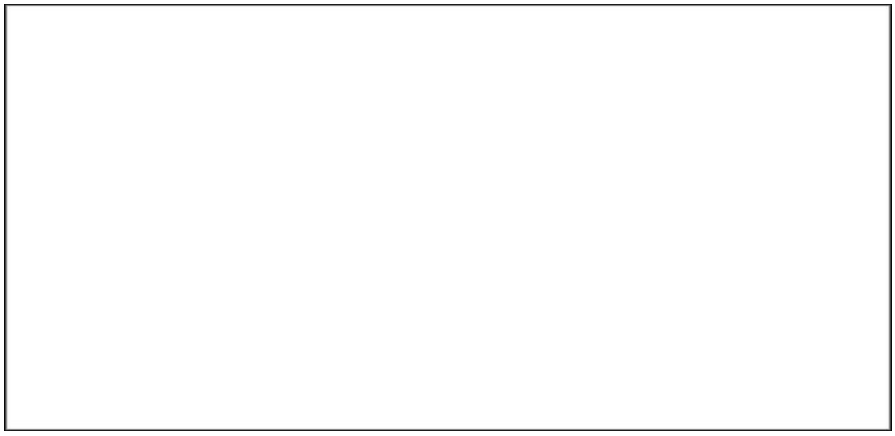
$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: OCTAVE Sustain CSR (102).
Data in this figure are [redacted].

Clinical remission

In OCTAVE Sustain, clinical remission rates with tofacitinib 10 mg or 5 mg were significantly higher than those with placebo in both the prior TNFi 'Yes' and 'No' subgroups (Figure 25 and Appendix E, Table 127). The difference between placebo and tofacitinib clinical remission rates was numerically higher in the prior TNFi 'No' subgroup than in the 'Yes' subgroup – this difference was more apparent in the tofacitinib 5-mg arm than in the 10-mg arm.

Among patients without prior TNFi exposure, the clinical remission rate at week 52 was [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] and [redacted] (both $p < 0.0001$). In the prior-TNFi subgroup, clinical remission rates were [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] ($p < 0.0001$) and [redacted] ($p = 0.0118$).

Figure 25 Proportion of patients in clinical remission in OCTAVE Sustain according to prior TNFi treatment (FAS, NRI, central and local reads)



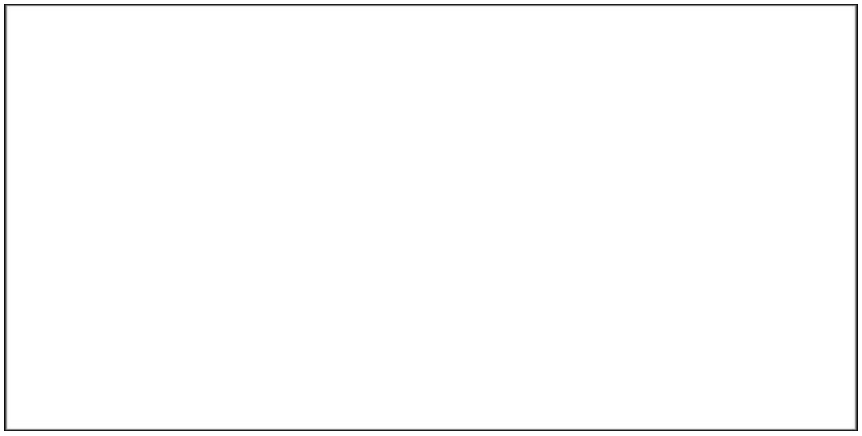
$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: OCTAVE Sustain CSR (102).
Data in this figure are [redacted].

Clinical response

In OCTAVE Sustain, clinical response rates with tofacitinib 10 mg or 5 mg were significantly higher than those with placebo in both the prior TNFi 'Yes' and 'No' subgroups (Figure 26 and Appendix E, Table 128). The difference between placebo and tofacitinib clinical response rates was numerically higher with both tofacitinib and placebo in the prior TNFi 'No' subgroup than in the 'Yes' subgroup, but the differences between tofacitinib and placebo were similar in both subgroups.

Among patients without prior TNFi exposure, the clinical response rate at week 52 was [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] and [redacted] (both $p < 0.0001$). In the prior-TNFi subgroup, clinical response rates were [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] and [redacted] (both $p < 0.0001$).

Figure 26 Proportion of patients with clinical response in OCTAVE Sustain according to prior TNFi treatment (FAS, NRI, central and local reads)



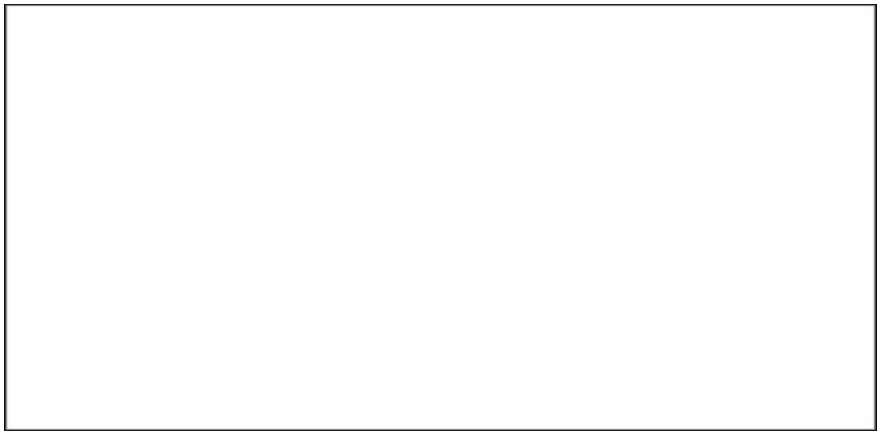
$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: OCTAVE Sustain CSR (102).
Data in this figure are [redacted].

Sustained corticosteroid-free remission among patients in remission at baseline

In OCTAVE Sustain, the proportion of patients who had sustained corticosteroid-free remission was higher with tofacitinib 10 mg or 5 mg than with placebo in both the prior TNFi 'Yes' and 'No' subgroups [redacted] (*p* values for these analyses should be treated with caution due to small subgroup numbers; Figure 27 and Appendix E, Table 129).

Among patients without prior TNFi exposure who were in remission at baseline, the corticosteroid-free remission rate at week 52 was [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] and [redacted] (both $p < 0.0001$). In the prior-TNFi subgroup, corticosteroid-free remission rates were [redacted] and [redacted] for tofacitinib 10 mg and 5 mg, respectively, compared with [redacted] for placebo, with corresponding absolute differences of [redacted] ($p = 0.0090$) and [redacted] ($p = 0.1032$).

Figure 27 Proportion of patients in remission at baseline who had sustained corticosteroid-free remission in OCTAVE Sustain according to prior TNFi treatment (FAS, NRI, central and local reads)



p < 0.05 for tofacitinib versus placebo.
Abbreviations: FAS, full analysis set; NRI, non-responder imputation.
Source: OCTAVE Sustain CSR (102).
Data in this figure are [redacted].

B.2.8 Meta-analysis

No pairwise meta-analysis was conducted. Head-to-head evidence is not available comparing tofacitinib with all of the comparators in the assessment scope; therefore, an NMA was conducted to estimate the relative efficacy of all relevant therapies (see section B.2.9).

B.2.9 Indirect and mixed treatment comparisons

Fill details of the methodology for the NMA are included in Appendix D.

B.2.9.1 Evidence network for Network Meta-Analysis (NMA)

Head-to-head RCTs between all comparators specified in the NICE scope have not been conducted; therefore, an NMA was undertaken to estimate the relative efficacy and safety between these treatments. NMA can provide comparative measures of effect for all relevant comparators in the absence of direct evidence and is most suitable when there are multi-arm trials included within networks. Use of an NMA in preference to pairwise meta-analysis allowed for the inclusion of all available and relevant evidence and allowed for more precise treatment effects to be calculated. The results from the NMA feed into the economic model described in section B.3, evaluating the cost effectiveness of tofacitinib against relevant comparators. This approach has been used in previous NICE STA submissions for biologics in ulcerative colitis (65, 66).

The primary goal of treatment for ulcerative colitis is to induce and maintain remission (section B.1.3.3): rates of clinical response and clinical remission are the most consistently reported outcomes across all studies, and are the most relevant efficacy parameter in ulcerative colitis to allow comparative analysis, in line with previous NICE technology appraisal and the key efficacy parameter in the cost-effectiveness model (see section B.3).

These data were synthesised using a multinomial model with probit link. For this, it was assumed that the numbers of patients who were reported in the trial publications as being in clinical response also included those patients who were in clinical remission (see section D.1.2.3.1). The proportion achieving mucosal healing was also well reported across the included RCTs and was deemed feasible for comparison using a binomial model with logit link, though this endpoint was not used in the cost-effectiveness model. Safety outcomes, including discontinuations due to adverse events, serious adverse events and serious infections, were also meta-analysed using binomial models with logit links and are presented in section B.2.10.8.

Full details of the methodology for the NMA are presented in Appendix D along with the SLR that was used to identify all studies that may have been relevant for indirect comparison with tofacitinib.

B.2.9.1.1 Selection of evidence contributing to the NMA

For RCTs to be eligible for inclusion in the NMA, they were required to have information about at least one of the following outcomes for either an induction (6–8 weeks) or maintenance (approximately 1 year) time point:

- Clinical response and/or clinical remission (induction and/or maintenance)
- Mucosal healing (induction and/or maintenance)
- Safety (induction; see section B.2.10.8)
    - Discontinuation due to adverse events (AEs)
    - Serious AEs
    - Serious infection

EMA-licensed doses of therapies specified in the scope were included. Where the drug license allows for dose increases during the maintenance phase, both the recommended doses and higher dose were included where they had been assessed in the clinical trials. Different doses and/or dosing regimens were treated as unique comparators.

The studies used in the base-case NMA are summarised in Table 22 and described in detail in Appendix D. All studies were connected to the network through a common direct comparison with placebo. All studies, except for one (99), were conducted in patients with moderately to severely active ulcerative colitis who had an inadequate response to or had failed to tolerate one or more of the following conventional therapies: oral or intravenous corticosteroids, azathioprine, and/or 6-mercaptopurine. Six studies also included patients who had an inadequate response or intolerance to prior TNFi therapies (83, 98, 99, 109). Thirteen studies reported data at the end of a short-term induction period, the length of which varied by treatment (tofacitinib and infliximab, 8 weeks; adalimumab, golimumab and vedolizumab, 6 weeks). Seven studies reported data at the end of a one-year maintenance period, though differences in maintenance study design are a major source of heterogeneity in the analysis (see section B.2.9.1.2).

The base case included all relevant trials, regardless of the country in which they were undertaken. In a sensitivity analysis, studies undertaken in Asia, which included predominantly Asian patients were excluded (110-114).

The base-case NMAs of clinical response, clinical remission and mucosal healing use locally read endoscopic outcomes from the OCTAVE trials as these are both more comparable to outcomes from other RCTs and likely to reflect use of tofacitinib in clinical practice (section B.2.3.1.2.4). Centrally read outcomes were used in a sensitivity analysis.

In order to reduce heterogeneity and increase the comparability of the dataset, the base case comprised separate analyses for patients with and without prior exposure to TNFi therapy. The TNFi-naïve subgroup analysis utilised data from trials in which all patients were TNFi-naïve and TNFi-naïve subgroup data from trials with a mixed population. The evidence for patients with prior TNFi exposure was more mixed, in that ULTRA-2 (109) reported outcomes for patients with any prior TNFi exposure whereas GEMINI 1 (83) reported outcomes only for patients with a prior TNFi failure. In the base case, the GEMINI 1 TNFi failure subgroup was synthesised with the TNFi-exposed subgroup from ULTRA 2 and the tofacitinib trials. In a sensitivity analysis, only TNFi failure subgroup data from the tofacitinib trials and GEMINI 1 were included and compared.

The decision to approach the analysis using subgroup analysis was informed by a number of factors, which are described in further detail in section B.2.9.4.

Table 22 Summary of the trials used to carry out the NMA

Trial / Comparator	Placebo	Tofacitinib (10 mg)	Tofacitinib (5 mg)	Adalimumab (160/80/40 mg)	Infliximab (5 mg/kg)	Golimumab (200/100 mg)	Golimumab (100 mg)	Golimumab (50 mg)	Vedolizumab Q4W	Vedolizumab Q8W	TNFi-exposure subgroups			Efficacy outcomes		Endo-scopy subscore		Predominantly Asian population
											Naïve	Exposed	Failed	Clinical resp/rem	Mucosal healing	Local	Central	
Induction phase																		
OCTAVE Induction 1 (98)	✓	✓									✓	✓	✓	✓	✓	✓	✓	
OCTAVE Induction 2 (98)	✓	✓									✓	✓	✓	✓	✓	✓	✓	
Sandborn 2012 (99)	✓	✓									✓	✓	✓	✓		✓		
ULTRA 1 (73)	✓			✓							✓			✓	✓	✓		
ULTRA 2 (109)	✓			✓							✓	✓		✓	✓	✓		
Suzuki 2014 (114)	✓			✓							✓			✓	✓	✓		✓
Mshimesh 2017 (113)				✓	✓						✓			✓	✓	✓		✓
ACT 1 (74)	✓				✓						✓			✓	✓	✓		
ACT 2 (74)	✓				✓						✓			✓	✓	✓		
Jiang 2015 (111)	✓				✓						✓			✓	✓	✓		✓
Kobayashi 2015 (112)	✓				✓						✓			✓	✓	✓		✓

Trial / Comparator	Placebo	Tofacitinib (10 mg)	Tofacitinib (5 mg)	Adalimumab (160/80/40 mg)	Infliximab (5 mg/kg)	Golimumab (200/100 mg)	Golimumab (100 mg)	Golimumab (50 mg)	Vedolizumab Q4W	Vedolizumab Q8W	TNFi-exposure subgroups			Efficacy outcomes		Endo-scopy subscore		Predominantly Asian population
											Naïve	Exposed	Failed	Clinical resp/rem	Mucosal healing	Local	Central	
PURSUIT-SC (75)	✓					✓					✓			✓	✓	✓		
GEMINI 1 (83)	✓								✓		✓		✓	✓	✓	✓		
Maintenance phase																		
Treat-through trial design																		
ULTRA 2 (109)	✓			✓							✓	✓		✓	✓	✓		
Suzuki 2014 (114)	✓			✓							✓			✓	✓	✓		✓
ACT 1 (74)	✓				✓						✓			✓	✓	✓		
Re-randomised responder trial design																		
OCTAVE Sustain (98)	✓	✓	✓								✓	✓	✓	✓	✓	✓	✓	
PURSUIT-M (115)	✓						✓	✓			✓			✓	✓	✓		
PURSUIT-J (110)	✓						✓				✓			✓	✓	✓		✓
GEMINI 1 (83)	✓								✓	✓	✓		✓	✓	✓	✓		

Abbreviations: TNF, tumour necrosis factor.

B.2.9.1.2 Impact of trial design on assessment of maintenance phase outcomes

The seven included studies presenting maintenance phase outcomes are diverse in terms of their study design. Broadly speaking, there are two study design types: treat-through trials and re-randomised responder trials. Trials with a treat-through design include ACT 1, ULTRA 2 and Suzuki 2014 (74, 109, 114). In these trials, patients are randomised at baseline and outcomes are measured at the end of an induction phase (6-8 weeks) and at the end of a maintenance phase (52–54 weeks).

Re-randomised responder trials, on the other hand, measured the outcomes at the end of a maintenance phase strictly among patients who achieved clinical response during induction. Induction phase clinical responders are re-randomised to placebo or to a maintenance dose of the intervention of interest and outcomes are measured at or around 1 year. Four included maintenance studies follow this design:

- **OCTAVE Sustain (98):** Patients responding to placebo (PBO) or tofacitinib (TOF) in OCTAVE Induction 1 or 2 were re-randomised to PBO or TOF (5 mg or 10 mg) for a further 52 weeks (total duration of induction and maintenance phases = 60 weeks)
- **GEMINI 1 (83):** Patients responding to vedolizumab (VED) (double-blind or open-label) were re-randomised to PBO or VED (Q8W or Q4W) for a further 46 weeks (total duration of induction and maintenance phases = 52 weeks)

- **PURSUIT-M (115):** Patients responding to golimumab (GOL) in PURSUIT-SC or PURSUIT-IV were re-randomised to PBO or GOL (50 or 100 mg) for a further 52 weeks (total duration of induction and maintenance phases = 60 weeks)
- **PURSUIT-J (110):** Patients responding to GOL in a 6-week open-label induction phase were randomised to PBO or GOL (100 mg) for a further 52 weeks (total duration of induction and maintenance phases = 60 weeks)

Simply combining the reported maintenance phase outcomes from these alternative trial design types would be inappropriate as it would violate the similarity and homogeneity assumptions necessary for network meta-analysis. Specifically, the populations allowed to enter the maintenance phases are different and could significantly bias estimates of relative efficacy. The placebo arms also lack comparability because some of the patients who receive placebo in the maintenance phase of re-randomised responder trials received active treatment during induction.

In order to make a valid comparison across these different trial types, the data from one trial type would need to be imputed to better match the other trial type. Two methods were considered, one which converted re-randomised responder trial data to better match a treat-through design and another which converted the treat-through trial data to better match the re-randomised responder trial design. The former method is described by Thorlund *et al.* (116) and a similar approach to the latter method was used in the NMA submitted by Takeda in TA342 (117).

Thorlund and colleagues' method requires significant imputation of missing data from the four re-randomised responder trials. The outcomes would also not have been as readily useable for the cost-effectiveness model, based on the assumptions applied for induction phase non-responders. Converting outcomes from the three treat-through trials into comparable outcomes of the four re-randomised responder trials was considered more robust, requiring less manipulation of observed data and less imputation of missing data. This was assumed to be more reflective of the way the drugs would be used in clinical practice and would better inform the economic analysis.

For the analysis, the observed data from the re-randomised responder trials (OCTAVE Sustain, GEMINI 1, PURSUIT-M and PURSUIT-J) were taken "as is" from the studies. The observed data from the treat-through trials (ACT 1, ULTRA 2 and Suzuki 2014) were adjusted, based on the assumption that the number of responders at the end of induction is a proxy for the total number of patients entering maintenance. Clinical response from the treat-through trials was based on the proportion achieving *sustained* clinical response, as this mitigates the risk of counting maintenance phase responders who were induction phase non-responders. Imputed inputs to the NMA of maintenance phase outcomes are further described in Appendix D.

Although it is known that placebo response and remission rates in ulcerative colitis clinical trials are greatly affected by the time the trial was conducted, favouring older trials over newer trials, this analysis was not able to adjust for the differences in placebo response rates across the included trials (20). It is therefore likely that the estimates of effect for

tofacitinib relative to other therapies may be underestimated; for example, for infliximab Jairath *et al*. (20) demonstrated lower placebo rates for clinical response and clinical remission in the pivotal studies compared with other biologics which were investigated in subsequent years..

B.2.9.2 Base-case NMA

This section presents the NMA for clinical response and clinical remission. The results for mucosal healing can be found in Appendix D. The results for adverse events can be found in section B.2.10.8.

B.2.9.2.1 Clinical response and clinical remission

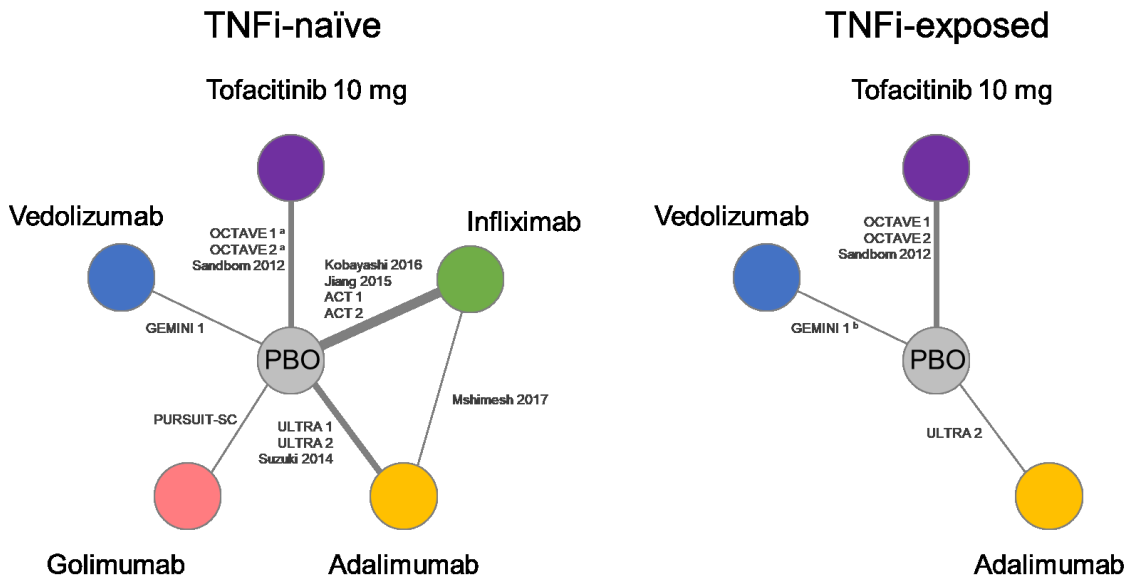
B.2.9.2.1.1 Evidence networks and model choice

Induction phase (8 weeks)

An NMA was used to compare the effects of TOF, VED, adalimumab (ADA), GOL and infliximab (INF) relative to PBO on clinical response and clinical remission in the induction phase. Data were available from 13 studies comparing two treatments. Figure 28 presents the network of evidence for the base-case induction phase NMA for patients naïve to TNFi therapy and for patients with prior TNFi exposure.

For the TNFi-naïve NMA, the fixed effect and random effects models were comparable, both in terms of their results and their fit (see Table 23). The model fit diagnostics were slightly better for the random effects model; thus, it was preferred. For the TNFi-exposed subgroup, the fixed effect model was preferred, as both the fixed and random effect models were comparable in terms of results and goodness of fit.

Figure 28 Base-case network of evidence for induction phase clinical response and clinical remission by TNFi-exposure subgroup



[a] Local read. [b] TNFi failures. Comparator doses: adalimumab 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6; golimumab 200 mg at week 0, 100 mg at week 2; infliximab 10 mg/kg; vedolizumab 300 mg at weeks 0 and 2. Abbreviations: PBO, placebo

Table 23 Model fit statistics for the induction phase NMA of clinical response and clinical remission (base case, multinomial probit)

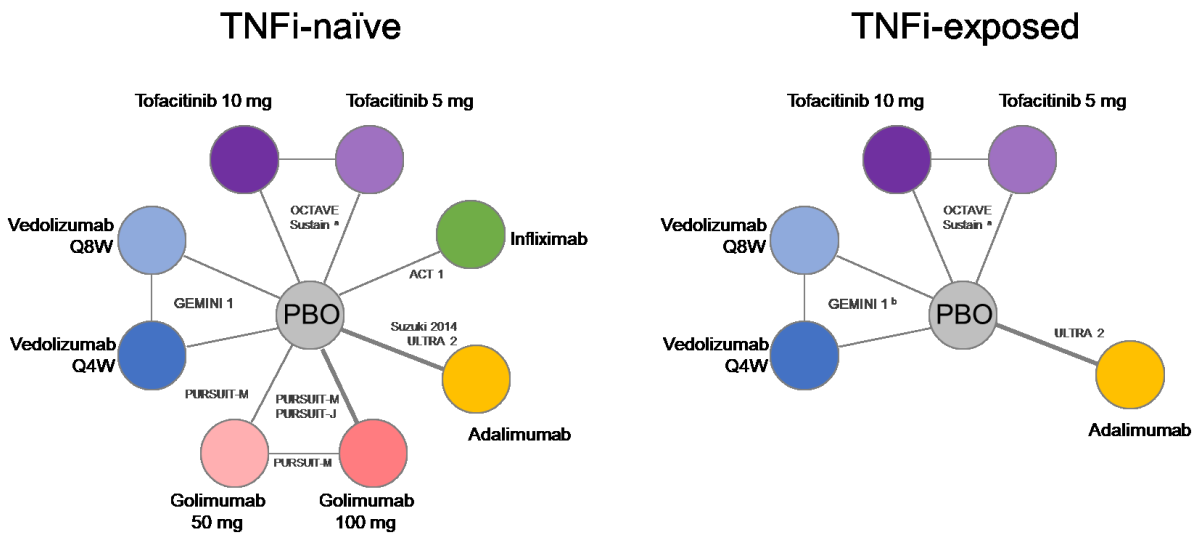
Phase	TNFi-exposure subgroup	Model type	Number of data points	Total residual deviance	DIC
Induction	TNFi-naïve	FE	[redacted]	[redacted]	[redacted]
		RE	[redacted]	[redacted]	[redacted]
	TNFi-exposed	**FE**	[redacted]	[redacted]	[redacted]
		RE	[redacted]	[redacted]	[redacted]

Abbreviations: DIC, Deviance Information Criterion; FE, fixed effects; RE, random effects; TNFi, tumour necrosis factor inhibitor. Bold text indicates preferred model.

Maintenance phase (8–52 weeks)

An NMA was used to compare the effects of TOF, VED, ADA, GOL and INF relative to PBO on clinical response and clinical remission in the maintenance phase. Data were available from seven studies comparing two treatments. Figure 29 presents the network of evidence for the base-case maintenance phase NMA for patients naïve to TNFi therapy and for patients with prior TNFi exposure.

Figure 29 Base-case network of evidence for maintenance phase clinical response and clinical remission by TNFi-exposure subgroup



[a] Local read. [b] TNFi failures. Adalimumab dose: 40 mg Q2W. Infliximab dose 5 mg/kg.
Abbreviations: PBO, placebo; Q4W, every 4 weeks; G8W, every 8 weeks.

For the TNFi-naïve NMA, the fixed effect and random effects models were comparable in terms of their fit (see Table 24). Though the model fit statistics indicate that the random effects model may better represent the data, the results were implausibly imprecise. Mean and median point estimates were similar across the two models, but by assuming random effects, no treatment was predicted to be significantly better than placebo. The results were based on a network that included only one trial for each of TOF, INF and VED and two trials for each of ADA and GOL; therefore, the evidence to support an assumption of random effects was limited. Despite its higher total residual deviance and DIC, the fixed effect model

was therefore preferred. For the TNFi-exposed NMA, both fixed and random effects models were attempted, but due to a paucity of data, only the fixed effect model could be run.

Table 24 Model fit statistics for the maintenance phase NMA of clinical response and clinical remission (base case, multinomial probit)

Phase	TNFi-exposure subgroup	Model type	Number of data points	Total residual deviance	DIC
Induction	TNFi-naïve	**FE**	[redacted]	[redacted]	[redacted]
		RE	[redacted]	[redacted]	[redacted]
	TNFi-exposed	**FE**	[redacted]	[redacted]	[redacted]
		RE	[redacted]		

Abbreviations: DIC, Deviance Information Criterion; FE, fixed effects; RE, random effects; TNFi, tumour necrosis factor inhibitor. Bold text indicates preferred model.

B.2.9.2.1.2 NMA results

Table 25 and Table 26 present the effects of each treatment relative to PBO on the probit scale as well as the odds ratios for clinical response and remission on the natural scale for the base case induction and maintenance phases, respectively. Odds ratios for tofacitinib compared with each other therapy are also presented along with the probabilities of achieving clinical response or clinical remission by the end of the induction phase and maintenance phase.

Induction

In the TNFi-naïve analysis, all treatments were associated with statistically significant beneficial treatment effects relative to PBO. There was a non-significant trend to indicate TOF is more efficacious than ADA and GOL and less efficacious than INF and VED.

In the TNFi-exposed analysis, TOF and VED were associated with statistically significant beneficial treatment effects relative to PBO. The greatest effect was associated with TOF, which was statistically significantly more efficacious than ADA. There was a non-significant trend to indicate TOF is more efficacious than VED in this population with prior TNFi exposure.

Maintenance

In the TNFi-naïve analysis, all treatments were associated with statistically significant beneficial treatment effects relative to PBO. The greatest effect was associated with TOF 10 mg followed by TOF 5 mg, which were statistically significantly more efficacious than ADA and GOL 50 mg. There was a non-significant trend to indicate TOF, at either a 5 mg or 10 mg dose, was more efficacious than maintenance therapy with INF, GOL 100 mg and VED (administered every 4 or 8 weeks).

In the TNFi-exposed analysis, TOF and VED were associated with statistically significant beneficial treatment effects relative to PBO. [redacted] Treatment with ADA was not found to be significantly better than PBO.

Taken together, the results for both TNFi subgroups in the induction and maintenance base-case analyses suggest that, based on the available data, tofacitinib is a very efficacious treatment within ulcerative colitis when compared with biologic therapies. This was confirmed a scenario analysis comparing tofacitinib to vedolizumab in the trial ITT populations (section B.2.9.3).

B.2.9.3 Sensitivity analyses to address uncertainties in the NMA

B.2.9.3.1 Sensitivity analyses conducted

Sensitivity analyses were performed to test alternatives to the base-case population inclusion criteria and assumptions (see section B.2.9.1):

- **Exclusion of Asian studies**: Studies undertaken in Asia and which included majority Asian patients were excluded. This sensitivity analysis is aligned with the base-case assumptions made in the NMA supporting TA329 (118).
- **Centrally read endoscopic subscores**: In this sensitivity analysis clinical response, clinical remission and mucosal healing rates were based on centrally read endoscopic subscores from the OCTAVE trials.
- **TNF-failure subgroup**: In this sensitivity analysis, only data for patients with prior TNFi failure were used, in contrast to the base case which included data from patients with prior TNFi exposure in ULTRA 2 and the OCTAVE trials.
- **Overall ITT analysis**: In this scenario analysis, clinical response and clinical remission outcomes from the overall ITT populations of included trials were synthesised, disregarding potential differences in treatment effect by prior TNFi-exposure. This scenario is described in detail in Appendix D.1.3.5.1.2.

B.2.9.3.2 Sensitivity analysis results

Full details and results of the sensitivity analyses performed are presented in Appendix D. The base case NMA results were relatively robust to changes in the choice of endoscopic subscore reading, to the exclusion of Asian studies and to the use of TNFi-failure subgroup or ITT data. For ease of comparison, the results of the base case and sensitivity analyses are presented in terms of each drug compared with placebo for the TNFi-naïve subgroup in Table 25 and for the TNFi-exposed subgroup in Table 26.

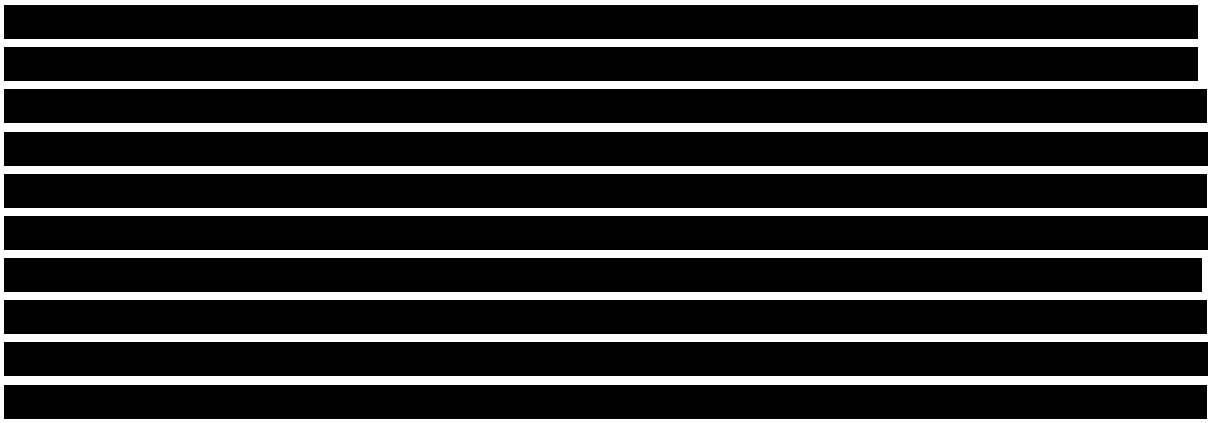


Table 25 Induction phase base-case NMA results – comparative effects and probabilities of achieving response and remission

Comparator	Comparator vs PBO			TOF vs comparator				
	Treatment effect, median (95% CrI)	Odds ratio, median (95%CrI)		Odds ratio, median (95%CrI)		Absolute probability		SUCRA [a]
	Probit scale	Clinical response	Clinical remission	Clinical response	Clinical remission	Clinical response	Clinical remission	
TNFi-naïve subgroup								
PBO								
TOF 10 mg								
INF 10 mg/kg								
ADA 160/80/40 mg [b]								
GOL 200/100 mg [c]								
VED 300 mg [d]								
TNFi-exposed subgroup								
PBO								
TOF 10 mg								
ADA 160/80/40 mg [b]								
VED 300 mg [d]								

[a] based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2.
Abbreviations: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; SUCRA, surface under cumulative ranking curve; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab.

Table 26 Maintenance phase base-case NMA results – comparative effects and probabilities of achieving response and remission

Comparator	Comparator vs PBO			TOF 5 mg vs comparator				
	Treatment effect, median (95% CrI)	Odds ratio, median (95%CrI)		Odds ratio, median (95%CrI)		Absolute probability		SUCRA [a]
	Probit scale	Clinical response	Clinical remission	Clinical response	Clinical remission	Clinical response	Clinical remission	
TNFi-naïve subgroup								
PBO				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF 5 mg	[redacted]	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
INF 5 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 40 mg Q2W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 50 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 100 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg Q8W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg Q4W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed subgroup								
PBO				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF 5 mg	[redacted]	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 40 mg Q2W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg Q8W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg Q4W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; SUCRA, surface under cumulative ranking curve; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab.

[a] based on treatment effect on probit scale

Table 27 Summary results of sensitivity analyses on clinical response and clinical remission for TNFi-naïve subgroup

Comparator	Treatment effect vs placebo, median (95% CrI), probit scale		
	Base-case NMA	Sensitivity analysis NMA using centrally read endoscopic subscores	Sensitivity analysis NMA excluding Asian studies
Induction phase			
TOF 10 mg	[redacted]	[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [a]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [b]	[redacted]	[redacted]	[redacted]
VED 300 mg [c]	[redacted]	[redacted]	[redacted]
Maintenance phase			
TOF 5 mg	[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]
INF 5 mg/kg	[redacted]	[redacted]	[redacted]
ADA 40 mg Q2W	[redacted]	[redacted]	[redacted]
GOL 50 mg	[redacted]	[redacted]	[redacted]
GOL 100 mg	[redacted]	[redacted]	[redacted]
VED 300 mg Q8W	[redacted]	[redacted]	[redacted]
VED 300 mg Q4W	[redacted]	[redacted]	[redacted]

[a] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [b] 200 mg at week 0, 100 mg at week 2. [c] At weeks 0 and 2.
Abbreviations: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab

Table 28 Summary results of sensitivity analyses on clinical response and clinical remission for TNFi-exposed subgroup

Comparator	Treatment effect vs placebo, median (95% CrI), probit scale		
	Base-case NMA	Sensitivity analysis NMA using centrally read endoscopic subscores	Sensitivity analysis NMA for TNFi-failure subgroup
Induction			
TOF 10 mg	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [a]	[redacted]	[redacted]	[redacted]
VED 300 mg [b]	[redacted]	[redacted]	[redacted]
Maintenance			
TOF 5 mg	[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	
ADA 40 mg Q2W	[redacted]	[redacted]	[redacted]
VED 300 mg Q8W	[redacted]	[redacted]	[redacted]
VED 300 mg Q4W	[redacted]	[redacted]	

[a] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [b] At weeks 0 and 2.
Abbreviations: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab.

B.2.9.4 Statistical assessment of heterogeneity

Careful consideration was given to potential sources of heterogeneity, including study design, interventions, outcome definitions and baseline characteristics (weight, disease severity, duration of disease, prior treatments, concomitant treatments). Analysis was only undertaken where it was judged that these factors were sufficiently similar across the network. Where enough data were available, distributions of these characteristics were compared across studies and treatment comparisons. This ensured that differences between the trials and comparisons were kept to a minimum. One patient characteristic was notably different across the evidence network: prior TNFi-exposure. Some studies included only TNFi-naïve patients and others included patients with and without prior TNFi exposure.

The decision to approach the NMAs using subgroup analysis was informed by a number of factors. First, subgroup analyses from OCTAVE Induction 1, OCTAVE Induction 2 and OCTAVE Sustain indicated that there may be an interaction between prior TNFi exposure and treatment effect, though this varied across outcomes in terms of level of significance and direction (see section B.2.7.2).

Second, previous technology appraisals have considered TNFi-naïve and TNFi-exposed populations separately. Archer *et al.* (118), in the NMA underpinning TA329, used data from the TNFi-naïve population rather than the ITT population in ULTRA 2 (109) "in order to increase comparability of the dataset" given that all other studies included only patients who were TNFi-naïve (73-75). The ITT population from ULTRA 2 was only included in a sensitivity analysis. The NMA presented by Takeda for TA342 was performed in the overall ITT population as well as by two TNFi-exposure subgroups (naïve and failures). In interpreting the results, the appraisal committee noted that because the NMA for the whole population would include a mixture of patients with and without prior TNFi exposure, and that these differences in patient characteristics may affect the results, the NMA in the overall ITT population would be subject to considerable uncertainty (66).

Based on these factors, we performed a single integrated induction phase NMA of clinical response and clinical remission with a shared between−trial heterogeneity parameter and an interaction term for prior TNFi exposure introduced in the treatment effect. The meta-regression followed methods recommended in NICE Decision Support Unit Technical Support Document (TSD) 3 (119). Our hypothesis was that the size of the treatment effect is not different in patients with and without prior TNFi exposure. The results of this model, using sub−group effects, was that the difference between populations is statistically significant. The covariate in the fixed effect model was −0.367 (95% CrI: −0.485, −0.246) and in the random effects model was −0.365 (95% CrI: −0.547, −0.169).

Though results for each subgroup could have been generated directly from the single, integrated NMA, we noted a number of limitations in that approach:

- Results for infliximab and golimumab in the TNFi-exposed subgroup were predicted from the data on other drugs in TNFi-exposed patients and from the data on infliximab and golimumab in TNFi-naïve patients. In the absence of any trial-based

observations for infliximab and golimumab in TNFi-exposed patients, there was no way to externally validate the NMA outputs for these drugs.

- Underpinning the analysis is the assumption that the placebo effect is the same across subgroups, and that the interaction term adjusted the treatment effects of all comparators by the same fixed amount. It is unclear whether this is supported by the evidence or clinical practice.

Guided by the evidence of a significant subgroup interaction, but considering the limitations of the single, integrated NMA, separate NMAs for each TNFi-subgroup were undertaken. This approach has been discussed with clinical experts, who indicated that addressing TNFi-naïve and TNFi-exposed subgroups separately is consistent with clinical practice.

For completeness, an analysis was performed on the overall intention-to-treat populations of each RCT. Potential differences in the treatment effects between patients with and without prior TNFi exposure were thus disregarded in this scenario. One induction phase analysis was performed, utilising all ITT evidence; one maintenance phase analysis was performed, utilising all ITT evidence from the re-randomised responder trials only. These results are presented in Appendix D and used to inform a scenario analysis in the economic evaluation (B.3.8.4).

B.2.9.5 Overview of NMA results

The NMA results have shown tofacitinib to be an efficacious induction and maintenance treatment in both patients with and without prior TNFi exposure. Among TNFi-naïve patients, tofacitinib is expected to generate a greater proportion of patients with clinical response, clinical remission and mucosal healing at the end of both induction and maintenance treatment than adalimumab and golimumab. Results for these comparisons were statistically significant in the maintenance phase, but not in the induction phase. Compared with infliximab and vedolizumab, tofacitinib showed comparative efficacy in induction and maintenance treatment across all outcomes analysed.

Among patients with prior TNFi exposure, tofacitinib is expected to be more efficacious than adalimumab in both induction and maintenance phases. Compared with vedolizumab, a non-significant trend suggests that treatment with tofacitinib will produce a higher proportion of patients with clinical response, clinical remission and mucosal healing during induction, and has comparable efficacy as maintenance therapy. Due to a lack of RCT data in the TNFi-exposed population, no comparisons could be made between tofacitinib and infliximab or golimumab.

These conclusions were robust to sensitivity analysis performed, including the use of more conservative clinical response and remission outcomes based on centrally read endoscopic subscores, the exclusion of studies with predominantly Asian patients and the substitution of subgroup data from a population with prior TNFi failure.

The additional analysis of comparing the intention-to-treat (ITT) population of tofacitinib and vedolizumab to conventional therapy confirmed the sub-group analysis findings, and also

demonstrated numerically better efficacy for tofacitinib compared to vedolizumab in both induction and maintenance treatment.

Taken together, the results of both TNFi subgroups and ITT in the induction and maintenance analyses suggest that, based on the available data, tofacitinib is a very efficacious treatment for moderately to severely active ulcerative colitis when compared with biologic thereapies.

B.2.10 Adverse reactions

Summary

- There is a substantial tofacitinib safety database incorporating its use in patients with ulcerative colitis (the OCTAVE clinical programme) as well as in other indications:
    - The OCTAVE programme comprises one Phase II and three Phase III trials, plus an ongoing long-term extension study. Tofacitinib has been evaluated in 1157 patients with moderate to severe ulcerative colitis, with a total exposure of 1986 patient-years and up to 4.4 years of treatment.
    - The tofacitinib clinical development programme for rheumatoid arthritis includes a total of 20 phase I, II and III clinical trials of up to 24 months' duration; two long-term extension studies; and two ongoing phase 3b/4 trials. It has gathered data from 7,065 patients with 22,875 patient-years exposure to tofacitinib, including patients with over 9 years on treatment.
- The safety profile of tofacitinib in ulcerative colitis is consistent with that seen in the rheumatoid arthritis programme
- The rates of adverse events (AEs) in the OCTAVE programme were similar across all treatment groups (between tofacitinib 5 mg and 10 mg groups and also compared to placebo) during both induction and maintenance therapy
- Common AEs were generally mild and manageable, and did not require treatment interruption or withdrawal.
- Serious adverse event (SAE) rates were not significantly different between treatment groups in the induction and maintenance studies; with event rates being numerically higher for placebo compared to both tofacitinib groups
- The rates of events leading to discontinuation were low, with worsening of ulcerative colitis the most common reason.
- AEs of special interest:
    - Infections of any severity were more common in the tofacitinib groups than in the placebo group, with most infections being mild or moderate in severity. The rate of serious infections was higher with tofacitinib than placebo in the induction studies, whereas in the maintenance study it was similar across all treatment groups.

- Most herpes zoster (HZ) infection events were non-serious, cutaneous, limited to one to two dermatomes, were not associated with post-herpetic neuralgia and did not lead to discontinuation. The risk of HZ was dose dependent but did not increase with longer treatment duration.
- Malignancy:
  - Non-melanoma skin cancer occurred in 13 patients treated with tofacitinib across all the studies.12 of these had prior thiopurine treatment, and 10 had previous TNF inhibitor treatment failure.
  - 15 patients treated with tofacitinib had cancer other than non-melanoma skin cancer, of which 14 had been previously treated with thiopurines and/or TNF inhibitor.
- Lipids and cardiovascular safety: tofacitinib treatment was associated with increases in serum lipid levels that were reversible on stopping treatment, however, LDL:HDL and TC:HDL ratios were unaffected. Major adverse cardiovascular events were infrequent.

- There were five deaths across all OCTAVE studies, one of which was assessed as related to the study drug.
- The comparative safety NMA demonstrates no significant differences in the safety profile compared to current bioliogic treatments in ulcerative colitis.

Safety results from the OCTAVE studies are reported in this section, with additional details provided in Appendix F.

In addition, tofacitinib Phase II study (80) safety results and OCTAVE Open trial treatment-emergent adverse events are summarised in Appendix F, Table 166 and Table 167, respectively, and were consistent with the Phase III trial results.

The tofacitinib 15 mg twice daily dose group was removed from the original protocol; hence, patients who were assigned to the tofacitinib 15 mg twice daily dose group were not included in the safety analysis sets. For completeness, safety data for the OCTAVE Induction 1 and 2 studies from these patients are summarised descriptively in Appendix F, Table 155.

B.2.10.1 Exposure data

The safety analysis set (SAS) included all patients who underwent randomisation and received at least one dose of the assigned treatment.

In both OCTAVE Induction 1 and 2, the planned total double-blind treatment period was 63 days. Most patients in each treatment group received at least 57 days of study drug: [redacted] out of [redacted] ([redacted]) and [redacted] out of [redacted] ([redacted]) in the tofacitinib 10 mg twice daily group in OCTAVE Induction 1 and 2, respectively, received study drug for at least 57 days, compared with [redacted] out of [redacted] ([redacted]) and [redacted] out of [redacted] ([redacted]) in the placebo groups. The median duration of treatment was [redacted] days in both tofacitinib 10 mg and placebo groups (100, 101).

In OCTAVE Sustain, the planned total double-blind treatment period was 371 days. The mean and median durations of treatment were [redacted] and [redacted] days, respectively, for the

tofacitinib 5 mg twice daily group, [redacted] and [redacted] days, respectively, for the tofacitinib 10 mg twice daily group, and [redacted] and [redacted] days, respectively, for the placebo group. (102).

B.2.10.2 Common adverse events

The most common adverse events (AEs) in the OCTAVE studies were worsening ulcerative colitis, nasopharyngitis, arthralgia and headache (Table 29). The frequencies of these events were generally similar across groups, with the exception of worsening ulcerative colitis, which was more frequent in the OCTAVE Sustain placebo group than in the tofacitinib groups, and nasopharyngitis, which in OCTAVE Sustain was more common with tofacitinib than with placebo (98).

Full details of all treatment-emergent adverse events affecting ≥ 2% of patients in any group by system organ class and preferred term are shown in Appendix F, Table 156, Table 157 and Table 158.

B.2.10.3 Serious adverse events

Serious adverse events (SAEs) were defined as any untoward medical occurrence at any dose that:

- Resulted in death;
- Was life-threatening (immediate risk of death);
- Required inpatient hospitalisation or prolongation of existing hospitalisation;
- Resulted in persistent or significant disability/incapacity (substantial disruption of the ability to conduct normal life functions); or
- Resulted in a congenital anomaly/birth defect.

Causality was subject to independent determination.

A full listing of SAEs according to system organ class in the OCTAVE trials is shown in Appendix F, Table 159, Table 160 and Table 161.

In OCTAVE Induction 1, SAEs occurred in 3.4% of patients treated with tofacitinib 10 mg, compared with 4.1% in the placebo group. In OCTAVE Induction 2, the corresponding percentages were 4.2% and 8.0%. In OCTAVE Sustain, SAEs occurred in 5.1% the patients in the tofacitinib 5 mg group, 5.6% in the tofacitinib 10 mg group and 6.6% in the placebo group (98). The most frequent SAE was ulcerative colitis, and most SAEs were related to ulcerative colitis (100-102).

B.2.10.4 Events leading to discontinuation

- Rates of events leading to discontinuation were low, with worsening of ulcerative colitis the most common reason (section B.2.4.3) (98).
- These rates were comparable between placebo and tofacitinib 10 mg in the OCTAVE Induction 1 and 2 studies.
- In OCTAVE Sustain, adverse events leading to discontinuation were more common in the placebo group than in the tofacitinib groups.

Table 29 Summary of adverse events in OCTAVE Induction 1 and 2 and in OCTAVE Sustain (SAS)

Safety event	OCTAVE Induction 1		OCTAVE Induction 2		OCTAVE Sustain		
	Placebo (N = 122)	Tofacitinib 10 mg (N = 476)	Placebo (N = 112)	Tofacitinib 10 mg (N = 429)	Placebo (N = 198)	Tofacitinib 5 mg (N = 198)	Tofacitinib 10 mg (N = 196)
Adverse events, n (%)	73 (59.8)	269 (56.5)	59 (52.7)	232 (54.1)	149 (75.3)	143 (72.2)	156 (79.6)
Serious adverse events, n (%)	5 (4.1)	16 (3.4)	9 (8.0)	18 (4.2)	13 (6.6)	10 (5.1)	11 (5.6)
Most frequent adverse events, n (%) [a]							
Worsening ulcerative colitis	5 (4.1)	11 (2.3)	6 (5.4)	13 (3.0)	71 (35.9)	36 (18.2)	29 (14.8)
Nasopharyngitis	9 (7.4)	34 (7.1)	4 (3.6)	21 (4.9)	11 (5.6)	19 (9.6)	27 (13.8)
Arthralgia	6 (4.9)	14 (2.9)	6 (5.4)	11 (2.6)	19 (9.6)	17 (8.6)	17 (8.7)
Headache	8 (6.6)	37 (7.8)	9 (8.0)	33 (7.7)	12 (6.1)	17 (8.6)	6 (3.1)
Infections, n (%)							
Any infection	19 (15.6)	111 (23.3)	17 (15.2)	78 (18.2)	48 (24.2)	71 (35.9)	78 (39.8)
Serious infection	0	6 (1.3)	0	1 (0.2)	2 (1.0)	2 (1.0)	1 (0.5)
Herpes zoster	1 (0.8)	3 (0.6)	0	2 (0.5)	1 (0.5)	3 (1.5)	10 (5.1)
Adverse events of special interest, n							
Intestinal perforation [b]	0	1	1	0	0	0	0
Cancer other than non-melanoma skin cancer [c]	0	0	0	0	1 [d]	0	0
Non-melanoma skin cancer [c]	0	1	0	1	1	0	3
Cardiovascular events [c]	0	2	0	2	0	1	1
Adverse events leading to discontinuation, n (%) [e]	2 (1.6)	18 (3.8)	8 (7.1)	17 (4.0)	37 (18.7)	18 (9.1)	19 (9.7)
Abnormal laboratory test results, n (%) [f]							
N for laboratory data	122	471	111	424	198	198	195
Total cholesterol >1.3× ULN	11 (9.0)	80 (17.0)	6 (5.4)	73 (17.2)	16 (8.1)	54 (27.3)	44(22.6)
Low-density lipoprotein >1.2× ULN	11 (9.0)	91 (19.3)	12 (10.8)	92 (21.7)	37 (18.7)	62 (31.3)	55 (28.2)
High-density lipoprotein <0.8× LLN	2 (1.6)	6 (1.3)	1 (0.9)	7 (1.7)	12 (6.1)	9 (4.5)	3 (1.5)
Triglycerides >1.3× ULN	1/ (0.8)	15 (3.2)	2 (1.8)	12 (2.8)	7 (3.5)	9 (4.5)	15 (7.7)
Creatine kinase >2× ULN, n/total N (%)	2/122 (1.6)	45/474 (9.5)	10/112 (8.9)	40/425 (9.4)	14/198 (7.1)	37/198 (18.7)	54/195 (27.7)
Addition or increase in dose of lipid lowering agent, n (%)	0	4 (0.8)	1 (0.9)	2 (0.5)	3 (1.5)	2 (1.0)	8 (4.1)

[a] Most frequent adverse events in OCTAVE Sustain. [b] Determined based on MedDRA preferred term. [c] Determined based on external adjudication. [d] invasive ductal breast carcinoma. [e] Including patients who discontinued treatment because of worsening ulcerative colitis. [f] Laboratory data were missing for some patients.
Abbreviations: LLN, lower limit of normal; SAS, safety analysis set; ULN, upper limit of normal.
Source: Sandborn *et al.* 2017 (98).

B.2.10.5 Adverse events of special interest

AEs of special interest in the OCTAVE trials were infections, herpes zoster infections (HZ), malignancies, gastrointestinal perforations and cardiovascular events; these are summarised in Table 29. Full details of all events of special interest are shown in Appendix F, Table 162, Table 163, Table 164 and Table 165.

Infections

In all the OCTAVE trials most infections were mild or moderate in severity, and the most frequently occurring infection across all the studies in the programme was nasopharyngitis.

In OCTAVE Induction 1 and 2, infections of any severity were more common in the tofacitinib 10 mg groups (23.3% and 18.2%, respectively) than in the placebo groups (15.6% and 15.2%). Similarly, in OCTAVE Sustain, infections occurred in 39.8% in the tofacitinib 10 mg group, 35.9% of the patients in the tofacitinib 5 mg group, and 24.2% in the placebo group (Table 29) (98). Serious infections (defined as infections that met SAE reporting criteria) were infrequent in the OCTAVE programme, with no apparent dose dependency in the risk:

- In OCTAVE Induction 1 and 2, serious infections occurred in six patients (1.3%) and one patient (0.2%), respectively, in the tofacitinib 10 mg groups; no patient in the placebo group had a serious infection.
- In OCTAVE Sustain, serious infections occurred in two patients (1.0%) in the tofacitinib 5 mg group, 1 (0.5%) in the tofacitinib 10 mg group, and 2 (1.0%) in the placebo group. This suggests the risk of serious infections did not increase with duration of tofacitinib treatment; based on the rate of serious infections in OCTAVE Sustain, use of induction data in the economic model is a conservative approach, and may bias the analysis against tofacitinib.
- In the overall cohort of patients treated with tofacitinib across all the programme studies (P2, P3 and ongoing LTE study, n = 1157 patients) the incident rate (IR; patients with events per 100 patient-years) of serious infection events was 1.87 (95% CI 1.32, 2.56). This is similar to the rate observed in the ORAL trial programme in rheumatoid arthritis (section B.2.10.7).

There were no cases of adjudicated tuberculosis (TB) with tofacitinib (all doses) across the phase 2, phase 3 and open-label LTE studies (98, 120).

There was no apparent clustering into specific types of serious infection, with only 4 safety events occurring more than once (excluding HZ) (Appendix F, Table 168).

It is also noteworthy that patients who developed a serious infection during the studies as defined in OCTAVE trial protocols were automatically withdrawn from the study, regardless of whether the infection was manageable or not.

Based on the safety data the risks and benefits of treatment should be considered prior to initiating tofacitinib in patients with recurrent infections, with a history of a serious or an

opportunistic infection, who have resided or travelled in areas of endemic mycoses, or who have underlying conditions that may predispose them to infection.

Patients should be closely monitored for the development of signs and symptoms of infection during and after treatment with tofacitinib, and treatment should be interrupted if a patient develops a serious infection, an opportunistic infection, or sepsis.

Herpes Zoster (HZ) infections

Most HZ infection events were non-serious, cutaneous, limited to one to two dermatomes, were not associated with post-herpetic neuralgia and did not lead to discontinuation. The risk of HZ was dose dependent but did not increase with longer treatment duration.

Across all the studies, including OCTAVE Open, 74 out of 1157 patients treated with tofacitinib developed HZ. In the induction studies (phase 2 and phase 3) similar proportions of patients developed HZ with tofacitinib compared to placebo. In OCTAVE Sustain, HZ was more frequent in the tofacitinib 10 mg group (10 patients; 5.1%) than in the tofacitinib 5 mg group (3 patients; 1.5%) or the placebo group (1 patient; 0.5%) (102).

HZ is a safety signal that was identified during the tofacitinib rheumatoid arthritis programme and the management of the risk of serious and important infections is addressed in the current draft SmPC to the EMA, consistent with the current SmPC for rheumatoid arthritis, which includes effective routine risk minimisation measures.

Malignancies

Malignancies occurred infrequently with tofacitinib treatment in the OCTAVE clinical programme. The reporting of malignancies is divided into non-melanoma skin cancer (NMSC) and malignancies excluding NMSC;

- NMSC: In the overall cohort of patients treated with tofacitinib across all the studies (1157 patients), 15 patients developed NMSC. Of these 15 patients, 7 reported a prior history of NMSC, 14 had been exposed to azathioprine or mercaptopurine and 14 had failed treatment with TNFis. It should be noted that patients with IBD may also have an increased incidence of NMSC. This increased risk may be related to the immune dysfunction associated with IBD or the concomitant therapy. An increased risk of NMSC has been associated with past or concurrent use of thiopurines (121, 122).
- Malignancies excluding NMSC: In the overall cohort of patients treated with tofacitinib across all studies, 13 out of 1157 patients developed malignancies, with 3 occurring more than 28 days after the last dose of tofacitinib. All 12 patients had previous treatment with thiopurines; 10 had previous TNFi treatment. The types of malignancies reported in the OCTAVE programme were generally consistent with those reported for IBD (123) and no clustering of malignancies into specific types of cancer was observed.

The management of the potential risk of malignancies and management of the risk of NMSC is addressed in the current draft SmPC to the EMA, and is consistent with the current SmPC for rheumatoid arthritis.

Gastrointestinal perforations

A total of three patients treated with tofacitinib across all the studies had a gastrointestinal (GI) perforation:

- In OCTAVE Induction 1, one patient in the tofacitinib 10 mg group had a serious adverse event of GI perforation and underwent colectomy. Causality was assessed as relating to the study drug and the patient was permanently discontinued.
- In OCTAVE Induction 2, one patient in the placebo group had a serious adverse event of GI perforation.
- In OCTAVE Open, 2 patients had GI perforation; neither were assessed as related to the study drug.

Serum lipids and cardiovascular safety

Tofacitinib treatment was associated with serum increases in total cholesterol (TC), high-density lipoprotein cholesterol (HDL) and low-density lipoprotein cholesterol (LDL) in patients with ulcerative colitis. These plateaued after 4 weeks and reversed when treatment was stopped. LDL:HDL and TC:HDL ratios were unaffected. Major adverse cardiovascular events (MACE) were infrequent, occurring in four patients across the clinical programme. Three of these patients had pre-existing multiple cardiovascular risk factors. The incidence rate of MACE in the tofacitinib trial programme was 0.20/100 PY (Table 30). The data do not suggest an increasing risk of developing MACE with longer duration of tofacitinib treatment. These results are similar to those reported in the tofacitinib rheumatoid arthritis programme and for other UC agents.

The management and monitoring for cardiovascular risk factors forms part of the current draft SmPC to the EMA, and are consistent with the current SmPC for rheumatoid arthritis. This states that assessment of lipid parameters should be performed after 8 weeks following initiation of tofacitinib therapy. Patients should be managed according to clinical guidelines for the management of hyperlipidaemia. Increases in total and LDL cholesterol associated with tofacitinib may be decreased to pre-treatment levels with statin therapy.

B.2.10.6 Deaths

There have been 5 deaths across the OCTAVE programme:

- In OCTAVE Induction 1, one patient treated with tofacitinib 10 mg died from dissecting aortic aneurysm. The event was assessed as not related to the study drug (98).
- There were no deaths in OCTAVE Induction 2 or OCTAVE Sustain (98).
- There were four deaths in OCTAVE Open, all in the 10 mg tofacitinib group. Of these, three deaths occurred > 28 days after the last dose of tofacitinib and were related to malignancies. Tofacitinib was considered to play a contributory role in one event (hepatic angiosarcoma).

B.2.10.7 Safety outcomes with tofacitinib in rheumatoid arthritis ORAL trial data

The safety profile of tofacitinib in ulcerative colitis is consistent that observed in the rheumatoid arthritis programme.

The ORAL tofacitinib clinical development programme for rheumatoid arthritis consists of 20 phase 1, 2 and 3 clinical trials of up to 24 months duration; two long-term extension (LTE) studies (one of which is ongoing) with up to 114 months of observation; and 2 phase 3b/4 trials: ORAL Surveillance and ORAL shift. The safety profile of tofacitinib has been evaluated in 7065 patients with 22,875 patient-years exposure to tofacitinib (phase 1, 2 and 3 and LTE studies).

Table 30 presents the overall safety findings of the OCTAVE trials in ulcerative colitis for the Phase II and III programme, including the OCTAVE Open long-term extension data, cut-off date 29th September 2017, alongside the extensive safety data from the ORAL trials in rheumatoid arthritis, cut-off date 2nd March 2017.

Table 30 Cumulative incidence rate (per 100 patient-years) for death and safety events of special interest comparing the tofacitinib ulcerative colitis and rheumatoid arthritis program

Safety Event	OCTAVE trial programme (Phase II, Phase III and LTE) N= 1157 PY= 1986		ORAL trial programme (Phase II, Phase III and LTE) N=7061 PY= 22,875	
	n (%)	IR (95% CI)	n (%)	IR (95% CI)
Death (all-cause)	5 (0.4%)	0.24 (0.08–0.57)	59 (0.8%)	0.25 (0.19–0.32)
Serious infection	38 (3.3%)	1.87 (1.32–2.56)	576 (8.6%)	2.48 (2.28–2,69)
Opportunistic infections	22 (2.0%)	1.09 (0.69–1.66)	90 (1.3%)	0.39 (0.31–0.47)
Non-herpes zoster OI	4 (0.4%)	0.20 (0.05–0.50)	34 (0.5%)	0.15 (0.10–0.20)
Herpes Zoster infections	74 (6.4%)	3.80 (2.99–4.77)	782 (11.1%)	3.63 (3.38–3.90)
Serious herpes zoster infections	5 (0.4%)	0.24 (0.08–0.57)	57 (0.8%)	0.24 (0.18–0.32)
Malignancy (excl. NMSC)	13 (1.2%)	0.84 (0.45–1.44)	177 (2.5%)	0.76 (0.65–0.88)
NMSC	15 (1.3%)	0.74 (0.42–1.23)	129 (1.8%)	0.56 ((0.46–0.66)
MACE	4 (0.4%)	0.20 (0.05–0.50)	85 (1.3%)	0.38 (0.30–0.47)
GI perforations (all cases)	4 (0.4%)	0.20 (0.05–0.50)	28 (0.4%)	0.12 (0.08–0.17)

Abbreviations: GI, gastrointestinal; IR, incidence rate; LTE, long-term extension; MACE, major cardiovascular events; NMSC, non-melanoma skin cancer; OI, opportunistic infection; PY, patient-year.

Source: Pfizer trial database analysis; UC Cohort data cut-off date 29th September 2017, RA Cohort data cut-off date 2nd March 2017. Source: FDA briefing document, March 2018 (124). Although there are differences in the disease, demographics, concomitant, medications, exposure time and co-morbidities between these indications. the overall safety profile is similar across indications.

Of particular interest, the incidence rate of opportunistic infections in the ulcerative colitis programme was numerically higher than in the rheumatoid arthritis programme; this was primarily attributable to an increased rate of HZ opportunistic infections. The higher incidence rate of HZ opportunistic infection reported in the ulcerative colitis program may have reflected a period effect, with increased vigilance in monitoring, requests for additional information by Pfizer and reporting by study sites as the risk of HZ associated with tofacitinib treatment became better recognised over time during the rheumatoid arthritis trial programme.

B.2.10.8 Comparison of tofacitinib safety outcomes with approved biologics in ulcerative colitis

Pfizer conducted a systematic review to identify all relevant clinical data from the published literature regarding the clinical effectiveness and safety of treatments in ulcerative colitis. Full details of the methodology and a full summary of the included and excluded studies, including the PRISMA flow diagram and reasons for exclusion, are also provided in Appendix D and section B.2.9.

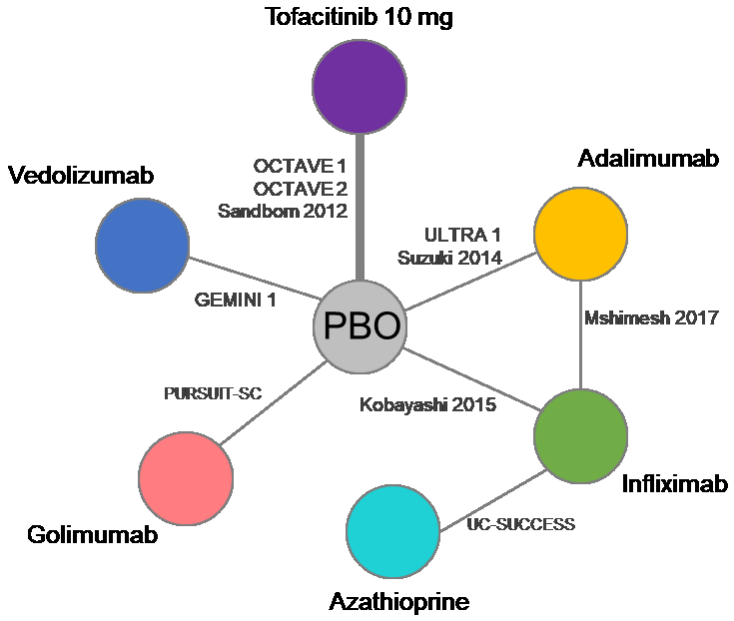
Analyses of safety outcomes from both induction and maintenance phases of studies were assessed for their feasibility and appropriateness. Induction phase safety endpoints were considered similar enough across the studies to allow for synthesis; however maintenance phase endpoints were subject to a number of limitations which could lead to biased estimates of relative safety. Briefly, differences between maintenance phase study designs (see section B.2.9.1.2) meant that a single, coherent comparison between all treatments could not be made. Even among the re-randomised responder trials, there were differences in who was eligible for inclusion (all induction phase responders or only responders to intervention therapies) and the potential for lingering effects of active treatments from induction to impact the assessment of safety outcomes for placebo arms in maintenance.

A set of 3 NMAs was used to compare the safety of tofacitinib (TOF), vedolizumab (VED), adalimumab (ADA), golimumab (GOL) and infliximab (INF), relative to PBO on discontinuation due to AEs, serious AEs (SAEs) and serious infections in the induction phase. Data were available from 10 studies comparing two treatments. Figure 30 presents the network of evidence for the base case for all 3 outcomes.

To maximise statistical power, especially in light of the rarity of analysed safety events, data from all patients were combined into a single analysis based on the assumption that the prior TNFi exposure has no influence on the safety outcomes.

B.2.10.8.1 Evidence networks and model choice

Figure 30 Base-case network of evidence for induction phase safety outcomes (discontinuation due to AEs, serious AEs and serious infections)



Abbreviations: PBO, placebo

For the analysis on discontinuation due to AEs, the fixed effect and random effects models were comparable, both in terms of their results and their fit (see Table 31). Given the similarity across the two models, the fixed effect model was preferred on the basis of lower DIC. For the analysis of serious AEs, the model fit was slightly better for the fixed effects model, thus it was preferred. For the outcome of serious infections, the fixed effect and random effects models were comparable, both in terms of their results and their fit, but the model fit diagnostics were slightly better for the random effects model, thus it was preferred.

Table 31 Model fit statistics for the induction phase NMA of safety outcomes (base case, binomial logit)

Outcome	Model type	Number of data points	Total residual deviance	DIC
Discontinuation due to AEs	**FE**	[redacted]	[redacted]	[redacted]
	RE	[redacted]	[redacted]	[redacted]
Serious AEs	**FE**	[redacted]	[redacted]	[redacted]
	RE	[redacted]	[redacted]	[redacted]
Serious infections	FE	[redacted]	[redacted]	[redacted]
	RE	[redacted]	[redacted]	[redacted]

Abbreviations: DIC, Deviance Information Criterion; FE, fixed effects; RE, random effects

B.2.10.8.2 Safety NMA Results

Table 32, Table 33 and Table 34 present the effects of each treatment relative to PBO on the logit scale as well as the odds ratios for each safety outcome (discontinuation due to AEs, serious AEs, serious infections, respectively) on the natural scale for the base case induction phase. Odds ratios for tofacitinib compared to each other therapy are also

presented along with the probabilities of each event occurring by the end of the induction phase.

Table 32 Induction phase base-case NMA results – comparative effects and probabilities of discontinuing due to AEs

Comparator	Comparator vs PBO		TOF vs comparator	Absolute probability, median (95% CrI)	SUCRA [a]
	Treatment effect (logit scale), median (95% CrI)	Odds ratio, median (95% CrI)	Odds ratio, median (95% CrI)		
PBO			[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]		[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] Based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2. **Abbreviations**: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; SUCRA, surface under cumulative ranking curve; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab.

Table 33 Induction phase base-case NMA results – comparative effects and probabilities of serious AEs

Comparator	Comparator vs PBO		TOF vs comparator	Absolute probability, median (95% CrI)	SUCRA [a]
	Treatment effect (logit scale), median (95% CrI)	Odds ratio, median (95% CrI)	Odds ratio, median (95% CrI)		
PBO			[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]		[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] Based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2. **Abbreviations**: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; SUCRA, surface under cumulative ranking curve; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab.

Table 34 Induction phase base-case NMA results – comparative effects and probabilities of serious infections

Comparator	Comparator vs PBO		TOF vs comparator	Absolute probability, median (95% CrI)	SUCRA [a]
	Treatment effect (logit scale), median (95% CrI)	Odds ratio, median (95% CrI)	Odds ratio, median (95% CrI)		
PBO			[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]		[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] Based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2. **Abbreviations**: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; SUCRA, surface under cumulative ranking curve; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab.

B.2.10.8.3 Summary of safety NMA results

[redacted].

Results of sensitivity analyses of safety outcomes are presented in Appendix D and do not differ substantially from the base case.

B.2.11 Ongoing studies

The OCTAVE Open study (NCT01470612) is ongoing, and additional data may be available within the next 12 months. In addition, preliminary results from a Phase IIIb/IV study of tofacitinib in patients with ulcerative colitis in stable remission (NCT03281304) may be available within the next 12 months.

B.2.12 Innovation

The cost-effectiveness analysis described in section B.3 models the benefits of tofacitinib based on the rates of clinical remission and clinical response in the OCTAVE trials. In addition to the utility gains associated with clinical remission and clinical response, the OCTAVE trials have demonstrated a number of benefits of tofacitinib that may not be included in the incremental cost-effectiveness analysis. Tofacitinib offers a new mechanism of action in ulcerative colitis and is an oral therapy. As a small molecule, tofacitinib is likely to avoid the issues related to immunogenicity seen with biologics, which has clinically important implications as outlined hereafter.

Tofacitinib is the first therapy in its class and offers a new mechanism of action in ulcerative colitis

Tofacitinib is a JAK inhibitor, the first in a new class of treatments that offers a novel mechanism of action for patients with moderately to severely active ulcerative colitis who do not respond adequately to conventional therapies or biologics.

Tofacitinib is an oral therapy that offers patients an alternative to current parenteral treatments

Biologics for moderately to severely active ulcerative colitis are given either as infusions or by subcutaneous injection. Patients with chronic conditions have been shown to prefer oral treatments to injectable or intravenous therapies (125-127). In a Delphi survey of patients with ulcerative colitis (n = 20) and physicians (n = 22), both groups identified oral administration as a highly relevant factor contributing to patient comfort and to medication adherence (128).

Tofacitinib is a small molecule and as such should not be associated with issues relating to immunogenicity

Efficacy for biologics, which are large proteins, is likely to be reduced over time due to anti-drug antibody formation, therefore therapeutic drug monitoring is common in clinical practice, often leading to dose escalations to adjust for the reduced drug trough-levels in order to recapture and maintain response to treatment. As a small molecule, tofacitinib is not likely to have the same issues with immunogenicity as the large proteins (92). It has been shown from pharmacokinetic studies of tofacitinib that plasma levels are similarly stable in patients who have remitted and those who have not achieved remission (97). Therefore, therapeutic drug monitoring is not a requirement for tofacitinib, unlike the biologic agents.

Tofacitinib is a synthetic small molecule given as monotherapy

The development of anti-drug antibodies to current biologic treatment has been mitigated by using biologics in combination with immunomodulatory (IM) agents (for example, infliximab and azathioprine), which may reduce the immunogenicity of TNFis, and therefore result in improved efficacy responses (96). However, potential synergistic efficacy benefits of the biologic+IM combination are also accompanied by an increase in safety events compared to biologics monotherapy.

Tofacitinib offers patients the opportunity to stop treatment and restart with similar efficacy

In the OCTAVE Open study, a significant number of patients who had received tofacitinib 10 mg after receiving placebo reached remission. Tofacitinib may be given to patients after a treatment interruption without the expectation of a reduced response. These long-term benefits of dose flexibility and treatment interruptions of tofacitinib may not be sufficiently captured in the economic analysis, and are likely to further increase cost-effectiveness for tofacitinib due to additional cost savings.

Tofacitinib provides rapid improvements in ulcerative colitis symptoms

Rapid onset of action is important to patients: in a survey of the preferences of 100 Canadian patients with ulcerative colitis, speed of symptom relief was rated as the most important medication attribute (129). In the OCTAVE Induction trials, there was a statistically significant improvement in partial Mayo score with tofacitinib as early as week 2 (see section B.2.6.1.1.3); in addition, changes in EQ-5D scores were significantly better with tofacitinib than with placebo after only 2 weeks (see section B.2.6.1.2). The benefits to patients of these rapid improvements in ulcerative colitis symptoms may not be captured in the QALY calculation.

B.2.13 Interpretation of clinical effectiveness and safety evidence

B.2.13.1 Principal findings from the OCTAVE clinical studies

The efficacy of tofacitinib for the treatment of moderately to severely active ulcerative colitis was demonstrated in three Phase III trials: OCTAVE Induction 1 and 2, and the OCTAVE Sustain maintenance trial. Tofacitinib demonstrated rapid and sustained improvements in ulcerative colitis symptoms, regardless of prior treatment with TNFi therapies.

As an induction therapy, tofacitinib was associated with significantly higher rates of remission, mucosal healing and clinical response than placebo. At 8 weeks, more than twice as many patients achieved remission with tofacitinib than with placebo (see section B.2.6.1.1.1), and approximately 30% of tofacitinib-treated patients had mucosal healing (see section B.2.6.1.1.2). In addition, tofacitinib maintenance therapy was significantly more efficacious than placebo, as demonstrated in OCTAVE Sustain (see section B.2.6.2.1).

Responses to tofacitinib were rapid and sustained. In OCTAVE Induction 1 and 2, statistically significant differences from placebo in partial Mayo score were seen as early as week 2 (see section B.2.6.1.1.4). In OCTAVE Sustain, among patients in remission after 8 weeks of induction therapy, half were in remission at week 52 (tofacitinib 10 mg, 56.4%; tofacitinib 5 mg, 46.2%; placebo, 10.2%; both $p < 0.0001$; see section B.2.6.2.1.1), and more than seven times as many achieved sustained corticosteroid-free remission, compared with placebo (tofacitinib 10 mg, 47.3%; tofacitinib 5 mg, 35.4%; placebo, 5.1%; both $p < 0.0001$; see section B.2.13.1).

In addition, interim data from OCTAVE Open demonstrated that a substantial number of patients with an initial response to tofacitinib induction therapy that were lost after randomisation to placebo in OCTAVE Sustain were able to recapture a response to treatment ([redacted]) of patients in this group were in remission after 12 months of treatment in OCTAVE Open; see section B.2.6.3.4).

Patients treated with tofacitinib in the OCTAVE trials experienced significant improvements in HRQoL compared with placebo, demonstrated across a range of quality of life measures collected in the Induction and Maintenance phase: EQ-5D, IBDQ and SF-36 (see section B.2.6.2.2).

Overall, the clinical outcomes were highly consistent, both among the OCTAVE trials and the Phase II study of tofacitinib versus placebo, providing robust evidence for treating moderately to severely active ulcerative colitis, including across subgroups, such as TNFi-experienced or TNFi-naïve patients (see section B.2.7).

Tofacitinib was well tolerated and the safety profile of tofacitinib is in general similar to that of TNFi (section B.2.10.8) and consistent with that of tofacitinib in rheumatoid arthritis (section B.2.10.7), which has extensive long-term data of 9.5 years, translating into 22,875 patient-years to date. It is acknowledged that tofacitinib appears to increase the risk of herpes zoster infections, although data suggest that the risk does not increase with prolonged tofacitinib exposure. In addition, a potential elevated risk for NMSC was also identified during the

ulcerative colitis trial programme compared to the rheumatoid arthritis programme. However, IBD is associated with an increased underlying risk of developing a series of conditions, including serious infections and malignancies. For example, the elevated risk of serious infections and opportunistic infections in patients with IBD has been attributed to factors such as older age, use of systemic corticosteroids, thiopurines, TNFi agents and immunomodulatory treatment, particularly when these agents are used in combination (130-134). More specifically, patients with IBD have a higher risk of developing herpes zoster infections than healthy individuals or those without IBD. Immunologic dysregulation due to the presence of IBD and immunomodulation produced by IBD therapeutics further increases the risk of shingles in IBD (135, 136). Additionally, the use of TNFi agents, thiopurines, and corticosteroids were identified as independent risk factors for development of herpes zoster infection (137, 138).

Similarly, the increased risk of individual malignancies, such as colorectal cancer, lymphoma, cervical dysplasia, cholangiocarcinoma, has been well documented in IBD (139-144). Patients with IBD may also have an increased incidence of NMSC, which has also been associated with past or concurrent treatment with thiopurines (121, 122).

Nevertheless, the management of the potential risk of serious infections and malignancies is adequately addressed in the current draft SmPC to the EMA, and is consistent with the current SmPC for rheumatoid arthritis , which includes effective routine risk minimisation measures.

B.2.13.2 Strengths and limitation of the clinical evidence base for tofacitinib

The clinical evidence provided by the OCTAVE trials demonstrates the efficacy and safety of tofacitinib in the treatment of moderately to severely active ulcerative colitis. All of the OCTAVE trials met their primary endpoints, and demonstrated rapid and sustained improvements in ulcerative colitis symptoms and HRQoL with tofacitinib compared with placebo.

A strength of the tofacitinib clinical programme is the use of a stringent endpoint as the primary endpoint – remission, defined as a total Mayo score of ≤ 2, with no subscore > 1 and a rectal bleeding subscore of 0, is a stricter endpoint than clinical response and clinical remission, which were used in studies of biological therapies for ulcerative colitis. Because the primary endpoints of the OCTAVE trials are not directly comparable to those in studies of other therapies, clinical response and clinical remission, which were secondary endpoints of the OCTAVE trials, are used in the economic model.

A further strength of the OCTAVE studies is the use of both central and local endoscopy reads – centrally assessed endoscopic subscores ensure consistency of analysis for the trial endpoints, while locally read subscores reflect the likely findings with tofacitinib in clinical practice. In addition, patients were re-randomised between induction and maintenance treatment, as recommended by draft EMA guidance on ulcerative colitis trials (145).

The main efficacy outcome assessed in the OCTAVE studies is remission, a stringent endpoint that requires both symptomatic improvement and endoscopic evidence of mucosal

healing. The secondary endpoint of mucosal healing is also regarded as an important therapeutic endpoint in clinical practice; achieving mucosal healing is associated with sustained clinical remission, a reduced need for corticosteroids and a decreased risk of surgery being required (104). In addition, sustained corticosteroid-free remission, a key secondary endpoint in OCTAVE Sustain, is regarded as an important clinical endpoint. Although corticosteroids may be used for induction of remission, because of their side-effect profile they are not typically used for long-term management of ulcerative colitis, making corticosteroid-free remission an important goal (105).

The patient-reported outcome measures included in the OCTAVE trials include the validated, disease-specific IBDQ, as recommended by draft EMA guidance (145), and the EQ-5D, a standardised and validated generic instrument that is recommended by NICE (146).

The OCTAVE trials were conducted at 313 sites worldwide, including five in the UK (98). The trials included patients with moderately to severely active ulcerative colitis, the majority had extensive colitis of pancolitis, and more than half had received previous treatment with a TNFi agent. The results achieved in this broad population are expected to be applicable to patients in England.

Limitations of the clinical evidence base for tofacitinib include the short duration of follow-up in the induction trials, which limits the evaluation of induction therapy beyond 8 weeks. However, OCTAVE Sustain provides data for up to 52 weeks in patients with a clinical response at 8 weeks, and OCTAVE Open demonstrates the efficacy of a longer period of treatment with tofacitinib in patients without a response at week 8. A further limitation is that although the efficacy and safety of tofacitinib was assessed for up to 52 weeks in OCTAVE Sustain, data on the long-term safety and efficacy of tofacitinib are based on the open-label OCTAVE Open study, which does not include a control arm.

As with other clinical trials in ulcerative colitis, a limitation of the OCTAVE studies is the lack of direct comparisons with active comparators. This limitation has been addressed by conducting an NMA to allow indirect comparisons with all of the comparators in the NICE decision problem. One limitation of the NMA is that no adjustment was possible for differences among trials in placebo response and remission rates, which are known to be affected by the time at which the trial was conducted (20). It is therefore likely that this analysis underestimates the relative efficacy of tofacitinib.

B.3 Cost effectiveness

Model methodology:

- A Markov cohort model was developed to evaluate the cost-effectiveness of tofacitinib in moderately to severely active ulcerative colitis (UC) from the perspective of the NHS and PSS.
- The model structure, methods, and assumptions reflect the approach taken by the Assessment Group in NICE TA329, published in February 2015 (65).
- The model consisted of a patient lifetime cohort analysis, using 8-week cycles, and utilises 9 health states defined by the type of treatment and level of disease control.
- Aligned with the NICE scope, the cost-effectiveness analysis compared tofacitinib with biologic therapies, TNFi (infliximab, adalimumab and golimumab) and vedolizumab, and conventional therapies (without biological treatments).
- A network meta-analysis (NMA) for efficacy and safety of tofacitinib and comparators was conducted to inform the economic analysis (sections B.2.9 and B.2.10.8).
- Due to data limitations, it was necessary to establish subgroups defined by prior TNFi exposure to compare tofacitinib versus NICE scope comparators, and consisted of:
    - Biologic-naïve patients; and
    - Patients with prior exposure to biologics

Base Case Analysis:

The base case analysis considered a tofacitinib dose of 10 mg twice daily induction and 5 mg twice daily maintenance inclusive of a Patient Access Scheme (PAS) price for tofacitinib (PAS0139), and PAS prices for comparators, where publicly available.

Biologic-naïve population:

- In the deterministic analysis the ICER for tofacitinib vs conventional therapy (CT) was £8,554.03 per QALY. Tofacitinib dominated adalimumab, golimumab and infliximab, while vedolizumab generated an additional marginal QALY of [redacted], resulting in an ICER of £615,056.62 per QALY.
- Probabilistic sensitivity analyses (PSA) showed that the mean ICER of 1,000 simulations for tofacitinib compared to CT (£5,433.94 per QALY) was consistent with the deterministic ICER. At £20,000 per QALY threshold, tofacitinib had an 80.5% probability of being the most cost-effective treatment, followed by conventional therapy (13.7%).

Biologic-experienced population:

- In the deterministic analysis the ICER tofacitinib vs conventional treatments was £10,301.85 per QALY. When compared with tofacitinib, vedolizumab generated marginal additional [redacted] QALYs, with an ICER of £7.8 million per QALY.
- Probabilistic sensitivity analyses showed that the mean ICER of 1,000 simulations for tofacitinib (£10,926.30 per QALY) was consistent with the deterministic ICER. In addition, vedolizumab was dominated by tofacitinib. At £20,000 per QALY threshold, tofacitinib had a 56.3% probability of being the most cost-effective treatment, followed by conventional therapy (43.1%).

Scenario Analysis:

Intention-to-treat analysis: The analysis considers a tofacitinib dose of 10 mg twice daily induction and 5 mg twice daily maintenance at the PAS price for tofacitinib (section B.2.9).

- In the deterministic analysis the ICER for tofacitinib vs conventional treatments was £7,805.06 per QALY. When compared with tofacitinib, vedolizumab was dominated.

Tofacitinib [redacted]-mix – Biologic-naive population: The analysis considers a tofacitinib dose of 10 mg twice daily induction and a mix of [redacted] of 5 mg twice daily and [redacted] of 10 mg twice daily maintenance dose at the PAS price for tofacitinib (see SmPC and sections B.2.9 and B.3.5.1).

- In the biologic naïve population, all the comparators were dominated by tofacitinib, and the ICER of tofacitinib versus conventional therapy was £12,627.81 per QALY.

Tofacitinib [redacted]-mix – Biologic-experienced population: The analysis considers a tofacitinib dose of 10 mg twice daily induction and a mix of [redacted] of 5 mg twice daily and [redacted] of 10 mg twice daily maintenance dose at PAS price for tofacitinib (see SmPC and sections B.2.9 and B.3.5.1).

- In the prior-exposed population, vedolizumab was dominated by tofacitinib and the ICER of tofacitinib vs conventional therapy was £13,946.75.

Conclusion:

- Results of the base-case analysis demonstrated tofacitinib to be a cost-effective treatment option at conventional willingness to pay thresholds in the deterministic and probabilistic analyses.
- Probabilistic sensitivity analyses were consistent with the deterministic results, showing a > 50% probability of tofacitinib being cost-effective at the £20,000 threshold.
- Sensitivity analysis, additional scenario analysis, and probabilistic sensitivity analysis demonstrated that the model results were robust to input range and assumption changes.

These results confirm that tofacitinib 5mg and 10mg BD represent a cost-effective use of NHS resources in moderately to severely active ulcerative colitis, irrespective of prior biologic-exposure.

B.3.1 Published cost-effectiveness studies

An SLR was performed to identify all relevant published economic evaluations of tofacitinib or any other therapy in moderately to severely active UC. Fifty-three publications met the inclusion criteria and can be broken down as follows: 6 full-text publications reported on 5 UK-based economic evaluations and 7 abstracts reported on a further 5 UK-based studies. Forty publications were identified describing non-UK economic evaluations, with 12 published as full-text articles and 28 as abstracts only. A summary of the published UK-based cost-effectiveness studies identified in the SLR as well as analyses developed to inform the recent NICE technology appraisals is presented in Table 35; these are described in detail in Appendix G, Table 175.

Full details of the process and methods used to identify and select the relevant cost-effectiveness evidence, including PRISMA flow diagram, summary of studies, critical appraisal and quality assessments, are described in Appendix G.

Table 35 Summary list of published cost-effectiveness studies

Study (year of publication)	Type of model	Interventions
Included published economic evaluations with UK NHS perspective: full publications		
Buckland *et al.* (2008) (147)	Decision tree model	HD MZL, SD MZL
Tsai *et al.* (2008) (148)	Decision tree (induction) and Markov model (maintenance)	INF, SoC
Tappenden *et al.* (2016) (149); Archer *et al.* (2015) (118)	Decision tree (induction) and Markov model (maintenance)	Surgery, ADA, INF, GOL, CT
Essat *et al.* (2016) (85)	Decision tree and Markov model	ACA, VED, CT
Wilson *et al.* (2017) (150)	Decision tree (induction) and Markov model (maintenance)	ADA, GOL, INF, VED
Included published economic evaluations with UK NHS perspective: abstract only		
Ali *et al.* (2012) (151)	Markov model	ADA, SoC, surgery
Mukhekat *et al.* (2014) (152)	Markov model	Continue vs discontinue 5-ASA
Wilson *et al.* (2015) (153) Wilson *et al.* (2016) (154)	Decision tree (induction) and Markov model (maintenance)	ADA, GOL, INF, VED
Wilson *et al.* (2015) (155); Wilson *et al.* (2016) (156)	Decision tree (induction) and Markov model (maintenance)	VED, CT
Yang *et al.* (2014) (157)	Markov model	ADA, SoC
Previous NICE Technology Appraisals		
TA329 (2015) (65)	Markov model	ADA, GOL, INF, colectomy
Vedolizumab NICE TA 342 (2015) (66)	Decision tree and Markov model	ADA, GOL, INF, CT, surgery

Abbreviations: ADA, adalimumab; AZA, azathioprine; CT, conventional therapy; GOL, golimumab; HD, high dose; INF, infliximab; ITT, intention to treat; MZL, mesalazine; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; SoC, Standard of Care; VED, vedolizumab

B.3.2 Economic analysis

A *de novo* model was developed to determine the cost effectiveness of tofacitinib compared with the comparators in the NICE scope for the treatment of adults with moderately to severely active ulcerative colitis. A cost-utility analysis was conducted, considering the UK NHS and Personal Social Services perspective, consistent with the NICE reference case. The model was conceptualised based on the information identified in the literature search described in Appendix G.

The objective of the *de novo* model was:

1. To accurately reflect clinical practice in the UK
2. To accommodate all possible comparisons of treatment strategies within the analysed population as defined by the NICE scope

B.3.2.1 Patient population

In line with the current appraisal scope the analysis considers people with moderately to severely active ulcerative colitis who are intolerant of, or whose disease has had an inadequate response or loss of response to conventional therapy (oral corticosteroids and/or immunosuppressants) or a TNFi.

Within this population, the clinical evidence suggested that prior exposure to biologic treatment may be a significant treatment effect modifier (see section B.2.7.2 and section B.2.9.4). Moreover, it is now part of normal clinical practice to consider previous biologic treatment when deciding on the most appropriate treatment (see section B.3.3.5). Furthermore, a previous economic analysis submitted as part of a NICE TA 342 separately considered patients naïve to biologic treatment and patients with prior exposure (66).

Consequently, the appraisal analysis of moderately to severely active UC patients considers two subgroups in the base-case:

1. Biologic-naïve patients, and
2. Patients with prior exposure to biologics

The appraisal population was separated in the two subgroups, reflecting baseline patient characteristics (described in Table 36) response and remission rates, and available treatment strategies.

Table 36 Patients baseline characteristics (OCTAVE Induction trials pooled analysis)

	Tofacitinib 10 mg BID	Placebo
Biologic-naïve patients		
N	417	104
Age, mean (SD)	41.1 (13.5)	43.2 (13.9)
Male	59.71%	61.54%
Patient weight, mean (SD)	74.8 (17)	73.7 (15)
Patients with prior exposure to biologics		
N	488	130
Age, mean (SD)	41.3 (14.1)	39.4 (14.5)
Male	58.81%	52.31%
Patient weight, mean (SD)	72.6 (16.5)	72.3 (17.6)

Abbreviation: SD, standard deviation

Please note, In a key scenario Pfizer also explored the intention-to-treat population, based on a NMA, directly comparing tofacitinib with vedolizumab, to allow comparison of the total moderately to severely UC clinical trial populations (section B.3.7). The average age, proportion of males, and weight were used (weighted by the patient sample size).

B.3.2.2 Model structure

A Markov model was developed in Microsoft Excel 2016® with outcomes evaluated by cohort analysis.

An 8-week Markov model cycle was implemented. The choice of cycle length was based on the induction and maintenance phase assessment intervals in the clinical trials of tofacitinib

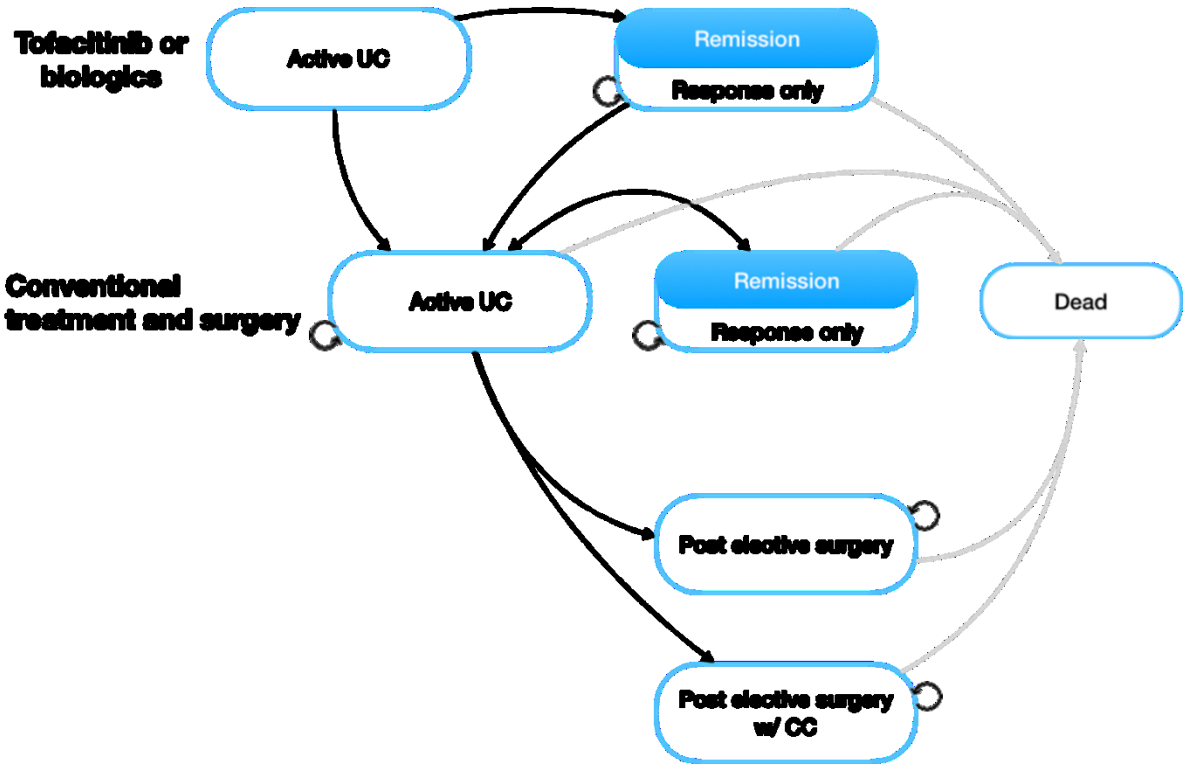
and other comparators, which informed the NMA (see section B.2.9.1). Prior to the selection of the cycle length, the induction treatment cost of all comparators was calculated to ensure that even for shorter induction durations (6 weeks), the respective treatment induction cost would not be overestimated for any comparator. Within the framework of a Markov model, a fixed cycle length was assumed for the duration of the model time horizon to allow for a continuous sequence of biologic treatments. When evidence was provided for longer timeframes (e.g. maintenance after one year of treatment), risks were adjusted to the 8-week cycle length.

The base-case analysis assumed a patient lifetime horizon. Given the relatively short duration of the cycle length, half-cycle correction was not implemented. An annual discounting rate of 3.5% was applied to both costs and quality-adjusted life years (QALY) (146).

Model schematic

A schematic diagram illustrating the structure of the model is shown in Figure 31. The model consisted of 9 health states defined by the type of treatment (biologic, non-biologic, surgery) and their level of disease control (active ulcerative colitis, response-no-remission and remission).. Patients who responded to treatment were separated to remission and response-no-remission. The definitions of response and remission followed those of 'clinical response' and 'clinical remission' in the clinical trials (see section B.3.3.1). Patients without either response or remission were defined as having active (moderately to severely) ulcerative colitis.

Figure 31 Schematic of cost-effectiveness model



Abbreviations: CC, colectomy complications; UC, ulcerative colitis.

In summary the model included the following health states:

- active ulcerative colitis, response-no-remission, and remission states, for biologic treatments or tofacitinib, as well as for non-biologic conventional treatment
- two post-surgery health states: with and without long-term complications
- an absorbing state (dead).

The surgical operation was modelled as a transient event rather than a health state. The proportion of the cohort who survived the operation would transition to the post-surgery health states.

Induction phase

The model assumed that patients entered with active ulcerative colitis and would undergo treatment induction with tofacitinib or a biologic (TNFi or vedolizumab). Based on the NMA output (section B.2.9.2), at week 8 of treatment, patients were categorised as non-responders, responders only, and remission. Non-responders were assumed to remain with active ulcerative colitis and discontinued treatment, and transitioned to conventional (non-biologic) therapy.

Maintenance phase

Patients who achieved response or remission in the induction phase entered the maintenance phase of the model, and continued to receive treatment with the same biologic until loss of response, acute exacerbation event or death. For the cohort who remained on treatment (responders), the ordered categorical results of the NMA were used to determine the proportion of patients achieving remission. The remaining patients were assumed to have responded, but not achieved remission.

Loss of response was informed by the NMA results of the maintenance clinical trials follow-up after approximately one year of treatment. If patients receiving biological therapy or tofacitinib lost their response at any point they were assumed to transition to conventional therapy.

A similar approach was followed during conventional therapy. Patients who did not respond or discontinued conventional therapy were assumed to remain with active UC.

Details on the above assumptions and how these align with other economic analyses are presented in B.3.3.1.

Elective surgery

The analysis assumed that a proportion of the cohort who did not respond, or who discontinued conventional therapy, would undergo elective colectomy. This assumption is aligned with in the recent TA329 AG approach (65).

A perioperative risk of complications and mortality risk was assumed for patients undergoing colectomy.

Acute exacerbation events

The analysis assumed that a proportion of patients would suffer ulcerative colitis-related acute exacerbation *events* and would require emergency surgery (transition lines not shown in the figure). The base-case analysis assumed that patients in remission were protected from exacerbations. Patients in all other health states remained at risk of acute events and emergency surgery. Sensitivity analysis considered additional scenarios: only patients with active ulcerative colitis, no patients, or all patients at risk of exacerbations.

Similar to elective surgery, a perioperative risk of complications and mortality risk was assumed for patients undergoing colectomy. Although flares of symptoms can be life-threatening (see section B.1.3.1), no additional mortality risk due to exacerbations was assumed in the model.

Post-surgery

Following colectomy, elective or emergency, patients were allocated to post-surgery states: without or with long-term complications (assumed to be represented by chronic pouchitis).

Death

Except for perioperative mortality, all patients had a probability of dying from other causes.

Table 37 Features of the economic analysis

	Previous appraisals		Current appraisal	
Factor	**TA329 [a]**	**TA342**	**Chosen values**	**Justification**
Model mathematical framework	Markov model	Decision tree in induction phase, and Markov model in maintenance phase	Markov model	This framework allows the modelling of recurrent risks, such as response to treatment after induction and maintenance
Time horizon	Patient lifetime	10 years	Patient lifetime	Since UC is a chronic condition, a patient lifetime time horizon allows the calculation of all relevant costs and quality of life impairment
Cycle length	8 weeks (induction) and 26 weeks (maintenance)	Induction (decision tree): 6 weeks, maintenance (Markov model): 8 weeks	8 weeks	The choice of cycle length was based on the maintenance phase assessment intervals in the clinical trials of tofacitinib and other comparators. A fixed cycle length was required throughout the model to allow flexibility of adding a sequence of biologic treatments.
Treatment waning effect and discontinuation	Treatment effect was assumed to be maintained with ongoing treatment. During the maintenance phase, patients receiving biological therapy were assumed to continue receiving the same biological treatment for as long as they continued to maintain response/remission. If patients receiving biological therapy lost their response at any point they were assumed	Treatment effect was assumed to be maintained with ongoing treatment. Within the model discontinuation of treatment was due to a lack of response by the end of the induction phase or due to adverse events. In addition, it was assumed that treatment with a biologic was limited to one year and all patients on therapy at week 54 of the model would switch to conventional therapy”	Treatment effect was assumed to be maintained with ongoing treatment. During the maintenance phase, patients receiving treatment were assumed to continue receiving the drug for as long as they maintained response/remission. If patients receiving biological therapy lost their response at any point they were assumed to transit to the active UC health state	Follows the approach taken in the independent economic analysis in TA329

	Previous appraisals		Current appraisal	
Factor	**TA329** [a]	**TA342**	**Chosen values**	**Justification**
	to transit to the active UC health state			
Source of utilities	Health state utilities (EQ-5D) for pre and post-surgical states from Woehl *et al.* 2008 (158) Disutility associated with chronic pouchitis from Arseneau *et al.* 2006 (159)	EQ-5D from GEMINI I trial for pre-surgery health states Surgery and post-surgery health states utilities from Punekar and Hawkin 2010, and Woehl *et al.* 2008 (158)	Baseline utility was estimated based on age and gender of the general population (160). Health state utilities (EQ-5D) for pre and post-surgical states from Woehl *et al.* 2008 (158).	Consistent with scenario analyses presented in previous TAs
Source of resource use	Tsai *et al.* 2008 (148)	Tsai *et al.* 2008, Buchanan *et al.*, 2011 (148, 161)	Tsai *et al.* 2008 (148)	Consistent with structure of economic model and previous TAs
Source of costs	BNF and NHS Reference Costs 2012/13 (162, 163)	NHS reference costs, BNF for drug costs (162, 163)	NHS reference costs (164), eMIT and MIMS for drug costs (165, 166)	Consistent with the NICE reference case
Pharmacological treatment adverse events	No AEs were considered	Serious infection, tuberculosis, lymphoma, hypersensitivity and injection site reaction included	Serious infection	Evidence on the incidence of serious infections was available for all drugs and their impact on costs and QALYs could be reasonably quantified
Mortality	Perioperative mortality associated with colectomy and other-cause mortality (corresponding to the general population mortality).	Mortality was applied as a baseline other-cause mortality rate, with state-specific relative risks to reflect an excess risk of death due to UC. A perioperative mortality risk was not included in the model	Perioperative mortality associated with colectomy and other-cause mortality (corresponding to the general population mortality).	No definitive evidence on the impact of UC on patient survival Regarding pre-surgery health states, the assumption is consistent with the evidence on standardised mortality ratios, indicating little difference in the risk of death between patients with UC and the general population (52). Regarding the perioperative risk, the approach is consistent with TA329

[a] Information reflects the independent economic analysis designed by the Analysis Group

Abbreviations: ADA: adalimumab; AEs: adverse events; AG: assessment group; eMIT: electronic market information tool; BNF: British National Formulary; GOL: golimumab; IBD: inflammatory bowel disease; INF: infliximab; MIMS, Monthly Index of Medical Specialties; NHS: national health system; NR: not reported; PSSRU, Personal Social Services Research; TA: technology appraisal; UC: ulcerative colitis

B.3.2.3 Intervention technology and comparators

In England and Wales, it is anticipated that tofacitinib will be used in the NHS by patients currently eligible for TNFi or vedolizumab treatment. In line with the NICE scope the tofacitinib comparators include:

- TNFi (infliximab, adalimumab and golimumab)
- Vedolizumab
- Conventional therapies, without biological treatments

Table 38 and Table 39 present details on the intervention, biologic comparators dose regimens and stopping rules, and details on the conventional treatment dose regimens and assumed patient usage, respectively. Regarding conventional treatment, the rationale for the regimens considered is described in Appendix M.1.1.

Table 38 Comparator treatment dose regimens and stopping rules

Treatment	Dosing instruction	Stopping rule – SmPC [a]	Clinical trial induction visit	Administrations and dose in model	
				Induction	Maintenance (8-week cycle period)
Tofacitinib (Xeljanz)	[redacted]	[redacted]	8 weeks (OCTAVE)	[redacted]	[redacted]
Adalimumab (Humira)	Injection, initially 160 mg, then 80 mg at week 2, and 40 mg every other week thereafter [c]	Available data suggest that clinical response is usually achieved within 2-8 weeks of treatment. Humira therapy should not be continued in patients failing to respond within this time period.	8 weeks (ULTRA)	160 mg + 80 mg + 40 mg x 3	4 mg x 4
Golimumab (Simponi)	Injection, initially 200 mg, then 100 mg at week 2, and 50 mg every 4 weeks thereafter [d]	Available data suggest that clinical response is usually achieved within 12-14 weeks of treatment (after 4 doses). Continued therapy should be reconsidered in patients who show no evidence of therapeutic benefit within this time period.	6 weeks (PURSUIT)	200 mg +100 mg + 50 mg	50 mg x 2
Infliximab (Remicade, Remsima,	By IV infusion, 5 mg/kg, repeated 2 weeks and 6	Continued therapy should be carefully reconsidered in	8 weeks (ACT)	5 mg/kg x 3	5 mg/kg

Treatment	Dosing instruction	Stopping rule – SmPC [a]	Clinical trial induction visit	Administrations and dose in model	
				Induction	Maintenance (8-week cycle period)
Inflectra)	weeks after initial infusion, then every 8 weeks	patients who show no evidence of therapeutic benefit within the first 12 weeks of treatment or after dose adjustment.			
Vedolizumab (Entyvio)	By IV infusion, 300 mg repeated 2 weeks and 6 weeks after initial infusion, then every 8 weeks [e]	Continued therapy for patients with ulcerative colitis should be carefully reconsidered if no evidence of therapeutic benefit is observed by Week 10.	6 weeks (GEMINI)	300 mg x3	300 mg

[a] The stopping rule in the NICE guidance follows the SmPC stopping rule
[b] A scenario with a maintenance dose of 10 mg twice daily was also explored in the analysis. For full SmPC wording please refer to appendix C.
[c] A scenario with an elevated maintenance dose (40 mg QW) received by 27% of the patients was explored in the analysis
[d] A scenario with a maintenance dose of 100 mg every 4 weeks was also explored in the analysis
[e] A scenario with a maintenance dose of 300 mg every 4 weeks was also explored in the analysis

Table 39 Conventional treatment dose regimens and assumed patient usage

Treatment	Dosing instruction	Patient usage
Aminosalicylates		
Balsalazide	1.5 g twice daily, adjusted according to response (maximum 6 g per day)	13%[a]
Mesalazine	1.2 to 2.4 g once daily	13%[a]
Olsalazine	500 mg twice daily	13%[a]
Sulfalazine	0.5 to 1 g twice daily	13%[a]
Corticosteroids		
Hydrocortisone	1 metered application once daily on alternate days	4%
Prednisolone	Initially 20–40 mg daily until remission occurs, followed by reducing dose	44%
Immunomodulators		
Azathioprine	2.0 to 2.5 mg/kg daily	46%

Source: BNF, RCP national audit report 2016 (3).
Usage of conventional therapy treatments at the initiation of a biologic (ADA, GOL, INF, INF biosimilar, VEDO) was assumed to be representative of usage in active disease and, therefore, after failure of biologics
[a] An equal distribution of the 5-ASA frequency of use among the different therapies was assumed

B.3.2.4 Treatment strategies

In the biologic-naïve population, the model compared six strategies, consisting of a biologic treatment or tofacitinib, followed by conventional therapy and surgery (Table 40). After a

single line of biologic treatment or tofacitinib patients received conventional therapy followed by surgery. Conventional therapy alone was also included as a comparator.

Table 40 Comparator strategies in the biologic-naïve population

Line of therapy	Comparator strategies					
1	Adalimumab	Golimumab	Infliximab	**Tofacitinib**	Vedolizumab	Conventional
2	Conventional	Conventional	Conventional	Conventional	Conventional	

In the population with prior biologic exposure, the model was assumed to start at second-line treatment (i.e. that all hypothetical patients enter the model having previously had exposure to a biologic therapy) and compared three strategies, consisting of tofacitinib or vedolizumab, followed by conventional therapy and surgery (Table 41). Here too, conventional therapy alone was included as a comparator.

Table 41 Comparator strategies in the biologic-prior-exposure and ITT population

Line of therapy	Comparator strategies		
1	**Tofacitinib**	Vedolizumab	Conventional
2	Conventional	Conventional	

B.3.2.5 Treatment continuation

As described in section B.3.2.2, continuation of treatment was dependent on response at the end of the induction period.

Regarding the TNFi treatments and vedolizumab, there were differences between the length of the induction period in the clinical trials and the recommended stopping rules in the SmPC (Table 38). To achieve consistency with the meta-analysis of the clinical evidence, the economic analysis used the clinical trial induction phase as a guide for the stopping rule for these treatments. A maximum of 8 weeks was assumed as the common induction phase for all biologic treatments, in line with the OCTAVE trial and previous economic models (118).

Once patients entered the maintenance phase (by achieving response during induction), continuation of treatment was assumed to be dependent on continued clinical response. This was consistent with the approach in previous economic analyses (118).

The risk of discontinuation due to adverse events, or other causes was considered low (see section B.2.10), and is likely to be outweighed by discontinuation due to lack of efficacy. Synthesis of discontinuation due to lack of efficacy with the other causes was likely to lead to overestimation of the overall risk and therefore, discontinuation due to AEs or other causes was not included in the model.

The model projections of the time-on-treatment (see section M.2. in Appendix M) were reviewed by clinical experts who confirmed that the estimates were plausible.

B.3.3 Clinical parameters and variables

B.3.3.1 Treatment effectiveness: clinical response and remission

The definition of clinical response in the model was a decrease from baseline Mayo score of ≥ 3 points and ≥ 30%, with a decrease in rectal bleeding subscore of ≥ 1 point or absolute rectal bleeding subscore of ≤ 1. The definition of clinical remission was a Mayo score of ≤ 2 and no individual subscore exceeding 1 point.

Table 42 Mayo score (OCTAVE trials pooled analysis) per health states, mean (SD)

Health state	Tofacitinib 5 mg BID	Tofacitinib 10 mg BID	Placebo	Total
Baseline				
Active UC	NA	9.0 ± 1.4	9.0 ± 1.5	9.0 ± 1.4
End of induction phase				
Response-no-remission	[redacted]	[redacted]	[redacted]	[redacted]
Remission	[redacted]	[redacted]	[redacted]	[redacted]
Active UC	[redacted]	[redacted]	[redacted]	[redacted]
End of maintenance phase				
Response-no-remission	[redacted]	[redacted]	[redacted]	[redacted]
Remission	[redacted]	[redacted]	[redacted]	[redacted]
Active UC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BID, twice a day; NA, not applicable; SD, standard deviation; UC, ulcerative colitis

B.3.3.1.1 Induction phase patient allocation

The proportion of patients achieving clinical response and clinical remission during induction, (i.e. one 8-week Markov cycle) was informed by the NMA of the clinical trial evidence (section B.2.9.2.1.1 and Table 43). The output of the NMA was transformed from the probit to the natural scale. For *k* the treatment and *j* the category (remission) we used:

$P_{kj} = \Phi(\theta_{kj})$, where Φ is the inverse of the normal cumulative distribution function.

Table 43 Clinical response and remission at induction

		Treatment effect: Median (CrI; 2.5%, 97.5%)	Proportion of patients in:		
			Response (incl. remission)	Response-no-remission	Remission
Biologic-naïve population					
Anchor		[redacted]			
Treatment effect	Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]
	Golimumab	[redacted]	[redacted]	[redacted]	[redacted]
	Infliximab	[redacted]	[redacted]	[redacted]	[redacted]
	Conventional (Placebo)	[redacted]	[redacted]	[redacted]	[redacted]
	Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]
	Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]
Cut-off (z)	Remission	[redacted]			

		Treatment effect: Median (CrI; 2.5%, 97.5%)	Proportion of patients in:		
			Response (incl. remission)	Response-no-remission	Remission
Prior exposure to biologic treatment					
Anchor		[redacted]			
Treatment effect	Conventional (Placebo)	[redacted]	[redacted]	[redacted]	[redacted]
	Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]
	Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]
Cut-off (z)	Remission	[redacted]			

B.3.3.1.2 Maintenance phase patient transitions

As described in B.3.2.2 patients were assumed to remain on treatment for as long as they sustained clinical response. In previous economic analyses, transitions between health states were informed by

a) (TA329); an NMA of response and remission data for weeks 8-32 and 32-52 (118)

b) (TA342); a calibration of the response and remission probabilities to fit the 1-year NMA estimates (117)

During the model conceptualisation, both approaches were considered. In (a) the assessment group (AG) had access to data for all comparators regarding the proportion of patients starting in and transitioning between response and remission health states. In our analysis the same evidence was not available. The outcomes of the maintenance phase NMA were based on initial response to treatment. It was not possible to determine a separate meta-analysis for responders and remitters at 8 weeks. Therefore, the definition of separate transition probabilities from and to response-no-remission and remission, required calibration to fit the proportion of the cohort to the target NMA estimates, as tried in (b).

The methods used in (b) were criticised by the ERG for discarding the empirical trial data. To improve on the methods, we attempted to use individual patient level data from the OCTAVE trial programme (induction and maintenance phase trials) as part of the calibration. The patient level data were used in an attempt to determine the baseline risk for those separate transitions (from response and remission), and to then synthesise it with the NMA relative risks for the other comparators. However, the result of that synthesis did not accurately predict the target data at the end of 52 weeks, and the data required further calibration with associated uncertainty.

An alternative approach was then pursued for the economic analysis. Since the maintenance NMA assumed that patients had at least responded to treatment, the model transition probabilities were based on response. Following the first 8-weeks allocation of patients to responders and remitters (induction), it was assumed that patients would continue treatment if they sustained response. Consequently, the risk of no-response (complement of treatment response) was assumed to be the same with the risk of treatment discontinuation. This

assumption was similar to the approach in the independent economic analysis by the NICE AG in TA329 (65).

The output of the NMA was transformed from the probit to the natural scale. For *k* the treatment and *j* the category (remission) we used:

$P_{kj} = \Phi(\theta_{kj})$, where Φ is the inverse of the normal cumulative distribution function.

Assuming a constant risk, the probability of no-response was adjusted for the 8-week cycle length. The resulting 8-week risk of no-response (and consequently, discontinuation) was applied every cycle for the entire duration patients remained on treatment.

To determine the timeframe of the probability of no-response, all maintenance phase duration follow-up was considered (Table 44). Since, the economic model would use meta-analysed data, the same duration was assumed for all comparator evidence. Given that tofacitinib is the technology appraised, the economic analysis used the OCTAVE trial programme duration for induction (8 weeks) and maintenance (52 weeks). That is, the output of the maintenance phase NMA was assumed to reflect sustained clinical response over 60 weeks of treatment; 8 weeks in induction and 52 weeks in maintenance. Therefore, the 52-week risk of no-response, from the maintenance phase NMA, was adjusted for the length of the Markov cycle (8 weeks): $P_{8week} = 1 - (1 - P_{52week})^{0.154}$ (Table 45).

By applying the above transition probability of discontinuation to all responders at the beginning of each cycle the model calculated the cohort of patients remaining on treatment. To separate the cohort of patients sustaining response and remission, the model used the categorical results from the NMA at the end of 1 year on treatment (Table 45). Furthermore, the model assumed that the observed allocation at one-year of treatment remained the same in all subsequent cycles.

Table 44 Duration of induction and maintenance phases in the trials considered

	Induction period	Maintenance period	Total trial duration
Pfizer model	8 weeks	52 weeks	60 weeks
Tofacitinib (OCTAVE Induction 1 and 2, Sandborn 2012, OCTAVE Sustain) (98, 99)	8 weeks	52 weeks	60 weeks
Adalimumab (ULTRA 1 and 2, Suzuki 2014, Mshimesh 2017) (73, 109, 114)	8 weeks	44 weeks	52 weeks
Golimumab (PURSUIT-SC, PURSUIT-M, PURSUIT-J) (75, 110, 115)	6 weeks	54 weeks	60 weeks
Infliximab (ACT 1 and 2, Jiang 2015, Kobayashi 2015, Mshimesh 2017) (74, 111) (112, 113)	8 weeks	46 weeks	54 weeks
Vedolizumab (GEMINI 1) (83)	6 weeks	46 weeks	52 weeks

Table 45 Clinical response and remission at maintenance, risk of no-response and proportions of remission and response-no-remission

		Treatment effect: Median (CrI; 2.5%, 97.5%)	Proportion of patients in:			Probability of no-response [a] (8 weeks)	Proportion of patients [a]	
			Response (incl. remission)	Response-no-remission	Remission		Response-no-remission	Remission
Index and calculations		A	B	C	D	1-[1-(1-B)]$^{0.154}$	C/B	D/B
Biologic-naïve population								
Anchor		[redacted]						
Treatment affect	Adalimumab							
	Golimumab 50 mg							
	Golimumab 100 mg							
	Infliximab							
	Conventional (placebo)							
	Tofacitinib 5 mg							
	Tofacitinib 10 mg							
	Vedolizumab Q8W							
	Vedolizumab Q4W							
Cut-off (z)	Remission							
Prior exposure to biologic treatment								
Anchor		[redacted]						
Treatment affect	Adalimumab (assumed the same for infliximab and golimumab)							
	Conventional (placebo)							
	Tofacitinib 5 mg							
	Tofacitinib 10 mg							
	Vedolizumab Q8W							
	Vedolizumab Q4W							
Cut-off (z)	Remission							

[a] Used in the economic model

Abbreviations: CrI, credible interval; mg, milligram; Q8W, every eight weeks; Q4W, every four weeks.

The above method, using a combination of probabilities of response and patient proportions, ensured internal consistency between the NMA results and the model projections of patients in remission and response after the end of one year with treatment (see Appendix M, section M.2.)

There was no evidence to test the validity of extrapolating the same risks beyond the first year of treatment. Nevertheless, clinical experts confirmed that, based on the model projected average time-on-treatment, the extrapolation estimates beyond one year were plausible (see section B.3.3.5).

Following discontinuation of the first treatment, patients would receive conventional therapy. For conventional therapy after a biologic or tofacitinib, the efficacy from the NMA conducted on the biologic-exposed subgroup was used (Table 45).

As in the TA329 AG model, it was assumed that patients in the conventional treatment group, and those who have previously achieved but lost response to biological therapy, would continue receiving conventional therapy irrespective of whether they achieve response to that conventional therapy.

B.3.3.2 Surgery and surgery complications

A proportion of patients who did not respond to conventional therapy (last pharmacological line) was assumed to all undergo colectomy. To inform the economic analysis on the probability of colectomy and the following complications, a focused search was conducted in Medline. A full search strategy has been included in Appendix M, section M.3.

The literature search identified 310 unique studies, of which 5 were of most relevance from a UK perspective, and are discussed in more detail hereafter for their applicability to support the economic analysis.

Misra *et al.* (167) was a retrospective population-based study using the Hospital Episode Statistics (HES) database. It included records between 1 April 1997 and 31 March 2012; N=73,318. The aim of the study was to compare the difference in colectomy rates for UC, dependent on the ethnic background of the patient. Chhaya *et al.* (168) was a large (N = 1,766), population-based cohort of incident cases of UC in the United Kingdom between 1989 and 2009 using data from the Clinical Practice Research Datalink (CPRD).

The remaining three studies comprised a cost-effectiveness analysis (149) and two retrospective cohort analyses (169, 170) with relatively small samples (N=38 to 143).

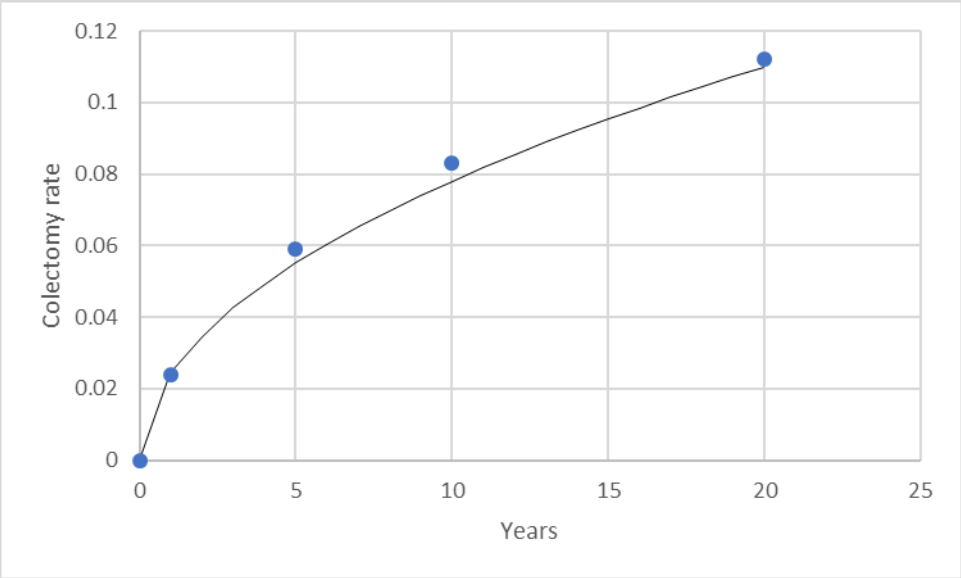
B.3.3.2.1 Colectomy rates

Misra *et al.* (167) reported that the colectomy rate excluding cases arising from colorectal cancer was 6.9% (n = 5,044/73,318) over 15 years. Of the 5,044 patients undergoing colectomy, 4,037 had elective and 1,481 had emergency colectomy.

Chhaya *et al.* (168) reported a cumulative risk of 2.4%, 5.9%, 8.3% and 11.2% at 1, 5, 10 and 20 years since diagnosis, suggesting a steady increase after the first 5 years since diagnosis (Figure 32). It follows that in the economic analysis, where patients entered the

model with an average time since diagnosis of 3 years, a time-independent probability was an appropriate assumption. A similar assumption was implemented in TA329 and TA342.

Figure 32 Cumulative probability of colectomy from Chhaya *et al.* 2015 (168)



Neither Misra *et al.* nor Chhaya *et al.* reported information on disease severity. Chhaya *et al.* attempted to adjust for severity by including 'early steroid use' in the regression model since this was "an established surrogate marker for a severe disease phenotype" (168).

For comparison with data used in previous cost-effectiveness analyses, Table 46 presents the 8-week risk of colectomy from Misra *et al.*, Chhaya *et al.* and data used in TA329 and TA342.

Table 46 Colectomy rates from several sources in the literature

	Reference	Description of evidence	Follow-up since diagnosis (years)	Calculated risk per cycle
Base case	Misra 2016 (167)	UK HES data retrospective analysis of records between 1 April 1997 and 31 March 2012; N=73,318. Rate excluded patients undergoing colectomy for CRC	15	0.0731%
Alternative sources	Chhaya 2015 (168)	UK CPRD data analysis of incident cases of UC between 1989 and 2009; N= 1,766	20	0.0913%
	Solberg 2009 (171) Used in TA329 AG model	South-eastern Norway population-based cohort analysis of patients with IBD between 1990 and 1994; N=843	10	0.1572%
	Frolkis 2013 (172) Used in TA342	Meta-analysis of population-based studies; rates from studies that reported on UC patients	1	0.77%
			10	0.2606%

Abbreviations: AG, assessment group; CPRD, Clinical Practice Research Database; CRC, colorectal cancer; IBD, inflammatory bowel disease; UC, ulcerative colitis.

We concluded that the Misra *et al.* 2016 study was the most relevant for the economic model; it reported results of a UK population, included a larger and more contemporary cohort compared with the alternative sources, excluded CRC-related surgery cases, and

provided a split for elective and emergency rates. For these reasons it was selected for use in the base-case analysis. Sensitivity analysis considered a range from no risk (0%) to the highest risk from Table 46 reflecting long-term observations (the 1-year risk was not considered to be long-term): the 10-year observation from Frolkis *et al.* 2013 (172) (used in TA342).

From Misra *et al.* 2016 the 15-year cumulative risk of elective surgery was assumed to be 5.5% (0.058% at every 8-week cycle). The emergency colectomy risk was assumed to be 2% (0.021% at every 8-week cycle).

B.3.3.2.2 Perioperative complications and mortality

To complete the analysis with data for colectomy complications, the National clinical audit of 2013 for inpatient care for adults with ulcerative colitis was consulted (173). This was the fourth inpatient care report published from the UK IBD audit reporting on national- and hospital-level findings on the quality of care provided to people admitted to hospital between 1 January 2013 and 31 December 2013 primarily for the treatment of ulcerative colitis. Note that a subsequent publication did not report information of perioperative complications (3).

In the 2014 publication, perioperative complications were reported for 32% and 35% of patients undergoing elective and non-elective surgery respectively. Wound infection was the most common complication; 8% and 9% respectively for elective and non-elective surgery (173). Sensitivity analysis tested a range from no complications (0%) to double the reported values (64% and 70%).

In the 2014 national clinical audit, the overall mortality rate was reported to be relatively low compared with previous versions of the audit (Table 47) (173). However, it was unclear what the perioperative mortality risk was. In an economic analysis by Archer *et al.* (118, 149), the risk of death (3.5% per operation) was based on the 2012 publication of the UK IBD audit: 28 deaths among 807 elective and emergency surgical episodes in adult patients with UC. Our analysis assumed the risk of mortality was 2.8%, per operation, based on the (19%) reduction seen in overall mortality between round 3 and round 4 of the audit (Table 47). Sensitivity analysis considered a range from no risk (0%) to the 2012 risk (3.5% per operation) used in Archer *et al.* (118, 149).

Table 47 Mortality during admission over three versions of the IBD audit

Audit round	Overall mortality % (n/N)	Reduction in overall mortality	Perioperative % (n/N)
Round 2 (2008−2010)	1.54% (46/2981)		
Round 3 (2010−2012)	0.92% (28/3049)	19%	3.5% (28/807)
Round 4 (2012−2014) [a]	0.75% (30/3987)		2.8% (calculation)

[a] Restricted to first admission per patient

B.3.3.2.3 Post-surgery complications

A proportion of patients are expected to experience long-term complications following colectomy.

Of the 5 UK studies identified in our search, Tappenden *et al.* (149) used a Japanese study to inform the model of the probability of chronic surgery-related complications. Arai *et al.* (174) included 296 patients with UC who underwent restorative proctocolectomy and reported on the overall incidence of complications (early and late).

A further review of the 119 studies from our Medline search identified two studies with useful information on post-surgery complications. One study was from Japan, based on 284 patients with UC who underwent a total proctocolectomy and IPAA. Suzuki *et al.* 2012 (175) and included a Kaplan-Meier curve showing the cumulative risk for developing pouchitis and reported that the risk was 10.7% at 1 year, 17.2% at 2 years, 24.0% at 5 years, and 38.2% at 10 years. Overall, 64 of the 244 patients (26.2%) developed idiopathic pouchitis. Another study, conducted in Leuven, Belgium, included an analysis of 173 patients who underwent proctocolectomy with IPAA for UC or IBDU (176). It reported that during a median follow-up of 6.5 years (IQR 3.4–9.9), 80 patients (46%) developed at least 1 episode of acute pouchitis.

The economic analysis used the study from Belgium in the base-case, on the assumption that it was more relevant to the population in the UK. The 6.5-year risk reported by Ferrante *et al.* was converted to an 8-week risk (1.5%) and applied to all patients who survived surgery. Sensitivity analysis considered a range from the Suzuki *et al.* (175) risk (0.7% every 8 weeks) to a value equal to the distance between the Ferrante and Suzuki values over the base-case risk (2.1% every 8 weeks).

B.3.3.3 Treatment safety: serious adverse events

The economic analysis considered events with a substantial impact on costs and HRQoL. These were assumed to be those that were reported as serious events; often defined as life-threatening, or which lead to hospitalisation or other medical emergencies.

A network meta-analysis was conducted on the total SAEs reported in the comparator clinical trials (section B.2.10.8). We noted that, of the total SAEs, the most common were GI events, events related to UC, or “worsening of disease”. The definition of those events across all clinical trials was unclear; they may have been exacerbation episodes or general disease worsening. The patient condition related to UC was already considered in the economic model with the definition of health states based on clinical response and clinical remission corresponding to Mayo scores. Therefore, the total SAE statistics were not considered further in the economic analysis.

The immunomodulatory or immunosuppressive effects associated with these treatments may predispose patients to serious infections. To reflect the risk of serious infections in the economic model an NMA was conducted including incidence rates for all comparators (section B.2.10.8). The model used the results of the NMA to obtain risks of the events in the induction and maintenance phase. Table 48 presents the transformation of the NMA output to 8-week probabilities for inclusion in the induction phase in the economic model. We noted that the duration of the induction phase is different for some clinical trials (Table 44). For simplification of the model input, it was assumed that the observed duration of serious infections was the same for all comparators.

Table 48 Induction phase serious infections

		Treatment effect: Median (CrI; 2.5%, 97.5%)	Incidence of event
Anchor		[redacted]	
Treatment effect	Adalimumab	[redacted]	[redacted]
	Golimumab	[redacted]	[redacted]
	Infliximab	[redacted]	[redacted]
	Conventional (Placebo)	[redacted]	[redacted]
	Tofacitinib	[redacted]	[redacted]
	Vedolizumab	[redacted]	[redacted]

Serious infection was a rare adverse event in the tofacitinib clinical trials. When meta-analysed this led to significant uncertainty in the precision estimates and a wide 95% CrI. If the NMA CrI was to be used in deterministic sensitivity analysis, the risk of serious infection would be shown as the most important variable in the model; this would be a misleading result, given the rareness of the event. Instead in the deterministic sensitivity analysis the risk of serious infection was varied from no difference from placebo to a 50% risk increase. The NMA output was used in the probabilistic sensitivity analysis (PSA).

To avoid double-counting of disutility and costs from co-occurrence of events, no other, less common, SAEs were considered. Sensitivity analysis considered no serious infection during the maintenance phase.

B.3.3.4 Mortality risk

UC treatment was assumed have no effect on overall mortality. Age-dependent all-cause mortality risks obtained from UK life tables (177) were applied as a background risk of death to all patients in pre- and post-surgery health states. To reflect the patient population in the model, the gender-specific mortality risk was combined into a blended rate, using the proportion of female patients across in each subgroup (see Table 36)

B.3.3.5 Clinical expert assessment

Details of the clinical expert assessment are presented in section B.3.10.

B.3.4 Measurement and valuation of health effects

Health effects in the current analysis were expressed in QALYs, in accordance with the NICE reference case.

B.3.4.1 Health-related quality of life data from clinical trials

EQ-5D-3L data were collected in the OCTAVE clinical trial programme at baseline, and then at visits on weeks 2, and 8 in OCTAVE 1 and OCTAVE 2. In OCTAVE Sustain at baseline, and then visits at weeks 4, 8, 16, 24, 32, 40 and 52. Details are presented in appendix M.3.

An analysis on patient level EQ-5D index score data was conducted to estimate the change in EQ-5D over time, based on the destination health state of the patients in each clinical trial. For instance, for remitters at week 8 in OCTAVE 1 and OCTAVE 2, the analysis looked back at their EQ-5D at week 4 and at baseline.

The data suggested that there is homogeneity in the mean EQ-5D index based on health state membership at the end of the trial(s).

To inform the economic analysis with EQ-5D from the OCTAVE trial programme, the baseline EQ-5D was used for the active UC health state (Table 239).

For the response-no-remission and remission health states the maintenance phase data were used. To approximate the EQ-5D across all visits, for each group of patients, defined by the trial arm (tofacitinib 5mg, 10 mg, placebo) and biologic exposure (naïve and exposed), the area under the curve (AUC) of the series of measurements at all visits was calculated by splitting the area in a series of trapeziums. The AUC was then calculated as the sum of all the individual trapeziums.

For the utility weight post-surgery, we assumed the same difference from remission as observed in Woehl et al. (118) where remission was 0.87 and post-colectomy 0.71; that is, reduction of 18.4%.

To inform the model with the precision across the health state utilities, the minimum and maximum values of the averages across the trial arms were used for active UC, response-no-remission and remission. For post-surgery the upper bound was assumed to be equal to remission and the lower bound equal to active UC.

Table 49 Utility values used in the cost-effectiveness model scenario with OCTAVE EQ-5D

Health state	Assumed utility	Range (Min-Max)	Number of observations	Comments / assumption
Active UC	[redacted]	[redacted]	[redacted]	Baseline of OCTAVE 1 and OCTAVE 2
Response-no-remission	[redacted]	[redacted]	[redacted]	Approximated see Table 238
Remission	[redacted]	[redacted]	[redacted]	
Post-surgery	[redacted]	[redacted]	[redacted]	Assumed 18.4% lower than remission (118)

B.3.4.2 Mapping

No mapping was needed to assess health state utility values as EQ-5D data were collected in the OCTAVE clinical trial programme.

B.3.4.3 Health-related quality of life studies

An SLR was conducted to identify relevant health utility elicitation/validation studies. Details of the search strategy, inclusion criteria and individual study results are described in Appendix H. Table 50 summarizes the EQ-5D utility values presented in the included studies which report EQ-5D utilities for multiple relevant health states. Utility values used in previous NICE submissions are also presented. Included studies which report on EQ-5D utility values for single UC health states are summarised in Appendix H.

Table 50 A summary of EQ-5D utility values by health state, as identified in SLR and previous technology appraisals

Study	Disease states, mean (SD)				Post-surgery, mean (SD)				Comments
	Severe	Moderate	Mild	Remission	IPAA	Ileo-stomy	No compli-cations	Compli-cations	
Swinburn 2012 (178)	0.45	0.68	0.8	0.9	0.59				EQ-5D-3L (UK)
Woehl 2008 (118)	0.41 (0.34)		0.76 (0.18)	0.87 (0.15)	0.71 (0.29)	0.72 (0.35)			EQ-5D-3L (UK)
van Assche 2016 (179) van Assche 2015 (180)	0.61 (0.22)	0.70 (0.19)	0.80 (0.15)	0.86 (0.17)					EQ-5D-5L (Europe)
Kosmas 2015 (38)	0.52			0.88			0.90	0.71	EQ-5D-5L (UK)
							0.74		
Gibson 2014 (181)	0.68 (0.19)		0.78 (0.18)	0.81 (0.18)					EQ-5D-5L (Australia)
Vaizey 2014 (26) Vaizey 2013 (26)	0.66 (0.24)		0.77 (0.11)	0.86 (0.15)					EQ-5D-5L (UK)
Casellas 2005 (182)	Median: 0.5 (IQR 0.5-0.7)		Median: 0.72 (IQR 0.5-0.8)	Median: 1.00 (IQR 0.8-1.0)					EQ-5D-3L (Spain)
Marteau 2009 (183)	0.660 (SE 0.03)	0.775 (0.013)		Baseline: 0.945 (0.023) At 12 months: 0.940 (SE 0.001)					EQ-5D-3L (UK tariff) mapping from UC-DAI
Poole 2010 (184) Poole 2009 (185)	Observed: 0.70 Mapped: 0.630	Observed: 0.811 Mapped: 0.801		Observed: 0.944 Mapped: 0.939					EQ-5D mapping (UK tariff) from UC-DAI
van Der Valk 2012 (186, 187)					0.85 (0.19)	0.85 (0.17)			EQ-5D-3L (Netherlands)
Kuruvilla 2012, (188)					0.9 (0.1)	0.9 (0.1)			EQ-5D-3L (USA)
Archer 2016 (118)	0.41		0.76	0.87			0.70		Based on Woehl 2008 (158)
Vedolizumab manufacturer submission (TA342) (117)	0.68		0.8	0.86	0.42		0.60	0.42	Disease health states were based on a post-hoc analysis of EQ-5D-3L data gathered in the pivotal phase 3 vedolizumab RCT (GEMINI 1) and surgery and post-surgery health states were based on values from a cost-effectiveness analysis of strategies for acute, severe UC (Punekar and Hawkins 2010) (189)

Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; IPAA, ileal pouch-anal anastomosis; SD, standard deviation.

B.3.4.4 Adverse reactions

B.3.4.4.1 Pharmaceutical treatments

Serious infection was the only adverse event considered in the analysis (section B.3.3.3). The HRQoL impact associated with serious infection has been modelled by applying a utility multiplier from the literature to the utility of patients experiencing the event (190). Diamantopoulos *et al.* 2014 estimated this reduction of utility based on a study by Sisk *et al.* 1997 (191). The multiplier (0.9858) was calculated using a utility for pneumonia (0.21) and adjusting it for the expected duration of the event (7 days) and the baseline age and gender of the Sisk *et al.* cohort (190). Sensitivity analysis considered a range from no reduction in the patient utility to double the proportional reduction.

B.3.4.4.2 Surgery

The systematic literature review identified one study (published as abstract) reporting the mean EQ-5D-5L of patients who had surgery over one year ago; 0.90 vs 0.71 ($p< 0.001$) with and without complications respectively (38). In line with that finding, the appraisal analysis considered a reduction of 21% of the post-surgery utility weight for patients who suffered from long-term complications. Sensitivity analysis considered a range from no reduction in the patient utility to double the proportional reduction.

B.3.4.5 Health-related quality of life data used in the cost-effectiveness analysis

B.3.4.5.1 Baseline utility

To reflect the chronic condition of the disease the health state utilities were synthesised with a baseline utility. This was assumed to reflect the natural decline of patients' physical and mental functions due to age and other co-morbidities.

The baseline utility value was taken from a model by Ara and Brazier (160). The regression model was based on data from four consecutive Health Surveys for England. The data included self-reported health status and EQ-5D and were used to generate mean health state utility values for cohorts with or without prevalent health conditions.

$$U_{Base}(age, gender) = 0.9508566 + 0.0212126 * Male – 0.0002587 * Age – 0.0000332 * Age^2$$

Note that for the age and gender values of U_{Base} the analysis used the model population age and gender (Table 36).

Sensitivity analysis assumed a constant baseline utility using the value of remission from Woehl *et al.* (118).

B.3.4.5.2 Disease utility weights

To align with previous economic evaluations that used evidence from a UK population, the base-case analysis used the data provided from a study by Woehl et al. (118). Using Woehl *et al*. also allows for comparability of this economic analysis with the previous TAs.

A scenario where the OCTAVE EQ-5D utility weights are used in the model was also conducted.

The utilities used in the base-case analysis are presented in Table 51.

Table 51 Summary of utility values for cost-effectiveness analysis

	Utility value: Mean (95% CI or range)	Reduction from baseline	Reference in submission (section and page number)	Justification
Baseline	Varies by age and gender	N/A	B.3.4.5.1 (page 137)	To reflect patient physical and mental functions due to age and other co-morbidities
Active UC	0.41 (0.36, 0.46)	53%	B.3.4.3 (page 135)	To allow comparability of the economic analysis results with previous economic evaluations
Response-no-remission	0.76 (0.73, 0.79)	13%		
Remission	0.87 (0.85, 0.89)	0%		
Post-colectomy without long-term complications	0.71 (0.67, 0.75)	18%		
Post-colectomy with long- term complications	0.64 (0.47, 0.71)	35%	B.3.4.4.2 (page 137)	Assumed the same reduction observed in EQ-5D-5L between no complications and with complications (38)
Serious infection reduction	N/A	1.42% (0, 2.84%)	B.3.4.4.1 (page 137)	Assuming pneumonia was a reliable proxy for all infections

A utility decrement, or multiplier (φ), was estimated based on the difference between the general population utility $U_{GenPop}(Age, Gender)$ and the utility of the health state or event U_{HS}:

$\varphi_{HS} = U_{HS}/U_{GenPop}$

To calculate the general population utility from Woehl *et al.*, using the model by Ara and Brazier (160) would result in a lower utility (0.84) than the value for remission (0.87). To ensure internal consistency with the data presented by Woehl *et al.*, the remission utility was assumed to be the same as the general population.

In the economic model the utility decrements were multiplied at each cycle with the baseline utility, based on the proportions of patients and their state membership: $U_{Base} * \varphi_{HS}$.

B.3.5 Cost and healthcare resource use identification, measurement and valuation

B.3.5.1 Intervention and comparators' costs and resource use

Identification of relevant cost and healthcare resource data is described in Appendix I.

Cost and healthcare resource use inputs considered in the base-case analysis comprised of drug acquisition, administration costs, costs associated with adverse events and conventional therapy. Only direct medical costs were included in the model. Costs were sourced from the 2016/17 NHS reference costs (164), the electronic Market Information Tool

(eMIT) (166), the Monthly Index of Medical Specialties (MIMS) (165), the Personal Social Services Research Unit (PSSRU) (192) and published literature. The rationale for the costs used is described in detail in Appendix M.5.

B.3.5.1.1 Intervention and biologic treatment costs

Drug acquisition costs were derived from the eMIT database (166) or from the online version of MIMS (165). Unit costs for each comparator are summarised in Table 52. Total drug costs were estimated for 8-week cycles. For infliximab, the drug cost was calculated based on patient characteristics in the OCTAVE trials, as described in Appendix M.5.

Table 52 Drug acquisition costs

Drug	Pack size	Dose (mg)	Pack cost	Cost per dose	Induction total cost	Maintenance total cost [e]
Tofacitinib 5 mg [a]	56	5	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib 10 mg	56	10				
Adalimumab	2	40 mg/0.8 mL	£704.28	£352.14	£2,112.84	£1,408.56
Golimumab 50 mg [b]	1	50 mg/1 mL	£762.97	£762.97	£2,288.91	£1,525.94
Golimumab 100 mg [b]	1	100 mg/1 mL	£762.97	£762.97	£2,288.91	£1,525.94
Remicade [c]	1	100 mg	£419.62	£419.62	£5,680.90	£1,893.63
Inflectra and Remsima [c, d]	1	100 mg	£377.66	£377.66	£5,154.05	£1,718.02
Vedolizumab Q8W [a]	1	300 mg	£2,050.00	£2,050.00	£6,562.11	£2,187.37
Vedolizumab Q4W [a]	1	300 mg	£2,050.00	£2,050.00	£6,562.11	£4,374.74

[a] A confidential simple discount patient access scheme (PAS) is in place
[b] Golimumab was approved by NICE under a PAS in which the cost of the 100 mg/1 mL formulation is available at the same price as the 50 mg/0.5 mL formulation.
[c] Infliximab cost is calculated using the fitted distribution approach and tofacitinib patients' characteristics (OCTAVE 1 and 2)
[d] Infliximab biosimilars approved in the UK (Remsima, Inflectra) are available at the same list price
[e] Costs are calculated per 8-week cycles

B.3.5.1.2 Conventional therapy costs

Drug acquisition costs were derived from the eMIT database (166) or from the online version of MIMS (165) if not available in eMIT. Unit costs as well as costs per cycle and usage are summarised in Table 53.

When conventional treatment was used on its own, the therapy mix was informed by a recent national audit of the Royal College of Physicians (RCP) on inflammatory bowel diseases (IBD) (3). The evidence on concomitant medication for ulcerative colitis treatment when co-administered with the initial treatment of a biologic was assumed to be reflective of the usage when in active ulcerative colitis. When conventional treatment was used as concomitant medication, the therapy mix was informed by data from the same source (3), but from the treatment follow-up at 3 months. Since immunomodulator use is not recommended for concomitant use with tofacitinib, azathioprine was excluded from the concomitant therapies of tofacitinib.

The average cost of conventional treatment was estimated at £52.18 per 8-week cycle. The average cost of concomitant treatment with azathioprine was estimated at £49.40 per 8-

week cycle. Sensitivity analysis considered the lowest and highest costs from the list of available treatments.

Table 53 Conventional therapy costs

Drug	Pack size	Strength (g/ mg) [b]	Pack cost	Cost per dose	Total cost per cycle	Total annual cost	Usage [c]	Usage as concomitant therapy [d]
Balsalazide	130	750 mg	£30.42	£0.23	£52.42	£340.70	12.6%	11.6%
Mesalazine	168	400 mg	£54.9	£0.33	£54.90	£356.85	12.6%	11.6%
Olsalazine	60	500 mg	£161	£2.68	£300.53	£1,953.47	12.6%	11.6%
Sulfasalazine	112	500 mg	£6.87	£0.06	£6.87	£44.66	12.6%	11.6%
Prednisolone	30	5 mg	£0.91	£0.03	£6.79	£44.17	44.1%	19.9%
Hydrocortisone	14 [a]	20.8 g	£9.33	£0.67	£18.66	£121.29	3.8%	0.6%
Azathioprine	56	50 mg	£2.17	£0.04	£6.51	£42.32	46.4%	37.2% / 0% [e]

[a] Hydrocortisone is available as a 20.8 g foam in aerosol delivering approximately 14 applications (165)
[b] See assumptions in Table 38
[c] Proportion of use of in conventional treatment as part of the conventional therapy mix
[d] Proportion of use of conventional treatments as concomitant therapy to biologics and tofacitinib
[e] Immunomodulator are not recommended in concomitant use with Tofacitinib
Abbreviations: mg = milligrams

B.3.5.1.3 Treatment administration costs

Tofacitinib is given orally and requires no resources for training or administration.

Adalimumab and golimumab are administered as a subcutaneous injection. All patients were assumed to be able to self-administer subcutaneous injections in the base case. This assumption reflects the expected zero cost to the NHS for injection support due to home-care and support schemes offered by the manufacturers.

Infliximab and vedolizumab are administered as intravenous (IV) infusions by a health care professional. The cost of IV administration was assumed to be equal to the cost of an outpatient visit and was based on the mean of a consultant- and a non-consultant led non-admitted face-to-face follow-up appointment. Unit costs were taken from the 2016-17 NHS Reference Cost values (164) and estimated to be £137.37. Unit costs and calculations are detailed in Table 54.

Table 54 Treatment administration costs

Currency Code	No. of attendances	National Average Unit Cost	Source/assumptions
Consultant led (CL) - Non-Admitted Face-to-Face Attendance, Follow-up	845,935	£141	NHS Reference Cost 2016-17 (164), CL WF01A (Gastroenterology)
Non-consultant led (NCL) - Non-Admitted Face-to-Face Attendance, Follow-up	94,264	£107	NHS Reference Cost 2016-17 (164), NCL WF01A (Gastroenterology)
Outpatient visit	£137.37 (IQR £70.2 to £161.72)		NHS Reference Cost 2016-17 (164), weighted average of the number of attendances and the unit cost of CL and NCL WF01A The IQR was determined as the range between the lowest and highest limit of the CL and NCL costs.

Abbreviations: CL, consultant led; IQR, interquartile range; NCL, non-consultant led

B.3.5.1.4 Monitoring costs

No additional treatment-related monitoring costs for tofacitinib were assumed. This was confirmed by clinical expert opinion (see section B.3.3.5).

B.3.5.2 Health state unit costs and resource use

Unit cost and annual resource use for health states of active UC, response-no-remission, and remission are presented in Table 55. Annual cost per health states are presented in table Table 56.

The resource use for outpatient visits, treatment monitoring and hospitalisation was based on a UK cost-effectiveness model by Tsai *et al.* (148), which was the most relevant reference identified in the SLR. Tsai *et al.* reported annual resource use for each of the model's health states as estimated by a panel of UK gastroenterologists. The Mayo scores by health states from Tsai *et al.* aligned with the health states defined in the analysis, which were based on observations from the OCTAVE trials. From OCTAVE, remission Mayo score was 0.9-1.3, response/no remission 3.8-4.2, and no response 8.5. In Tsai *et al.* (148), remission was defined as a Mayo score of 0-2, mild (corresponding to response without remission) 3-5 and moderate to severe (corresponding to no response) 6-12.

Disease monitoring included regular outpatient visits, blood tests, and endoscopy. Tsai *et al.* (148) included hospitalisation episodes for standard of care or infliximab in their calculations. A clinical expert advised that hospitalisation would increase as the patient health state worsens (see section B.3.3.5). The estimated annual 0.3 hospitalisation for standard care was increased to 1.20 for the response without remission health state and to 1.50 for the active UC state.

As Tsai *et al.* (148) reported only a value for each health state, a range was assumed for sensitivity analysis, using the adjacent health states as low or high limits. For example, the response-no-remission resource use uses the active UC resource use as the high limit and the remission resource use as the low limits. For remission and post-surgery without complications the low limit was assumed to be no resource (0%) and the high limit was set to that of response-no remission.

The cost of hospitalisation was calculated as the weighted average of all the attendances of the non-elective inpatient entries from the NHS reference costs (£2,985).

Unit costs were taken from the 2016-17 NHS Reference Costs (164).

B.3.5.3 Adverse reaction unit costs and resource use

Serious infections were included in the base-case analysis and were assumed to include a hospitalisation event. The cost of a serious infection was considered to be a weighted average of six types of infection: sepsis, tuberculosis, pneumonia, skin and soft tissue infection, bone and joint infection and urinary tract infection. Weights were based on the number of finished consultant episodes described in the NHS reference costs for the

relevant Healthcare Resource Group (HRG). Costs associated with each adverse event were obtained from 2016/17 NHS reference costs (164).

Table 55 Annual medical resource use by health state

Resource item	Unit cost	Resource use per patient per year (148)				
		Active UC	Response without remission	Remission	Post-colectomy	
					no complications	With complications
Mayo score in Tsai *et al.* 2008		6–12	3–5	0–2	NA	NA
Mayo score in OCTAVE		8.5	3.8–4.2	0.9–1.3	NA	NA
Outpatient visit (specialist)[a]	£137.37	6.50 (4.5–8.5)	4.50 (2–6.5)	2 (0–4.5)	1.50 (0–4.5)	1.75 (1.5–2)
Blood tests	£3.06	6.50 (3.9–9.1)	3.90 (3.25–6.5)	3.25 (0–3.9)	1.50 (0–3.9)	3.25 (1.5–5)
Endoscopy	£277.29	2.00 (0.5–3.5)	0.50 (0.2–2)	0.20 (0–0.5)	1.25 (0–1.3)	0.65 (0–1.3)
Hospitalisation episodes	£2,984.71	1.50 (1.2–1.8)	1.20 (0.3–1.5)	0.30 (0–1.2)	0 (0–1.2)	3.25 (0–6.5)

[a] See Table 54 for the outpatient visit unit cost calculation
Abbreviations: NA, not applicable

Table 56. Annual cost by health states

Health state	Annual cost
Active UC	£5,944.46
Response-no-remission	£4,350.41
Remission	£1,235.56
Post-colectomy without complications	£557.26
Post-colectomy with complications	£10,130.91

Abbreviations: UC, ulcerative colitis

Table 57 Unit costs of treatment for adverse events

Adverse event	Adverse event sub-type	Unit cost	Weights	Mean cost	Source
Serious infection	Sepsis	£2658.76	79,532	£2,538.97	NHS reference costs 2016/17: Weighted average of WJ06A to WJ06J (non-elective inpatient long-stay) (164)
	Tuberculosis	£3,752.97	2,635		NHS reference costs 2016/17: weighted average of DZ14F to DZ14J (non-elective inpatient long-stay) (164)
	Pneumonia	£2,499.45	317,020		NHS reference costs 2016/17: weighted average of DZ11K to DZ11V and DZ23H to DZ23N (non-elective inpatient long-stay) (164)
	Soft tissue infection	£1,856.74	13,132		NHS reference costs 2016/17: weighted average of HD21D to HD21H (non-elective inpatient long-stay) (164)
	Bone and joint infections	£4,687.08	10,957		NHS reference costs 2016/17: weighted average of HD25D to HD25H (non-elective inpatient long-stay) (164)
	Urinary tract infection	£2,452.81	171,440		NHS reference costs 2016/17: weighted average of LA04H to LA04S (non-elective inpatient long-stay) (164)

Abbreviations: NHS, national health service

B.3.5.4 Miscellaneous unit costs and resource use

B.3.5.4.1 Colectomy operation costs

The cost associated with colectomy operation was determined from the 2016-17 NHS Reference Cost (164) assuming a weighted average of elective inpatient costs for proximal (FF32C) and distal colon procedures (FF33B): £6,090.52 and £7,294.56 for the health states without and with complications, respectively. Calculations are detailed in Table 58.

Table 58 Costs of colectomy operation and perioperative complications

Currency Code	No. of attendances	National Average Unit Cost	Source/assumptions
Proximal Colon Procedures, 19 years and over, with CC Score 0-2	4867	£6,085.07	NHS Reference Cost 2016-17 (164), EL FF32C
Distal Colon Procedures, 19 years and over, with CC Score 0-2	2029	£6,103.58	NHS Reference Cost 2016-17 (164), EL FF33B
Cost of colectomy without complication	**£6,090.52 (IQR £4,994.07 to £7,112.52)**		**NHS Reference Cost 2016-17 (164), weighted average of the number of attendances and the unit cost of EL FF32C and FF33B The IQR was determined as the range between the lowest and highest limit of the EL FF32C and FF33B and NCL costs.**
Proximal Colon Procedures, 19 years and over, with CC Score 3-5	1563	£6,803.58	NHS Reference Cost 2016-17 (164), EL FF32B
Proximal Colon Procedures, 19 years and over, with CC Score 6+	514	£8,484.21	NHS Reference Cost 2016-17 (164), EL FF32A
Distal Colon Procedures, 19 years and over, with CC Score 3+	592	£7,557.94	NHS Reference Cost 2016-17 (164), EL FF33A
Colectomy operation with complications	**£7,294.56 (IQR £5,544.42 to £9,742.12)**		**NHS Reference Cost 2016-17 (164), weighted average of the number of attendances and the unit cost of EL FF32B, FF32A and FF33A The IQR was determined as the range between the lowest and highest limit of the EL FF32B, FF32A and FF33A costs.**

Abbreviations: CC: complication and comorbidity; EL: elective inpatient

B.3.6 Summary of base-case analysis inputs and assumptions

B.3.6.1 Summary of base-case analysis inputs

Table 59 Summary of variables applied in the economic model

Parameter	Variable	Mean or median value	Precision around the mean / median	Probabilistic distribution and parameterisation	Reference to section in submission
Model parameters					
Model settings	Discount rate (costs and effects)	3.50%	Fixed	No sampling	B.3.2.2 (p.106)
Patient characteristics - biologic-naïve population	Age	41.1 years	Fixed	No sampling	B.3.2.1 (p.104)
	Weight	74.8 kg			
	Female	40.30%			
	Time since diagnosis	8.16 years			
Patient characteristics - biologic-exposed population	Age	41.3 years	Fixed	No sampling	
	Weight	72.6 kg			
	Female	59%			
	Time since diagnosis	8.16 years			
Efficacy and safety					
Induction remission cut-off; probit scale*	Biologic-naïve	[redacted]	[redacted]	Direct use of NMA output (CODA)	B.3.3.1 (p. 120)
	Biologic-exposed	[redacted]	[redacted]		
Induction response and remission in biologic-naïve population, probit scale*	Anchor	[redacted]	[redacted]	Direct use of NMA output (CODA)	
	Adalimumab	[redacted]	[redacted]		
	Golimumab	[redacted]	[redacted]		
	Infliximab	[redacted]	[redacted]		
	Conventional (Placebo)	[redacted]	[redacted]		
	Tofacitinib	[redacted]	[redacted]		
	Vedolizumab	[redacted]	[redacted]		
Induction response and remission in biologic-exposed population, probit scale*	Anchor	[redacted]	[redacted]	Direct use of NMA output (CODA)	
	Conventional (Placebo)	[redacted]	[redacted]		
	Tofacitinib	[redacted]	[redacted]		
	Vedolizumab	[redacted]	[redacted]		
Maintenance remission cut-off; probit scale*	Biologic-naïve population	[redacted]	[redacted]	Direct use of NMA output (CODA)	
	Biologic-exposed	[redacted]	[redacted]		
Maintenance response and remission in biologic-naïve population, probit scale*	Anchor	[redacted]	[redacted]	Direct use of NMA output (CODA)	
	Adalimumab	[redacted]	[redacted]		
	Golimumab 50 mg	[redacted]	[redacted]		
	Golimumab 100 mg	[redacted]	[redacted]		
	Infliximab	[redacted]	[redacted]		
	Conventional (placebo)	[redacted]	[redacted]		
	Tofacitinib 5 mg	[redacted]	[redacted]		
	Tofacitinib 10 mg	[redacted]	[redacted]		

Parameter	Variable	Mean or median value	Precision around the mean / median	Probabilistic distribution and parameterisation	Reference to section in submission
	Vedolizumab Q8W	[redacted]	[redacted]		
	Vedolizumab Q4W	[redacted]	[redacted]		
Maintenance response and remission in biologic-exposed population, probit scale*	Anchor	[redacted]	[redacted]	Direct use of NMA output (CODA)	
	Conventional (placebo)	[redacted]	[redacted]		
	Tofacitinib 5 mg	[redacted]	[redacted]		
	Tofacitinib 10 mg	[redacted]	[redacted]		
	Vedolizumab Q8W	[redacted]	[redacted]		
	Vedolizumab Q4W	[redacted]	[redacted]		
Colectomy rates	Elective colectomy rate	0.00058	0 to 0.0026	Beta using the number of observations (n) and the total sample (N)	B.3.3.2.4 (p.31)
	Emergency colectomy rate	0.00021	0 to 0.0026		
Perioperative complication and mortality	Perioperative mortality risk	0.02843	0 to 0.0347	Beta using the number of observations (n) and the total sample (N)	B.3.3.2.5
	Perioperative elective surgery complications	0.3167	0 to 0.64		
	Perioperative emergency surgery complications	0.347	0 to 0.7		
Post-surgery complications	Long-term complications	0.01458	0.0074 to 0.0218	Beta using the number of observations (n) and the total sample (N)	B.3.3.2.6
Serious infection Treatment effects, probit scale (median)*	Placebo	[redacted]	[redacted]	Direct use of NMA output (CODA)	B.3.3.3 (p.41)
	Tofacitinib	[redacted]	[redacted]		
Utility					
EQ-5D weighted average per health states (Woehl 2008 with time dependent multiplier)	Active UC	0.41	0.36 to 0.46	Beta using the standard the precision parameters from the sources	B.3.4 (p.132)
	Response-no-remission	0.76	0.73 to 0.79		
	Remission	0.87	0.85 to 0.89		
	Post-colectomy without long-term complications	0.71	0.67 to 0.75		
	Post-colectomy with long- term complications	0.56	0.41 to 0.71		
Adverse events and post-surgery complications reduction in utility weight	Post-surgery complication reduction	0.7889	0.5778 to 1		
	Serious infection	0.9858	0.9716 to 1		
Cost and resource use					

Parameter	Variable	Mean or median value	Precision around the mean / median	Probabilistic distribution and parameterisation	Reference to section in submission
Drug costs (PAS price for tofacitinib and golimumab, list prices for other therapies)	Adalimumab	£352.14 per dose	Fixed	No sampling	B.3.5.1.1, Table 45 (p.136)
	Golimumab	£762.97 per dose			
	Infliximab (Remicade)	£419.62 per dose			
	Infliximab (biosimilars)	£377.66 per dose			
	Tofacitinib	[redacted]			
	Vedolizumab	£2,050.00 per dose			
Conventional therapy drug costs	Balsalazide	£0.23 per dose	Fixed	No sampling	B.3.5.1.2, Table 47 (p.140)
	Mesalazine	£0.33 per dose			
	Olsalazine	£2.68 per dose			
	Sulfasalazine	£0.06 per dose			
	Prednisolone	£0.03 per dose			
	Hydrocortisone	£0.67 per dose			
	Azathioprine	£0.04 per dose			
Administration costs	Infliximab and vedolizumab	£137.37	£70.2 to £161.72	Gamma using the interquartile range	B.3.5.1.3 (p.141)
Healthcare resource use costs	Outpatient visit (specialist)	£137.37	£70.20 to £161.72		B.3.5.2 (p.142)
	Blood tests	£3.06	£2.22 to £3.60		
	Endoscopy	£277.29	£149.39 to £399.65		
	Hospitalisation episodes	£2,984.71	£2,381.80 to £3,434.28		
Outpatient visit (specialist)	Consultant led (CL) - Non-Admitted Face-to-Face Attendance, Follow-up	£141	£108.33 to £161.72		B.3.5.1.3 (p.141)
	Non-consultant led (NCL) - Non-Admitted Face-to-Face Attendance, Follow-up	£107	£70.2 to £127.2		
Resource use (per year): active UC	Outpatient visit (specialist)	6.5	4.5 to 8.5	No sampling	B.3.5.2 (p.142)
	Blood tests	6.5	3.9 to 9.1		
	Endoscopy	2	0.5 to 3.5		
	Hospitalisation episodes	1.5	1.2 to 1.8		
Resource use (per year): response-no-remission	Outpatient visit (specialist)	4.5	2 to 6.5		
	Blood tests	3.9	3.25 to 6.5		
	Endoscopy	0.5	0.2 to 2		
	Hospitalisation episodes	1.2	0.3 to 1.5		
Resource use (per year): remission	Outpatient visit (specialist)	2	0 to 4.5		
	Blood tests	3.25	0 to 3.9		

Parameter	Variable	Mean or median value	Precision around the mean / median	Probabilistic distribution and parameterisation	Reference to section in submission
	Endoscopy	0.2	0 to 0.5		
	Hospitalisation episodes	0.3	0 to 1.2		
Resource use (per year): post-colectomy without complications	Outpatient visit (specialist)	1.5	0 to 4.50		
	Blood tests	1.5	0 to 3.90		
	Endoscopy	1.25	0 to 1.30		
	Hospitalisation episodes	0	0 to 1.20		
Resource use (per year): post-colectomy with complications	Outpatient visit (specialist)	1.75	1.5 to 2		
	Blood tests	3.25	1.5 to 5		
	Endoscopy	0.65	0 to 1.3		
	Hospitalisation episodes	3.25	0 to 6.5		
Adverse event costs (per event)	Serious infection	£2,538.97	£1,078.02 to £11,470.56	Gamma using the lowest and highest interquartile range values of all the relevant events	B.3.5.3 (p.143)
Colectomy operation costs without complications	Elective inpatient costs for proximal colon procedures	£6,085.07	£5,071.814 to £6,688.2	Gamma using the lowest and highest interquartile range values of all the relevant codes	B.3.5.4 (p.144)
	Elective inpatient costs for distal colon procedures	£6,103.58	£4,994.07 to £7,112.52		
	Colectomy operation costs without complications (weighted average of proximal and distal colon procedures)	£6,090.52	£4,994.07 to £6,688.2		
Colectomy operation costs with complications	Elective inpatient costs for proximal colon procedures (CC score 3-5)	£6,803.58	£5,544.42 to £7,663.82	Gamma using the lowest and highest interquartile range values of all the relevant codes; truncated to not exceed the probabilistic value without complications	
	Elective inpatient costs for proximal colon procedures CC score 6+)	£8,484.21	£6,241.99 to £9,742.12		
	Elective inpatient costs for distal colon procedures CC score 3+)	£7,557.94	£5,677.29 to £8,628.36		
	Colectomy operation costs with complication (weighted average of proximal and distal colon procedures)	£7,294.56	£5,544.42 to £9,742.12		

B.3.6.2 Assumptions

Table 60 Assumptions in the base case analysis

Parameter	Assumptions	Consistent with prior TAs?	Justification
Time horizon	Lifetime	Yes	UC is a chronic condition; a patient lifetime time horizon allows the calculation of all relevant costs and quality of life impairment.
Cycle length	8-weeks	Yes	A fixed cycle length of 8 weeks was assumed for the duration of the model time horizon to allow for a continuous sequence of treatments.
Treatment efficacy	Clinical response and clinical remission	Yes	Used in the clinical trials and in clinical practice and consistent with Archer 2016 (118)
Dis-continuation	Discontinuation due to lack of efficacy (response)	Yes	Consistent with clinical practice and NICE recommendations (B.3.3.5)
Health states	Defined by response and/or remission; Patients who responded to treatment were separated to remission and response-no-remission. The model assumed that the observed allocation at one-year of treatment remained the same in all consecutive cycles.	Yes	Defined based on the clinical trials used to reflect treatment efficacy in the economic analysis.
Elective surgery	A proportion of patients that do not respond or discontinue conventional treatment will undergo colectomy	Yes	Consistent with clinical practice (B.3.3.5)
Emergency surgery	In remission patients are protected from exacerbation; from all other states patients may undergo emergency surgery to manage the exacerbations	Yes	Consistent with clinical practice (B.3.3.5)
Background mortality	UC treatment was assumed not to have any effect on overall mortality.	Yes	Additional risks of death excluded as they are very likely to be small (118).
Adverse events	Serious infections were included in the base-case analysis using the output of the NMA. The range of events with tofacitinib was assumed to be from 0 to a 50% increase from the base-case value	No	Serious infection is an event often associated with immunosuppressive treatments. Serious infection is a rare adverse event leading to a large 95% CrI in the NMA. This led to misleading results in the sensitivity analysis. Instead the range 0 – 50% increase was used for tofacitinib. The NMA output (95% CrI limits) was used in the PSA.
Perioperative risk of complications and mortality	Perioperative complications were included for patients undergoing elective and non-elective surgery	Yes	Consistent with clinical practice and Archer 2016 (118).
Surgery long-term complications	The model assumed pouchitis reflects the long-term complications	Yes	Consistent with Archer 2016 (118).
Continuation of conventional treatment	Patients in the conventional treatment group and those who have previously achieved but lost response to biological therapy were assumed to continue receiving conventional therapy irrespective of whether they achieve response to that conventional therapy	Yes	Simplifying assumption; consistent with Archer

Parameter	Assumptions	Consistent with prior TAs?	Justification
Time-independent risk of surgery	A time-independent probability was considered as an appropriate assumption in the analysis	Yes	Evidence in the literature combined with the population in the model (>3 years duration of disease)
Baseline utility	The model health state utilities were synthesised with a baseline utility.	No	This was assumed to reflect the natural decline of patients' physical and mental functions due to age and other co-morbidities and reflects the chronic nature of the condition
Biologic treatments	Golimumab formulation	Yes	It was assumed that the 100 mg vials of golimumab were used in induction (2x100 mg vial at week 0 and 1x100 mg vial at week 2) and the 50 mg vials were used for the maintenance dose (1x50 mg vial Q4W)
Conventional therapy mix	The RCP audit data for concomitant medication at treatment initiation was used	No	Assumption in absence of evidence on the conventional treatment mix. The evidence at treatment initiation were assumed to be reflective of active UC
Conventional therapy treatments	Hydrocortisone was considered as a topical treatment (rectal foam); prednisolone was assumed to represent the oral corticosteroid treatment group as the proportion of use of budesonide is low and beclomethasone is used as an add-on treatment to 5-ASA.	No	Simplifying assumption
	Azathioprine was assumed to represent the immunomodulator group	No	
Concomitant medication	The proportion of use of conventional treatments as concomitant treatments to biologics at 3-months follow-up was used. Azathioprine was excluded from concomitant use with tofacitinib	Yes	The evidence at 3-months follow-up were assumed to be reflective of continuous concomitant use.
Administration cost for injections	No administration cost was assumed for self-injection treatment	Yes	Consistent with clinical practice
Hospitalisation	An increased frequency of hospitalisation was assumed with severity of condition	No	Consistent with clinical practice (B.3.3.5)
Cost of serious infection	The cost of a serious infection was considered to be a weighted average of six types of infection: sepsis, tuberculosis, pneumonia, skin and soft tissue infection, bone and joint infection and urinary tract infection	No	Simplifying assumption in the absence of other evidence

Abbreviations: 5-ASA: 5-aminosalicylates; KOL: key opinion leader; NHS: national health system; QoL: quality of life; RCP: royal college of physicians; TNF: tumour necrosis factor; UC: ulcerative colitis

B.3.7 Base-case results

The economic analysis results are presented below for biologic-naïve and biologic-exposed patients.

All results include tofacitinib confidential discount.

B.3.7.1 Base-case incremental cost-effectiveness analysis results

Clinical outcomes from the model and disaggregated results of the base-case incremental cost-effectiveness analysis are presented in Appendix J.

B.3.7.1.1 Biologic-naïve patients

For biologic-naïve patients, in the incremental cost-effectiveness analysis (Table 61), tofacitinib dominated adalimumab, golimumab and infliximab. The ICER of tofacitinib compared to conventional treatment was £8,554.03 per QALY. When compared with tofacitinib, vedolizumab generated an additional [redacted] QALYs, with an ICER of £615,056.62 per QALY.

Table 61 Biologic-naïve patients: full incremental cost-effectiveness results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental	ICER (£/QALY) TOF vs comparator
	QALYs	Costs (£)	QALYs	Costs (£)		
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£8,554.03
Adalimumab	[redacted]	[redacted]	-	-	Dominated	Dominated
Golimumab	[redacted]	[redacted]	-	-	Dominated	Dominated
Infliximab	[redacted]	[redacted]	-	-	Dominated	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£8,554.03	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£615,056.62	£615,056.62

Abbreviations: ICER, incremental cost-effectiveness ratio; N/A, not applicable; QALY, quality-adjusted life year; TOF: tofacitinib.

B.3.7.1.2 Biologic experienced patients

For biologic experienced patients (Table 62), the ICER for tofacitinib ICER compared with conventional treatment was £10,301.85 per QALY. Compared with tofacitinib, vedolizumab generated a marginal additional [redacted] QALYs, with an ICER of £7.8 million per QALY.

Table 62 Biologic-exposed patients: full incremental cost-effectiveness results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental	ICER (£/QALY) TOF vs comparator
	QALYs	Costs (£)	QALYs	Costs (£)		
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£10,301.85
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£10,301.85	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£7,838,238.48	£7,838,238.48

Abbreviations: ICER, incremental cost-effectiveness ratio; N/A, not applicable; QALY, quality-adjusted life year; TOF: tofacitinib.

B.3.7.2 Key scenarios

B.3.7.2.1 Intention-to-treat population analysis

The ITT population NMA results were considered in this scenario. When comparing the ITT population across all clinical trials, the studies for tofacitinib and vedolizumab showed similarities in the distribution of TNFi-naïve and exposed patients. As such tofacitinb compared with vedolizumab represented the least confounded pair when considering clinical trial design, trial population characteristics and feasibility of the evidence network to allow an ITT scenario analysis. Cost-effectiveness results are presented in Table 65 and deterministic sensitivity analysis results in Table 66.

In the ITT population, tofacitinib was associated with an ICER of £7,805.06 per QALY when compared with conventional therapy, while vedolizumab was dominated by tofacitinib, this result was robust to parameters variation in the one way sensitivity analysis (Table 64).

Table 63. Overall ITT population: deterministic results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental
	QALYs	Costs (£)	QALYs	Costs (£)	
Conventional	[redacted]	[redacted]	N/A	N/A	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£7,805.06

Abbreviations: ICER, incremental cost-effectiveness ratio; N/A, not applicable; QALY, quality-adjusted life year.

Table 64. Overall ITT population: sensitivity analysis results (ICER vedolizumab versus tofacitinib)

	ICER (£/QALY)	
Parameter	**Low limit**	**High limit**
Response/remission treatment effect - induction	£4,153,408.22	Dominated
Remission (z) - induction	Dominated	Dominated
Response/remission treatment effect - maintenance	Dominated	Dominated
Remission (z) - maintenance	Dominated	Dominated
Colectomy risk (No risk - Frolkis 10y)	Dominated	Dominated
Periorative mortality risk (0 - 3%)	Dominated	Dominated
Periorative complications (No risk - double the risk)	Dominated	Dominated
Post-operative pouchitis (0.7 - 2%)	Dominated	Dominated
Serious infection risk	Dominated	Dominated
Post-surgery complication utility weight reduction (0% - 40%)	Dominated	Dominated
Pre-surgery health state utilities	Dominated	Dominated
Post-surgery health state utilities	Dominated	Dominated
Serious infection utility reduction (0% - 3%)	Dominated	Dominated
Health-state related resource use per patient per year	Dominated	Dominated
Conventional treatment costs (min-max)	Dominated	Dominated
OP administration cost (£70 - £161)	Dominated	Dominated
OP visit + blood test costs	Dominated	Dominated
Hospitalisation cost	Dominated	Dominated
Endoscopy cost	Dominated	Dominated

	ICER (£/QALY)	
Parameter	**Low limit**	**High limit**
Colectomy cost	Dominated	Dominated
Serious infection costs	Dominated	Dominated

Abbreviations: ICER, incremental cost-effectiveness ratio; OP, outpatient; QALY, quality adjusted life year.

B.3.7.2.2 Tofacitinib maintenance dose mix

Clinical trials evaluated maintenance doses of 10 mg and 5 mg. This scenario analysis assumed that [redacted] of the patients received 5 mg maintenance dose and [redacted] received 10 mg (see section B.3.10.1 and Appendix M.7). The results of this scenario are presented for both naïve and prior-exposed populations.

B.3.7.2.2.1 TNFi-naïve population

In the biologic naïve population, when assuming a mix maintenance dose of 5 and 10 mg ([redacted]/[redacted]), the total QALYs and costs for tofacitinib increased (Table 65). Tofacitinib dominated all biologic comparators and was associated with an ICER of £12,627.81 per QALY versus conventional therapy.

Table 65. Biologic naïve population: tofacitinib [redacted]/[redacted] maintenance dose mix

Strategy	Total		Incremental		ICER (£/QALY) fully incremental
	QALYs	Costs (£)	QALYs	Costs (£)	
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£12,627.81

Abbreviations: ICER, incremental cost-effectiveness ratio; N/A, not applicable; QALY, quality-adjusted life year.

B.3.7.2.2.2 TNFi-exposed population

The deterministic results of this analysis are presented in Table 66. Tofacitinib was associated with an ICER of £13,946.75 per QALY versus conventional therapy and dominated vedolizumab.

Table 66. Prior exposed population: tofacitinib [redacted]/[redacted] maintenance dose mix

Strategy	Total		Incremental		ICER (£/QALY) fully incremental
	QALYs	Costs (£)	QALYs	Costs (£)	
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£13,946.75

Abbreviations: ICER, incremental cost-effectiveness ratio; N/A, not applicable; QALY, quality-adjusted life year.

B.3.8 Sensitivity analyses

B.3.8.1 Summary of sensitivity analysis inputs

One-way sensitivity analysis was undertaken to assess the impact of key variables on the model outcomes. Mean values and limits tested in the deterministic sensitivity analysis are presented in the summary of the base-case analysis (Table 59). A brief summary is presented below.

B.3.8.1.1 Patient proportions and probabilities

Efficacy and safety parameters were tested using the high and low limits of the 95% CrI from the NMA.

The risk of colectomy was explored considering the absence of risk as the lower limit and the 10-year colectomy risk of 0.7% reported in Frolkis et al. 2013 (172). The perioperative mortality risk parameter was explored between no risk to a risk of 3.4% as reported in the IBD audit 2012 (193). The risk of perioperative complications was tested from no risk to a 50% increase of the mean value. The lower limit considered for the post-operative pouchitis risk parameter was the value reported in Suzuki *et al.* 2012 (175) and the highest limit assumed a 50% increase of the mean risk. The risk of serious infection with tofacitinib was explored from no risk to a 50% increased risk, to avoid misleading results in the sensitivity analysis due to the uncertainty around the median value generated from the NMA (see section B.3.3.3).

B.3.8.1.2 Utility weights

Pre- and post-surgery health state utilities were varied around the mean value reported in Woehl *et al.* 2008. Each health state utility was varied around the 95% CI values, calculated from the standard deviation (SD) and the total sample (N). The utility reduction associated with serious infection was explored from a 0% (1) to a 3% (0.9716) reduction. The reduction of utility weights due to post-surgery complications tested with lower and higher limits determined from Kosmas et al. 2015 (0% to 40% reduction of the mean value) (38).

B.3.8.1.3 Costs and resource use

In the active disease health state, the lower limit of each health care resource use value was assumed to be similar to the mean resource use in the response-no-remission health state. To calculate the upper limit, the distance between the lower limit and the mean was used as a proxy.

In the response-no-remission health state, the low and high limits were determined by the mean values in remission and active UC.

In the remission health state, the low limit assumed no resource and the high limit used the response-no-remission mean.

The resource use associated with post-colectomy without complications was varied from no resource use to the mean resource use in the response-no-remission health state with the

exception of endoscopy which was assumed to be equal to the upper limit of the post-colectomy with complication endoscopy resource use.

The lower resource use values for post colectomy with complications was assumed to be equal to the mean resource use of the post-colectomy health state without complications. To derive the higher limit, we assumed a value equal to the distance between the mean and the lower limit.

The cost of conventional therapy was tested assuming a range of costs from the lowest to the highest. That is, all patients receive the treatment with the lowest cost per 8 weeks and all receive the treatment with the highest cost per 8 weeks.

All health care resource use associated costs were tested around the mean value considering the interquartile range reported in the NHS reference costs 2016-17 (164).

B.3.8.2 Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was undertaken with 1,000 samples. The stability of the sample results was tested and is presented in Appendix M.6; stability of results was achieved at approximately 400 samples onwards.

A full list of all parameters included in the PSA is presented in section B.3.6.1, Table . Uncertainty in the response and remission estimates was incorporated by using 1,000 of the simulated treatment effects from the NMA. By using the stimulated outputs, the joint posterior distribution and any correlation of treatment effects from the NMA were preserved in the analysis. For the remaining variables probability distributions were based on precision estimates from the data sources, such as confidence interval, interquartile ranges and the like.

A summary of the probabilistic results is presented in Table 67 and Table 68 for the biologic-naïve and exposed groups respectively, and included tofacitinib's confidential discount. In the probabilistic sensitivity analysis, we observed an increase in the total QALYs and costs of all the biologic comparators and tofacitinib, compared with the results of the base-case deterministic analysis. The differences between the probabilistic and deterministic results were attributed to the use of 1,000 random values from NMA output (Coda). The average of 1,000 values, used in PSA, was different to the median of all the NMA samples, used in the base-case. Nevertheless, the conclusions from the PSA results remained broadly the same with the deterministic base-case results (B.3.7).

In the biologic-naïve population, tofacitinib dominated the TNFi treatments. The ICER of tofacitinib in the comparison with conventional therapy was £5,433.94 per QALY. The ICER of vedolizumab was £424,327.10 per QALY when compared to tofacitinib. In the prior exposed population, tofacitinib had an ICER of £10,926.30 per QALY compared with conventional therapy. The total QALYs of tofacitinib and vedolizumab were marginally different in the deterministic sensitivity analysis; with the increase in the total QALYs and costs in the probabilistic analysis, the cost-effectiveness of vedolizumab changed and it was dominated by tofacitinib.

A graphical representation of the simulations is presented in Figure 33 and Figure 34 for the biologic-naïve and exposed groups respectively. The multiple cost-effectiveness acceptability curves are presented in Figure 35 and Figure 36.

In the biologic-naïve population, at a willingness-to-pay threshold of £20,000 per QALY gained, tofacitinib had the highest probability of being cost-effective (80.5%) followed by conventional therapy (13.7%). At £30,000, tofacitinib had a 87.1% probability of being the most cost-effective of the treatment options.

In the prior exposed population, tofacitinib had the highest probability of being cost-effective (56.3%) followed by conventional therapy (43.1%) while vedolizumab had the lowest probability to be cost-effective (0.6%) at a £20,000 per QALY threshold. Tofacitinib still remained to have the highest likelihood to be cost-effective at a £30,000 per QALY gained threshold (70.5%).

Table 67 Biologic-naïve patients: results of probabilistic sensitivity analysis

Strategy	Total		Incremental		ICER (£/QALY) fully incremental	ICER (£/QALY) TOF vs comparator,
	QALYs	Costs (£)	QALYs	Costs (£)		
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£5,433.94
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£5,433.94	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£424,327.10	£424,327.10

Abbreviations: ICER, incremental cost-effectiveness ratio, as cost per QALY; N/A, not applicable; QALY, quality-adjusted life year; TOF, tofacitinib.

Table 68 Biologic-exposed patients: results of probabilistic sensitivity analysis

Strategy	Total		Incremental		ICER (£/QALY) fully incremental	ICER (£/QALY) TOF vs comparator,
	QALYs	Costs (£)	QALYs	Costs (£)		
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£10,926.30
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£10,926.30	N/A

Abbreviations: ICER, incremental cost-effectiveness ratio, as cost per QALY; N/A, not applicable; QALY, quality-adjusted life year; TOF, tofacitinib.

Figure 33 Biologic-naïve population: PSA Scatterplot on cost-effectiveness plane

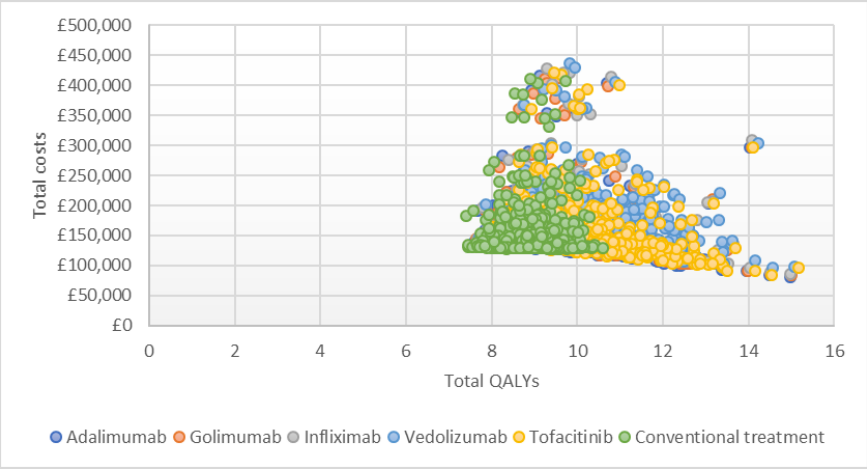
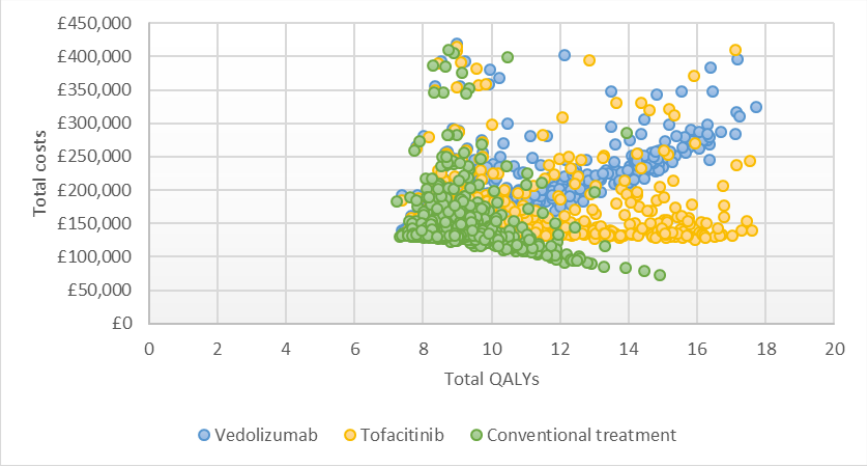


Figure 34 Biologic-exposed population: PSA Scatterplot on cost-effectiveness plane



Abbreviations: QALYs, quality-adjusted life years.

Figure 35 Biologic-naïve population: cost-effectiveness acceptability curve

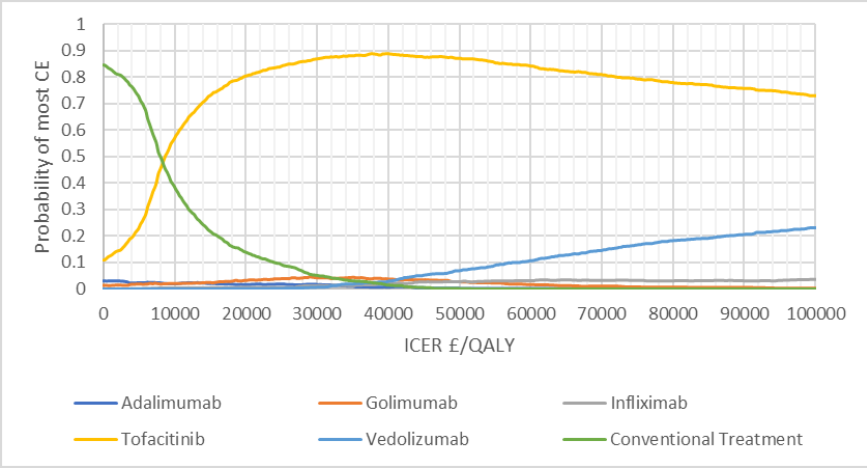
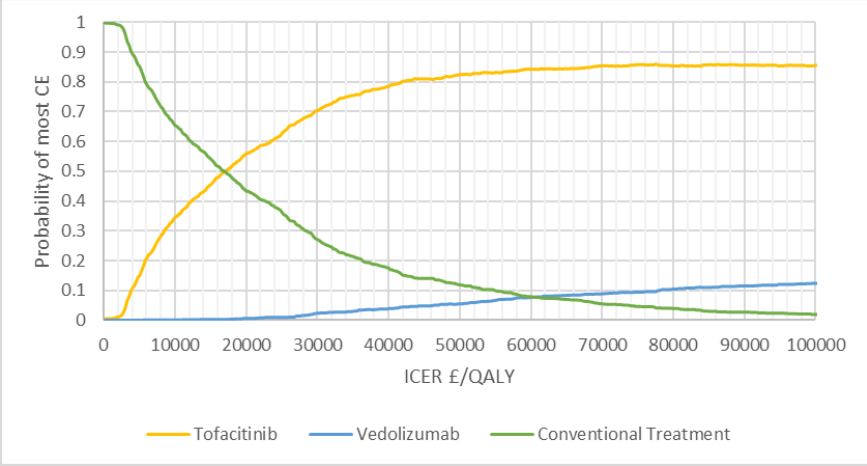


Figure 36 Biologic-exposed population: cost-effectiveness acceptability curve



Abbreviations: CE, cost-effective; ICER, incremental cost-effectiveness ratio, as cost per QALY; QALY, quality-adjusted life year.

B.3.8.3 Deterministic sensitivity analysis

The results of the deterministic sensitivity analysis are presented with a tornado diagram for the comparison of tofacitinib versus conventional treatment.

B.3.8.3.1 Biologic-naïve patients

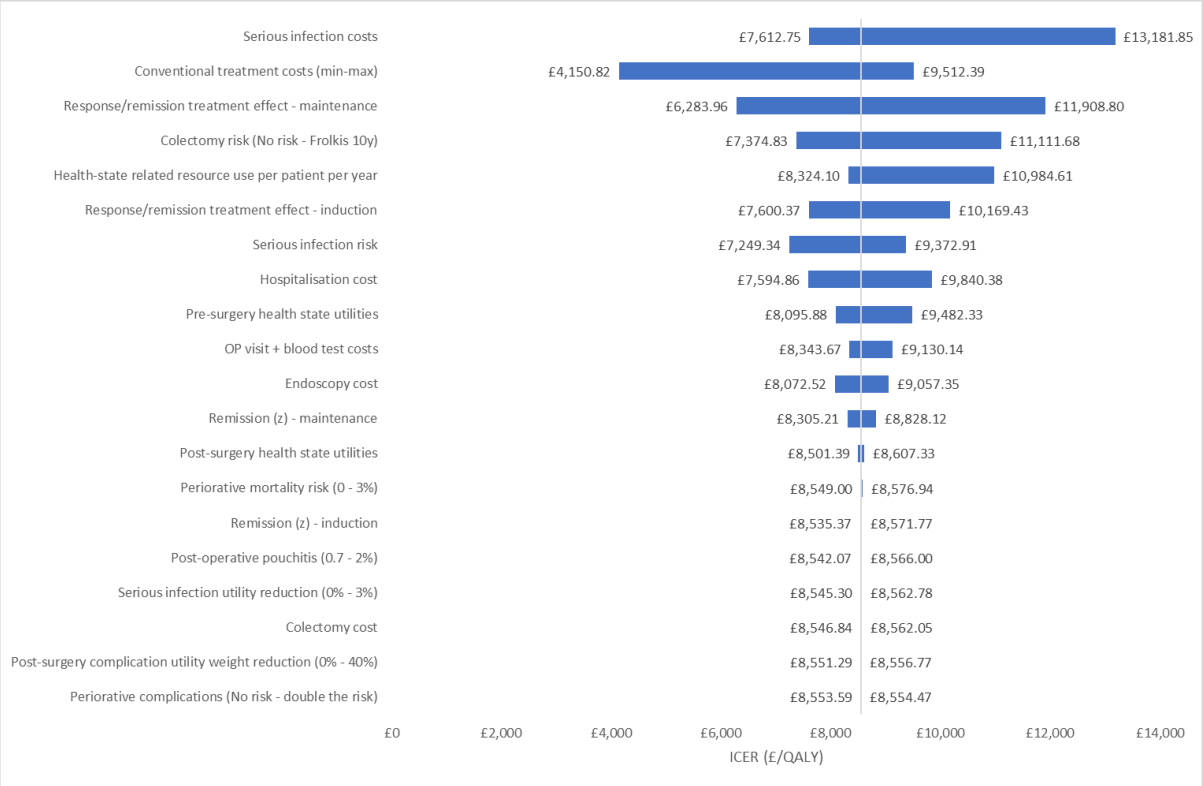
Deterministic sensitivity analysis results for biologic-naïve patients are presented in Table 69 and Figure 37. The ICER was most sensitive to changes in the serious infections cost, the comparator cost (conventional treatment) and the response estimates during the maintenance phase.

Table 69 Biologic-naïve deterministic sensitivity analysis results (comparison with conventional therapy)

	ICER (£/QALY)	
Base case	£8,554.03	
Parameter	**Low limit**	**High limit**
Response/remission treatment effect - induction	£7,600.37	£10,169.43
Remission (z) – induction	£8,535.37	£8,571.77
Response/remission treatment effect - maintenance	£6,283.96	£11,908.80
Remission (z) – maintenance	£8,305.21	£8,828.12
Colectomy risk (No risk - Frolkis 10y)	£7,374.83	£11,111.68
Perioperative mortality risk (0 - 3%)	£8,576.94	£8,549.00
Perioperative complications (No risk - double the risk)	£8,553.59	£8,554.47
Post-operative pouchitis (0.7 - 2%)	£8,566.00	£8,542.07
Serious infection risk	£7,249.34	£9,372.91
Post-surgery complication utility weight reduction (0% - 40%)	£8,556.77	£8,551.29
Pre-surgery health state utilities	£8,095.88	£9,482.33
Post-surgery health state utilities	£8,501.39	£8,607.33
Serious infection utility reduction (0% - 3%)	£8,545.30	£8,562.78
Health state related resource use per patient per year	£8,324.10	£10,984.61
Conventional treatment costs (min-max)	£9,512.39	£4,150.82
OP administration cost (£70 - £161)	£8,554.03	£8,554.03
OP visit + blood test costs	£9,130.14	£8,343.67
Hospitalisation cost	£9,840.38	£7,594.86
Endoscopy cost	£9,057.35	£8,072.52
Colectomy cost	£8,562.05	£8,546.84
Serious infection costs	£7,612.75	£13,181.85

Abbreviations: INMB, incremental net monetary benefit; OP, outpatient.

Figure 37 Biologic-naïve patients: Tornado diagram of deterministic sensitivity analysis (comparison with conventional therapy)



Abbreviations: ICER, incremental cost-effectiveness ratio; OP, outpatient; QALY, quality adjusted life year.

B.3.8.3.2 Biologic experienced patients

Deterministic sensitivity analysis results for biologic experienced patients are presented in Table 70 and Figure 38. The model sensitivity was similar to that observed in the biologic-naïve population analysis.

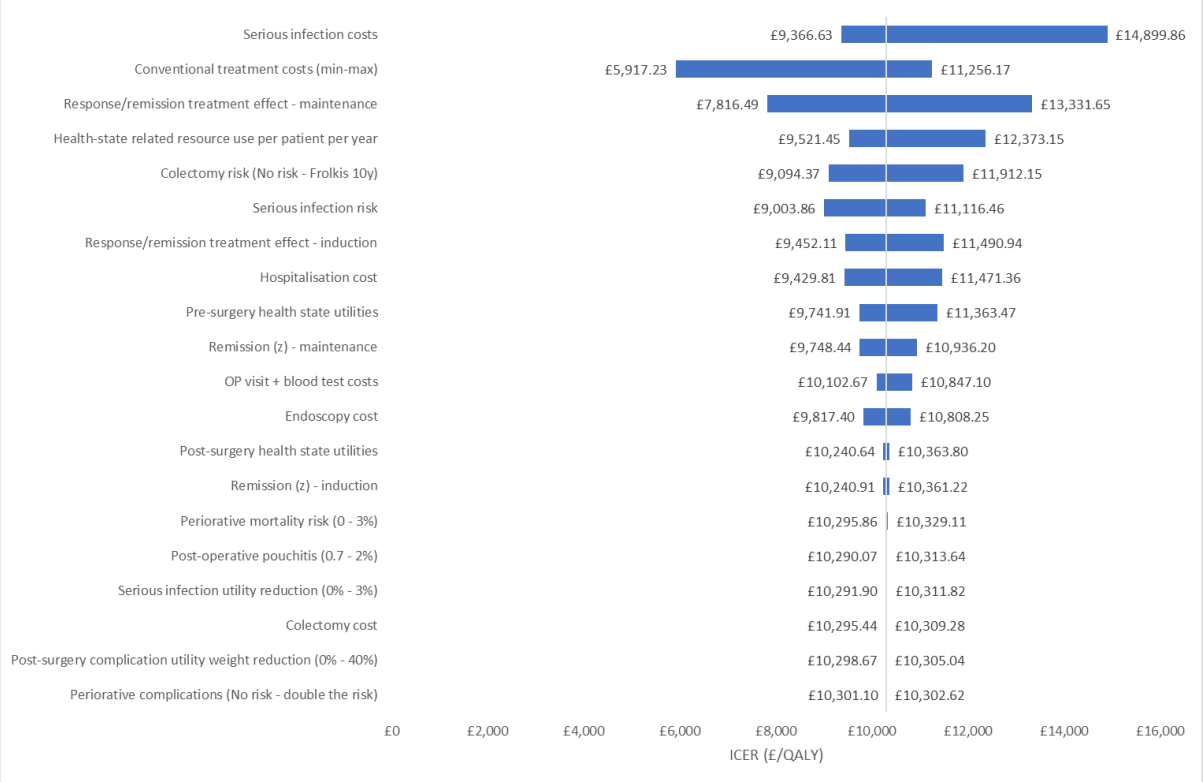
Table 70 Biologic-exposed deterministic sensitivity analysis results (comparison with conventional treatment)

	ICER (£/QALY)	
Base case	£10,301.85	
Parameter	**Low limit**	**High limit**
Response/remission treatment effect - induction	£9,452.11	£11,490.94
Remission (z) - induction	£10,240.91	£10,361.22
Response/remission treatment effect - maintenance	£7,816.49	£13,331.65
Remission (z) - maintenance	£9,748.44	£10,936.20
Colectomy risk (No risk - Frolkis 10y)	£9,094.37	£11,912.15
Perioperative mortality risk (0 - 3%)	£10,329.11	£10,295.86
Perioperative complications (No risk - double the risk)	£10,301.10	£10,302.62
Post-operative pouchitis (0.7 - 2%)	£10,313.64	£10,290.07
Serious infection risk	£9,003.86	£11,116.46
Post-surgery complication utility weight reduction (0% - 40%)	£10,305.04	£10,298.67

	ICER (£/QALY)	
Base case	£10,301.85	
Parameter	**Low limit**	**High limit**
Pre-surgery health state utilities	£9,741.91	£11,363.47
Post-surgery health state utilities	£10,240.64	£10,363.80
Serious infection utility reduction (0% - 3%)	£10,291.90	£10,311.82
Health state related resource use per patient per year	£9,521.45	£12,373.15
Conventional treatment costs (min-max)	£11,256.17	£5,917.23
OP administration cost (£70 - £161)	£10,301.85	£10,301.85
OP visit + blood test costs	£10,847.10	£10,102.67
Hospitalisation cost	£11,471.36	£9,429.81
Endoscopy cost	£10,808.25	£9,817.40
Colectomy cost	£10,309.28	£10,295.44
Serious infection costs	£9,366.63	£14,899.86

Abbreviations: INMB, incremental net monetary benefit; OP, outpatient.

Figure 38 Biologic-exposed patients: Tornado diagram of deterministic sensitivity analysis (comparison with conventional therapy)



Abbreviations: ICER, incremental cost-effectiveness ratio; OP, outpatient; QALY, quality adjusted life year.

B.3.8.4 Additional scenario analyses

Further scenario analyses were undertaken to assess the impact of key variables on the model outcomes (Table 71). Note that these analyses were not part of the tornado diagram because they did not reflect a range around the base-case values.

Table 71 Summary of scenarios explored

Scenario and cross reference		Scenario detail	Brief rationale
1	Used a fixed baseline utility instead of age-adjusted (B.3.4.5.1, page 137).	This scenario used a fixed baseline utility from Woehl et al. 2008 (remission value) (158)	This scenario tested the sensitivity of the model on the assumption that patient quality of life stays constant over time.
2	Used the OCTAVE trials utility weights (B.3.4.1, page 134)	This scenario considered the EQ-5D utility from the OCTAVE clinical trials	EQ-5D data were collected as part of the Phase III clinical trials of tofacitinib (98, 100-102).
3	Used Swinburn 2012 utility weights (B.3.4.3, page 135)	This scenario considered the EQ-5D utility from the Swinburn 2012 study	Swinburn *et al.* 2012 (178) health state utilities were considered in Archer cost-effectiveness analysis (118). It was included here for comparison with previous analyses.
4	Acute exacerbations/emergency surgery from any state (B.3.2.2, page 117)	Assumed that patients can undergo emergency surgery from any health state	Due to the uncertainty on the likely protection from acute events based on the level of response/remission, this scenario explored the assumption that patients can undergo emergency surgery regardless of state membership
5	Acute exacerbations/emergency surgery from active UC (B.3.2.2, page 117)	Assumed that patients can undergo emergency surgery only in the active disease health state	As above but assuming response to treatment offers the same level of protection from acute events, as remission
6	No acute exacerbations / emergency surgery (B.3.2.2, page 117)	This scenario explored the absence of emergency surgery	As above, but assuming no emergency surgery in the model
7	Central read NMA results (B.2.9.4, page 96)	This scenario considered the central read outcomes (response and remission rates) from the OCTAVE trial program.	Central read was the primary endpoint in OCTAVE trials (98).
8	Discounting every cycle (B.3.2.2, page 117)	This scenario considered discounting effectiveness and costs every cycle instead of every year	It tested the sensitivity of the model when the discounting of outcomes is applied every 8 weeks.
9	Adalimumab maintenance dose mix of 73% 40 mg Q2W and 27% 40 mg QW (B.3.5.1.1, page 139)	This scenario considered that 73% of patients receive adalimumab 40 mg Q2W and 23% receive 40 mg QW. In the absence of any evidence for the efficacy of the increased dose, only the cost of adalimumab was changed.	Dose escalation of adalimumab was considered in Archer *et al*. (118)
10	Golimumab 100 mg Q4W in maintenance phase (B.3.5.1.1, page 139)	This scenario explored a maintenance dose of 100 mg of golimumab Q4W	A 100 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients, such as those who have experienced a

Scenario and cross reference		Scenario detail	Brief rationale
			decrease in their response
11	Vedolizumab 300 mg Q4W in maintenance phase (B.3.5.1.1, page 139)	This scenario explored a maintenance dose of 300 mg Q4W of vedolizumab	A 300 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients who have a body weight ≥ 80 kg

Abbreviations: EQ-5D, 5-dimension EuroQol questionnaire; ITT, intention to treat; NMA, network meta-analysis; QW, every week; Q2W, every 2 weeks; Q4W, every 4 weeks; UC, ulcerative colitis.

The results of the scenario analyses are presented in Table 72 and Table 76 for the TNFi-naïve and exposed populations, respectively. Regarding the dose escalation scenarios, the cost-effectiveness results for adalimumab, golimumab and vedolizumab are presented separately in Table 73, Table 74 and Table 75 for the TNFi-naïve patients. The results for vedolizumab are presented in Table 77 for the TNFi-exposed population.Overall in both populations, results were mainly sensitive to changes in utilities (scenarios 2 and 3).

Table 72 Biologic-naïve population: scenario analysis

		Incr. QALYs	Incr. costs (£)	ICER (£/QALY)
Base case		[redacted]	[redacted]	£8,554.03
Scenarios				
1	Change the baseline utility from age-adjusted to fixed (B.3.4.5.1, page 137).	[redacted]	[redacted]	£8,760.18
2	Use OCTAVE trials utility weights (B.3.4.1, page 134)	[redacted]	[redacted]	£15,507.53
3	Use Swinburn 2012 utility weights (B.3.4.3, page 135)	[redacted]	[redacted]	£11,931.99
4	Acute exacerbations/emergency surgery from any state (B.3.2.2, page 117)	[redacted]	[redacted]	£8,194.24
5	Acute exacerbations/emergency surgery from active UC (B.3.2.2, page 117)	[redacted]	[redacted]	£8,651.84
6	No acute exacerbations / emergency surgery (B.3.2.2, page 117)	[redacted]	[redacted]	£8,709.58
7	Central read NMA results (B.2.9.4, page 96)	[redacted]	[redacted]	£9,468.72
8	Discounting every cycle (B.3.2.2, page 117)	[redacted]	[redacted]	£8,606.29

Abbreviations: EOW, every other week; EW, every week; ICER, incremental cost-effectiveness ratio; Incr, incremental; ITT, intention to treat; NMA, network meta-analysis; Q4W, every four weeks; QALY, quality-adjusted life year; UC, ulcerative colitits

Table 73. Biologic naïve population: scenario 9, deterministic results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental
	QALYs	Costs (£)	QALYs	Costs (£)	
Conventional	[redacted]	[redacted]	N/A	N/A	–
Adalimumab	[redacted]	[redacted]	**–**	**–**	**Dominated**
Golimumab	[redacted]	[redacted]	–	–	Dominated
Infliximab	[redacted]	[redacted]	–	–	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£8,554.03
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£615,056.62

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

Table 74. Biologic naïve population: scenario 10, deterministic results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental
	QALYs	Costs (£)	QALYs	Costs (£)	
Conventional	[redacted]	[redacted]	N/A	N/A	–
Adalimumab	[redacted]	[redacted]	–	–	Dominated
Golimumab	[redacted]	[redacted]	**–**	**–**	**Dominated**
Infliximab	[redacted]	[redacted]	–	–	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£8,554.03
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£615,056

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

Table 75. Biologic naïve population: scenario 11, deterministic results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental
	QALYs	Costs (£)	QALYs	Costs (£)	
Conventional	[redacted]	[redacted]	N/A	N/A	–
Adalimumab	[redacted]	[redacted]	–	–	Dominated
Golimumab	[redacted]	[redacted]	–	–	Dominated
Infliximab	[redacted]	[redacted]	–	–	Dominated
Vedolizumab	[redacted]	[redacted]	**–**	**–**	**Dominated**
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£8,554.03

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

Table 76 Biologic-exposed population: scenario analysis

		Incr. QALYs	Incr. costs (£)	ICER (£ per QALY)
Base -case		[redacted]	[redacted]	£10,301.85
Scenarios				
1	Change the baseline utility from age-adjusted to fixed (B.3.4.5.1, page 137).	[redacted]	[redacted]	£10,589.16
2	Use OCTAVE trials utility weights (B.3.4.1, page 134)	[redacted]	[redacted]	£18,275.50
3	Use Swinburn 2012 utility weights (B.3.4.3, page 135)	[redacted]	[redacted]	£14,487.42
4	Acute exacerbations/emergency surgery from any state (B.3.2.2, page 117)	[redacted]	[redacted]	£9,961.81
5	Acute exacerbations/emergency surgery from active UC (B.3.2.2, page 117)	[redacted]	[redacted]	£10,475.41
6	No acute exacerbations / emergency surgery (B.3.2.2, page 117)	[redacted]	[redacted]	£10,593.24
7	Central read NMA results (B.2.9.4, page 96)	[redacted]	[redacted]	£10,792.84
8	Discounting every cycle (B.3.2.2, page 117)	[redacted]	[redacted]	£10,398.27

Abbreviations: EOW, every other weeks; EW, every weeks; ICER, incremental cost-effectiveness ratio; Incr, incremental; ITT, intention to treat; NMA, network meta-analysis; Q4W, every four weeks; QALY, quality-adjusted life year; UC, ulcerative colitits

Table 77. Biologic prior-exposed population: scenario 11, deterministic results

Strategy	Total		Incremental		ICER (£/QALY) vs conventional treatment	ICER (£/QALY) vs tofacitinib
	QALYs	Costs (£)	QALYs	Costs (£)		
Conventional	[redacted]	[redacted]	N/A	N/A	N/A	£10,301.85
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£10,301.85	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	**£84,579.85**	**Dominated**

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

B.3.8.5 Summary of sensitivity analyses results

Overall, the sensitivity analysis showed that the economic model results were robust across a range of input parameters and assumptions. As in the base-case analysis, the PSA cost-effectiveness frontier comprised conventional therapy and tofacitinib in both the biologic-naïve and prior exposed populations. The results of the PSA indicated that tofacitinib had a probability of more than 55% of being cost-effective at £20,000 per QALY threshold and a probability of over 70% at £30,000 per QALY threshold in the TNFi-exposed group. In the TNFi-naïve population, the probability of cost-effectiveness under these thresholds increased to 80% and 87% for £20,000 and £30,000 per QALY gained, respectively. The thresholds at which tofacitinib was estimated to be the most cost-effective strategy was £8,500 for the TNFi-naïve population and £17,250 for the prior TNFi-exposure population.

The deterministic sensitivity analysis results showed consistent results across the two populations: TNFi-naïve and TNFi-exposed. In the comparison with conventional treatment the most important driver of cost-effectiveness was the use of the OCTAVE trials EQ-5D data for the model health states. The main reason for this change in the ICER was the smaller difference between the utlity of active UC and response or remission, in the data from the OCTAVE trials compared with the data from Woehl *et al*. 2008 (158) (base case). The reason that serious infection costs had such an impact on the cost-effectiveness results is because of the very wide range of the values tested in sensitivity analysis: £700 (lower quarter value for soft tissue infection) to £11,000 (upper quarter value for pneumonia or sepsis). Furthermore, in the comparison of tofacitinib with conventional treatment, testing the two extreme cost values (£6.79 for prednisolone – £300.53 for olansazine) resulted in a substantial variation of the ICER. Finally, the 2.5% and 97.5% bounds of the NMA results for response in maintenance were important drivers of the cost-effectiveness results because they determined treatment continuation.

B.3.9 Subgroup analysis

No further subgroups were considered.

B.3.10 Validation of cost-effectiveness analysis

Validation of various clinical and economic inputs and assumptions was performed by engaging with UK clinical experts, statisticians and health economists as summarised in Table 78.

Table 78 Validation of the *de novo* cost-effectiveness analysis

Validation performed by	Format	Date(s)	Key aspects covered (list not exhaustive)
UK Consultant Gastroenterologist	Continuous engagement throughout	ongoing	UK treatment pathway, Clinical outcomes and clinical data, clinical assumptions, UK-specific input parameters
3x UK Consultant Gastroenterologist,	Advisory board	April 2018	Clinical data and data gaps; UK treatment pathway, clinical assumptions; model

Validation performed by	Format	Date(s)	Key aspects covered (list not exhaustive)
3x UK Health Economists	meeting		structure and assumptions, model health states; time on treatment; resource-use, costs, and utility estimates
UK Health economics expert as external reviewer	Continuous engagement throughout	ongoing	Technical model validation and critique of model methds, assumptions and inputs, and subsequent model reports
UK statistician as external reviewer	Continuous engagement throughout	ongoing	Review of SLR outputs, and review and critique of the Network Meta-Analysis (NMA) feasibility, NMA methods, assumptions and analyses, and subsequent NMA reports
Symmetron Ltd	Quality control	April 2018	Checked input data against sources, reviewed model programming

B.3.10.1 Clinical validation

In summary during a UK advisory board meeting clinical and model assumptions were validated with clinicians and health economists. They confirmed that the model assumptions and predictions were plausible and that the structure of the model reflected clinical practice in England and Wales. Furthermore, clinical expert opinion was sought to estimate, validate, and guide assumptions pertaining to the healthcare resource use inputs, as well as on how to interpret NICE guidelines for the definition of the treatment strategies.

With regard to the model methods, the following items were validated.

OCTAVE study patient population:

The experts suggested that the OCTAVE study baseline characteristics are well balanced and broadly reflect UK practice, although the disease duration is higher than seen in clinical practice (6–7yrs for OCTAVE trials vs ~2–4yrs in clinical practice). However, clinicians suggested that it appears that the OCTAVE trial patients would better reflect the harder to treat patients in clinical practice, who are likely to have failed a series of prior treatments.

Subgroup analysis by prior TNFi-exposure

Although the OCTAVE trials were not powered for the stratified subgroups, clinical experts suggested that the separation by prior TNFi-exposure is clinically relevant as the patient treatment history is an important decision criterion for selecting the most appropriate treatment. Clinical experts specifically listed the history and risk of immunogenicity as the main limitation of current biologics available for treating ulcerative colitis.

Time on treatment and discontinuation rates

When presented with the time on treatment output and average discontinuation rates for tofacitinib and biologics from the economic model, clinical experts explained that this is reflective of clinical practice, anticipating the vast majority (~80%) of patients to have failed or discontinued biologic treatment within the first 2–3 years, and only a few patients would maintain the same treatment for longerlonger periods of time, potentially up to 8 years.

Monitoring costs applicable to tofacitinib

The expert panel reflected on the OCTAVE clinical data presented and the draft SmPC, and suggested that no additional monitoring would be required in clinical practice that is not already in place for immunomodulators, TNFi and vedolizumab treatments.

Health-state unit costs and resource use, including rate of hospitalisation

The clinical experts, including a specialist in colonoscopy, confirmed the resource use presented by Tsai *et al*. (148), would be the most reflective in clinical practice, including hospitalisation episodes, for active disease, response and remission. However, the experts suggested that the cost assumptions (as listed in Table 55) appear to be on the lower end for each of the resource items.

Emergency surgery

Clinical experts confirmed that acute flares are largely unpredictable and can affect any patient regardless of treatment at any time, necessitating emergency surgery.

Quality of Life

Both clinicians and health economist'seconomist's recommended use of Woehl *et al.* (158) within the base case analysis as it has been integrally used in previous NICE technology appraisals. The panel felt that a randomised clinical trial setting and the OCTAVE re-randomisation design are likely to impact placebo EQ-5D values, limiting it's representation of active UC as a proxy, and therefore suggested that OCTAVE data should be used only in a scenario. The panel stated, that the overall conclusions are unlikely to change.

Tofacitinib 10mg twice daily as a maintenance dose

Clinical experts welcomed the flexibility of the up and down dosing, including interruptions without a risk of immunogenicity. The advisers stated that it is reasonable to assume [redacted] of patients may benefit from 10mg twice daily maintenance dosing, as this would be broadly reflective of clinical practice based on their experience with current biologic therapies. The clinical advisors highlighted that this is unlikely to be limited to patients with prior-TNF-exposure only, as the whole patient history and risk factors are taken into account when deciding on a treatment.

B.3.10.2 Internal validation of de novo cost-effectiveness analysis

The analysis builds on methods from previous appraisals and translates effectively the clinical trial evidence into the economic model. In appendix M, section M.2 the proportion of the cohort in response and remission predicted by the model, was plotted against the NMA estimates for response and remission at induction and maintenance phase. Note that the maintenance phase rates were calculated as dependent on response at induction.

B.3.10.3 Quality control

Several quality control measures were undertaken to validate the model findings included in this submission. Internal quality control was undertaken by the developers of the model on behalf of the manufacturer. The model results were compared with the MTA 329 model (118), and any identified discrepancies were clarified and resolved. A second modeler, not

involved in the programming, reviewed the model code and formulae, and conducted extreme value analysis to verify the model results. The lead modeler scrutinised the programming and references. In addition, the model was critiqued by an external independent health economist with a full review of model structure, parameter inputs, and core assumptions.

B.3.11 Interpretation and conclusions of economic evidence

This was a cost-effectiveness analysis of tofacitinib for the treatment of moderately to severely ulcerative colitis, considering two subgroup populations; biologic-naïve and biologic-exposed. The analysis was based on a comprehensive evidence review and an NMA of the available evidence from randomised clinical trials. The NMA provided evidence for the allocation of patients between response, remission, and no-response at 8 weeks; and subsequent continuation of response, and remission after 1 year of treatment.

The structure of the economic model expanded on previous economic evaluations (118) and updated the assumptions, where possible, with contemporary evidence from UK sources. One of the strengths of the economic analysis is the use of an age-dependent (and gender-dependent) baseline utility, reflecting the natural decline of patients' physical and mental health due to age and other co-morbidities. Given that the average patient was predicted to stay on biologic treatment or tofacitinib for 1–2 years, before discontinuing to conventional treatment, a fixed utility over the remaining 40 years would overestimate the accumulated model QALY results.

Furthermore, as part of the update of the input, the economic analysis used evidence from a large, UK, retrospective population-based study using the HES database for colectomy rates (167). Moreover, the analysis considered serious infections, often associated with the immunomodulatory or immunosuppressive effects of biologic treatments. The incidence of serious infections in the clinical trials was meta-analysed and used to populate the economic model.

In comparison with previous economic evaluations, the appraisal model generated comparable QALY results for patients on biologic treatment: [redacted]s in the appraisal model versus 11 QALYs in Archer *et al.* (118). The lower total QALYs in the appraisal model was attributed to the implementation of the baseline utility; which was lower (0.85 QALYs) than what was assumed by Archer *et al.* (0.87 QALYs in remission) (118).

The appraisal model generated more costs over lifetime ([redacted]) compared with the Archer *et al.* analysis (£74,000 – £97,000) (118). This difference was caused by the higher resource use frequency, in particular hospitalisation while in active ulcerative colitis (section B.3.3.5).

In the base-case analysis when evaluating the cost effectiveness of tofacitinib compared with the biologic treatments, the lower dose regimens were assumed for maintenance. Scenario analysis considered a mixed population receiving 5 mg ([redacted]) and 10 mg ([redacted]) tofacitinib doses during maintenance. The cost-effectiveness of tofacitinib remained within the £20,000 per QALY threshold in both TNFi-naïve and exposed populations.

Further scenarios considered higher doses for the biologic treatments; changing both the treatment cost and efficacy outcomes for golimumab and vedolizumab, and only cost of treatment for adalimumab. The conclusions of the base-case analysis were not changed.

Finally, an ITT population scenario was considered and the comparison of tofacitinib and vedolizumab was presented – being the least confounded by differences in the trial populations. In this scenario vedolizumab was dominated by tofacitinib.

The results of this analysis are expected to be applicable to clinical practice in England and Wales. The health state definition was based on Mayo scores, used to identify treatment response and continuation in clinical practice. Furthermore, most of the evidence for unit costs, resource use, as well as the disease utility weights were obtained from UK sources.

In conclusion, the results of the economic analysis suggested that, under the £20,000 per QALY threshold, tofacitinib was a cost-effective treatment option for patients with moderately to severely ulcerative colitis. Deterministic one-way sensitivity analysis, additional scenario analysis, and probabilistic sensitivity analysis suggested that the model results were robust to input range and assumption changes.

B.4 References

1. Travis SP, Stange EF, Lemann M, Oresland T, Bemelman WA, Chowers Y, et al. European evidence-based Consensus on the management of ulcerative colitis: Current management. J Crohns Colitis. 2008;2(1):24-62.
2. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology. 2012;142(1):46-54.e42; quiz e30.
3. Royal College of Physicians. National clinical audit of biological therapies - UK inflammatory bowel disease (IBD) audit. 2016 [Available from: https://www.hqip.org.uk/resources/national-clinical-audit-of-biological-therapies-uk-inflammatory-bowel-disease-ibd-audit/.
4. Cosnes J, Gower-Rousseau C, Seksik P, Cortot A. Epidemiology and natural history of inflammatory bowel diseases. Gastroenterology. 2011;140(6):1785-94.
5. Danese S, Fiocchi C. Ulcerative colitis. N Engl J Med. 2011;365(18):1713-25.
6. Burisch J, Jess T, Martinato M, Lakatos PL. The burden of inflammatory bowel disease in Europe. Journal of Crohn's and Colitis. 2013;7(4):322-37.
7. Rubin GP, Hungin APS, Chinn DJ, Dwarakanath D. Quality of life in patients with established inflammatory bowel disease: a UK general practice survey. Alimentary Pharmacology & Therapeutics. 2004;19(5):529-35.
8. Mowat C, Cole A, Windsor A, Ahmad T, Arnott I, Driscoll R, et al. Guidelines for the management of inflammatory bowel disease in adults. Gut. 2011;60(5):571-607.
9. Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. Lancet. 2017;389(10080):1756-70.
10. Dignass A, Eliakim R, Magro F, Maaser C, Chowers Y, Geboes K, et al. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 1: definitions and diagnosis. J Crohns Colitis. 2012;6(10):965-90.
11. Magro F, Gionchetti P, Eliakim R, Ardizzone S, Armuzzi A, Barreiro-de Acosta M, et al. Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: Definitions, Diagnosis, Extra-intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-anal Pouch Disorders. Journal of Crohn's and Colitis. 2017;11(6):649-70.
12. Langholz E, Munkholm P, Davidsen M, Binder V. Course of ulcerative colitis: analysis of changes in disease activity over years. Gastroenterology. 1994;107(1):3-11.
13. Henriksen M, Jahnsen J, Lygren I, Sauar J, Kjellevold O, Schulz T, et al. Ulcerative colitis and clinical course: results of a 5-year population-based follow-up study (the IBSEN study). Inflammatory bowel diseases. 2006;12(7):543-50.
14. Langholz E, Munkholm P, Davidsen M, Nielsen OH, Binder V. Changes in extent of ulcerative colitis: a study on the course and prognostic factors. Scandinavian journal of gastroenterology. 1996;31(3):260-6.
15. Fakhoury M, Negrulj R, Mooranian A, Al-Salami H. Inflammatory bowel disease: clinical aspects and treatments. Journal of inflammation research. 2014;7:113-20.
16. Silverberg MS, Satsangi J, Ahmad T, Arnott ID, Bernstein CN, Brant SR, et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Canadian journal of gastroenterology = Journal canadien de gastroenterologie. 2005;19 Suppl A:5a-36a.
17. National Institute for Health and Care Excellence. Ulcerative colitis: management. Clinical guideline 166. 2013 [Updated 19 Jan 2018. Available from: https://www.nice.org.uk/guidance/cg166.

18. Walsh AJ, Bryant RV, Travis SP. Current best practice for disease activity assessment in IBD. Nature reviews Gastroenterology & hepatology. 2016;13(10):567-79.
19. Mohammed Vashist N, Samaan M, Mosli MH, Parker CE, MacDonald JK, Nelson SA, et al. Endoscopic scoring indices for evaluation of disease activity in ulcerative colitis. Cochrane Database of Systematic Reviews. 2018(1).
20. Jairath V, Zou GY, Parker CE, MacDonald JK, AlAmeel T, Al Beshir M, et al. Placebo response and remission rates in randomised trials of induction and maintenance therapy for ulcerative colitis. The Cochrane database of systematic reviews. 2017;9:Cd011572.
21. D'Haens G, Sandborn WJ, Feagan BG, Geboes K, Hanauer SB, Irvine EJ, et al. A review of activity indices and efficacy end points for clinical trials of medical therapy in adults with ulcerative colitis. Gastroenterology. 2007;132(2):763-86.
22. Colombel JFR, P. Reinisch, W. Esser, D. Wang, Y. Lang, Y. Marano, C. W. Strauss, R. Oddens, B. J. Feagan, B. G. Hanauer, S. B. Lichtenstein, G. R. Present, D. Sands, B. E. Sandborn, W. J. Early mucosal healing with infliximab is associated with improved long-term clinical outcomes in ulcerative colitis. Gastroenterology. 2011;141(4):1194-201.
23. Lewis JD, Chuai S, Nessel L, Lichtenstein GR, Aberra FN, Ellenberg JH. Use of the noninvasive components of the Mayo score to assess clinical response in ulcerative colitis. Inflammatory bowel diseases. 2008;14(12):1660-6.
24. Carlsen K, Munkholm P, Burisch J. Evaluation of Quality of Life in Crohn's Disease and Ulcerative Colitis: What Is Health-Related Quality of Life? In: Baumgart DC, editor. Crohn's Disease and Ulcerative Colitis: From Epidemiology and Immunobiology to a Rational Diagnostic and Therapeutic Approach. Cham: Springer International Publishing; 2017. p. 279-89.
25. Ghosh S, Mitchell R. Impact of inflammatory bowel disease on quality of life: Results of the European Federation of Crohn's and Ulcerative Colitis Associations (EFCCA) patient survey. J Crohns Colitis. 2007;1(1):10-20.
26. Vaizey CJ, Gibson PR, Black CM, Nicholls RJ, Weston AR, Gaya DR, et al. Disease status, patient quality of life and healthcare resource use for ulcerative colitis in the UK: an observational study. Frontline Gastroenterology. 2014;5(3):183-9.
27. Rocchi A, Benchimol EI, Bernstein CN, Bitton A, Feagan B, Panaccione R, et al. Inflammatory bowel disease: a Canadian burden of illness review. Canadian journal of gastroenterology = Journal canadien de gastroenterologie. 2012;26(11):811-7.
28. Sammut J, Scerri J, Xuereb RB. The lived experience of adults with ulcerative colitis. Journal of clinical nursing. 2015;24(17-18):2659-67.
29. Gray WN, Graef DM, Schuman SS, Janicke DM, Hommel KA. Parenting stress in pediatric IBD: relations with child psychopathology, family functioning, and disease severity. Journal of developmental and behavioral pediatrics : JDBP. 2013;34(4):237-44.
30. Dibley L, Norton C. Experiences of fecal incontinence in people with inflammatory bowel disease: self-reported experiences among a community sample. Inflammatory bowel diseases. 2013;19(7):1450-62.
31. Kemp K, Griffiths J, Lovell K. Understanding the health and social care needs of people living with IBD: a meta-synthesis of the evidence. World J Gastroenterol. 2012;18(43):6240-9.
32. Norton C, Dibley LB, Bassett P. Faecal incontinence in inflammatory bowel disease: associations and effect on quality of life. J Crohns Colitis. 2013;7(8):e302-11.
33. Jelsness-Jorgensen LP, Bernklev T, Henriksen M, Torp R, Moum BA. Chronic fatigue is associated with impaired health-related quality of life in inflammatory bowel disease. Aliment Pharmacol Ther. 2011;33(1):106-14.
34. Pickard AS, Yang Y, Lee TA. Comparison of health-related quality of life measures in chronic obstructive pulmonary disease. Health and Quality of Life Outcomes. 2011;9:26-.
35. Baker F, Haffer Samuel C, Denniston M. Health-related quality of life of cancer and noncancer patients in Medicare managed care. Cancer. 2003;97(3):674-81.

36. Lacson E, Jr., Xu J, Lin SF, Dean SG, Lazarus JM, Hakim RM. A comparison of SF-36 and SF-12 composite scores and subsequent hospitalization and mortality risks in long-term dialysis patients. Clinical journal of the American Society of Nephrology : CJASN. 2010;5(2):252-60.
37. Vogelaar L, Spijker AV, van der Woude CJ. The impact of biologics on health-related quality of life in patients with inflammatory bowel disease. Clin Exp Gastroenterol. 2009;2:101-9.
38. Kosmas CE, Kerr C, Marsh SE, Gentile B, Doll HA, Finney-Hayward T. Health Related Quality of Life (HRQoL) In Patients Who Have Undergone Colectomy for Ulcerative Colitis: Impacts of Complications Post-Surgery. Value Health. 2015;18(7):A630.
39. Kuhn F, Klar E. Surgical Principles in the Treatment of Ulcerative Colitis. Viszeralmedizin. 2015;31(4):246-50.
40. Bernstein CN, Fried M, Krabshuis JH, Cohen H, Eliakim R, Fedail S, et al. World Gastroenterology Organization Practice Guidelines for the diagnosis and management of IBD in 2010. Inflammatory bowel diseases. 2010;16(1):112-24.
41. Vavricka SR, Schoepfer A, Scharl M, Lakatos PL, Navarini A, Rogler G. Extraintestinal Manifestations of Inflammatory Bowel Disease. Inflammatory bowel diseases. 2015;21(8):1982-92.
42. Vavricka SR, Brun L, Ballabeni P, Pittet V, Prinz Vavricka BM, Zeitz J, et al. Frequency and risk factors for extraintestinal manifestations in the Swiss inflammatory bowel disease cohort. The American journal of gastroenterology. 2011;106(1):110-9.
43. Gan SI, Beck PL. A new look at toxic megacolon: an update and review of incidence, etiology, pathogenesis, and management. The American journal of gastroenterology. 2003;98(11):2363-71.
44. Lewis JD, Deren JJ, Lichtenstein GR. Cancer risk in patients with inflammatory bowel disease. Gastroenterology clinics of North America. 1999;28(2):459-77, x.
45. Bonovas S, Fiorino G, Lytras T, Nikolopoulos G, Peyrin-Biroulet L, Danese S. Systematic review with meta-analysis: use of 5-aminosalicylates and risk of colorectal neoplasia in patients with inflammatory bowel disease. Aliment Pharmacol Ther. 2017;45(9):1179-92.
46. Velayos FS, Loftus EV, Jr., Jess T, Harmsen WS, Bida J, Zinsmeister AR, et al. Predictive and protective factors associated with colorectal cancer in ulcerative colitis: A case-control study. Gastroenterology. 2006;130(7):1941-9.
47. Beaugerie L, Svrcek M, Seksik P, Bouvier AM, Simon T, Allez M, et al. Risk of colorectal high-grade dysplasia and cancer in a prospective observational cohort of patients with inflammatory bowel disease. Gastroenterology. 2013;145(1):166-75.e8.
48. Van Assche G, Dignass A, Bokemeyer B, Danese S, Gionchetti P, Moser G, et al. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis Part 3: Special situations. Journal of Crohn's and Colitis. 2013;7(1):1-33.
49. Dinesen LC, Walsh AJ, Protic MN, Heap G, Cummings F, Warren BF, et al. The pattern and outcome of acute severe colitis. J Crohns Colitis. 2010;4(4):431-7.
50. Cottone M, Scimeca D, Mocciaro F, Civitavecchia G, Perricone G, Orlando A. Clinical course of ulcerative colitis. Dig Liver Dis. 2008;40 Suppl 2:S247-52.
51. Mehta SJ, Silver AR, Lindsay JO. Review article: strategies for the management of chronic unremitting ulcerative colitis. Aliment Pharmacol Ther. 2013;38(2):77-97.
52. Jess T, Gamborg M, Munkholm P, Sorensen TI. Overall and cause-specific mortality in ulcerative colitis: meta-analysis of population-based inception cohort studies. The American journal of gastroenterology. 2007;102(3):609-17.
53. Roberts SE, Williams JG, Yeates D, Goldacre MJ. Mortality in patients with and without colectomy admitted to hospital for ulcerative colitis and Crohn's disease: record linkage studies. BMJ. 2007;335(7628):1033.
54. Cohen RD, Yu AP, Wu EQ, Xie J, Mulani PM, Chao J. Systematic review: the costs of ulcerative colitis in Western countries. Aliment Pharmacol Ther. 2010;31(7):693-707.

55. van der Valk ME, Mangen MJ, Leenders M, Dijkstra G, van Bodegraven AA, Fidder HH, et al. Healthcare costs of inflammatory bowel disease have shifted from hospitalisation and surgery towards anti-TNFalpha therapy: results from the COIN study. Gut. 2014;63(1):72-9.
56. Bassi A, Dodd S, Williamson P, Bodger K. Cost of illness of inflammatory bowel disease in the UK: a single centre retrospective study. Gut. 2004;53(10):1471-8.
57. van der Valk ME, Mangen MJ, Severs M, van der Have M, Dijkstra G, van Bodegraven AA, et al. Evolution of Costs of Inflammatory Bowel Disease over Two Years of Follow-Up. PLoS One. 2016;11(4):e0142481.
58. Lindsay JO, Bergman A, Patel AS, Alesso SM, Peyrin-Biroulet L. Systematic review: the financial burden of surgical complications in patients with ulcerative colitis. Aliment Pharmacol Ther. 2015;41(11):1066-78.
59. Harbord M, Eliakim R, Bettenworth D, Karmiris K, Katsanos K, Kopylov U, et al. Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 2: Current Management. Journal of Crohn's and Colitis. 2017;11(7):769-84.
60. Chris P. Steroids and 5-aminosalicylic acids in moderate ulcerative colitis: addressing the dilemma. Therapeutic Advances in Gastroenterology. 2012;6(1):33-8.
61. Probert C. Steroids and 5-aminosalicylic acids in moderate ulcerative colitis: addressing the dilemma. Therap Adv Gastroenterol. 2013;6(1):33-8.
62. Westwood N, Travis SP. Review article: what do patients with inflammatory bowel disease want for their clinical management? Aliment Pharmacol Ther. 2008;27 Suppl 1:1-8.
63. Shah SC, Colombel JF, Sands BE, Narula N. Mucosal Healing Is Associated With Improved Long-term Outcomes of Patients With Ulcerative Colitis: A Systematic Review and Meta-analysis. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2016;14(9):1245-55.e8.
64. Choi CR, Al Bakir I, Ding NJ, Lee GH, Askari A, Warusavitarne J, et al. Cumulative burden of inflammation predicts colorectal neoplasia risk in ulcerative colitis: a large single-centre study. Gut. 2017.
65. National Institute for Health and Care Excellence. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy. . 2015 [Updated 19 Jan 2018. Available from: https://www.nice.org.uk/guidance/ta329.
66. National Institute for Health and Care Excellence. Vedolizumab for treating moderately to severely active ulcerative colitis. 2015 [Updated 23 Jan 2018. Available from: https://www.nice.org.uk/guidance/ta342.
67. Hoivik ML, Moum B, Solberg IC, Cvancarova M, Hoie O, Vatn MH, et al. Health-related quality of life in patients with ulcerative colitis after a 10-year disease course: results from the IBSEN study. Inflammatory bowel diseases. 2012;18(8):1540-9.
68. Hydrocortisone sodium succinate injection Prescribing Information. Lake Forest, IL: Hospira, Inc; 2011. RL-2526.
69. Medrol Prescribing Information. NY, NY: Pharmacia and Upjohn Company LLC; 2016. ID 618054084.
70. Faubion WA, Jr., Loftus EV, Jr., Harmsen WS, Zinsmeister AR, Sandborn WJ. The natural history of corticosteroid therapy for inflammatory bowel disease: a population-based study. Gastroenterology. 2001;121(2):255-60.
71. Gisbert J P, Linares P M, McNicholl A G, MatÉ J, GomollÓN F. Meta-analysis: the efficacy of azathioprine and mercaptopurine in ulcerative colitis. Alimentary Pharmacology & Therapeutics. 2009;30(2):126-37.
72. Sandborn WJ. Azathioprine: state of the art in inflammatory bowel disease. Scandinavian journal of gastroenterology Supplement. 1998;225:92-9.
73. Reinisch W, Sandborn WJ, Hommes DW, D'Haens G, Hanauer S, Schreiber S, et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. Gut. 2011;60(6):780-7.

74. Rutgeerts P, Sandborn WJ, Feagan BG, Reinisch W, Olson A, Johanns J, et al. Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis. New England Journal of Medicine. 2005;353(23):2462-76.
75. Sandborn WJ, Feagan BG, Marano C, Zhang H, Strauss R, Johanns J, et al. Subcutaneous golimumab induces clinical response and remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2014;146(1):85-95; quiz e14-5.
76. Nielsen OH, Ainsworth MA. Tumor necrosis factor inhibitors for inflammatory bowel disease. N Engl J Med. 2013;369(8):754-62.
77. Olesen CM, Coskun M, Peyrin-Biroulet L, Nielsen OH. Mechanisms behind efficacy of tumor necrosis factor inhibitors in inflammatory bowel diseases. Pharmacology & therapeutics. 2016;159:110-9.
78. Grinspan A, Kornbluth A. Positioning Therapy for Ulcerative Colitis. Curr Gastroenterol Rep. 2015;17(8):29.
79. Iborra M, Perez-Gisbert J, Bosca-Watts MM, Lopez-Garcia A, Garcia-Sanchez V, Lopez-Sanroman A, et al. Effectiveness of adalimumab for the treatment of ulcerative colitis in clinical practice: comparison between anti-tumour necrosis factor-naive and non-naive patients. Journal of gastroenterology. 2017;52(7):788-99.
80. Khanna R, Sattin BD, Afif W, Benchimol EI, Bernard EJ, Bitton A, et al. Review article: a clinician's guide for therapeutic drug monitoring of infliximab in inflammatory bowel disease. Aliment Pharmacol Ther. 2013;38(5):447-59.
81. Remicade® Prescribing Information. Horsham, PA: Janssen Biotech, Inc; 2013.
82. Bressler B, Marshall JK, Bernstein CN, Bitton A, Jones J, Leontiadis GI, et al. Clinical practice guidelines for the medical management of nonhospitalized ulcerative colitis: the Toronto consensus. Gastroenterology. 2015;148(5):1035-58.e3.
83. Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel J-F, Sandborn WJ, et al. Vedolizumab as Induction and Maintenance Therapy for Ulcerative Colitis. New England Journal of Medicine. 2013;369(8):699-710.
84. Hanauer SB. How will new and future therapies change our management of IBD? Northwestern University Feinberg School of Medicine. Advances in Inflammatory Bowel Diseases Scientific Meeting December 2015, p 27.
85. Essat M, Tappenden P, Ren S, Bessey A, Archer R, Wong R, et al. Vedolizumab for the Treatment of Adults with Moderate-to-Severe Active Ulcerative Colitis: An Evidence Review Group Perspective of a NICE Single Technology Appraisal. PharmacoEconomics. 2016;34(3):245-57.
86. Hodge JA, Kawabata TT, Krishnaswami S, Clark JD, Telliez JB, Dowty ME, et al. The mechanism of action of tofacitinib - an oral Janus kinase inhibitor for the treatment of rheumatoid arthritis. Clinical and experimental rheumatology. 2016;34(2):318-28.
87. Murray PJ. The JAK-STAT signaling pathway: input and output integration. Journal of immunology (Baltimore, Md : 1950). 2007;178(5):2623-9.
88. Danese S, Grisham M, Hodge J, Telliez JB. JAK inhibition using tofacitinib for inflammatory bowel disease treatment: a hub for multiple inflammatory cytokines. American journal of physiology Gastrointestinal and liver physiology. 2016;310(3):G155-62.
89. Coskun M, Salem M, Pedersen J, Nielsen OH. Involvement of JAK/STAT signaling in the pathogenesis of inflammatory bowel disease. Pharmacological research. 2013;76:1-8.
90. Galien R. Janus kinases in inflammatory bowel disease: Four kinases for multiple purposes. Pharmacological reports : PR. 2016;68(4):789-96.
91. O'Sullivan LA, Liongue C, Lewis RS, Stephenson SE, Ward AC. Cytokine receptor signaling through the Jak-Stat-Socs pathway in disease. Molecular immunology. 2007;44(10):2497-506.
92. Boland BS, Sandborn WJ, Chang JT. Update on Janus kinase antagonists in inflammatory bowel disease. Gastroenterology clinics of North America. 2014;43(3):603-17.
93. Yanai H, Hanauer SB. Assessing response and loss of response to biological therapies in IBD. The American journal of gastroenterology. 2011;106(4):685-98.

94. Ordas I, Feagan BG, Sandborn WJ. Therapeutic drug monitoring of tumor necrosis factor antagonists in inflammatory bowel disease. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2012;10(10):1079-87; quiz e85-6.
95. Vermeire S, Gils A, Accossato P, Lula S, Marren A. Immunogenicity of biologics in inflammatory bowel disease. Therap Adv Gastroenterol. 2018;11:1756283x17750355.
96. Panaccione RG, S. Middleton, S. Márquez, J. R. Scott, B. B. Flint, L. Hoogstraten, H. J. Chen, A. C. Zheng, H. Danese, S. Rutgeerts, P. Combination therapy with infliximab and azathioprine is superior to monotherapy with either agent in ulcerative colitis. Gastroenterology [Internet]. 2014; 146(2):[392-400.e3 pp.]. Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/777/CN-00958777/frame.html http://www.gastrojournal.org/article/S0016-5085(13)01526-6/pdf.
97. Mukherjee A, D'Haens GR, Sandborn WJ, Tsuchiwata S, Vong C, Deng C, et al. DOP074 Pharmacokinetics and exposure-response of tofacitinib in a Phase 3 maintenance study in ulcerative colitis patients. Journal of Crohn's and Colitis. 2017;11(suppl_1):S69-S70.
98. Sandborn WJ, Su C, Sands BE, D'Haens GR, Vermeire S, Schreiber S, et al. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2017;376(18):1723-36.
99. Sandborn WJ, Ghosh S, Panes J, Vranic I, Su C, Rousell S, et al. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. N Engl J Med. 2012;367(7):616-24.
100. Inc. P. A Multicentre, Randomised, Double-Blind, Placebo-Controlled, Parallel-Group Study of Oral CP-690,550 as an Induction Therapy in Subjects With Moderate to Severe Ulcerative Colitis: OCTAVE Induction 1 (Clinical Study Report). 25 May 2016. Data on file.
101. Inc. P. A Multicentre, Randomised, Double-Blind, Placebo-Controlled, Parallel-Group Study of Oral CP-690,550 as an Induction Therapy in Subjects With Moderate to Severe Ulcerative Colitis: OCTAVE Induction 2 (Clinical Study Report). 6 May 2016. Data on file.
102. Inc. P. A Multicentre, Randomised, Double-Blind, Placebo-Controlled, Parallel-Group Study of Oral CP-690,550 as a Maintenance Therapy in Subjects With Moderate to Severe Ulcerative Colitis: OCTAVE Sustain (Clinical Study Report). 16 December 2016. Data on file.
103. Inc. P. A Multicentre, Randomised, Double-Blind, Placebo-Controlled, Parallel-Group Study of Oral CP-690,550 in Subjects With Moderate to Severe Ulcerative Colitis: OCTAVE Open (Interim Clinical Study Report). 20 December 2016. Data on file.
104. Walsh A, Palmer R, Travis S. Mucosal healing as a target of therapy for colonic inflammatory bowel disease and methods to score disease activity. Gastrointestinal endoscopy clinics of North America. 2014;24(3):367-78.
105. Khan HMW, Mehmood F, Khan N. Optimal management of steroid-dependent ulcerative colitis. Clinical and Experimental Gastroenterology. 2015;8:293-302.
106. Sandborn W, Feagan B, Marano C, Strauss R, Han C, Johanns J, et al. P275 Clinical response is a meaningful endpoint in ulcerative colitis clinical studies. Journal of Crohn's and Colitis. 2013;7:S119.
107. Papp KA, Fonjallaz P, Casset-Semanaz F, Krueger JG, Wittkowski KM. Approaches to reporting long-term data. Current medical research and opinion. 2008;24(7):2001-8.
108. Stark RG, Reitmeir P, Leidl R, Konig HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. Inflammatory bowel diseases. 2010;16(1):42-51.
109. Sandborn WJ, van Assche G, Reinisch W, Colombel JF, D'Haens G, Wolf DC, et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2012;142(2):257-65 e1-3.
110. Hibi T, Imai Y, Senoo A, Ohta K, Ukyo Y. Efficacy and safety of golimumab 52-week maintenance therapy in Japanese patients with moderate to severely active ulcerative colitis: a phase 3, double-blind, randomized, placebo-controlled study-(PURSUIT-J study). Journal of gastroenterology. 2017;52(10):1101-11.

111. Jiang XL, Cui HF, Gao J, Fan H. Low-dose Infliximab for Induction and Maintenance Treatment in Chinese Patients With Moderate to Severe Active Ulcerative Colitis. J Clin Gastroenterol. 2015;49(7):582-8.
112. Kobayashi T, Suzuki Y, Motoya S, Hirai F, Ogata H, Ito H, et al. First trough level of infliximab at week 2 predicts future outcomes of induction therapy in ulcerative colitis-results from a multicenter prospective randomized controlled trial and its post hoc analysis. Journal of gastroenterology. 2016;51(3):241-51.
113. Mshimesh BAR. Efficacy and safety of adalimumab versus infliximab in patients suffered from moderate to severe active ulcerative colitis. Asian Journal of Pharmaceutical and Clinical Research. 2017;10(3):300-7.
114. Suzuki Y, Motoya S, Hanai H, Matsumoto T, Hibi T, Robinson AM, et al. Efficacy and safety of adalimumab in Japanese patients with moderately to severely active ulcerative colitis. Journal of gastroenterology. 2014;49(2):283-94.
115. Sandborn WJ, Feagan BG, Marano C, Zhang H, Strauss R, Johanns J, et al. Subcutaneous golimumab maintains clinical response in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2014;146(1):96-109 e1.
116. Thorlund K, Druyts E, Toor K, Jansen JP, Mills EJ. Incorporating alternative design clinical trials in network meta-analyses. Clin Epidemiol. 2015;7:29-35.
117. National Institute for Health and Care Excellence. Vedolizumab for treating moderately to severely active ulcerative colitis - manufacturer's submission. 2015 [Updated 3 April 2018. Available from: https://www.nice.org.uk/guidance/ta342/documents/ulcerative-colitis-moderate-to-severely-active-vedolizumab-id691-evaluation-report-part-12.
118. Archer R, Tappenden P, Ren S, James MMS, Harvey R, Basarir H, et al. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (Including a review of TA140 and TA262): Clinical effectiveness systematic review and economic model. Health Technology Assessment. 2016;20(39).
119. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity--subgroups, meta-regression, bias, and bias-adjustment. Med Decis Making. 2013;33(5):618-40.
120. Winthrop KL, Loftus Jr EV, Baumgart DC, Reinisch W, Thorpe AJ, Nduaka CI, et al. P487 Tofacitinib for the treatment of ulcerative colitis: Analysis of infection rates from the OCTAVE clinical programme. Journal of Crohn's and Colitis. 2018;12(supplement_1):S351-S2.
121. Long MD, Martin CF, Pipkin CA, Herfarth HH, Sandler RS, Kappelman MD. Risk of melanoma and nonmelanoma skin cancer among patients with inflammatory bowel disease. Gastroenterology. 2012;143(2):390-9.e1.
122. Peyrin-Biroulet L, Khosrotehrani K, Carrat F, Bouvier AM, Chevaux JB, Simon T, et al. Increased risk for nonmelanoma skin cancers in patients who receive thiopurines for inflammatory bowel disease. Gastroenterology. 2011;141(5):1621-28.e1-5.
123. Bernstein CN, Blanchard JF, Kliewer E, Wajda A. Cancer risk in patients with inflammatory bowel disease: a population-based study. Cancer. 2001;91(4):854-62.
124. Pfizer I. FDA Briefing Document - Gastrointestinal Drug Advisory Committee Meeting - XELJANZ (tofacitinib). 2018 [Available from: https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/GastrointestinalDrugsAdvisoryCommittee/UCM599512.pd. Accessed: 4 May 2018.]
125. Augustovski F, Beratarrechea A, Irazola V, Rubinstein F, Tesolin P, Gonzalez J, et al. Patient preferences for biologic agents in rheumatoid arthritis: a discrete-choice experiment. Value Health. 2013;16(2):385-93.
126. Barclay N, Tarallo M, Hendrikx T, Marett S. Patient Preference for Oral Versus Injectable and Intravenous Methods of Treatment for Rheumatoid Arthritis. Value in Health. 2013;16(7):A568.

127. Krüger K, Alten R, Schiffner-Rohe J, Behmer OS, Schiffhorst G, Rellecke J, et al. THU0350 Patient Preferences in the Choice of Disease Modifying Anti-Rheumatic Drugs. Annals of the Rheumatic Diseases. 2015;74(Suppl 2):322.
128. Casellas F, Ginard D, Riestra S. Patient satisfaction in the management of mild-to-moderate ulcerative colitis: Results of a Delphi study among patients and physicians. Digestive and Liver Disease. 2016;48(10):1172-9.
129. Gray JR, Leung E, Scales J. Treatment of ulcerative colitis from the patient's perspective: a survey of preferences and satisfaction with therapy. Aliment Pharmacol Ther. 2009;29(10):1114-20.
130. Ford AC, Peyrin-Biroulet L. Opportunistic infections with anti-tumor necrosis factor-alpha therapy in inflammatory bowel disease: meta-analysis of randomized controlled trials. The American journal of gastroenterology. 2013;108(8):1268-76.
131. Naganuma M, Kunisaki R, Yoshimura N, Takeuchi Y, Watanabe M. A prospective analysis of the incidence of and risk factors for opportunistic infections in patients with inflammatory bowel disease. Journal of gastroenterology. 2013;48(5):595-600.
132. Toruner M, Loftus EV, Jr., Harmsen WS, Zinsmeister AR, Orenstein R, Sandborn WJ, et al. Risk factors for opportunistic infections in patients with inflammatory bowel disease. Gastroenterology. 2008;134(4):929-36.
133. Afif W, Loftus EV, Jr. Safety profile of IBD therapeutics: infectious risks. Gastroenterology clinics of North America. 2009;38(4):691-709.
134. Cottone M, Kohn A, Daperno M, Armuzzi A, Guidi L, D'Inca R, et al. Advanced age is an independent risk factor for severe infections and mortality in patients given anti-tumor necrosis factor therapy for inflammatory bowel disease. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2011;9(1):30-5.
135. Dave M, Purohit T, Razonable R, Loftus EV, Jr. Opportunistic infections due to inflammatory bowel disease therapy. Inflammatory bowel diseases. 2014;20(1):196-212.
136. Long MD, Martin C, Sandler RS, Kappelman MD. Increased risk of herpes zoster among 108 604 patients with inflammatory bowel disease. Aliment Pharmacol Ther. 2013;37(4):420-9.
137. Gupta G, Lautenbach E, Lewis JD. Incidence and risk factors for herpes zoster among patients with inflammatory bowel disease. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2006;4(12):1483-90.
138. Winthrop KL, Baddley JW, Chen L, Liu L, Grijalva CG, Delzell E, et al. Association between the initiation of anti-tumor necrosis factor therapy and the risk of herpes zoster. Jama. 2013;309(9):887-95.
139. Beaugerie L, Itzkowitz SH. Cancers Complicating Inflammatory Bowel Disease. New England Journal of Medicine. 2015;372(15):1441-52.
140. Mir-Madjlessi SH, Farmer RG, Easley KA, Beck GJ. Colorectal and extracolonic malignancy in ulcerative colitis. Cancer. 1986;58(7):1569-74.
141. Herrinton LJ, Liu L, Weng X, Lewis JD, Hutfless S, Allison JE. Role of thiopurine and anti-TNF therapy in lymphoma in inflammatory bowel disease. The American journal of gastroenterology. 2011;106(12):2146-53.
142. Rungoe C, Simonsen J, Riis L, Frisch M, Langholz E, Jess T. Inflammatory bowel disease and cervical neoplasia: a population-based nationwide cohort study. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2015;13(4):693-700.e1.
143. Kotlyar DS, Lewis JD, Beaugerie L, Tierney A, Brensinger CM, Gisbert JP, et al. Risk of lymphoma in patients with inflammatory bowel disease treated with azathioprine and 6-mercaptopurine: a meta-analysis. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2015;13(5):847-58.e4; quiz e48-50.

144. Karlen P, Lofberg R, Brostrom O, Leijonmarck CE, Hellers G, Persson PG. Increased risk of cancer in ulcerative colitis: a population-based cohort study. The American journal of gastroenterology. 1999;94(4):1047-52.
145. European Medicines Agency. Guideline on the development of new medicinal products for the treatment of Ulcerative Colitis (draft). 2016 [Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/07/WC500211431.pdf. Accessed: 24 January 2018.]
146. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013 London: National Institute for Health and Care Excellence. 2013 [Available from: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781. Accessed: 8 March 2017.]
147. Buckland A, Bodger K. The cost-utility of high dose oral mesalazine for moderately active ulcerative colitis. Alimentary Pharmacology and Therapeutics. 2008;28(11-12):1287-96.
148. Tsai HH, Punekar YS, Morris J, Fortun P. A model of the long-term cost effectiveness of scheduled maintenance treatment with infliximab for moderate-to-severe ulcerative colitis. Alimentary Pharmacology and Therapeutics. 2008;28(10):1230-9.
149. Tappenden P, Ren S, Archer R, Harvey R, James MM, Basarir H, et al. A Model-Based Economic Evaluation of Biologic and Non-Biologic Options for the Treatment of Adults with Moderately-to-Severely Active Ulcerative Colitis after the Failure of Conventional Therapy. PharmacoEconomics. 2016;34(10):1023-38.
150. Wilson MR, Bergman A, Chevrou-Severac H, Selby R, Smyth M, Kerrigan MC. Cost-effectiveness of vedolizumab compared with infliximab, adalimumab, and golimumab in patients with ulcerative colitis in the United Kingdom. European Journal of Health Economics. 2017:1-12.
151. Ali T, Skup M, Yang M, Wu E, Mulani P, Chao J. Cost-effectiveness of adalimumab in moderately to severely active ulcerative colitis. American Journal of Gastroenterology. 2012;107:S641.
152. Mukhekat V, Ahmad I, Merali Y, Aldulaimi D. The application of a markov model for evauating asa therapy for ulcerative colitis. Gut. 2014;63:A73.
153. Wilson M, Kerrigan M, Smyth M, Bergman A, Chevrou-Severac H, Selby R. Cost-effectiveness of vedolizumab compared with infliximab, adalimumab, and golimumab for treatment of moderately-to-severely active ulcerative colitis in the United Kingdom. United European Gastroenterology Journal. 2015;1):A608.
154. Wilson M, Kerrigan M, Smyth M, Chevrou-Severac H, Bergman A, Selby R. Cost-effectiveness of vedolizumab compared with infliximab, adalimumab, and golimumab for treatment of moderately-to-severely active ulcerative colitis in the United Kingdom. Gut. 2016;65:A260-A1.
155. Wilson M, Kerrigan M, Smyth M, Chevrou-Severac H, Bergman A, Selby R. Cost-effectiveness of vedolizumab compared with conventional therapy for treatment of moderately-to-severely active ulcerative colitis in the United Kingdom. United European Gastroenterology Journal. 2015;1):A608.
156. Wilson M, Kerrigan M, Smyth M, Chevrou-Severac H, Bergman A, Selby R. Cost-effectiveness of vedolizumab compared with conventional therapy for treatment of moderately-to-severely active ulcerative colitis in the United Kingdom. Gut. 2016;65:A260.
157. Yang H, Curry A, Wang S, Yang M, Skup M, Chao J, et al. Cost-effectiveness of adalimumab in moderately to severely active ulcerative colitis (sub-acute) in the UK. United European Gastroenterology Journal. 2014;1):A374-A5.
158. Woehl AH, A.B.; McEwan, P. The relation between disease activity, quality of life and health utility in patients with ulcerative colitis. Gut. 2008;57(Suppl 1):A153.
159. Arseneau KO, Sultan S, Provenzale DT, Onken J, Bickston SJ, Foley E, et al. Do Patient Preferences Influence Decisions on Treatment for Patients With Steroid-Refractory Ulcerative Colitis? Clinical Gastroenterology and Hepatology. 2006;4(9):1135-42.

160. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value Health. 2010;13(5):509-18.
161. Buchanan J, Wordsworth S, Ahmad T, Perrin A, Vermeire S, Sans M, et al. Managing the long term care of inflammatory bowel disease patients: The cost to European health care providers. Journal of Crohn's and Colitis. 2011;5(4):301-16.
162. NHS Reference Costs [Internet]. Available from: https://www.gov.uk/government/collections/nhs-reference-costs.
163. British national formulary, March 2018 [Internet]. Available from: https://bnf.nice.org.uk/.
164. Department of Health. NHS Reference Costs, 2016–17.
165. Monthly Index of Medical Specialties, March 2018 [Internet]. Available from: https://www.mims.co.uk/.
166. Department of Health and Social Care. eMIT national database, March 2018. 2018 [Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit. Accessed: 3 April 2018.]
167. Misra R, Askari A, Faiz O, Arebi N. Colectomy Rates for Ulcerative Colitis Differ between Ethnic Groups: Results from a 15-Year Nationwide Cohort Study. Can J Gastroenterol Hepatol. 2016;2016:8723949.
168. Chhaya V, Saxena S, Cecil E, Chatu S, Subramanian V, Curcin V, et al. The impact of timing and duration of thiopurine treatment on colectomy in ulcerative colitis: a national population-based study of incident cases between 1989-2009. Aliment Pharmacol Ther. 2015;41(1):87-98.
169. Moody GA, Jayanthi V, Probert CS, Mac Kay H, Mayberry JF. Long-term therapy with sulphasalazine protects against colorectal cancer in ulcerative colitis: a retrospective study of colorectal cancer risk and compliance with treatment in Leicestershire. Eur J Gastroenterol Hepatol. 1996;8(12):1179-83.
170. Russo EA, Harris AW, Campbell S, Lindsay J, Hart A, Arebi N, et al. Experience of maintenance infliximab therapy for refractory ulcerative colitis from six centres in England. Aliment Pharmacol Ther. 2009;29(3):308-14.
171. Solberg IC, Lygren I, Jahnsen J, Aadland E, Hoie O, Cvancarova M, et al. Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (IBSEN Study). Scandinavian journal of gastroenterology. 2009;44(4):431-40.
172. Frolkis AD, Dykeman J, Negron ME, Debruyn J, Jette N, Fiest KM, et al. Risk of surgery for inflammatory bowel diseases has decreased over time: a systematic review and meta-analysis of population-based studies. Gastroenterology. 2013;145(5):996-1006.
173. Royal College of Physicians. National clinical audit of biological therapies - UK inflammatory bowel disease (IBD) audit. 2014 [Available from: https://data.gov.uk/dataset/uk-inflammatory-bowel-disease-ibd-audit-round-4-2012-2014.
174. Arai K, Koganei K, Kimura H, Akatani M, Kitoh F, Sugita A, et al. Incidence and outcome of complications following restorative proctocolectomy. Am J Surg. 2005;190(1):39-42.
175. Suzuki H, Ogawa H, Shibata C, Haneda S, Watanabe K, Takahashi K-i, et al. The Long-Term Clinical Course of Pouchitis After Total Proctocolectomy and IPAA for Ulcerative Colitis. Diseases of the Colon & Rectum. 2012;55(3):330-6.
176. Ferrante M, Declerck S, De Hertogh G, Van Assche G, Geboes K, Rutgeerts P, et al. Outcome after proctocolectomy with ileal pouch-anal anastomosis for ulcerative colitis. Inflammatory bowel diseases. 2008;14(1):20-8.
177. Office for National Statistics. National life tables: England; release September 2017. 2017 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandreferencetables. Accessed: 11/03/2018.]
178. Swinburn P, Elwick H, Bean K, Curry A, Patel S, Bodger K, et al. The impact of surgery on health related quality of life in ulcerative colitis. Gut. 2012;61:A237.

179. Van Assche G, Peyrin-Biroulet L, Sturm A, Gisbert JP, Gaya DR, Bokemeyer B, et al. Burden of disease and patient-reported outcomes in patients with moderate to severe ulcerative colitis in the last 12 months - Multicenter European cohort study. Dig Liver Dis. 2016;48(6):592-600.
180. Van Assche G, Peyrin-Biroulet L, Ding Q, Kachroo S, Black CM, Fan T, et al. Health-related quality of life and productivity loss among moderate to severe ulcerative colitis patients treated with conventional therapies in Europe: the UC cares (ulcerative colitis condition, attitude, resources and educational study). United European Gastroenterology Journal. 2015;1):A612-A3.
181. Gibson PR, Vaizey C, Black CM, Nicholls R, Weston AR, Bampton P, et al. Relationship between disease severity and quality of life and assessment of health care utilization and cost for ulcerative colitis in Australia: A cross-sectional, observational study. Journal of Crohn's and Colitis. 2014;8(7):598-606.
182. Casellas F, Arenas JI, Baudet JS, Fabregas S, Garcia N, Gelabert J, et al. Impairment of health-related quality of life in patients with inflammatory bowel disease: A spanish multicenter study. Inflammatory bowel diseases. 2005;11(5):488-96.
183. Marteau P, Poole C, Nielsen S, Currie C. Quality of life (health-related utility) in adults with ulcerative colitis in remission vs. mild/moderate and severe relapse: Findings from the podium study. Canadian Journal of Gastroenterology Conference: Canadian Digestive Diseases Week. 2009;23(no pagination).
184. Poole CD, Connolly MP, Nielsen SK, Currie CJ, Marteau P. A comparison of physician-rated disease severity and patient reported outcomes in mild to moderately active ulcerative colitis. Journal of Crohn's and Colitis. 2010;4(3):275-82.
185. Poole CD, Nielsen SK, Currie CJ, Marteau P, Masure JA. Quality of life (health-related utility) in adults with ulcerative colitis in remission vs. mild/moderate and severe relapse: Findings from the Podium study. Gastroenterology. 2009;1):A203.
186. Van Der Valk M, Mangen MJ, Dijkstra G, Van Bodegraven A, Fidder H, De Jong D, et al. Is there a difference in quality of life or costs between ulcerative colitis patients with a pouch or an ileostomy? Journal of Crohn's and Colitis. 2012;6:S87.
187. Van Der Valk ME, Mangen MJJ, Dijkstra G, Van Bodegraven AA, Fidder H, De Jong DJ, et al. Is there a difference in quality of life and costs between ulcerative colitis patients with a pouch or an ileostomy? Gastroenterology. 2012;1):S263.
188. Kuruvilla K, Osler T, Hyman NH. A comparison of the quality of life of ulcerative colitis patients after IPAA vs Ileostomy. Diseases of the Colon and Rectum. 2012;55(11):1131-7.
189. Punekar YS, Hawkins N. Cost-effectiveness of infliximab for the treatment of acute exacerbations of ulcerative colitis. European Journal of Health Economics. 2010;11(1):67-76.
190. Diamantopoulos A, Finckh A, Huizinga T, Sungher DK, Sawyer L, Neto D, et al. Tocilizumab in the treatment of rheumatoid arthritis: a cost-effectiveness analysis in the UK. PharmacoEconomics. 2014;32(8):775-87.
191. Sisk JE, Moskowitz AJ, Whang W, Lin JD, Fedson DS, McBean AM, et al. Cost-effectiveness of vaccination against pneumococcal bacteremia among elderly people. JAMA. 1997;278(16):1333-9.
192. Curtis L, Burns A. Unit Costs of Health & Social Care 2016 Canterbury, Kent: Personal Social Services Research Unit 2016 [Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2016/index.php. Accessed: 2 March 2018.]
193. Royal College of Physicians. Report of the results for the national clinical audit of adult inflammatory bowel disease inpatient care in the UK London: RCP. 2012 [
194. Sandborn WJR, P. Feagan, B. G. Reinisch, W. Olson, A. Johanns, J. Lu, J. Horgan, K. Rachmilewitz, D. Hanauer, S. B. Lichtenstein, G. R. de Villiers, W. J. S. Present, D. Sands, B. E. Colombel, J. F. Colectomy Rate Comparison After Treatment of Ulcerative Colitis With Placebo or Infliximab. Gastroenterology. 2009;137(4):1250-60.

195. Feagan BGR, W. Rutgeerts, P. Sandborn, W. J. Yan, S. Eisenberg, D. Bala, M. Johanns, J. Olson, A. Hanauer, S. B. The effects of infliximab therapy on health-related quality of life in ulcerative colitis patients. American Journal of Gastroenterology. 2007;102(4):794-802.
196. Armuzzi ADP, B. Lupascu, A. Fedeli, P. Leo, D. Mentella, M. C. Vincenti, F. Melina, D. Gasbarrini, G. Pola, P. Gasbarinni, A. Infliximab in the treatment of steroid-dependent ulcerative colitis. European Review for Medical and Pharmacological Sciences. 2004;8(5):231-3.
197. Feagan BG GG, Wild G, et al. . Treatment of ulcerative colitis with a humanized antibody to the alpha4beta7 integrin. New England Journal of Medicine. 2005;352(24):2499-507.
198. Loftus ES, C. Panaccione, R. Sandborn, W. Smyth, M. James, A. Xu, J. Abhyankar, B. Corticosteroid dose reduction in ulcerative colitis patients treated with vedolizumab. Gut. 2016;65:A254-A5.
199. Loftus Jr EVS, C. A. Panaccione, R. Sandborn, W. J. Smyth, M. Green, A. Xu, J. Abhyankar, B. Corticosteroid dose reduction in ulcerative colitis patients treated with vedolizumab. United European Gastroenterology Journal. 2015;1):A255-A6.
200. Sandborn WC, J. F. Panaccione, R. Lasch, K. Sankoh, S. Abhyankar, B. Deep clinical remission in patients with Ulcerative Colitis: Evaluating the effects of vedolizumab on various combinations of endoscopic and patient-reported outcomes. Journal of Crohn's and Colitis. 2015;9:S237-S8.
201. Sandborn WC, J. F. Panaccione, R. Lasch, K. Mody, R. Green, A. Abhyankar, B. Deep remission as a predictor of clinical outcomes in vedolizumab-treated patients with ulcerative colitis. Journal of Crohn's and Colitis. 2015;9:S299-S300.
202. Sandborn WC, J. F. Panaccione, R. Lasch, K. Sankoh, S. Abhyankar, B. Deep remission with vedolizumab in patients with ulcerative colitis: Evaluating various combinations of endoscopic and patient-reported outcomes. American Journal of Gastroenterology. 2014;109:S483.
203. Feagan BGP, H. Colombel, J. F. Rubin, D. T. James, A. Mody, R. Lasch, K. Effects of vedolizumab on health-related quality of life in patients with ulcerative colitis: results from the randomised GEMINI 1 trial. Alimentary Pharmacology and Therapeutics. 2017;45(2):264-75.
204. Sands BES, I. Farraye, F. A. Cheifetz, A. S. Abhyankar, B. Sankoh, S. Smyth, M. Efficacy and safety of retreatment with vedolizumab in patients with ulcerative colitis. Journal of Crohn's and Colitis. 2015;9:S37.
205. Yajnik VK, N. Dubinsky, M. Axler, J. James, A. Abhyankar, B. Lasch, K. Efficacy and Safety of Vedolizumab in Ulcerative Colitis and Crohn's Disease Patients Stratified by Age. Advances in Therapy. 2017;34(2):542-59.
206. Yajnik VK, N. Dubinsky, M. Axler, J. Green, A. Abhyankar, B. Lasch, K. Efficacy and safety of vedolizumab with advancing age in patients with ulcerative colitis: Results from the GEMINI 1 study. Journal of Crohn's and Colitis. 2015;9:S363-S4.
207. Feagan BS, B. Sankoh, S. Milch, C. Fox, I. Efficacy of vedolizumab in ulcerative colitis by prior treatment failure in gemini i, a randomized, placebo-controlled, double-blind, multicenter trial. Inflammatory bowel diseases. 2012;18:S1-S2.
208. Feagan BGR, D. T. Danese, S. Vermeire, S. Abhyankar, B. Sankoh, S. James, A. Smyth, M. Efficacy of Vedolizumab Induction and Maintenance Therapy in Patients With Ulcerative Colitis, Regardless of Prior Exposure to Tumor Necrosis Factor Antagonists. Clinical Gastroenterology and Hepatology. 2017;15(2):229-39.e5.
209. Feagan BS, C. Melmed, G. Isaacs, K. Lasch, K. Rosario, M. James, A. Abhyankar, B. Efficacy of vedolizumab with and without continued immunosuppressant use in gemini 1 and gemini 2. Gut. 2016;65:A244-A5.
210. Colombel JFLJ, E. V. Siegel, C. A. Lewis, J. D. Smyth, M. Sankoh, S. Abhyankar, B. Efficacy of vedolizumab with concomitant corticosteroid or immunomodulator

use in patients with ulcerative colitis from GEMINI 1. Journal of Crohn's and Colitis. 2015;9:S296-S7.
211. Feagan BGS, W. Smyth, M. D. Sankoh, S. Parikh, A. Fox, I. Effects of continued vedolizumab therapy for ulcerative colitis in week 6 induction therapy nonresponders. Gastroenterology. 2014;1):S-590.
212. Feagan BC, J. F. Rubin, D. Mody, R. Sankoh, S. Lasch, K. Improvements in health-related quality of life in patients with ulcerative colitis treated with vedolizumab. Journal of Crohn's and Colitis. 2014;8:S51-S2.
213. Feagan BG RP, Sands BE, et al; GEMINI 1 Study Group,. Induction and maintenance therapy with vedolizumab, a novel biologic therapy for ulcerative colitis. Gastroenterology and Hepatology. 2014;10(1):64-6.
214. Feagan BGR, P. J. Sands, B. E. Colombel, J. Sandborn, W. J. Colombel, J. Hanauer, S. B. Van Assche, G. A. Axler, J. Kim, H. Danese, S. Fox, I. Milch, C. Sankoh, S. Wyant, T. Xu, J. Parikh, A. Induction therapy for ulcerative colitis: Results of gemini i, a randomized, placebo-controlled, double-blind, multicenter phase 3 trial. Gastroenterology. 2012;1):S160.
215. Sands BEC, R. Isaacs, K. Fedorak, R. N. Abhyankar, B. Sankoh, S. Smyth, M. Infusion-related reactions with vedolizumab treatment in patients with UC or CD during the GEMINI 1 and GEMINI 2 clinical trials. Journal of Crohn's and Colitis. 2015;9:S392-S3.
216. Colombel JFS, B. E. Feagan, B. G. Loftus, E. V. Sankoh, S. Fox, I. Parikh, A. Milch, C. Integrated safety analysis of vedolizumab for the treatment of ulcerative colitis or crohn's disease. Gastroenterology. 2013;1):S113.
217. Sands BH, S. Colombel, J. F. Danese, S. Abreu, M. Ahuja, V. Ponich, T. Hilmi, I. Sankoh, S. Smyth, M. Abhyankar, B. Fox, I. Feagan, B. Reductions in corticosteroid use in patients with ulcerative colitis or crohn's disease treated with vedolizumab. American Journal of Gastroenterology. 2013;108:S503.
218. Colombel JFS, C. A. Abhyankar, B. Loftus Jr, E. V. Lewis, J. D. Sankoh, S. Smyth, M. Milch, C. Safety of vedolizumab alone or with concomitant corticosteroids and/or immunosuppressants in patients with ulcerative colitis or crohn's disease. United European Gastroenterology Journal. 2014;1):A82.
219. Bokemeyer BS, A. Axler, J. L. Curtis, R. I. Ehehalt, R. Schreiber, S. Geransar, P. James, A. Kaviya, A. Khalid, J. M. Wolf, D. C. Feagan, B. G. Sustained remission with vedolizumab in patients with moderately to severely active ulcerative colitis: A GEMINI 1 post hoc analysis of week 14 remitters. Gastroenterology. 2017;152 (5 Supplement 1):S604.
220. Sandborn WS, B. Rutgeerts, P. Sankoh, S. Rosario, M. Milch, C. Fox, I. Sustained therapeutic benefit of vedolizumab throughout 1 year in ulcerative colitis in GEMINI I, a randomized, placebo-controlled, double-blind, multicenter trial. Journal of Crohn's and Colitis. 2013;7:S138-S9.
221. D'Haens GRC, J. F. Dubinsky, M. Abhyankar, B. James, A. Lasch, K. The efficacy of vedolizumab by disease activity and prior tumor necrosis factor a antagonist failure in patients with ulcerative colitis or Crohn's disease: Post HOC analyses from the gemini 1 and gemini 2 studies. Gastroenterology. 2016;1):S804-S5.
222. D'Haens GC, J. F. Dubinsky, M. Abhyankar, B. James, A. Lasch, K. The efficacy of vedolizumab by disease activity and prior tumour necrosis factor-alpha antagonist failure in patients with ulcerative colitis or Crohn's disease: Post-hoc analyses from the GEMINI 1 and GEMINI 2 studies. Journal of Crohn's and Colitis. 2016;10:S58.
223. Lam MCWB, B. Vedolizumab for ulcerative colitis and Crohn's disease: Results and implications of GEMINI studies. Immunotherapy. 2014;6(9):963-71.
224. Feagan BR, P. Sands, B. Sandborn, W. Colombel, J. F. Hanauer, S. Van Assche, G. Axler, J. Kim, H. J. Danese, S. Fox, I. Milch, C. Sankoh, S. Wyant, T. Xu, J. Parikh, A. Vedolizumab maintenance therapy for ulcerative colitis:

Results of gemini I, a randomized, placebo-controlled, double-blind, multicenter phase 3 trial. American Journal of Gastroenterology. 2012;107:S609-S10.
225. Kobayashi TS, Y. Motoya, S. Hirai, F. Ogata, H. Ito, H. Sato, N. Ozaki, K. Watanabe, M. Hibi, T. First trough level of infliximab at week 2 predicts future outcomes of induction therapy in ulcerative colitis: A post-hoc analysis of a multicenter prospective randomized controlled trial. American Journal of Gastroenterology. 2015;110:S787.
226. Suzuki YM, S. Hirai, F. Ogata, H. Ito, H. Sato, N. Ozaki, K. Watanabe, M. Hibi, T. Infliximab therapy for Japanese patients with ulcerative colitis: Efficacy, safety, and association between serum infliximab levels and early response in a randomized, double-blind, placebo-controlled study. Journal of Crohn's and Colitis. 2015;9:S372-S3.
227. Sandborn WS, B. E. D'Haens, G. R. Vermeire, S. Schreiber, S. Danese, S. Panes, J. Feagan, B. G. Reinisch, W. Niezychowski, W. Friedman, G. Lawendy, N. Yu, D. Woodworth, D. A. Mukherjee, A. Healey, P. J. Zhang, H. Su, C. Efficacy and safety of oral tofacitinib as induction therapy in patients with moderate to severe ulcerative colitis: Results from two phase 3 randomized controlled trials. Gastroenterology. 2016;1):S157.
228. Panes JS, C. Marren, A. Yu, D. Woodworth, D. Zhang, H. Healey, P. Improvement in patient-reported outcomes in 2 Phase 3 studies of tofacitinib in patients with moderately to severely active ulcerative colitis. Journal of Crohn's and Colitis. 2016;10:S283-S4.
229. Danese SS, W. J. Panes, J. Zhang, H. Woodworth, D. Marren, A. Su, C. Onset of efficacy of tofacitinib for induction therapy in patients with active ulcerative colitis in two multinational, phase 3 clinical trials. Digestive and Liver Disease. 2017;49 (Supplement 2):e90-e1.
230. Sandborn WJD, S. Panes, J. Zhang, H. Woodworth, D. Marren, A. Su, C. Onset of efficacy of tofacitinib for induction therapy in patients with active ulcerative colitis in two multinational, phase 3 clinical trials. American Journal of Gastroenterology. 2016;111:S260-S1.
231. Feagan BV, S. Sandborn, W. J. Reinisch, W. Tarabar, D. Su, C. Niezychowski, W. Zhang, H. Woodworth, D. Yu, D. Sands, B. Tofacitinib for induction therapy in patients with active ulcerative colitis in two phase 3 clinical trials: Results by local and central endoscopic assessments. American Journal of Gastroenterology. 2016;111:S319-S20.
232. Sands BDH, G. Sandborn, W. J. Hibi, T. Su, C. Niezychowski, W. Ghosh, S. Zhang, H. Yu, D. Woodworth, D. Healey, P. Marren, A. Panes, J. Tofacitinib has induction efficacy in moderately to severely active ulcerative colitis, regardless of prior TNF inhibitor therapy. American Journal of Gastroenterology. 2016;111:S261.
233. Sandborn WJS, B. E. Danese, S. D'Haens, G. R. Vermeire, S. Schreiber, S. Feagan, B. G. Reinisch, W. Friedman, G. Woodworth, D. A. Zhang, H. Lawendy, N. Niezychowski, W. Su, C. Panes, J. Efficacy and safety of oral tofacitinib as maintenance therapy in patients with moderate to severe ulcerative colitis: Results from a phase 3 randomized controlled trial. Gastroenterology. 2017;152 (5 Supplement 1):S199.
234. Panes JR, D. T. Vermeire, S. Lindsay, J. O. Sands, B. E. Su, C. Friedman, G. Zhang, H. Kayhan, C. Manuchehri, A. Healey, P. J. Maintenance of quality of life improvement in a phase 3 study of tofacitinib for patients with moderately to severely active ulcerative colitis. Gastroenterology. 2017;152 (5 Supplement 1):S601-S2.
235. Probert CSH, S. D. Schreiber, S. Kuhbacher, T. Ghosh, S. Arnott, I. D. Forbes, A. Infliximab in moderately severe glucocorticoid resistant ulcerative colitis: a randomised controlled trial. Gut. 2003;52(7):998-1002.
236. Sandborn WJF, B. G. Marano, C. W. Strauss, R. Johanns, J. Zhang, H. Colombel, J. Reinisch, W. Gibson, P. Collins, J. Jarnerot, G. A. Rutgeerts, P. J. A phase 2/3 randomized, placebo-controlled, double-blind study to evaluate the safety and

efficacy of subcutaneous golimumab induction therapy in patients with moderately to severely active ulcerative colitis: Pursuit sc. Gastroenterology. 2012;1):S161.
237. Sandborn WF, B. Marano, C. Strauss, R. Johanns, J. Zhang, H. Colombel, J. F. Reinisch, W. Gibson, P. Collins, J. Tarabar, D. Hebzda, Z. Rutgeerts, P. Assessment of fecal markers and clinical outcomes in patients with moderately to severely active ulcerative colitis: Results from pursuit-SC induction. American Journal of Gastroenterology. 2014;109:S502-S3.
238. Feagan BG, P. Marano, C. Strauss, R. Han, C. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Collins, J. Jarnerot, G. Rutgeerts, P. Sandborn, W. Impact of golimumab sc on disease specific and generic hrqol in patients with moderately to severely active UC: Pursuit-SC induction. Inflammatory bowel diseases. 2012;18:S43.
239. Feagan BG, P. Marano, C. Strauss, R. Han, C. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Collins, J. Jarnerot, G. Rutgeerts, P. Sandborn, W. Impact of subcutaneously administered golimumab on disease specific and generic health-related quality of life in patients with moderately to severely active ulcerative colitis: Results from PURSUIT-SC induction. Journal of Crohn's and Colitis. 2013;7:S99-S100.
240. Rutgeerts PF, B. Marano, C. Strauss, R. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Gibson, P. R. Collins, J. Jarnerot, G. Sandborn, W. J. Phase 2/3 randomized, placebo-controlled, double-blind study of SC golimumab induction in moderate to Severe UC. Journal of Gastroenterology and Hepatology. 2013;28:591.
241. Sandborn WF, B. Colombel, J. F. Reinisch, W. Gibson, P. Rutgeerts, P. Weng, H. Yao, R. Marano, C. Zhang, H. Strauss, R. Relationship between clinical outcomes and disease duration, extent, and severity in patients with ulcerative colitis who received 6 weeks of treatment with golimumab. United European Gastroenterology Journal. 2014;1):A538-A9.
242. Anonymous. Safety and Efficacy of Subcutaneous Golimumab Induction Therapy in Patients with Moderately to Severely Active Ulcerative Colitis. Clinical Advances in Hematology and Oncology. 2012;10 (10):25.
243. Gibson PF, B. Marano, C. Strauss, R. Han, C. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Collins, J. Jarnerot, G. Rutgeerts, P. Sandborn, W. Subcutaneously administered golimumab induction therapy improves health states measured by EQ-5D in patients with moderately to severely active ulcerative colitis: Results from PURSUIT-SC. Journal of Crohn's and Colitis. 2013;7:S65.
244. Sandborn W, Colombel JF, Feagan B, Marano C, Strauss R, Han C, et al. Early and sustained remission after treatment with subcutaneously administered golimumab is associated with normalized health-related quality of life in patients with moderate to severe ulcerative colitis: Post-HOC analysis from pursuit induction and maintenance trials. American Journal of Gastroenterology. 2013;108:S519.
245. Sandborn WR, W. Feagan, B. Marano, C. Strauss, R. Han, C. Johanns, J. Zhang, H. Gibson, P. Collins, J. Jarnerot, G. Rutgeerts, P. Colombel, J. F. Early responses in patient-reported outcomes are predictors for mucosal healing in patients with moderately to severely active ulcerative colitis treated with golimumab: Results from pursuit SC induction. American Journal of Gastroenterology. 2013;108:S519.
246. Rutgeerts PF, B. G. Marano, C. W. Padgett, L. Strauss, R. Johanns, J. Adedokun, O. J. Guzzo, C. Zhang, H. Colombel, J. F. Reinisch, W. Gibson, P. R. Sandborn, W. J. Randomised clinical trial: A placebo-controlled study of intravenous golimumab induction therapy for ulcerative colitis. Alimentary Pharmacology and Therapeutics. 2015;42(5):504-14.
247. Sandborn WJF, B. G. Marano, C. Zhang, H. Strauss, R. Johanns, J. Adedokun, O. J. Guzzo, C. Colombel, J. Reinisch, W. Gibson, P. R. Collins, J.

Jarnerot, G. Rutgeerts, P. Subcutaneous Golimumab Maintains Clinical Response in Patients With Moderate-to-Severe Ulcerative Colitis. Gastroenterology.
248. Sandborn WF, B. Marano, C. Strauss, R. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Gibson, P. Collins, J. Jarnerot, G. Rutgeerts, P. A phase 3 randomized, placebo-controlled, double-blind study to evaluate the safety and efficacy of subcutaneous golimumab maintenance therapy in patients with moderately to severely active ulcerative colitis: Pursuit-maintenance. American Journal of Gastroenterology. 2012;107:S673-S4.
249. Rutgeerts PF, B. G. Marano, C. Strauss, R. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Gibson, P. Collins, J. Jarnerot, G. Sandborn, W. Phase 3 study to evaluate the SC golimumab maintenance therapy in moderate to severe UC: Pursuit-maintenance. Journal of Gastroenterology and Hepatology. 2013;28:593.
250. Rutgeerts PR, W. Feagan, B. Sandborn, W. Tarabar, D. Habzda, Z. Weng, H. Yao, R. Zhang, H. Adedokun, O. Marano, C. Strauss, R. Weight-based maintenance dosing of golimumab in patients with moderate to severe ulcerative colitis. United European Gastroenterology Journal. 2014;1):A529-A30.
251. Panes JS, C. Bushmakin, A. G. Cappelleri, J. C. Mamolo, C. Healey, P. Tofacitinib in active ulcerative colitis: Analysis of efficacy based on patient-reported outcomes. Journal of Crohn's and Colitis. 2014;8:S51.
252. Panes JS, C. Bushmakin, A. G. Cappelleri, J. C. Mamolo, C. Healey, P. Randomized trial of tofacitinib in active ulcerative colitis: Analysis of efficacy based on patientreported outcomes. BMC Gastroenterology. 2015;15 (1) (no pagination)(9).
253. Panes JS, W. J. Zhang, H. Yu, D. Niezychowski, W. Su, C. Evaluation of the relationship between fecal calprotectin concentrations and clinical and endoscopic outcome measures in a phase 2 study of tofacitinib, an oral janus kinase inhibitor, in active ulcerative colitis. Journal of Crohn's and Colitis. 2013;7:S107S-8.
254. Panes JS, C. Bushmakin, A. Cappelleri, J. Healey, P. Direct and indirect effects of tofacitinib on treatment satisfaction in patients with ulcerative colitis. Journal of Crohn's and Colitis. 2015;9:S380-S1.
255. Panaccione RG, S. Middleton, S. Marquez, J. Flint, L. J. van Hoogstraten H Chen, A. Zheng, H. Danese, S. Rutgeerts, P. Improvement in patient quality of life during treatment with infliximab, azathioprine, or combination infliximab+azathioprine for moderate-to-severe ulcerative colitis. American Journal of Gastroenterology. 2013;108:S544-S5.
256. Panaccione RG, S. Middleton, S. Velazquez, J. R. M. Khalif, I. Flint, L. Van Hoogstraten, H. J. Zheng, H. Danese, S. Rutgeerts, P. Infliximab, azathioprine, or infliximab + azathioprine for treatment of moderate to severe ulcerative colitis: The UC SUCCESS trial. Journal of Crohn's and Colitis. 2011;5 (1):S8.
257. Feagan BS, W. Lazar, A. Thakkar, R. Skup, M. Yang, M. Chao, J. Mulani, P. Adalimumab induction dose reduces the risk of hospitalizations and colectomies in patients with ulcerative colitis during the first 8 weeks of therapy. American Journal of Gastroenterology. 2012;107:S647.
258. Feagan BGS, W. J. Lazar, A. Thakkar, R. B. Huang, B. Reilly, N. Chen, N. Yang, M. Skup, M. Mulani, P. Chao, J. Adalimumab therapy is associated with reduced risk of hospitalization in patients with ulcerative colitis. Gastroenterology. 2014;146(1):110-8.e3.
259. Feagan BS, W. Yang, M. Lomax, K. Lazar, A. Thakkar, R. Mulani, P. Chao, J. Adalimumab therapy reduces hospitalization and colectomy rates in patients with ulcerative colitis: Data from controlled trials. American Journal of Gastroenterology. 2011;106:S452.
260. Feagan BS, M. Ozbay, A. B. Lazar, A. Thakkar, R. Chen, N. Chao, J. Sandborn, W. Adalimumab therapy reduces hospitalization rates in patients with moderate ulcerative colitis. American Journal of Gastroenterology. 2014;109:S480.

261. Sandborn WF, B. Yang, M. Skup, M. Thakkar, R. Lazar, A. Mulani, P. Chao, J. Adalimumab therapy reduces hospitalization rates in patients with ulcerative colitis among initial responders. Inflammatory bowel diseases. 2012;18:S40.
262. Panaccione RC, J. F. Sandborn, W. Reinisch, W. D'Haens, G. R. Hommes, D. W. Ghosh, S. Pappalardo, B. Petersson, J. Robinson, A. M. Lazar, A. Zhou, Q. Thakkar, R. Durable clinical remission and response in adalimumab-treated patients with ulcerative colitis. Gastroenterology. 2015;1):S264.
263. Colombel JFS, W. J. Ghosh, S. Wolf, D. C. Panaccione, R. Feagan, B. Reinisch, W. Robinson, A. M. Lazar, A. Kron, M. Huang, B. Skup, M. Thakkar, R. B. Four-year maintenance treatment with adalimumab in patients with moderately to severely active ulcerative Colitis: Data from ULTRA 1, 2, and 3. American Journal of Gastroenterology. 2014;109(11):1771-80.
264. Sandborn WJVA, G. Reinisch, W. Colombel, J. D'Haens, G. Wolf, D. C. Kron, M. Tighe, M. Lazar, A. Thakkar, R. B. Induction and maintenance of clinical remission by adalimumab in patients with moderate-to-severe ulcerative colitis. Journal of Crohn's and Colitis. 2011;5 (1):S68.
265. Sandborn WVA, G. Reinisch, W. Colombel, J. F. D'Haens, G. Wolf, D. Kron, M. Tighe, M. Lazar, A. Thakkar, R. Induction and maintenance of clinical remission by adalimumab in patientswith moderate-to-severe Ulcerative Colitis. Inflammatory bowel diseases. 2011;17:S3-S4.
266. Hanauer SS, W. J. Colombel, J. F. Reinisch, W. Vermeire, S. Petersson, J. Pappalardo, B. Robinson, A. M. Lazar, A. Zhou, Q. Thakkar, R. B. Rapid changes in laboratory parameters and early response to adalimumab: Pooled analysis from ULTRA 1 and ULTRA 2. Journal of Crohn's and Colitis. 2015;9:S289-S90.
267. Sandborn WJW, D. C. Van Assche, G. Yang, M. Mulani, P. M. Reilly, N. Lazar, A. Thakkar, R. B. Chao, J. Rapid onset of adalimumab and long-term efficacy among week-8responders in adults with moderate to severe active Ulcerative Colitis. Inflammatory bowel diseases. 2011;17:S4.
268. Lewis JDG, S. Sandborn, W. Van Assche, G. A. D'Haens, G. R. Lazar, A. Eichner, S. Huang, B. Robinson, A. Thakkar, R. Rates of "patient-defined" remission with adalimumab in patients with ulcerative colitis: Subanalysis of ultra 1 and ultra 2. Gastroenterology. 2013;1):S232.
269. Colombel JFG, S. Sandborn, W. Van Assche, G. Reinisch, W. Lazar, A. Eichner, S. Huang, B. Robinson, A. Thakkar, R. Safety of adalimumab in global clinical trials of ulcerative colitis patients. Inflammatory bowel diseases. 2013;19:S78-S9.
270. Reinisch WS, W. J. Hommes, D. W. D'Haens, G. R. Hanauer, S. B. Schreiber, S. Panaccione, R. Fedorak, R. N. Tighe, M. B. Huang, B. Lazar, A. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis. Gastroenterology. 2010;1):S114-S5.
271. Reinisch WS, W. Thakkar, R. Lazar, A. Huang, B. Mulani, P. Chao, J. Adalimumab induction therapy improves health-related quality of life in patients with moderately to severely active ulcerative colitis. American Journal of Gastroenterology. 2010;105:S441-S2.
272. Feagan BS, W. Yang, M. Skup, M. Thakkar, R. Lazar, A. Mulani, P. Chao, J. Adalimumab therapy reduces hospitalization and colectomy rates in patients with ulcerative colitis among initial responders. American Journal of Gastroenterology. 2012;107:S642-S3.
273. Sandborn WJVA, G. Thakkar, R. B. Lazar, A. Kron, M. Yang, M. Chao, J. Mulani, P. M. Adalimumab improves health-related quality of life for 52 weeks in patients with ulcerative colitis. Journal of Crohn's and Colitis. 2011;5 (1):S87-S8.
274. Sandborn WT, R. Lazar, A. Skup, M. Yang, M. Chao, J. Mulani, P. Van Assche, G. Time-in-remission analysis of adalimumab vs. placebo for the treatment of ulcerative colitis. American Journal of Gastroenterology. 2012;107:S641-S2.

275. Colombel JFS, W. Van Assche, G. D'Haens, G. Wolf, D. Kron, M. Lazar, A. Robinson, A. Chao, J. Thakkar, R. Benefit-risk assessment of adalimumab as maintenance treatment for ulcerative colitis: Near analysis of week 8 responders in ultra 2. American Journal of Gastroenterology. 2012;107:S643-S4.
276. D'Haens GVA, G. Wolf, D. Sandborn, W. Colombel, J. F. Lazar, A. Kron, M. Robinson, A. Thakkar, R. Mucosal healing in ulcerative colitis patients with week 8 response to adalimumab: Subanalysis of ultra 2. Inflammatory bowel diseases. 2012;18:S21.
277. Sandborn WJC, J. F. D'Haens, G. Van Assche, G. Wolf, D. Kron, M. Lazar, A. Robinson, A. M. Yang, M. Chao, J. D. Thakkar, R. One-year maintenance outcomes among patients with moderately-to-severely active ulcerative colitis who responded to induction therapy with adalimumab: Subgroup analyses from ULTRA 2. Alimentary Pharmacology and Therapeutics. 2013;37(2):204-13.
278. Sandborn WJDH, G. R. Colombel, J. F. Van Assche, G. A. Wolf, D. C. Kron, M. Lazar, A. Robinson, A. Chao, J. Thakkar, R. One-year response and remission rates in ulcerative colitis patients with week 8 response to adalimumab: subanalysis of ultra 2. Gastroenterology. 2012;1):S565.
279. Van Assche GW, D. D'Haens, G. Sandborn, W. Colombel, J. F. Lazar, A. Kron, M. Robinson, A. Thakkar, R. Reduced steroid usage in ulcerative colitis patients with week 8 response to adalimumab: Subanalysis of ultra 2. Inflammatory bowel diseases. 2012;18:S21-S2.
280. Ghosh SW, D. Sandborn, W. Colombel, J. F. Lazar, A. Eichner, S. Robinson, A. Thakkar, R. Sustained efficacy in patients with ulcerative colitis treated with adalimumab: Results from ULTRA 2. Journal of Crohn's and Colitis. 2013;7:S238.
281. Colombel JFP, S. Sandborn, W. D'Haens, G. Robinson, A. Petersson, J. Huang, B. Lazar, A. Thakkar, R. Time to remission and response in adalimumabtreated patients with moderately to severely active ulcerative colitis from ultra 2. United European Gastroenterology Journal. 2013;1):A219-A20.
282. Sandborn WDH, G. Colombel, J. F. Van Assche, G. Wolf, D. Kron, M. Lazar, A. Robinson, A. Chao, J. Thakkar, R. One-year response and remission rates in ulcerative colitis patients with week 8 response to adalimumab: Subanalysis of ULTRA 2. Journal of Crohn's and Colitis. 2012;6:S92.
283. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making. 2013;33(5):607-17.
284. MRC Biostatistics Unit. Frequently Asked Questions about DIC. [Available from: https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-dic/#q8. Accessed: 11 April 2018.]
285. Sutton AJ, Welton NJ, Cooper N. Evidence Synthesis for Decision Making in Healthcare: John Wiley & Sons; 2012.
286. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 5: the baseline natural history model. Med Decis Making. 2013;33(5):657-70.
287. Bratasiuk AM, Niroda AI. Efficiency comparison of tofacitinib and budesonid in treatment of nonspecific ulcerative colitis. Wiadomosci lekarskie (Warsaw, Poland : 1960). 2017;70(2):241.
288. Beilman CL, Thanh NX, Ung V, Ma C, Wong K, Kroeker KI, et al. Real-Life Treatment Paradigms Show Adalimumab Is Cost-Effective for the Management of Ulcerative Colitis. Canadian Journal of Gastroenterology and Hepatology. 2016;2016 (no pagination)(5315798).
289. Lofland JH, Mallow P, Rizzo J. Cost-per-remission analysis of infliximab compared to adalimumab among adults with moderate-to-severe ulcerative colitis. Journal of Medical Economics. 2013;16(4):461-7.
290. Moradi N, Tofighi S, Zanganeh M, Sari AA, Khedmat H, Zarei L. Economic evaluation of infliximab for treatment of refractory ulcerative colitis in iran: Cost-effectiveness analysis. Iranian Journal of Pharmaceutical Sciences. 2016;12(4):33-42.

291. Park KT, Tsai R, Perez F, Cipriano LE, Bass D, Garber AM. Cost-effectiveness of early colectomy with ileal pouch-anal anastamosis versus standard medical therapy in severe ulcerative colitis. Annals of Surgery. 2012;256(1):117-24.
292. Stawowczyk E, Kawalec P, Pilc A. Cost-utility analysis of 1-year treatment with adalimumab/standard care and standard care alone for ulcerative colitis in Poland. European Journal of Clinical Pharmacology. 2016;72(11):1319-25.
293. Stawowczyk E, Kawalec P, Pilc A. Cost-Utility Analysis of Infliximab with Standard Care versus Standard Care Alone for Induction and Maintenance Treatment of Patients with Ulcerative Colitis in Poland. Pharmacotherapy. 2016;36(5):472-81.
294. Stawowczyk E, Kawalec P, Pilc A. Cost-effectiveness analysis of 1-year treatment with golimumab/standard care and standard care alone for ulcerative colitis in Poland. PLoS ONE. 2016;11 (8) (no pagination)(e0160444).
295. Stern S, Ward AJ, Saint-Laurent Thibault C, Camacho F, Rahme E, Naessens D, et al. Cost-effectiveness of golimumab for the treatment of patients with moderate-to-severe ulcerative colitis in Quebec using a patient level state transition microsimulation. Journal of Medical Economics. 2017:1-11.
296. Toor K, Druyts E, Jansen JP, Thorlund K. Cost per remission and cost per response with infliximab, adalimumab, and golimumab for the treatment of moderately-to-severely active ulcerative colitis. Journal of Medical Economics. 2015;18(6):437-46.
297. Ung V, Thanh NX, Wong K, Kroeker KI, Lee T, Wang H, et al. Real-life treatment paradigms show infliximab is cost-effective for management of ulcerative colitis. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2014;12(11):1871-8 e8.
298. Xie F, Blackhouse G, Assasi N, Gaebel K, Robertson D, Goeree R. Cost-utility analysis of infliximab and adalimumab for refractory ulcerative colitis. Cost Effectiveness and Resource Allocation. 2009;7 (no pagination)(20).
299. Yokomizo L, Limketkai B, Park KT. Cost-effectiveness of adalimumab, infliximab or vedolizumab as first-line biological therapy in moderate-to-severe ulcerative colitis. BMJ Open Gastroenterology. 2016;3 (1) (no pagination)(e000093).
300. Beilman C, Nguyen T, Ung V, Ma C, Wong K, Kroeker K, et al. Cost-utility analysis shows adalimumab is cost-effective for the management of ulcerative colitis. Canadian Journal of Gastroenterology and Hepatology Conference. 2016(pagination).
301. Davies SEC, Inserra C, Ardizzone S, Muserra G, Medaglia M, Saporito T. Cost-utility analyses (CUA) of biological agents for refractory ulcerative colitis: The importance of local analysis. International Journal of Clinical Pharmacy. 2017;39 (1):314.
302. De Paula E, Blumer VD, Trevizam YC, Asano EW. Cost-minimization analysis of infliximab versus adalimumab in the treatment of Crohn's disease and ulcerative colitis. Value in Health. 2015;18 (3):A296.
303. Decimoni TM, G.; Patel, H.; Demuth, D. Cost per clinical outcomes with vedolizumab for the treatment of moderately to severely active ulcerative colitis in Brazil. United European Gastroenterology Journal. 2016;2(Supplement 1).
304. Ghosh S, Desjardins O, Skup M, Wang S, Yang M, Yang H, et al. Cost-effectiveness of adalimumab for the treatment of moderately to severely active ulcerative colitis in Canada. Value in Health. 2014;17 (3):A37.
305. Holubar S, Piazik B, Xu K, Dulai P, Tosteson A, Siegel C, et al. Cost-effectiveness of infliximab versus colectomy for severe ulcerative colitis: A markov analysis. Inflammatory bowel diseases. 2012;18:S57-S8.
306. Jones J, Worobetz L, Bedi A, Manns B. A cost-utility analysis of infliximab compared with surgery for the treatment of Ulcerative Colitis; a public payer perspective. Gastroenterology. 2009;1):A172.
307. Liu Y, Ozbay AB, Skup M, Reichmann W, Diener M, Chao J. Comparison of cost per responder and remitter in patients with inflammatory bowel disease in the united states: An indirect comparison of adalimumab and vedolizumab. American Journal of Gastroenterology. 2014;109:S479.

308. Liu Y, Reichmann B, Yan S, Macaulay D, Skup M, Mulani P, et al. Indirect comparison of adalimumab and infliximab in ulcerative colitis: Cost per remitter analysis. American Journal of Gastroenterology. 2013;108:S520.
309. Lofland J, Gunnarsson C, Mallow P, Rizzo J, Feagan B. Cost-per-responder analysis of infliximab compared to adalimumab among individuals with moderate to severe ulcerative colitis. American Journal of Gastroenterology. 2011;106:S418.
310. Mallow P, Rizzo J, Gathany T, Lofland J. Cost per clinical response, clinical remission, and mucosal healing of golimumab among patients with ulcerative colitis: Results from PURSUIT maintenance. Inflammatory bowel diseases. 2013;19:S52.
311. Mallow P, Rizzo J, Queener MK, Gathany T, Lofland J. Cost per clinical remission and clinical response of adalimumab, golimumab, and infliximab in the treatment of patients with ulcerative colitis: An induction analysis. American Journal of Gastroenterology. 2013;108:S515-S6.
312. Malone D, Hurwitz J, Lofland J, Nejadnik B, Vanderpoel J, Waters H. Cost-effectiveness of immunomodulators versus infliximab for adults with moderate-to-severe ulcerative colitis. Inflammatory bowel diseases. 2011;17:S53-S4.
313. Monfared AAT, Feagan BG. Cost-effectiveness of infliximab scheduled maintenance treatment for the management of patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2013;1):S637.
314. Moradi N, Tofighi S, Zanganeh M, Akbari Sari A, Zarei L. Economic evaluation of infliximab for treatment of refractory ulcerative colitis in Iran: Cost-effectiveness analysis. Value in Health. 2015;18 (7):A628.
315. Odes SH, Greenberg D, Vardi H, Friger M, Stockbrugger RW, Langholz E, et al. The cost-effectiveness of infliximab is different in ulcerative colitis and crohn's disease: Markov model in a population-based incident patient cohort. Gastroenterology. 2013;1):S639.
316. Odes SH, Vardi H, Greenberg D, Friger M, Stockbrugger RW, O'Morain CA, et al. Cost-effectiveness of episodic or maintenance infliximab treatment versus standard treatment in a community-based incidence cohort of adult ulcerative colitis patients with 10-years follow-up. Gastroenterology. 2012;1):S256.
317. Park KT, Yokomizo L. First-line biologic treatment for moderate-to-severe ulcerative colitis: A cost-per-remission decision analysis. American Journal of Gastroenterology. 2015;110:S777.
318. Park KT, Yokomizo L. Adalimumab versus infliximab in inducing mucosal healing in moderate-to-severe ulcerative colitis: A cost-per-remission decision analytic study. Gastroenterology. 2015;1):S199.
319. Pichardo-Pina CA, Sanchez-Casillas JL. Analysis of cost minimization of adalimumab compared with infliximab for the treatment of ulcerative colitis in the public market in Mexico. Value in Health. 2015;18 (7):A867.
320. Pyadushkina E, Derkach EV. Cost-Effectiveness of Vedolizumab for the Treatment of Adult Patients with Moderately-to-Severely Active Ulcerative Colitis in Russia. Value in Health. 2017;20(9):A635.
321. Schneider Y, Gold S, Saumoy M, Cohen-Mekelburg SA, Scherl E, Steinlauf AF. Utilizing a Markov Model to Determine the Cost-Effectiveness of Vedolizumab Compared to Adalimumab for patient's with Ulcerative Colitis after Loss of Response to Infliximab. Gastroenterology. 2017;152(5):S588-S9.
322. Thorlund K, Druyts E, Eapen S, Mills E. Cost-effectiveness of golimumab versus infliximab and adalimumab for the treatment of moderate to severe ulcerative colitis. Value in Health. 2014;17 (3):A38.
323. Thorlund K, Druyts E, Kanters S, Eapen S, Mills E. Cost-effectiveness of golimumab versus infliximab and adalimumab for the treatment of moderate to severe ulcerative colitis. Gastroenterology. 2014;1):S-202.
324. Ung V, Lee TW, Wang H, Kroeker KI, Nguyen TX, Ohinmaa A, et al. Long-term cost-effectiveness analysis of infliximab in the management of ulcerative colitis. Gastroenterology. 2011;1):S201.

325. Wang S, Yang H, Yang M, Bao Y. Cost-utility analysis of adalimumab for the treatment of moderate-to-severe ulcerative colitis in patients in Spain. Value in Health. 2015;18 (7):A670.
326. Wilson M, Mody R, Ursan I, Carr S. Cost-effectiveness of vedolizumab compared with conventional therapy and biologics for treatment of moderately to severely active ulcerative colitis in the United States. Gastroenterology. 2015;1):S198-S9.
327. Xie F, Blackhouse G, Assasi N, Hopkins R, Gaebel K, O'Reilly D, et al. Cost effectiveness analysis of anti-tnf-alpha; drugs for refractory ulcerative colitis. Value in Health. 2009;12 (3):A58.
328. Hanauer SB, Sandborn WJ, Dallaire C, Archambault A, Yacyshyn B, Yeh C, et al. Delayed-release oral mesalamine 4.8 g/day (800 mg tablets) compared to 2.4 g/day (400 mg tablets) for the treatment of mildly to moderately active ulcerative colitis: The ASCEND I trial. Canadian journal of gastroenterology = Journal canadien de gastroenterologie. 2007;21(12):827-34.
329. Hanauer SB, Sandborn WJ, Kornbluth A, Katz S, Safdi M, Woogen S, et al. Delayed-release oral mesalamine at 4.8 g/day (800 mg tablet) for the treatment of moderately active ulcerative colitis: the ASCEND II trial. The American journal of gastroenterology. 2005;100(11):2478-85.
330. Netten AC, L. Unit costs of health and social care 2006. Caterbury, UK: PSSRU Personal Social Services Research Unit; 2006.
331. Woehl A, Hawthorne AB, Morgan CL, Punekar Y, McEwan P. The epidemiology and health care resource use in patients with Crohn's disease: A population based UK study. Value in Health. 2007;10(6):A355.
332. Essat MT, P.; Ren, S.; Bessy, A.; Archer, R.; Wong, R.; Hoque, S.; Lobo, A. Vedolizumab for the treatment of adults with moderately to severely active ulcerative colitis: a single technology appraisal.: School of Health and Related Research (ScHARR); 2014.
333. Royal College of Physicians. National clinical audit of biological therapies: UK inflammatory bowel disease (IBD) audit. 2013 [
334. Panes J, Nos P, Domenech E, Marin-Jimenez I, Riestra S, Gisbert JP, et al. Quality of life in ulcerative colitis is already impaired by mild disease activity. United European Gastroenterology Journal. 2015;1):A422.
335. Vaizey C, Gibson PR, Black CM, Nicholls RJ, Weston AR, Gaya DR, et al. The relationship between disease severity, quality of life and health care resource utilization among United Kingdom patients with ulcerative colitis. Journal of Crohn's and Colitis. 2013;7:S259.
336. Louis E, Roughley A, Thakkar R, Yang M, Skup M, Chao J, et al. Impact of ulcerative colitis on patient quality of life in a real-world clinical setting. Gastroenterology. 2013;1):S421.
337. Van Assche G, Peyrin-Biroulet L, Fan T, Lara N, Lynam M, Rojas-Farreras S, et al. Impact of active ulcerative colitis on health related quality of life and productivity loss. Value in Health. 2014;17 (3):A40.
338. Leidl R, Reitmeir P, Konig HH, Stark R. The performance of a value set for the EQ-5D based on experienced health states in patients with inflammatory bowel disease. Value in Health. 2012;15(1):151-7.
339. Panes J, Domenech E, Peris MA, Nos P, Riestra S, Julia de Paramo B, et al. Association between disease activity and quality of life in ulcerative colitis: Results from the CRONICA-UC study. Journal of Gastroenterology and Hepatology (Australia). 2017;32(11):1818-24.
340. Kawalec P, Stawowczyk E. A physician based assessment of disease activity, quality of life and costs of ulcerative colitis in Poland. Value in Health. 2017;20 (5):A219.
341. Kawalec P. Disease activity, quality of life and indirect costs associated with ulcerative colitis in Poland. Value in Health. 2016;19 (7):A585.
342. Kim ESC, K. B. Park, K. S. Jang, B. I. Kim, K. O. Jeon, S. W. Jung, M. K. Kim, E. Y. Yang, C. H. Predictive factors of impaired quality of life in korean patients with

inactive inflammatory bowel disease: Association with functional gastrointestinal disorders and mood disorders. Journal of Clinical Gastroenterology. 2013;47(4):e38-e44.
343. Casellas FBDA, M. Iglesias, M. Robles, V. Nos, P. Aguas, M. Riestra, S. De Francisco, R. Papo, M. Borruel, N. Mucosal healing restores normal health and quality of life in patients with inflammatory bowel disease. European Journal of Gastroenterology and Hepatology. 2012;24(7):762-9.
344. Casellas Jorda F, Barreiro-de Acosta M, Iglesias M, Robles V, Nos P, Aguas M, et al. Restoration of health in patients with inflammatory bowel disease and mucosal healing. Journal of Crohn's and Colitis. 2012;6:S46.
345. Travis S, Feagan BG, Peyrin-Biroulet L, Panaccione R, Danese S, Lazar A, et al. Effect of Adalimumab on Clinical Outcomes and Health-related Quality of Life Among Patients With Ulcerative Colitis in a Clinical Practice Setting: Results From InspirADA. Journal of Crohn's & colitis. 2017;11(11):1317-25.
346. Travis S, Feagan B, Peyrin-Biroulet L, Panaccione R, Danese S, Lazar A, et al. Effect of adalimumab on clinical outcomes and health-related quality of life amongst patients with ulcerative colitis in a clinical practice setting: Results from INSPIRADA. Journal of Crohn's and Colitis. 2016;10:S391-S2.
347. Travis S, Feagan BG, Peyrin-Biroulet L, Panaccione R, Danese S, Lazar A, et al. Effect of adalimumab on clinical outcomes and health-related quality of life among patients with ulcerative colitis in a clinical practice setting: Results from INSPIRADA. Gastroenterology. 2016;1):S633.
348. Irving P, Probert C, Gaya D, Hamlin PJ, Sebastian S, Gillespie G, et al. Patient-reported quality of life during golimumab induction for moderate to severe ulcerative colitis in the United Kingdom: Results from the GO-COLITIS study. Gut. 2016;65:A87.
349. Irving P, Probert C, Gaya DR, Hamlin PJ, Sebastian S, Gillespie G, et al. Patient-reported quality of life during golimumab induction for moderate-to-severe ulcerative colitis in the United Kingdom: Results from the GO-COLITIS study. Journal of Crohn's and Colitis. 2016;10:S250-S1.
350. Irving P, Cummings F, Bloom SL, Lazar A, Robinson AM, Pappalardo BL, et al. Effect of adalimumab on patients with moderate to severe ulcerative colitis in UK clinical practice setting: Results from Inspirada. Gut. 2016;65:A86-A7.
351. Piercy J, Ozbay AB, Roughley A, Chao J, Skup M. Comparison of ulcerative colitis with crohn's disease from the perspective of patient burden. Gastroenterology. 2015;1):S839.
352. van Gennep S, Sahami S, Buskens CJ, van den Brink GR, Ponsioen CY, D'Hoore A, et al. Comparison of health-related quality of life and disability in ulcerative colitis patients following restorative proctocolectomy with ileal pouch-anal anastomosis versus anti-tumor necrosis factor therapy. Eur J Gastroenterol Hepatol. 2017;29(3):338-44.
353. Van Gennep S, Sahami S, Bemelman W, Buskens C, D'Hoore A, De Buck van Overstraeten A, et al. Disability and health-related quality of life in patients with moderate-severe ulcerative colitis: Impact of restorative proctocolectomy versus medical treatment with anti-tumour necrosis factor agents. Journal of Crohn's and Colitis. 2016;10:S193.
354. Arebi N, Karapiperis D, Tripoli S, Poonja N. Cost-effectiveness of surgery and infliximab for refractory ulcerative colitis: A single centre study using patient-level data. Journal of Crohn's and Colitis. 2013;7:S115-S6.
355. Erem HK, R. Ozdemir, Y. Aytac, E. Gurland, B. Manilich, E. Remzi, F. Preoperative quality of life and postoperative changes evaluation in colorectal surgery with "the knowledge program". Diseases of the Colon and Rectum. 2012;55 (5):e248-e9.
356. Brown C, Gibson PR, Hart A, Kaplan GG, Kachroo S, Ding Q, et al. Long-term outcomes of colectomy surgery among patients with ulcerative colitis. Springerplus. 2015;4:573.
357. Brown C, Gibson P, Hart A, Kaplan G, Hautamaki E, Flood E, et al. Long-term impacts of colectomy surgery among ulcerative colitis patients study (LOCUS): The final analysis. Journal of Crohn's and Colitis. 2013;7:S190.

358. Brown C, Gibson P, Hart A, Kaplan G, Hautamaki E, Flood E, et al. Long-term impacts of colectomy surgery among ulcerative colitis patients study (LOCUS): A preliminary analysis. Journal of Crohn's and Colitis. 2012;6:S122.
359. Kaplan G, Hart A, Brown C, Gibson P, Fan T, Stokes L, et al. Factors that influence ulcerative colitis patients' anxiety and depression following the colectomy surgery. Value in Health. 2013;16 (3):A215-A6.
360. Richards DMH, S. A. Irving, M. H. Scott, N. A. Patient quality of life after successful restorative proctocolectomy is normal. Colorectal Disease. 2001;3(4):223-6.
361. Holubar S, Cima R, Pemberton J, Loftus E, Larson D, O'Byrne M. Does referral bias influence reported outcomes after ileal pouch-anal anastomosis for chronic ulcerative colitis? A population-based study in olmsted county, minnesota. Diseases of the Colon and Rectum. 2010;53 (4):607-8.
362. Yarlas A, D'Haens G, Willshire D, Teynor M. Health-related quality of life varies as a function of remission status in patients with mildto-moderate ulcerative colitis receiving short-term and long-term daily therapy with multimatrix mesalazine. United European Gastroenterology Journal. 2015;1):A252.
363. Yarlas A, D'Haens G, Willshire D, Teynor M. Health-related quality of life varies as a function of remission status in patients with mild-to-moderate ulcerative colitis receiving short-term and long-term daily therapy with multimatrix mesalazine. Journal of Gastroenterology and Hepatology (Australia). 2015;30:144-5.
364. Hodgkins P, Yen L, Yarlas A, Karlstadt R, Solomon D, Kane S. Impact of MMX(R) mesalamine on improvement and maintenance of health-related quality of life in patients with ulcerative colitis. Inflammatory bowel diseases. 2013;19(2):386-96.
365. Rokke O, Iversen K, Olsen T, Ristesund SM, Eide GE, Turowski GE. Long-term followup of patients with active J-reservoirs after restorative proctocolectomy for ulcerative colitis with regard to reservoir function, mucosal changes, and quality of life. ISRN Gastroenterology. 2011;2011 (no pagination)(430171).
366. Rokke O, Iversen K, Olsen T, Ristesund S-M, Eide GE, Turowski GE. Long-term followup with evaluation of the surgical and functional results of the ileal pouch reservoir in restorative proctocolectomy for ulcerative colitis. ISRN gastroenterol. 2011;2011:625842.
367. Yen L, Yarlas A, Hodgkins P. Concordance among multiple health outcomes measures in responsiveness to multi-matrix mesalamine maintenance treatment for patients with quiescent ulcerative colitis. Value in Health. 2011;14 (3):A184.
368. Graff LA, Walker JR, Lix L, Clara I, Rawsthorne P, Rogala L, et al. The Relationship of Inflammatory Bowel Disease Type and Activity to Psychological Functioning and Quality of Life. Clinical Gastroenterology and Hepatology. 2006;4(12):1491-501.e1.
369. Hauser W, Dietz N, Steder-Neukamm U, Janke KH, Stallmach A. Biopsychosocial determinants of health-related quality of life after ileal pouch anal anastomosis for ultracerative colitis. Inflammatory bowel diseases. 2004;10(4):399-407.
370. Provenzale D, Shearin M, Phillips-Bute BG, Drossman DA, Li Z, Tillinger W, et al. Health-related quality of life after ileoanal pull-through: Evaluation and assessment of new health status measures. Gastroenterology. 1997;113(1):7-14.
371. Soetikno RM, Mrad R, Pao V, Lenert LA. Quality-of-life research on the Internet: feasibility and potential biases in patients with ulcerative colitis. Journal of the American Medical Informatics Association : JAMIA. 1997;4(6):426-35.
372. Oliveira A, Rosete M, Manso A, Leite J. Biological therapy or ileoanal pouch anastomosis in ulcerative colitis. Colorectal Disease. 2017;19 (Supplement 2):19.
373. Kitsenko Y, Tulina I, Tsarkov P. Does the time of ileal pouch creation in three-stage surgery for ulcerative colitis affect the quality of life? Colorectal Disease. 2017;19 (Supplement 2):45.
374. Kostrejova M, Duricova D, Lukas M, Pekny P, Bortlik M, Mudr R, et al. Long-term functional outcome and quality of life after restorative proctocolectomy. Gastroenterologie a Hepatologie. 2015;69:2S13.

375. Holt N, Pavli P, Subramaniam K. Quality of life in patients with ulcerative colitis. Journal of Gastroenterology and Hepatology (Australia). 2017;32:128.
376. Cekic C, Aslan F, Arabul M, Ipek S, Yuksel ES, Alper E, et al. The effect of intravenous iron treament on quality of life in Inflammatory Bowel Disease patients with non-anemic iron deficiency. Journal of Crohn's and Colitis. 2015;9:S336.
377. Kostrejova M, Duricova D, Lukas M, Pekny P, Bortlik M, Dosedel J, et al. Long-term functional outcome and quality of life after restorative proctocolectomy and ileal-pouch anal anastomosis in ulcerative colitis patients. United European Gastroenterology Journal. 2015;1):A222.
378. Cohan J, Rhee J, Finlayson E, Varma M. Proctocolectomy: Impact on relationship quality in ulcerative colitis patients and their partners. Diseases of the Colon and Rectum. 2015;58 (5):e118-e20.
379. Burisch J, Weimers P, Pedersen N, Cukovic-Cavka S, Vucelic B, Kaimakliotis I, et al. Health-related quality of life improves during one year of medical and surgical treatment in a European population-based inception cohort of patients with Inflammatory Bowel Disease - An ECCO-EpiCom study. Journal of Crohn's and Colitis. 2014;8(9):1030-42.
380. Koerdt S, Jehle EC, Kreis ME, Kasparek MS. Quality of life after proctocolectomy and ileal pouch-anal anastomosis in patients with ulcerative colitis. International Journal of Colorectal Disease. 2014;29(5):545-54.
381. Kozlowska KA, Baczyk G, Krokowicz P. Quality of life in patients with ulcerative colitis treated surgically. Przeglad Gastroenterologiczny. 2014;9(4):220-6.
382. Gudlaugsdottir K, Valsdottir EB, Stefansson TB. Quality of life after colectomy due to ulcerative colitis. Colorectal Disease. 2014;16:136.
383. Flores C, Calixto R, Francesconi CF. Inflammatory bowel disease: Low impact on scores of quality of life, depression and anxiety in patients attending a tertiary care center in Brazil. Gastroenterology. 2013;1):S653-S4.
384. Bours P, Bakia JM, Jonkers D, Romberg-Camps M, Masclee A, Van Heurn E, et al. Long-term quality of life and sexual function after ileal pouch-anal anastomosis in adults with ulcerative colitis or familial adenomatous polyposis. Gastroenterology. 2013;1):S641.
385. Strid H, Ohman L, Bajor A, Lasson A, Svedlund J, Simren M. Three-year clinical follow-up in a Swedish cohort of patients with newly diagnosed ulcerative colitis. Journal of Crohn's and Colitis. 2013;7:S143-S4.
386. Pavlides M, Rahman M, Cleland J, Gaunt R, Travis S, Mortensen N, et al. Outcomes after ileal pouch anal anastomosis in patients with primary sclerosing cholangitis. Journal of Crohn's and Colitis. 2013;7:S82.
387. Jonefjall B, Strid H, Ohman L, Svedlund J, Bergstedt A, Simren M. Characterization of IBS-like symptoms in patients with ulcerative colitis in clinical remission. Gastroenterology. 2012;1):S653.
388. Iglesias M, Barreiro-de Acosta M, Caamano F, Vazquez I, Lorenzo Gonzalez A, Domingez-Munoz JE. Influence of psychological variables on health-related quality of life on ulcerative colitis patients. Journal of Crohn's and Colitis. 2012;6:S49.
389. Ince M, Kirat HT, Geisler DP, Remzi FH, Kiran RP. The negative effects of surgery persist beyond the early postoperative period after laparoscopic colorectal resection. Techniques in Coloproctology. 2011;15(2):173-7.
390. Rubin DTP, J. Lindsay, J. O. Vermeire, S. Yarlas, A. Bayliss, M. Cappelleri, J. C. Maher, S. Bushmakin, A. G. Chen, L. A. Manuchehri, A. Healey, P. J. The SF-36 health survey distinguishes disease burden on functioning and well-being between patients with active vs inactive ulcerative colitis. Gastroenterology. 2017;152 (5 Supplement 1):S338.
391. Chen LA, Yarlas A, Bayliss M, Cappelleri JC, Maher S, Bushmakin AG, et al. A Systematic Review of the SF-36® Health Survey for Measuring Health-Related Quality of Life in Patients with Ulcerative Colitis. Gastroenterology. 2017;152(5):S338.

392. Blczyk G, Formanowicz D, Gmerek L, Krokowicz P. Health-related quality of life assessment among patients with inflammatory bowel diseases after surgery-review. Przeglad Gastroenterologiczny. 2017;12(1):6-16.
393. Schiergens TSH, V. Schobel, T. N. Englert, G. H. Kreis, M. E. Thasler, W. E. Werner, J. Kasparek, M. S. Long-term Quality of Life of Patients with Permanent End Ileostomy: Results of a Nationwide Cross-Sectional Survey. Diseases of the Colon and Rectum. 2017;60(1):51-60.
394. Malinowski KP, Kawalec P. Health utility of patients with Crohn's disease and ulcerative colitis: a systematic review and meta-analysis. Expert Review of Pharmacoeconomics and Outcomes Research. 2016;16(4):441-53.
395. Yarlas A, Bayliss M, Cappelleri JC, Maher SM, Bushmakin AG, Chen LA, et al. A systematic review of the SF-36 health survey for measuring health-related quality of life in patients with ulcerative colitis. Value in Health. 2016;19 (7):A514.
396. Cohan JNR, J. Y. Finlayson, E. Varma, M. G. Impact of surgery on relationship quality in patients with ulcerative colitis and their partners. Diseases of the Colon and Rectum. 2015;58(12):1144-50.
397. Brandsborg ST, A. Nicholls, J. Laurberg, S. Restorative proctocolectomy in patients with ulcerative colitis: A cross-sectional Danish population study on function and quality of life. Colorectal Disease. 2013;15(8):e453-e61.
398. Waljee AK, Higgins PDR, Waljee JF, Tujios SR, Saxena A, Brown LK, et al. Perceived and actual quality of life with ulcerative colitis: A comparison of medically and surgically treated patients. American Journal of Gastroenterology. 2011;106(4):794-9.
399. Portela F, Magro F, Ministro P, De Deus J, Cravo M, Cotter J, et al. Does medical or surgical option influence quality of life in UC patients with a previous severe flare? Journal of Crohn's and Colitis. 2011;5 (1):S135.
400. Tulchinsky H, Dotan I, Halpern Z, Klausner JM, Rabau M. A longitudinal study of quality of life and functional outcome of ulcerative colitis patients after proctocolectomy with ileal pouch anal anastomosis. Colorectal Disease. 2009;11:22.
401. Brown L, Waljee A, Higgins P, Waljee J, Morris A. Proximity to disease and preferences for treatment: Physician and patient perceptions of moderate ulcerative colitis and postcolectomy states. Diseases of the Colon and Rectum. 2010;53 (4):604-5.
402. Waljee AK, Waljee JF, Morris AM, Tujios SR, Higgins PD. Distorted perceptions of the utility of ulcerative colitis: Is colectomy a mind-altering experience? Gastroenterology. 2009;1):A26.
403. Dowson HMC, A. S. Ballard, K. Gage, H. Rockall, T. A. Systematic review of quality of life following laparoscopic and open colorectal surgery. Colorectal Disease. 2008;10(8):757-68.
404. Ansari RA, F. Razjouyan, H. Etemadi, A. Amjadi, H. Merat, S. Malekzadeh, R. Ulcerative colitis and irritable bowel syndrome: Relationships with quality of life. European Journal of Gastroenterology and Hepatology. 2008;20(1):46-50.
405. Berndtsson I, Lindholm E, Oresland T, Borjesson L. Long-term outcome after ileal pouch-anal anastomosis: Function and health-related quality of life. Diseases of the Colon and Rectum. 2007;50(10):1545-52.
406. Mokrowiecka A, Jurek K, Pinkowski D, Malecka-Panas E. The comparison of Health-Related Quality of Life (HRQL) in patients with GERD, peptic ulcer disease and ulcerative colitis. Advances in medical sciences. 2006;51:142-7.
407. Bernklev T, Jahnsen J, Lygren I, Henriksen M, Vatn M, Moum B. Health-related quality of life in patients with inflammatory bowel disease measured with the short form-36: Psychometric assessments and a comparison with general population norms. Inflammatory bowel diseases. 2005;11(10):909-18.
408. Berndtsson IEKL, E. Oresland, T. Hulten, L. Health-related quality of life and pouch function in continent ileostomy patients: A 30-year perspective. Diseases of the Colon and Rectum. 2004;47(12):2131-7.

409. Hjortswang H, Jarnerot G, Curman B, Sandberg-Gertzen H, Tysk C, Blomberg B, et al. The influence of demographic and disease-related factors on health-related quality of life in patients with ulcerative colitis. European Journal of Gastroenterology and Hepatology. 2003;15(9):1011-20.
410. Carmon E, Keidar A, Ravid A, Godlman G, Rabau M. The correlation between quality of life and functional outcome in ulcerative colitis patients after proctocolectomy ileal pouch anal anastomosis. Colorectal Disease. 2003;5(3):228-32.
411. Barton JG, Paden MA, Lane M, Postier RG. Comparison of postoperative outcomes in ulcerative colitis and familial polyposis patients after ileoanal pouch operations. American Journal of Surgery. 2001;182(6):616-20.
412. Muir AJ, Edwards LJ, Sanders LL, Bollinger RR, Koruda MJ, Bachwich DR, et al. A prospective evaluation of health-related quality of life after ileal pouch anal anastomosis for ulcerative colitis. American Journal of Gastroenterology. 2001;96(5):1480-5.
413. Williams JG, Cheung WY, Russell IT, Cohen DR, Longo M, Lervy B. Open access follow up for inflammatory bowel disease: Pragmatic randomised trial and cost effectiveness study. British Medical Journal. 2000;320(7234):544-8.
414. Tiainen J, Matikainen M. Health-related quality of life after ileal J-pouch-anal anastomosis for ulcerative colitis: Long-term results. Scandinavian journal of gastroenterology. 1999;34(6):601-5.
415. McLeod RS, Churchill DN, Lock AM, Vanderburgh S, Cohen Z. Quality of life of patients with ulcerative colitis preoperatively and postoperatively. Gastroenterology. 1991;101(5):1307-13.
416. Meijs S, Gardenbroek TJ, Sprangers MA, Bemelman WA, Buskens CJ, D'Haens GR, et al. Health-related quality of life and disability in patients with ulcerative colitis and proctocolectomy with ileoanal pouch versus treatment with anti-TNF agents. Journal of Crohn's & colitis. 2014;8(7):686-92.
417. Brown LK, Waljee AK, Higgins PD, Waljee JF, Morris AM. Proximity to disease and perception of utility: physicians' vs patients' assessment of treatment options for ulcerative colitis. Diseases of the Colon & Rectum. 2011;54(12):1529-36.
418. Romberg-Camps MJ, Bol Y, Dagnelie PC, Hesselink-van de Kruijs MA, Kester AD, Engels LG, et al. Fatigue and health-related quality of life in inflammatory bowel disease: results from a population-based study in the Netherlands: the IBD-South Limburg cohort. Inflammatory bowel diseases. 2010;16(12):2137-47.
419. Camilleri-Brennan J, Munro A, Steele RJ. Does an ileoanal pouch offer a better quality of life than a permanent ileostomy for patients with ulcerative colitis? J Gastrointest Surg. 2003;7(6):814-9.
420. Camilleri-Brennan J, Steele RJ. Objective assessment of quality of life following panproctocolectomy and ileostomy for ulcerative colitis. Ann R Coll Surg Engl. 2001;83(5):321-4.
421. Seifarth C, Borner L, Siegmund B, Buhr HJ, Ritz J-P, Grone J. Impact of staged surgery on quality of life in refractory ulcerative colitis. Surg Endosc. 2017;31(2):643-9.
422. Ishii H, Kawai K, Hata K, Shuno Y, Nishikawa T, Tanaka T, et al. Comparison of Functional Outcomes of Patients Who Underwent Hand-Sewn or Stapled Ileal Pouch-Anal Anastomosis for Ulcerative Colitis. Int Surg. 2015;100(7-8):1169-76.
423. Reinisch W, Sandborn WJ, Bala M, Yan S, Feagan BG, Rutgeerts P, et al. Response and remission are associated with improved quality of life, employment and disability status, hours worked, and productivity of patients with ulcerative colitis. Inflammatory bowel diseases [Internet]. 2007; 13(9):[1135-40 pp.]. Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/982/CN-00617982/frame.html.
424. Tulchinsky H, Dotan I, Halpern Z, Klausner JM, Rabau M. A longitudinal study of quality of life and functional outcome of patients with ulcerative colitis after proctocolectomy with ileal pouch-anal anastomosis. Diseases of the Colon and Rectum. 2010;53(6):866-73.

425. Pace DE, Seshadri PA, Chiasson PM, Poulin EC, Schlachta CM, Mamazza J. Early experience with laparoscopic ileal pouch-anal anastomosis for ulcerative colitis. Surgical Laparoscopy, Endoscopy and Percutaneous Techniques. 2002;12(5):337-41.
426. Black CM, Yu E, McCann E, Kachroo S. Dose escalation and healthcare resource use among ulcerative colitis patients treated with adalimumab in english hospitals: An analysis of real-world data. PLoS ONE. 2016;11 (2) (no pagination)(e0149692).
427. Black CM, McCann E, Yu E, Nixon M, Kachroo S. Dose escalation and healthcare resource use among ulcerative colitis patients treated with adalimumab in england. United European Gastroenterology Journal. 2015;1):A612.
428. Bodger K, Yen L, Szende A, McDermott J, Hodgkins P. Medical resource utilization and associated costs in patients with ulcerative colitis in the UK: A chart review analysis. Journal of Crohn's and Colitis. 2013;7:S183-S4.
429. Bodger K, Yen L, Szende A, Sharma G, Chen YJ, McDermott J, et al. Medical resource utilization and associated costs in patients with ulcerative colitis in the UK: A chart review analysis. European Journal of Gastroenterology and Hepatology. 2014;26(2):213-21.
430. Bodger K, Yen L, Szende A, Sharma G, McDermott J, Chen JY, et al. Medical resource utilisation and costs in patients with ulcerative colitis in the UK: A chart review analysis. Gut. 2012;61:A223-A4.
431. Ghosh N, Leonard S, Miles G, Premchand P. A model designed to calculate the cost of care for inflammatory bowel disease (IBD). Gut. 2014a;63:A234-A5.
432. Ghosh N, Leonard S, Miles G, Premchand P. A cost of care model for inflammatory bowel disease. Journal of Crohn's and Colitis. 2014b;8:S309-S10.
433. Ghosh N, Premchand P. A UK cost of care model for inflammatory bowel disease. Frontline Gastroenterology. 2015;6(3):169-74.
434. Ghosh S, Peyrin-Biroulet L, Casellas F, O'Shea C, Lee W.J, Pappalardo B, et al. Op018 - Healthcare resource utilization and quality of life in patients with ulcerative colitis by disease severity: baseline data from iconic. United European Gastroenterology Week 2017.
435. Soubieres A, Pollok R, Poullis A. Financial cost of emergency ulcerative colitis (UC) care in the UK in an increasingly cost-aware national health service. Gastroenterology. 2016;1):S796.
436. Soubieres AA, Poullis A. The true cost of Ulcerative Colitis (UC) in the UK - 24 months following the first nonelective admission. Gut. 2016;65:A139-A40.
437. Soubieres AA, Poullis A. The true cost of Ulcerative Colitis (UC) in the UK - 24 months prior to the first non-elective admission. Gut. 2016;65:A75-A6.
438. Travis S, Feagan B, Peyrin-Biroulet L, Panaccione R, Danese S, Lazar A, et al. The costs of care for patients with ulcerative colitis: Effect of adalimumab on health care resources utilisation in clinical practice from INSPIRADA. Journal of Crohn's and Colitis. 2016;10:S298-S9.
439. Watts G, Arasaradnam R, Travis S, Sharma A, Awasthi A, Limdi J, et al. Real-world use of adalimumab in UK patients with ulcerative colitis: An observational retrospective study. Journal of Crohn's and Colitis. 2016;10:S412.
440. Watts G, Arasaradnam R, Travis S, Sharma A, Awasthi A, Limdi J, et al. P512 - Real world use of adalimumab: effects on corticosteroid treatment and hospital resource utilisation in UK patients with moderate/severe ulcerative colitis. European Crohn's and Colitis Organisation 2017.
441. Van Assche G, Peyrin-Biroulet L, Fan T, Ding Q, Lara N, Lynam M, et al. Health care resource utilization among moderate to severe ulcerative colitis patients treated with conventional therapies in europe: the uc cares (ulcerative colitis condition, attitude, resources and educational study). United European Gastroenterology Journal. 2014;1):A527.
442. British National Formulary. Treatment summary: ulcerative colitis. [Available from: https://bnf.nice.org.uk/treatment-summary/ulcerative-colitis.html. Accessed: December 2017.]

443. Hatswell AJ, Porter J, Lee D, Hertel N, Latimer NR. The Cost of Costing Treatments Incorrectly: Errors in the Application of Drug Prices in Economic Evaluation Due to Failing to Account for the Distribution of Patient Weight. Value Health. 2016;19(8):1055-8.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Tofacitinib for moderately to severely active ulcerative colitis [ID1218]

Dear Jo,

The Evidence Review Group, Southampton Health Technology Assessments Centre (SHTAC), and the technical team at NICE have looked at the submission received on 15 May 2018 from Pfizer. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **Friday 22 June 2018.** Your response and any supporting documents should be uploaded to NICE Docs/Appraisals.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as [redacted] in turquoise, and all information submitted as [redacted] in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Aminata Thiam, Technical Lead (Aminata.thiam@nice.org.uk). Any procedural questions should be addressed to Thomas Feist, Project Manager (Thomas.feist@nice.org.uk).

Yours sincerely

Joanna Richardson
Technical Adviser – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for confidential information

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Section A: Clarification on effectiveness data

Decision problem

A1. **Priority question.** The decision problem (company submission [CS] Table 1 p.18) states that the population is "people with moderately to severely active ulcerative colitis" which is broader than the population specified in the final NICE scope ("people with moderately to severely active ulcerative colitis who are intolerant of, or whose disease has had an inadequate response or loss of response to conventional therapy (oral corticosteroids and/or immunosuppressants) or a TNF-alpha inhibitor"). We assume that this is a semantic inaccuracy, since this description ("people with moderately to severely active ulcerative colitis") has been applied to the final NICE scope column within CS Table 1. Please confirm that the decision problem fully matches the final NICE scope. If not, please provide a rationale.

A2. **Priority question.** The trial publication and protocol for the Phase II trial of tofacitinib (NCT00787202) does not mention that patients had to be intolerant of, or had an inadequate response to conventional therapy or a TNF-alpha inhibitor. Therefore the population does not appear to match the final NICE scope. Please clarify the population eligibility criteria for this trial.

A3. The proposed position of tofacitinib within the treatment pathway as shown in CS Figure 1 p.28 does not appear to match the population and comparators as specified in the final NICE scope (that is, 'conventional therapy without biological treatments' is listed as a comparator in the final NICE scope whereas only 'conventional therapy in combination with a biologic' is shown as a comparator for tofacitinib in step 2 of Figure 1). Please explain this discrepancy.

Literature searches

A4. **Priority question.** Please provide the clinical study report (CSR) for the tofacitinib Phase II trial.

A5. Please provide a list of the 137 references excluded from the systematic literature review (SLR).

A6. The reference numbers cited in Appendix D.1.2.1 Table 84 p.21 for the included publications equate to 96 and not 102. Please explain the discrepancy and provide the additional references if any are missing.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Clinical effectiveness trials

A7. **Priority question.** Please provide (in a similar format to CS Table 15 p.54) the baseline characteristics of the TNF inhibitor-naïve and TNF inhibitor-experienced subgroups in the two OCTAVE Induction trials, the OCTAVE Sustain trial and the Phase II trial. Where available, please also provide the baseline characteristics for these subgroups for the comparator trials included in each NMA.

A8. Please specify how many UK sites were participating in the Phase II trial of tofacitinib (NCT00787202).

A9. Please specify how many of the patients in each of the Phase II (NCT00787202) and Phase III (NCT01465763, NCT01465763, and NCT01458574**)** trials of tofacitinib were from the UK.

A10. Table 15 p.54 shows how the patients receiving 15 mg of tofacitinib from the OCTAVE Induction trials and who discontinued were assigned to OCTAVE Sustain.
   a. When was the 15 mg treatment of tofacitinib discontinued?
   b. What happened to the patients randomised to this treatment arm until they were assigned to OCTAVE Sustain – did they continue with tofacitinib and if so, at what dose?

Effectiveness outcomes

A11. Please provide the following details about endoscopy reading:
(a) How many central reading centres were there in each of the Phase II and Phase III trials of tofacitinib, and how many readers were there within each centre?
(b) Were central endoscopy readers blinded? If so, to which patient characteristics?
(c) Given that Phase II and Phase III trials of tofacitinib involved multiple countries and study centres, how was standardisation of endoscopy reading ensured, for both local and central reading?

A12. Please provide supporting evidence for the minimal clinically important differences (MCID) on the IBDQ scale and the WPAI-UC scale, including a justification for using 16 points on the IBDQ scale as a threshold for response and 170 points on the IBDQ scale as a threshold for remission.

Health-related quality of life outcome

A13. **Priority question**: The CS states that analysis of change from baseline EQ-5D in the OCTAVE Induction 1 and 2 and OCTAVE Sustain trials were conducted using a linear mixed-effects model with no imputation for missing data (CS section B.2.4.2

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 16 p.57). The formulas are specified in the CSRs, but no detail is given about the process used to select the co-variates or the fit of alternative specifications.
(a) Please describe and justify the choice of model structure for these analyses.
(b) Is the final set of covariates clinically plausible?

Adverse events

A14. The CS argues that, being a small molecule, tofacitinib will not have immunogenicity but this appears to be based on speculative reasoning and preclinical studies (e.g. as described in the Boland et al. paper, ref 92) rather than on evidence from long-term safety data. The EMA guideline on "Development of new medicinal products for the treatment of ulcerative colitis" states that for new biological therapy trials, one should investigate whether binding-antibodies and/or neutralising antibodies are developed and the impact of this on the long-term efficacy and safety of the product should be investigated (section 8.2 p.12). The CS does not provide any evidence or discussion relating to anti-drug antibodies (ADAs) for tofacitinib, either from the company's research programme on ulcerative colitis or that on rheumatoid arthritis. Please clarify whether ADAs have been measured in any of the studies on tofacitinib and if so please provide the results.

A15. Data on adverse events in the OCTAVE Induction 1 trial CSR appear to show that two patients in the tofacitinib arm had severe neutropenia (CSR Table 14.3.4.1.15) but this is not obviously reflected in CS Appendix Table 156 p.287. The data for neutropenia in the CSR for the OCTAVE Induction 2 trial (CSR Table 14.3.4.1.15) also do not seem to match those in the CS (CS Appendix Table 157 p.288). Please explain whether this is due to differences in the way adverse events have been classified in the CSRs and CS. Additionally, the SmPC reports that tofacitinib may be associated with anaemia, neutropenia and lymphopenia, but of these, only anaemia is listed as an adverse event in CS Tables 156 to 158. Please clarify how many patients in each trial had neutropenia and lymphopenia.

Network meta-analysis (NMA)

A16. **Priority question.** Please explain why the Phase II trial was considered appropriate to be included in the NMA but was not described in the same level of detail as the OCTAVE Induction 1 and 2 trials in the CS?

A17. **Priority question.** It is unclear how the baseline (placebo) response and remission are calculated. Please provide the values for *meanA* and *precA* used in the Multinomial probit model (CS Appendix section D.1.3.4.1 p.244). Please also include the data and priors in WinBUGs format for the Multinomial probit baseline model. The replication of this leads to trap error.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

A18. **Priority question.** Please provide your rationale for conducting the efficacy analyses on the multinomial probit rather than logit scale? The logit has the advantage that the coefficients are more interpretable.

A19. **Priority question.** We note the presence of closed loops in some of the evidence networks. Please provide the results of any inconsistency analysis as part as of the standard critique of the NMA as per NICE TSD 4 ("Inconsistency in network of evidence based on randomised controlled trials").

A20. Were there any attempts to adjust efficacy for differential lengths of follow-up in the induction and maintenance periods?

A21. Please provide the REsd for the following NMA models to understand the choice of model selection: (i) Maintenance for TNF-exposed; (ii) serious infections; and (iii) ITT Maintenance.

A22. Please provide further details of the single integrated induction phase meta-regression model. Only the coefficient of the interaction is provided in section B, no further details are forthcoming in CS Appendix D. Was a similar analysis conducted for the safety endpoints?

Section B: Clarification on cost-effectiveness data

B1. **Priority question**: Please explain and justify the simple approach to estimating utilities from OCTAVE EQ-5D data for use in scenario analysis (CS Table 49 p.140 and CS Appendix M Tables 238 and 239 p.427).

(a) The estimate for the Active UC state is based on a simple mean of baseline utility scores from OCTAVE Induction 1 and 2. This neglects data for patients with Active UC at week 8, trends between 0, 2 and 8 weeks for patients with no response in both arms and difference in baseline utility by trial and patient subgroup (Figure 54 and 56 p. 423 CS Appendix M). It also neglects data from patients with active disease at the end of OCTAVE Sustain.

(b) Utility estimates for the response-no-remission and remission health states were obtained using an area under the curve (AUC) analysis of EQ-5D data from OCTAVE Sustain based on week 52 response/remission status. This presumes that the average utility over the year is representative for the health state at the end of the year. Patients may have changed health status during the year.

B2. **Priority question**: A more appropriate method for analysis of the longitudinal EQ-5D data would be to use a mixed model to estimate utility scores at the end of the induction and maintenance periods (when clinical response and clinical remission were measured) with adjustment for previous utility scores, treatment group, patient

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

subgroup (biologic-naive or prior exposure) and possibly other baseline covariates. Please explain why this type of analysis was not used?

B3. Table 38 [redacted].

B4. It is stated in CS (section B.3.4.1 p.139) and CS Appendices (M.4 p. 422) that an analysis of patient level EQ-5D data was conducted to estimate change in EQ-5D utility over time, based on the destination health state: "For instance, for remitters at week 8 in OCTAVE Induction 1 and 2 trials, the analysis looked back at week 4 and at baseline".

(a) Is the reference to 'week 4' in this sentence an error (elsewhere it is stated that EQ-5D was measured at 0, 2 and 8 weeks in the induction trials)?

(b) Was the analysis conducted with the OCTAVE Sustain trial as well as the OCTAVE Induction 1 and 2 trials?

(c) Please explain the statistical methods used for the analysis and present the results. Did the analysis include any adjustment for co-variates? If so, how were they chosen? What do you mean by the 'non-responder imputation method'? How did you test for homogeneity in mean EQ-5D results at the end of the trial?

(d) Adjusting for baseline utility as a covariate is more efficient than using change from baseline scores (Manca et al. Health Economics 14(5): 487-496). Was this approach considered? If so, please present the results.

B5. Table 49 (CS p.140) reports utility estimates by health state from the OCTAVE trials used for scenario analysis in the economic model. Please report a measure of variance (standard error or confidence interval) for the active US, response no remission, and remission estimates.

B6. **Priority question**: The ERG is unable to replicate the following scenarios. Please provide further clarification on how you conducted these analyses.

a. ITT population analysis (described in CS section B.3.7.2 p.156)
b. Tofacitinib maintenance dose (described in CS section B.3.7.2 p.157)
c. Central read NMA results (CS section B.3.8.4, scenario 7 p.165)
d. Adalimumab maintenance dose (CS section B.3.8.4, scenario 9 p.165)
e. Golimumab maintenance dose (CS section B.3.8.4, scenario 10 p.165)
f. Vedolizumab maintenance dose (CS section B.3.8.4, scenario 11 p.165)

Single technology appraisal

Tofacitinib for moderately to severely active ulcerative colitis [ID1218]

Dear Jo,

The Evidence Review Group, Southampton Health Technology Assessments Centre (SHTAC), and the technical team at NICE have looked at the submission received on 15 May 2018 from Pfizer. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **Friday 22 June 2018.** Your response and any supporting documents should be uploaded to NICE Docs/Appraisals.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Aminata Thiam, Technical Lead (Aminata.thiam@nice.org.uk). Any procedural questions should be addressed to Thomas Feist, Project Manager (Thomas.feist@nice.org.uk).

Yours sincerely

Joanna Richardson
Technical Adviser – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for confidential information

Friday 22nd June 2018

Company response to ERG clarification questions (received 8th June 2018)

Dear Jo,

Thank you for the clarification questions and the opportunity to provide further detail to aid the evaluation of our evidence submission.

Within the scope of responding to the ERG questions, Pfizer also conducted two additional analyses, which are briefly summarised following this paragraph, with detailed analyses within the body of the responses. Please note, although Pfizer provided a robust company submission, applying sound methods and assumptions, the intention for providing the additional analyses is to address any potential outstanding uncertainty on the tofacitinib evidence. As anticipated, the overall conclusions from the CS in terms of comparative efficacy and cost-effectiveness remain unchanged.

1. Exclusion of the tofacitinib Phase II study from the efficacy base-case in the network meta-analysis, in reference to questions A2 and A16.
   - Compared with the base-case network, results remain unchanged, irrespective of inclusion or exclusion from the network of the Phase II trial data for tofacitinib.

2. The provision of an alternative meta-regression model for the OCTAVE Induction and Maintenance EQ-5D datasets, in reference to questions A13, B1, B2 and B4.
   - Compared with the OCTAVE EQ-5D data presented in the company submission (CS) and used in the economic scenarios as presented in the CS (Table 72 and 76), ICERs using EQ-5D derived from the alternative meta-regression model only marginally increased the ICER with all deterministic ICERs remaining around or below the £20,000 ICER threshold level.

Pfizer is confident that all questions have been adequately addressed and meet the ERGs expectations; nevertheless, please do not hesitate to contact me if you require any further information.

Please find below Pfizer's responses to the ERG's questions.

In addition, we have also provided various reference documents (see full list below), alongside the signed Appendix D (checklist for confidential information).

Sincerely,

Angela Helen Blake
Head of Health & Value UK

List of additional reference documents:

- *Clinical Study Report (CSR) for the tofacitinib Phase II study*
- *Statistical Analysis Plan (SAP) for the OCTAVE induction 1 and 2 and Sustain studies*
- *OCTAVE UC Clinical Research Organisation (CRO) Project Charter*
- *Post-hoc EQ-5D analysis tables (ID: SCSA3920265)*
- *Literature references pack, including 11 supportive references*

ERG Clarifications including Pfizer Responses

Section A: Clarification on effectiveness data

Decision problem

A1. **Priority question.** The decision problem (company submission [CS] Table 1 p.18) states that the population is "people with moderately to severely active ulcerative colitis" which is broader than the population specified in the final NICE scope ("people with moderately to severely active ulcerative colitis who are intolerant of, or whose disease has had an inadequate response or loss of response to conventional therapy (oral corticosteroids and/or immunosuppressants) or a TNF-alpha inhibitor"). We assume that this is a semantic inaccuracy, since this description ("people with moderately to severely active ulcerative colitis") has been applied to the final NICE scope column within CS Table 1. Please confirm that the decision problem fully matches the final NICE scope. If not, please provide a rationale.

Pfizer response:
Pfizer can confirm that this was a semantic inaccuracy in Table 1, page 18 of the CS, transcribed from the last version of the NICE draft scope. Pfizer acknowledges that this population description includes by definition patients treatment-naïve to conventional therapy, and as such would not be fully reflective of Xeljanz' licence and the final NICE scope, nor match the clinical and economic evidence presented in the company submission.

A2. **Priority question.** The trial publication and protocol for the Phase II trial of tofacitinib (NCT00787202) does not mention that patients had to be intolerant of, or had an inadequate response to conventional therapy or a TNF-alpha inhibitor. Therefore the population does not appear to match the final NICE scope. Please clarify the population eligibility criteria for this trial.

Pfizer response:
For inclusion in the Phase II trial patients had to have moderate to severe ulcerative colitis (defined as Mayo score ≥6). Patients were excluded if they were treatment-naïve (without previous exposure to treatment). Thus patients were only included if they continued to have moderate to severe disease despite previous treatment. The full lists of inclusion and exclusion criteria for the Phase 2 are listed in Appendix A – Tofacitinib Phase 2 Inclusion and Exclusion Criteriaof this document.

The distribution of patients by treatment arm listing the proportions of prior failed pharmacological treatments at baseline is presented in the Table 1.

Table 1 Baseline characteristics tofacitinib Phase II trial - Failed Drug Treatments for Ulcerative Colitis

Treatment	Placebo	Tofacitinib BID			
No. of subjects, n (%)	N=48	0.5mg N=31	3mg N=33	10mg N=33	15mg N=49
Aminosalicylates	14 (29.17)	9 (29.03)	11 (33.33)	8 (24.24)	18 (36.73)
Immunosuppressants	20 (41.67)	13 (41.94)	12 (36.36)	16 (48.48)	18 (36.73)
Steroids	15 (31.25)	12 (38.71)	8 (24.24)	5 (15.15)	13 (26.53)
TNF-i	12 (25)	2 (6.45)	6 (18.18)	6 (18.18)	10 (20.41)

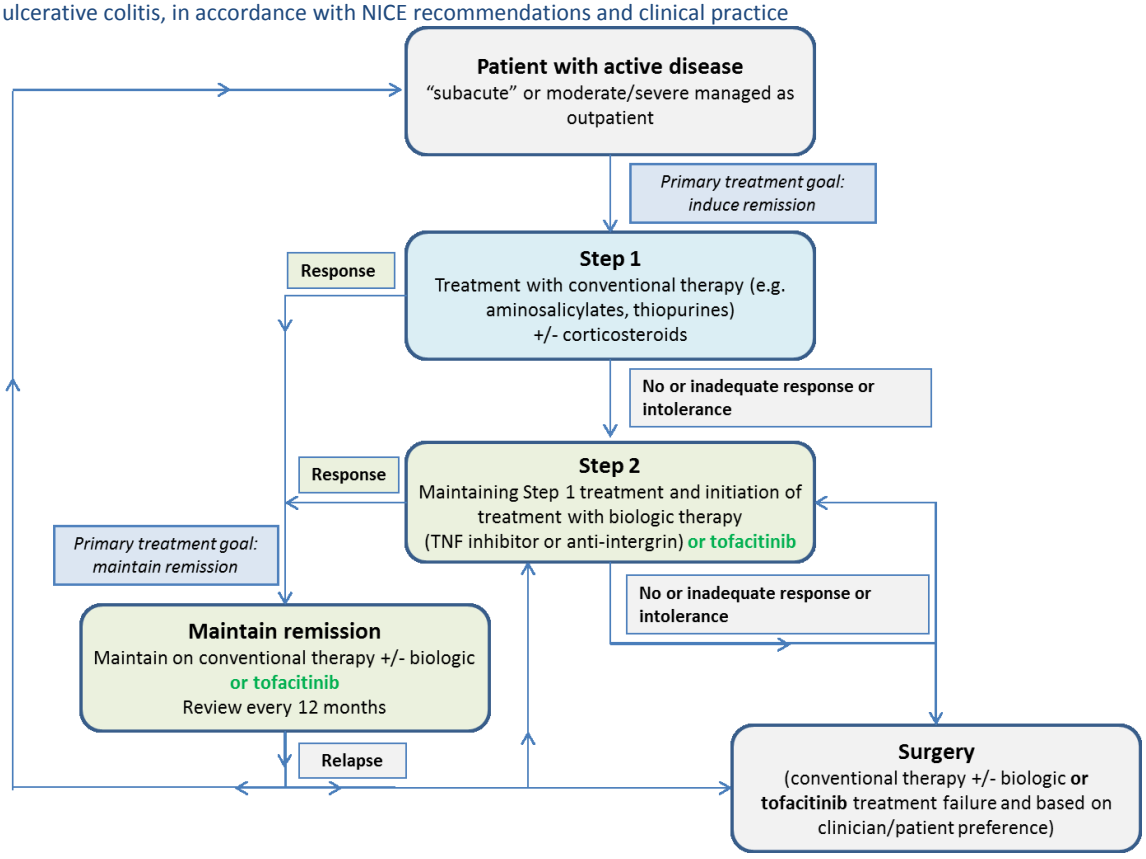
Source P2 CSR Table 13; N=total number of subjects treated with study treatment; BID = twice daily; TNF-i = Tumour necrosis factor inhibitor

A3. The proposed position of tofacitinib within the treatment pathway as shown in CS Figure 1 p.28 does not appear to match the population and comparators as specified in the final NICE scope (that is, 'conventional therapy without biological treatments' is listed as a comparator in the final NICE scope whereas only 'conventional therapy in combination with a biologic' is shown as a comparator for tofacitinib in step 2 of Figure 1). Please explain this discrepancy.

Pfizer response:
CS Figure 1 p.28 is a flowchart based on current NICE guidelines and clinical practice. It shows the clinical pathway for a patient with active ulcerative colitis, which was also validated with clinicians at an UK advisory board.
However, this flowchart is a simplification of the clinical pathway as the treatment of ulcerative colitis is dependent on multiple factors, the including patient's medical history and clinical decision making on the appropriateness of therapies, and therefore may not adequately capture the nuances of clinical practice when comparing to the NICE scope.

In order to better represent the NICE scope Pfizer has further simplified the flowchart, included as Figure 1 within this document. Figure 1 now plainly states that at Step 2 conventional therapy is maintained and advanced treatments, either biologics or tofacitinib, are being added. Furthermore, Pfizer decided to remove the "change of biologic" from Step 2, which accounted for the consideration of clinicians being able to substitute one biologic for another. In addition, the substitution of biologics is deemed a clinical decision and could also apply to the "Maintain remission" step. Lastly, although dose titration would still be applicable to both biologics and tofacitinib, removal of this detail has also aided the simplification of the flowchart.

Figure 1 Proposed position of tofacitinib within the treatment pathway for patients with moderately to severely active ulcerative colitis, in accordance with NICE recommendations and clinical practice



Literature searches

A4. **Priority question.** Please provide the clinical study report (CSR) for the tofacitinib Phase II trial.

Pfizer response:
The CSR for tofacitinib Phase II trial (NCT00787202) is provided within this response.

A5. Please provide a list of the 137 references excluded from the systematic literature review (SLR).

Pfizer response:
The full list of the 137 references that were identified during the systematic literature review and were subsequently excluded, are listed in Appendix B – List of excluded studies and reason for exclusion. In addition, Pfizer have also provided the rationales for exclusion.

A6. The reference numbers cited in Appendix D.1.2.1 Table 84 p.21 for the included publications equate to 96 and not 102. Please explain the discrepancy and provide the additional references if any are missing.

Pfizer response:
Pfizer reviewed Table 84 and counted 102 unique references. In order to present the information more clearly, Appendix C – List of included studieslists each reference separately with its corresponding reference number from the company submission.

Although Pfizer did not identify any missing references, we identified one referencing error. Reference 211 in the CS wrongly cites Feagan 2014, "*Effects of continued vedolizumab therapy for ulcerative colitis in week 6 induction therapy nonresponders*" (Gastroenterology. 2014;1: S-590), which is an excluded study. Instead, it should cite Feagan 2014, "*Health-Related Quality of Life in Patients With Ulcerative Colitis After Treatment With Vedolizumab: Results From the Gemini 1 Study*" (Gastroenterology. 2014; 1: S-590). Both abstracts appeared on the same page of the same of conference proceedings and were mixed as a result.

The incorrect reference now appears in the list of excluded studies (Appendix B – List of excluded studies and reason for exclusion):

Feagan BGS, W. Smyth, M. D. Sankoh, S. Parikh, A. Fox, I. Effects of continued vedolizumab therapy for ulcerative colitis in week 6 induction therapy nonresponders. Gastroenterology. 2014; 1: S-590.

The correct reference for inclusion should be:

Brian G. Feagan, Jean-Frederic Colombel, David T. Rubin, Reema Mody, Serap Sankoh,Karen Lasch. Health-Related Quality of Life in Patients With Ulcerative Colitis After Treatment With Vedolizumab: Results From the Gemini 1 Study. Gastroenterology. 2014; 1: S-590.

Clinical effectiveness trials

A7. **Priority question.** Please provide (in a similar format to CS Table 15 p.54) the baseline characteristics of the TNF inhibitor-naïve and TNF inhibitor-experienced subgroups in the two OCTAVE Induction trials, the OCTAVE Sustain trial and the Phase II trial. Where available, please also provide the baseline characteristics for these subgroups for the comparator trials included in each NMA.

Pfizer response

Table 2 presents the baseline characteristics by TNFi-naïve and TNFi-exposed subgroups for the pooled OCTAVE Induction trials.
Table 3 presents the baseline characteristics by TNFi-naïve and TNFi-exposed subgroups for the OCTAVE Sustain trial.
Table 4 presents the baseline characteristics by TNFi-naïve and TNFi-failure subgroups for the GEMINI 1 RCT (1).
No publication presents the baseline characteristics for TNFi-naïve and TNFi-exposed patients from the ULTRA 2 trial.
All other trials included in the NMA were conducted in TNFi-naïve populations.
Unfortunately, data tables on the baseline demographic characteristics by prior treatment subgroup from the tofacitinib phase 2 study were not readily available.

Table 2 Baseline Demographic and Disease Characteristics of the TNFi-naïve and TNFi-exposed patients in OCTAVE 1 and 2 trials (pooled)

	TNFi-naïve patients		TNFi-exposed patients	
Characteristic	**Placebo (N = 104)**	**Tofacitinib 10 mg (N = 417)**	**Placebo (N = 130)**	**Tofacitinib 10 mg (N = 488)**
Male sex, n (%) [a]	[redacted]	[redacted]	[redacted]	[redacted]
Age, years [b]	[redacted]	[redacted]	[redacted]	[redacted]
Duration of disease — years [b]				
Median	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
Extent of disease, n/total n (%) [c,d]				
Proctosigmoiditis	[redacted]	[redacted]	[redacted]	[redacted]
Left-sided colitis	[redacted]	[redacted]	[redacted]	[redacted]
Extensive colitis or pancolitis	[redacted]	[redacted]	[redacted]	[redacted]
Total Mayo score [b,e]	[redacted]	[redacted]	[redacted]	[redacted]
Partial Mayo score [b,e]	[redacted]	[redacted]	[redacted]	[redacted]
C-reactive protein, mg/litre [b]				
Median	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]
Oral glucocorticoid use at baseline — no. (%) [b]	[redacted]	[redacted]	[redacted]	[redacted]
Previous treatment failure, n (%) [c,f]				
TNF antagonist	[redacted]	[redacted]	[redacted]	[redacted]
Immunosuppressant [g]	[redacted]	[redacted]	[redacted]	[redacted]
White race, n (%) [h]	[redacted]	[redacted]	[redacted]	[redacted]
Weight, kg	[redacted]	[redacted]	[redacted]	[redacted]
Smoking status, n (%) [c,i]				
Never smoked	[redacted]	[redacted]	[redacted]	[redacted]
Current smoker	[redacted]	[redacted]	[redacted]	[redacted]
Former smoker	[redacted]	[redacted]	[redacted]	[redacted]

Plus–minus values are means ±SD. There were no significant differences between groups within each trial unless otherwise noted.

[e] The total Mayo score ranges from 0 to 12 and the partial Mayo score (i.e., the total Mayo score excluding the endoscopic subscore) ranges from 0 to 9, with higher scores indicating more severe disease.

[f] Previous treatment failure was determined by the investigator.

[g] Immunosuppressants included agents such as azathioprine and mercaptopurine and did not include biologic agents (e.g., TNF antagonists) or glucocorticoids.

[h] Unspecified race was treated as missing data.

Abbreviations: SD, standard deviation; TNFi, tumour necrosis factor inhibitor.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 3 Baseline Demographic and Disease Characteristics of the TNFi-naïve and TNFi-exposed patients in OCTAVE Sustain

Characteristic	TNFi-naïve patients			TNFi-exposed patients		
	Placebo (N = 106)	Tofacitinib 5 mg (N =108)	Tofacitinib 10 mg (N =96)	Placebo (N =92)	Tofacitinib 5 mg (N = 90)	Tofacitinib 10 mg (N = 100)
Male sex, n (%) [a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age, years [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Duration of disease — years [b]						
Median	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Extent of disease, n/total n (%) [c,d]						
Proctosigmoiditis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Left-sided colitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Extensive colitis or pancolitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total Mayo score [b,e]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Partial Mayo score [b,e]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
C-reactive protein, mg/litre [b]						
Median	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oral glucocorticoid use at baseline — no. (%) [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous treatment failure, n (%) [c,f]						
TNF antagonist	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Immunosuppressant [g]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
White race, n (%) [h]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weight, kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Smoking status, n (%) [c,i]						
Never smoked	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Current smoker	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Former smoker	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Plus–minus values are means ±SD. There were no significant differences between groups within each trial unless otherwise noted.

[e] The total Mayo score ranges from 0 to 12 and the partial Mayo score (i.e., the total Mayo score excluding the endoscopic subscore) ranges from 0 to 9, with higher scores indicating more severe disease.

[f] Previous treatment failure was determined by the investigator.

[g] Immunosuppressants included agents such as azathioprine and mercaptopurine and did not include biologic agents (e.g., TNF antagonists) or glucocorticoids.

[h] Unspecified race was treated as missing data.

Abbreviations: SD, standard deviation; TNFi, tumour necrosis factor inhibitor.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 4 Baseline Demographic and Disease Characteristics of the TNFi-naïve and TNFi-exposed patients in GEMINI 1

	TNFi-naïve patients			TNFi-failure patients		
Characteristics	**Placebo (n=76)**	**Vedolizumab cohort 1 (n=130)**	**Vedolizumab cohort 2 (n=258)**	**Placebo (n=63)**	**Vedolizumab cohort 1 (n=82)**	**Vedolizumab cohort 2 (n=222)**
Age, y, mean ± SD	40.5 ± 11.7	39.7 ± 13.1	40.6 ± 13.6	41.8 ± 13.1	39.7 ± 12.5	40.2 ± 13.2
Male sex, n (%)	47 (62)	69 (53)	151 (59)	35 (56)	50 (61)	122 (55)
Weight, kg, mean ± SD	70.0 ± 18.8	69.2 ± 16.6	72.7 ± 19.4	74.2 ± 16.4	74.9 ± 17.0	75.3 ± 19.8
BMI, kg/m^2, mean ± SD	24.3 ± 5.7	24.1 ± 4.7	25.1 ± 6.2	25.0 ± 4.5	25.6 ± 5.0	25.5 ± 6.1
Current smoker, n (%)	7 (9)	7 (5)	17 (7)	1 (2)	4 (5)	15 (7)
Disease duration, y, mean ± SD	6.1 ± 6.4	5.8 ± 5.2	6.4 ± 6.2	8.0 ± 7.6	6.4 ± 5.0	8.0 ± 7.0
Mayo Clinic score, mean ± SD	8.5 ± 1.5	8.4 ± 1.8	8.5 ± 1.7	8.6 ± 1.9	8.7 ± 1.8	8.6 ± 1.8
fCal, mg/g, mean ± SD	2714 ± 3408	2357 ± 3595	1493 ± 1980	2196 ± 3256	3008 ± 4270	1306 ± 1604
Disease localization, n (%)						
Proctosigmoiditis	10 (13)	14 (11)	43 (17)	8 (13)	10 (12)	23 (10)
Left-sided colitis	35 (46)	66 (51)	99 (38)	20 (32)	19 (23)	76 (34)
Extensive colitis	7 (9)	14 (11)	33 (13)	9 (14)	10 (12)	24 (11)
Pancolitis	24 (32)	36 (28)	83 (32)	26 (41)	43 (52)	99 (45)
Concomitant medications, n (%)						
CS only	28 (37)	42 (32)	98 (38)	27 (43)	30 (37)	81 (36)
IS only	10 (13)	24 (18)	68 (26)	6 (10)	5 (6)	37 (17)
CS and IS	16 (21)	31 (24)	33 (13)	8 (13)	13 (16)	33 (15)
No CS and IS	22 (29)	33 (25)	59 (23)	22 (35)	34 (41)	71 (32)
Prednisone-equivalent dose, mg, median (min, max)	20.0 (5.0–40.0)	20.0 (2.5–40.0)	20.0 (0.6–80.0)	15.0 (5.0–30.0)	20.0 (5.0–30.0)	20.0 (1.0–176.3)
Type of TNFi failure, n (%)						
Inadequate response	N/A	N/A	N/A	29 (46)	44 (54)	103 (46)
Loss of response	N/A	N/A	N/A	26 (41)	32 (39)	83 (37)
Intolerance	N/A	N/A	N/A	8 (13)	6 (7)	36 (16)

Source: Feagan BGR, D. T. Danese, S. Vermeire, S. Abhyankar, B. Sankoh, S. James, A. Smyth, M. Efficacy of Vedolizumab Induction and Maintenance Therapy in Patients With Ulcerative Colitis, Regardless of Prior Exposure to Tumor Necrosis Factor Antagonists. Clinical Gastroenterology and Hepatology. 2017;15(2):229-39.e5

Abbreviations: BMI, body mass index; CS, corticosteroid; fCal, fecal calprotectin; IS, immunosuppressant; N/A; not applicable; SD, standard deviation; TNFi, tumour necrosis factor inhibitor; y, years

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

A8. Please specify how many UK sites were participating in the Phase II trial of tofacitinib (NCT00787202).

Pfizer response:
A total of 3 UK sites participated in the Phase II trial (NCT00787202); however, only 2 of the sites randomised patients (page 2 of the CSR).

A9. Please specify how many of the patients in each of the Phase II (NCT00787202) and Phase III (NCT01465763, NCT01465763, and NCT01458574**)** trials of tofacitinib were from the UK.

Pfizer response:
Patient enrolment in UK sites for the tofacitinib in ulcerative colitis Phase II and III programme is presented in Table 5:

Table 5 Patient enrolment from UK sites in the tofacitinib in ulcerative colitis trial programme

Study	Patients randomized (n)					
	Tofacitinib 0.3mg	Tofacitinib 5mg	Tofacitinib 10mg	Tofacitinib 15mg	Placebo	**Total**
Phase II (NCT00787202)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OCTAVE Induction 1 (NCT01465763)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OCTAVE Induction 2 (NCT01458951)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OCTAVE Sustain (NCT01458574)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: A3921063 clinical study report Table 13.2.1.2; A3921094 clinical study report Table 14.1.1.3.4; A3921095 clinical study report Table 14.1.1.3.4; A3921096 clinical study report Table 14.1.1.3.3

A10. Table 15 p.54 shows how the patients receiving 15 mg of tofacitinib from the OCTAVE Induction trials and who discontinued were assigned to OCTAVE Sustain.
   a. When was the 15 mg treatment of tofacitinib discontinued?
   b. What happened to the patients randomised to this treatment arm until they were assigned to OCTAVE Sustain – did they continue with tofacitinib and if so, at what dose?

Pfizer response:
a. The 15mg BID arm of OCTAVE Induction 1 (Study 1094, NCT01465763) and OCTAVE Induction 2 (Study 1095, NCT01458951) were removed by a protocol amendment, dated 30 November 2012 (OCTAVE Induction 1 CSR, page 72, Protocol Amendment 3; OCTAVE Induction 2 CSR, page 72, Protocol Amendment 2). Please note, the study initiation dates of the Induction 1 and Induction 2 clinical studies were the 18th of April 2012 and 21st of June 2012 respectively.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

b. Patients who were randomized prior to the protocol amendment and who were still active in the studies at the time of the protocol amendment continued to receive blinded treatment as assigned at baseline for the treatment period. To confirm, patients assigned to 15mg BID continued to receive 15mg BID for the remainder of the Induction trial period. Please note that the tofacitinib 15 mg BID dose group was removed from the original protocol; hence, patients who were assigned to the tofacitinib 15 mg BID dose group were not included in the Induction analysis sets.

From both Induction studies, a total of 19 patients were eligible to enter OCTAVE Sustain at the end of the Induction studies, which corresponds to the 19 patients presented in Table 15 of the CS.

Effectiveness outcomes

A11. Please provide the following details about endoscopy reading:
(a) How many central reading centres were there in each of the Phase II and Phase III trials of tofacitinib, and how many readers were there within each centre?
(b) Were central endoscopy readers blinded? If so, to which patient characteristics?
(c) Given that Phase II and Phase III trials of tofacitinib involved multiple countries and study centres, how was standardisation of endoscopy reading ensured, for both local and central reading?

Pfizer response:

a. Central reading was not advocated in the Phase II study. In this study an appropriately trained endoscopist performed the flexible sigmoidoscopy/colonoscopy. The same endoscopist was to perform the endoscopy for both baseline and Week 8 visits, if possible. When this was not possible, the endoscopist who performed each procedure was clearly documented (page 53 of the Phase II CSR).

   In all the Phase III studies the video of the endoscopic appearance was read by both the study site investigator and a central reader. Central readers for the Phase III studies were identified by the Contract Research Organisation (Robarts Clinical Trials) from their existing pool of qualified central readers, who had expertise in endoscopic assessments. A minimum of four central readers conducted central review activities, based on flexible sigmoidoscopy and assessment of the worst severity within the recorded mucosa.

b. Central readers were blinded to site, patient, and visit identifiers (2, 3). To maintain masking of central readers, all videos were first reviewed by Robarts Image Services personnel (part of the Contract Research Organisation). Procedure IDs were assigned to each video during video processing to keep central readers blinded to both subject and visit details.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

c. The local investigators who performed endoscopy were required to be qualified gastroenterologists able to conduct endoscopy and interpret the results. The investigators underwent training at sign up for trial participation (3). This involved a reminder of the Mayo Endoscopic Score, its components and examples of the range of scores, as well as the essential elements required to ensure consistency in video capture and quality (3).

   As outlined above, central readers for the Phase III studies were identified by the Contract Research Organisation (Robarts Clinical Trials) and had expertise in endoscopic assessments. According to the CRO Project Charter, most central readers possessed over 20 years of experience in endoscopic assessment and the treatment of inflammatory bowel disease. All were practicing physicians and considered key opinion leaders in global IBD trials. None of the central readers were participating site investigators for the OCTAVE programme. Central readers were trained to ensure consistency in the review, assessment and scoring of endoscopy videos. They were provided with copies of study-specific documents for review, including the study protocols, a training summary and Mayo Endoscopic Subscore examples. Central readers were also provided with two sample endoscopic videos for review, including one that illustrated a video that could be assessed for central reading and one that could not. If a subscore submitted by a central reader was queried additional reviews by other central readers were allowed (2, 3).

A12. Please provide supporting evidence for the minimal clinically important differences (MCID) on the IBDQ scale and the WPAI-UC scale, including a justification for using 16 points on the IBDQ scale as a threshold for response and 170 points on the IBDQ scale as a threshold for remission.

Pfizer response:
The IBDQ is a health related quality of life measure used in both Crohn's Disease and Ulcerative Colitis. It is a 32-item questionnaire, evaluating general activities of daily living, specific intestinal function such as bowel habit and abdominal pain, as well as social performance, personal interactions, and emotional status. The total IBDQ score ranges from 32 to 224. For the total score and each domain, a higher score indicates better quality of life.

In 1994, the validity and reliability of the IBDQ was investigated in Crohn's Disease in the large Canadian Crohn's Relapse Prevention Trial (4). In this trial the IBDQ was validated against disease activity measures such as the Crohn's Disease Activity Index (CDAI). The results suggested that an increase in disease activity correlated with decrease in HRQoL as measured by the IBDQ. The study suggested that a change of 16 points on the IBDQ appeared to have a clinically important impact on health related quality of life. In their follow up publication in 1999 Irvine et al clarified that a clinically important mean change in score was observed to be a decrease of between 16 and 30 points, which corresponded to a relapse using the CDAI or a change in therapy by the attending physician (5). The IBDQ was

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

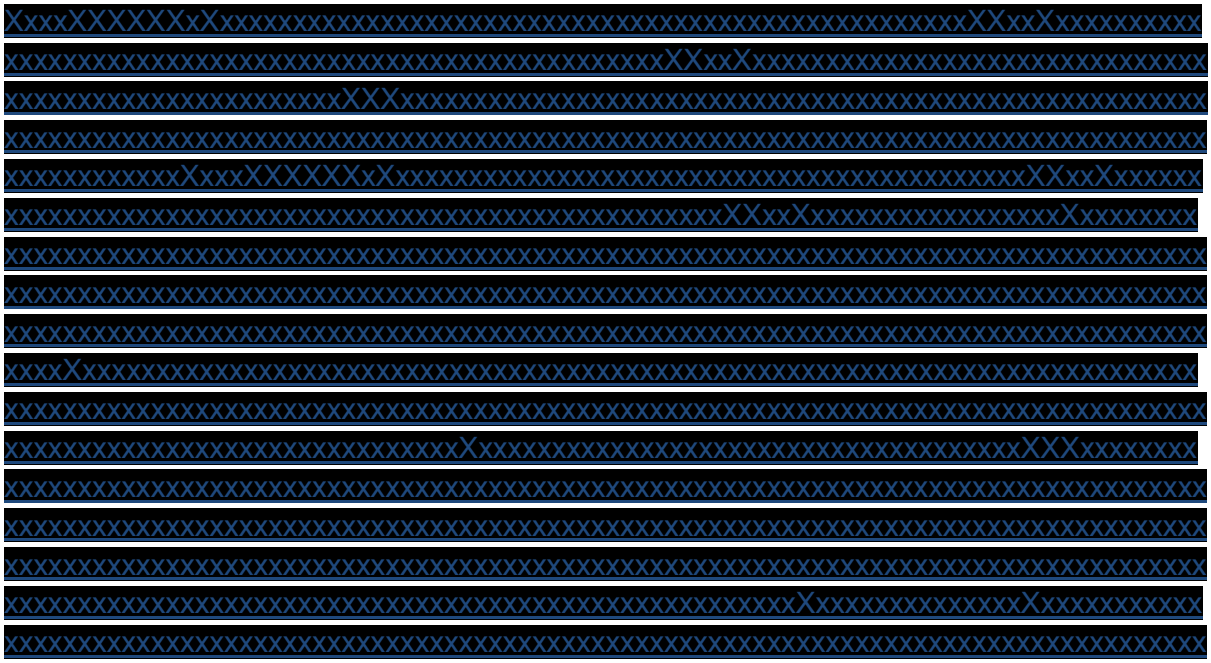
validated in ulcerative colitis by a group of UK clinicians in 1998 (6). Please note, in other ulcerative colitis clinical trials of biologic agents a minimal clinically important difference of ≥16 was also used, such as in the GEMINI I trial for vedolizumab in ulcerative colitis (7).

In the tofacitinib OCTAVE trials a minimal clinically important difference was not specified for WPAI-UC, only a change from baseline was recorded, therefore Pfizer has not specified this in the company submission either. Pfizer is not aware of any validated MCID for the WPAI having been published for ulcerative colitis although there is a suggestion extrapolated from Crohn's Disease that a 7% decrease is a minimal clinically important difference for the WPAI (8; 9).

Health-related quality of life outcome

A13. **Priority question**: The CS states that analysis of change from baseline EQ-5D in the OCTAVE Induction 1 and 2 and OCTAVE Sustain trials were conducted using a linear mixed-effects model with no imputation for missing data (CS section B.2.4.2 Table 16 p.57). The formulas are specified in the CSRs, but no detail is given about the process used to select the co-variates or the fit of alternative specifications.
(a) Please describe and justify the choice of model structure for these analyses.
(b) Is the final set of covariates clinically plausible?

Pfizer response for a) and b):
Pfizer have provided the statistical analysis plans (SAPs) for the ERG's review within this response. All three SAPs are to be treated as commercial in confidence.
As detailed within the SAPs the statistical models were pre-specified in order to analyse the change from baseline EQ-5D.



10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

[redacted]

For details of model fit and the significance levels of covariates in the linear mixed-effects model for OCTAVE Induction 1, OCTAVE Induction 2 and OCTAVE Sustain, please see Table 265a.7.3, Table 265a.7.4 and Table 265a.7.5 in the provided reference document Post-Hoc OCTAVE EQ-5D analysis tables (ID SCSA3920265) [CiC], respectively.

Adverse events

A14. The CS argues that, being a small molecule, tofacitinib will not have immunogenicity but this appears to be based on speculative reasoning and preclinical studies (e.g. as described in the Boland et al. paper, ref 92) rather than on evidence from long-term safety data. The EMA guideline on "Development of new medicinal products for the treatment of ulcerative colitis" states that for new biological therapy trials, one should investigate whether binding-antibodies and/or neutralising antibodies are developed and the impact of this on the long-term efficacy and safety of the product should be investigated (section 8.2 p.12). The CS does not provide any evidence or discussion relating to anti-drug antibodies (ADAs) for tofacitinib, either from the company's research programme on ulcerative colitis or that on rheumatoid arthritis. Please clarify whether ADAs have been measured in any of the studies on tofacitinib and if so please provide the results.

Pfizer response:
Pfizer can confirm that anti-drug antibodies were not measured for tofacitinib in the OCTAVE trials. However, the EMA's 'Guideline on the development of new medicinal products for the treatment of Ulcerative Colitis' quoted in the ERG clarification question does state in Section 8.2 that '*the administration of new biologicals (e.g., cytokines, anti-cytokines, monoclonal antibodies) may trigger the development of antibodies. Therefore, whether binding-antibodies and/or neutralising antibodies against these products are developed and the impact of this on the long-term efficacy and safety of the product should be investigated*.'
It should be noted that this is a specific recommendation for biological therapies and therefore does not apply to tofacitinib.
By definition a biologic is manufactured in a living system such as a microorganism, or plant or animal cells. Most biologics are very large, complex molecules or mixtures of molecules, typically >1kDA, with an extracellularly mode of action.
In contrast, tofacitinib is a chemically synthesized small molecule that works intracellularly, and has a molecular weight of ~ 500Da.

Immunogenicity describes the phenomenon specific to protein-based therapeutics, where recombinant proteins or monoclonal antibodies are recognized as antigens, stimulating a

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

humoral or cell-mediated immune response. Anti-drug antibodies to biological therapies can lead to the formation of multivalent immune complexes with the target drug, leading to rapid drug clearance and/or inactivation of the drug, with a consequential loss of therapeutic response (10). In gastroenterology, this is mitigated either by the addition of thiopurines or by significantly increasing the dose of the biological agent. Therefore, clinicians in gastroenterology measure serum drug levels in their patients receiving biologics who have lost their initial response to therapy in order to optimise their treatment.
As cited in the tofacitinib clinical submission Section B.2.12, pharmacokinetic and pharmacodynamic evidence for tofacitinib from the OCTAVE open-label, long term extension trial, showed that the plasma concentration of tofacitinib reaches steady state within 24 hours of the start of therapy, and remains stable over the course of maintenance treatment (11). Within dosing groups, small variations in plasma concentration at 52 weeks did not correlate with changes in remission status and no loss of efficacy due to low plasma concentration was identified. This suggests that although some patients do lose response to treatment with tofacitinib over time it is not due to drug clearance and low serum levels, i.e. the documented consequences of immunogenicity.

Therefore, given what is known about the phenomenon of immunogenicity as it relates to biologics, such as large proteins, it is reasonable to conclude that tofacitinib will not be subject to immunogenicity. This view has been widely accepted by the clinical and scientific community, and cited in multiple independent expert review articles (12, 13, 14).

A15. **(Part 1)** Data on adverse events in the OCTAVE Induction 1 trial CSR appear to show that two patients in the tofacitinib arm had severe neutropenia (CSR Table 14.3.4.1.15) but this is not obviously reflected in CS Appendix Table 156 p.287. The data for neutropenia in the CSR for the OCTAVE Induction 2 trial (CSR Table 14.3.4.1.15) also do not seem to match those in the CS (CS Appendix Table 157 p.288). Please explain whether this is due to differences in the way adverse events have been classified in the CSRs and CS.
(Part 2) Additionally, the SmPC reports that tofacitinib may be associated with anaemia, neutropenia and lymphopenia, but of these, only anaemia is listed as an adverse event in CS Tables 156 to 158. Please clarify how many patients in each trial had neutropenia and lymphopenia.

Pfizer response to Part 1:
Treatment emergent adverse events (TEAEs) and routine laboratory values were reported differently in the OCTAVE trials. Investigators were required to report AEs as they emerged throughout the studies providing they satisfied certain criteria for reporting (A3921094 and A3921095 protocols).
In the CS Appendix tables 156 and 157, the treatment emergent adverse events reported are for those that had an incidence of ≥2%. According to MedDRA preferred terms, confirmed neutropenia would fall under the 'Blood and Lymphatic System Disorders' System Organ Class, but the incidence of confirmed neutropenia in both Induction 1 and 2 studies

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

did not reach 2% therefore has not been included in these tables. The incidence rates for treatment emergent adverse events in the induction cohort for tofacitinib 10 mg was 2.6% for anaemia, 0.3% for lymphopenia and 0% for neutropenia Table 6.

Laboratory values, however, were reported separately to TEAEs. In the CSRs for Induction 1 and Induction 2, Tables 14.3.4.1.15 presents the number of patients reporting changes in ANC over 8 weeks (without regard to baseline abnormalities) in categories defined as:

- Mild = absolute neutrophil count: 1.5 to <2;
- Moderate = absolute neutrophil count: 1 to <1.5;
- Severe = absolute neutrophil count: 0.5 to < 1;
- Potentially life threatening = absolute neutrophil count: < 0.5.

Based on the above ANC counts the Induction protocols set out the monitoring and discontinuation rules (A3921095 protocol Appendix 3. Guidelines for Monitoring and Discontinuations):

1. *The following laboratory abnormalities require monitoring and re-testing ideally within 3-5 days:*
    - *Absolute neutrophil counts <1.2 x 109/L (<1200/mm3).*
2. *Treatment with CP-690,550 (tofacitinib) will be discontinued and the subject withdrawn from this study for:*
    - *Two sequential absolute neutrophil counts <0.75 x 109/L (<750/mm3).*

For Induction 1, predominantly cases of mild/moderate neutropenia were observed for both tofacitinib 10 mg and placebo. Two cases of severe neutropenia were recorded at week 8 for tofacitinib 10 mg (CSR Table 51). However, no cases led to additional monitoring and/or discontinuation as per protocol requirement (CSR Table 62).

Similarly, for Induction 2, predominantly cases of mild/moderate neutropenia were observed for both tofacitinib 10 mg and placebo (CSR Table 51). However, no cases led to additional monitoring and/or discontinuation as per protocol requirement (CSR Table 62).

Pfizer response to Part 2:
The Summary of Product Characteristics for Xeljanz (tofacitinib) lists anaemia (common), lymphopenia and neutropenia (uncommon) as adverse drug reactions (ADR). It should be noted that these ADRs reflect the extensive clinical trial programme for rheumatoid arthritis, as well as ulcerative colitis.

The reported incidence of lymphopenia, anaemia and neutropenia for the ulcerative colitis Phase 2 and Phase 3 Induction studies is listed in Table 6.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 6 Proportions and Incidence Rates for Lab Related Treatment Emergent Adverse Events

Type	Tofacitinib 10 mg		Placebo		% Diff (95% CI)
	N	n (%)	N	n (%)	
Anaemia	938	24 (2.6%)	282	10 (3.5%)	-1.0 (-3.86, 1.14)
Lymphopenia	938	3 (0.3%)	282	1 (0.4%)	-0.0 (-1.52, 0.71)
Neutropenia	938	0	282	0	0.0 (0.00,0.00)

Source: EMA submission Table 14.2.16.c1

Network meta-analysis (NMA)

A16. **Priority question.** Please explain why the Phase II trial was considered appropriate to be included in the NMA but was not described in the same level of detail as the OCTAVE Induction 1 and 2 trials in the CS?

Pfizer response:
Please also consider Pfizer response to ERG question A2 of this document, in conjunction to this question.

The Phase II trial of tofacitinib in ulcerative colitis was a small dose finding study, involving 194 patients in total, of whom only 33 received the licensed dose, 10 mg twice daily. The submission focussed on describing the larger phase III studies due to space constraints. The phase II trial publication met the inclusion criteria for the NMA, and was therefore included in the analysis, but is described only briefly in section B.2 of the CS.

A summary of the Phase II study is provided in Appendix D – Tofacitinib Phase II study summary

Results in the Phase II study tofacitinib 10 mg group were consistent with the results of the OCTAVE Induction 1 and 2 trials. The primary endpoint of clinical response at 8 weeks was achieved by 61% of patients (20/33) receiving tofacitinib 10 mg, similar to the results in OCTAVE Induction 1 (central reads, 59.9%; local reads, 60.7%) and 2 (central reads, 55.0%; local reads, 58.0%).

As an extension to the ERG question, Pfizer have explored the impact of the Phase II study within the network meta-analysis, which demonstrates that results remain unchanged, whether or not the Phase II trial for tofacitinib was included or excluded. The results are presented in Table 7.
For ease of comparison, Pfizer also included the Induction base-case NMA results (Table 25 from page 92 of the CS) in Appendix D – Tofacitinib Phase II study summaryof this document.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 7 Induction phase scenario NMA results excluding tofacitinib phase 2 study (Sandborn et al. 2012) – comparative effects and probabilities of achieving response and remission

Comparator	Comparator vs PBO			TOF vs comparator				
	Treatment effect, median (95% CrI)	Odds ratio, median (95%CrI)		Odds ratio, median (95%CrI)		Absolute probability		SUCRA [a]
	Probit scale	Clinical response	Clinical remission	Clinical response	Clinical remission	Clinical response	Clinical remission	
TNFi-naïve subgroup (random effects)								
PBO								
TOF 10 mg								
INF 10 mg/kg								
ADA 160/80/40 mg [b]								
GOL 200/100 mg [c]								
VED 300 mg [d]								
TNFi-exposed subgroup (fixed effect)								
PBO								
TOF 10 mg								
ADA 160/80/40 mg [b]								
VED 300 mg [d]								

[a] based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2.
Abbreviations: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; SUCRA, surface under cumulative ranking curve; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

A17. **Priority question.** It is unclear how the baseline (placebo) response and remission are calculated. Please provide the values for *meanA* and *precA* used in the Multinomial probit model (CS Appendix section D.1.3.4.1 p.244). Please also include the data and priors in WinBUGs format for the Multinomial probit baseline model. The replication of this leads to trap error.

Pfizer response:

The baseline placebo response and remission were not calculated manually. The baseline model, as given in TSD5 and shown in Appendix D.1.3.4.1, estimates the baseline probability of non-response (on the probit scale), based on the number of patients in non-response in the placebo arm of each study including placebo. This is then inputted as baseline risk into the probit model and used to anchor all the estimations.

The meanA and precA used in each probit model are shown in the Table 8.

Table 8 Summary of the meanA and precA values per multinomial probit model

Phase	Population	Scenario	meanA	precA
Induction	TNFi-naive	Base case	xxxxx	xx
		SA excluding Asian studies	xxxxx	x
		SA using centrally read endoscopy	xxxxx	xx
	TNFi-exposed	Base case	xxxxx	x
		SA using centrally read endoscopy	xxxxx	x
		SA using TNF failure only	xxxxx	x
	Overall ITT	SA using overall ITT populations	xxxxx	xx
Maintenance	TNFi-naive	Base case	xxxxx	x
		SA excluding Asian studies	xxxxx	x
		SA using centrally read endoscopy	xxxxx	x
	TNFi-exposed	Base case	xxxxx	xxx
		SA using centrally read endoscopy	xxxxx	xxx
		SA using TNF failure only	xxxxx	xxxx
	Overall ITT	SA using overall ITT populations from re-randomised responder trials only	xxxxx	x

Please note; Pfizer did not experience any trap errors for the multinomial probit baseline model. The data and initial values used to run the baseline model for each base case analysis are provided in Winbugs® format in Appendix E – NMA Winbug code

A18. **Priority question.** Please provide your rationale for conducting the efficacy analyses on the multinomial probit rather than logit scale? The logit has the advantage that the coefficients are more interpretable.

Pfizer response:

The classification of patients into responders and remitters is based on the Mayo score. The Mayo score is measured in a continuous manner in the studies, and patients are then classified into the three mutually exclusive categories "no response", "response without remission" and "remission" using pre-defined cut-offs.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

In this situation, a multinomial probit model is more efficient than several independent logit models on each category, as it prevents the occurrence of incompatible results across the different categories (for example here a predicted proportion of remitters that is greater than the predicted proportion of responders). In addition, it allows for the correlations between variables to be fully taken into account in the probabilistic sensitivity analysis (PSA), which makes the economic model more robust.

This choice of model follows the recommendation given in TSD2 for such outcomes (15). The code used for the analysis are the ones given in this TSD in example 6 (psoriasis example on the PASI outcome). This approach is also in line with the analysis undertaken by the Assessment Group for Multiple Technology Appraisal 329 (16). Whilst the CS for Single Technology Appraisal 342 analysed clinical response and clinical remission separately using a binomial likelihood model, the ERG considered the approach to be only partially appropriate (17). The ERG stated that these results "should be interpreted with caution, because these results were estimated without considering the dependence/correlation between response and remission. Ideally, the NMA should take account of the nature of the data i.e. ordered categorical. Use of these results in the economic model ignores this dependence and would potentially generate inappropriate samples for PSA." Pfizer agrees with this assessment.

A19. **Priority question.** We note the presence of closed loops in some of the evidence networks. Please provide the results of any inconsistency analysis as part as of the standard critique of the NMA as per NICE TSD 4 ("Inconsistency in network of evidence based on randomised controlled trials").

Pfizer response:
There is one closed loop in the induction networks, made of placebo, infliximab and adalimumab, due to the presence of the Mshimesh study comparing infliximab to adalimumab directly. The results of the inconsistency assessment for each outcome are shown in Table 9. This analysis is independent from the NMA, so the response-remission outcomes are analysed independently. No evidence of inconsistency was found.

Table 9 Results of the inconsistency assessment

Phase	Population	Outcome measure	Direct effect of infliximab versus adalimumab (Mshimesh 2017)	Indirect effect of infliximab versus adalimumab (Bucher method)	Inconsistency test p-value
Response	TNF naive	OR	[redacted]	[redacted]	[redacted]
		RR	[redacted]	[redacted]	[redacted]
Remission	TNF naive	OR	[redacted]	[redacted]	[redacted]
		RR	[redacted]	[redacted]	[redacted]
Mucosal healing	TNF naive	OR	[redacted]	[redacted]	[redacted]
		RR	[redacted]	[redacted]	[redacted]
Serious infections	Overall	OR	[redacted]	[redacted]	[redacted]
		RR	[redacted]	[redacted]	[redacted]
Serious AE	Overall	OR	[redacted]	[redacted]	[redacted]
		RR	[redacted]	[redacted]	[redacted]
Discontinuations due to AE	Overall	OR	[redacted]	[redacted]	[redacted]
		RR	[redacted]	[redacted]	[redacted]

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

The closed loops in the maintenance networks are due to the presence of 3-arm trials, which are internally consistent by definition. No inconsistency assessment was therefore performed, as this method is only applied to independent source of data.

A20. Were there any attempts to adjust efficacy for differential lengths of follow-up in the induction and maintenance periods?

Pfizer response:
In the induction phase, all studies report results at week 8, except GEMINI1 and PURSUIT-SC which report results at week 6. No attempt was made to account for this difference in the model, because these trials are the only sources of data for vedolizumab and golimumab and as such, it would have been impossible to properly estimate what difference was due to the treatment effect and what difference was the effect of an earlier measure.

A similar phenomenon was observed in the maintenance phase. The only studies reporting data at a different time point (54 weeks instead of 52 weeks) are ACT1, PURSUIT-M and PURSUIT-J, which are the only sources of data for infliximab and golimumab.

The Assessment Group for MTA 329 did not adjust for differences between trial phase duration, nor did the CS for STA 342. The ERG and Appraisal Committee for STA 342 agreed that the differences were not expected to have an impact on the results. Clinical experts consulted during the development of Pfizer's CS also confirmed that differences in the evaluation time-points were negligible and unlikely result in meaningful differences.

A21. Please provide the REsd for the following NMA models to understand the choice of model selection: (i) Maintenance for TNF-exposed; (ii) serious infections; and (iii) ITT Maintenance.

Pfizer response:

i. The model fit statistics for the base case maintenance phase analysis of patients with prior TNFi exposure are presented in Table 24 of the CS. Here it indicates that the random effects model could not be run due to convergence issues, therefore no standard deviation (SD) of the random effects (RE) model can be provided.
ii. For the serious infections RE NMA model the median SD is [redacted].
iii. For the ITT maintenance (re-randomised studies only) RE NMA model the median SD [redacted]).

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

A22. Please provide further details of the single integrated induction phase meta-regression model. Only the coefficient of the interaction is provided in section B, no further details are forthcoming in CS Appendix D. Was a similar analysis conducted for the safety endpoints?

Pfizer response:
Table 10 presents a full set of results from the random effects single integrated induction phase NMA of clinical response and clinical remission. As mentioned in the CS page 95, the covariate in the random effects model was −0.365 (95% CrI: −0.547, −0.169) indicating that there is a statistically significant difference in effect between patients with and without prior TNFi exposure.

As noted in the submission, Pfizer has a number of reservations regarding the integrated analysis and thus these results should be interpreted with caution.

- Underpinning the analysis is the assumption that the placebo effect is the same across subgroups, and that the interaction term adjusted the treatment effects of all comparators by the same fixed amount. The relative effects of each treatment versus other active treatments are also the same in both subgroups and it is unclear whether this is supported by the evidence or clinical practice.
- Results for infliximab and golimumab in the TNFi-exposed subgroup were predicted from the data on other drugs in TNFi-exposed patients and from the data on infliximab and golimumab in TNFi-naïve patients. In the absence of any trial-based observations for infliximab and golimumab in TNFi-exposed patients, there was no way to externally validate the NMA outputs for these drugs.
- The model is relatively weak because gaps in the data (i.e. a lack of data on all treatments in both TNFi-exposure subgroups) make it difficult to estimate it properly. For example, with evidence available for infliximab and golimumab only in TNFi-naïve patients, the model struggles to attribute observed differences to the treatment effect or the subgroup effect.

The model was considered reliable enough to check for the significance of the interaction covariate, but not considered reliable to produce robust results for each treatment in each TNFi-exposure subgroup and was therefore not used in the economic analysis.

No similar analysis was attempted for the safety endpoints because no safety data was presented by TNFi-exposure subgroup in the comparator trials.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Table 10 Results of single integrated Induction phase NMA with meta-regression – comparative effects and probabilities of achieving response and remission

Comparator	Comparator vs PBO			TOF vs comparator				
	Treatment effect, median (95% CrI)	Odds ratio, median (95%CrI)		Odds ratio, median (95%CrI)		Absolute probability		SUCRA [a]
	Probit scale	Clinical response	Clinical remission	Clinical response	Clinical remission	Clinical response	Clinical remission	
TNFi-naïve subgroup (random effects)								
PBO				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed subgroup (random effects)								
PBO				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2.
Abbreviations: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; SUCRA, surface under cumulative ranking curve; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Section B: Clarification on cost-effectiveness data

B1. **Priority question**: Please explain and justify the simple approach to estimating utilities from OCTAVE EQ-5D data for use in scenario analysis (CS Table 49 p.140 and CS Appendix M Tables 238 and 239 p.427).

(a) The estimate for the Active UC state is based on a simple mean of baseline utility scores from OCTAVE Induction 1 and 2. This neglects data for patients with Active UC at week 8, trends between 0, 2 and 8 weeks for patients with no response in both arms and difference in baseline utility by trial and patient subgroup (Figure 54 and 56 p. 423 CS Appendix M). It also neglects data from patients with active disease at the end of OCTAVE Sustain.

(b) Utility estimates for the response-no-remission and remission health states were obtained using an area under the curve (AUC) analysis of EQ-5D data from OCTAVE Sustain based on week 52 response/remission status. This presumes that the average utility over the year is representative for the health state at the end of the year. Patients may have changed health status during the year.

Pfizer response:

(a)

As outlined in Section B.3.10.1 of the CS Pfizer explored the OCTAVE EQ-5D data with experts and is was suggested that based on the clinical trial design the OCTAVE EQ-5D data for non-responders is likely to be biased. The re-randomisation design of the OCTAVE Sustain trial substantially limits the assessment of the active disease EQ-5D utility as all patients entering the OCTAVE Sustain trial were required to be OCTAVE induction treatment responders. Therefore no non-responder data post week 8 induction phase is available for neither tofacitinib nor placebo within a randomised Phase 3 maintenance trial setting. This is likely to negatively impact the EQ-5D utility values at the end of OCTAVE Sustain as a true reflection of the active UC utility state values.

In order to use tofacitinib clinical trial data, the experts felt that the clinically most representative option was to use the EQ-5D baseline values to inform the active UC health state utilities. To further address any uncertainty in interpretation of the OCTAVE EQ-5D utility values and to increase comparability with previous NICE technology appraisal in moderately to severely ulcerative colitis, Pfizer used Woehl et al 2008 utilities as the base-case health state values for the economic analyses in the CS.

(b)

The objective was to populate the model with utility values from the OCTAVE program for use in sensitivity analysis, based on the inherent OCTAVE Sustain trial design issues (as discussed above). The AUC approach was used to fit the summary statistics obtained from the clinical trial within the economic model structure.

In response to the ERG comments, a regression analysis has been conducted using a linear mixed effect model, grouping the patients by efficacy endpoints:

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

- Pooled analysis of OCTAVE 1 and 2 trials
    - Week 8 non-clinical responders
    - Week 8 clinical responders but not clinical remitters
    - Week 8 clinical remitters
- OCTAVE Sustain trial
    - Week 52 non-clinical responders
    - Week 52 clinical responders but not clinical remitters
    - Week 52 clinical remitters

The model was defined as:

- OCTAVE Induction 1 and 2: Treatment + Prior treatment with TNFi therapy + Corticosteroid use at baseline + Geographic region + Week + Treatment*Week + Baseline EQ-5D with subjects as random effect
- OCTAVE Sustain: EQ-5D Score = Treatment + Induction Treatment + Baseline Remission Status + Week + Treatment*Week + Baseline EQ-5D with subjects as random effect.

The results of the regression analysis are presented in the provided reference document Post-Hoc OCTAVE EQ-5D analysis tables (ID SCSA3920265) [CiC], Table 265a.7.1 and Table 265a7.6 for induction and Table 265a.7.2 and Table 265a.7.7 for maintenance.

However, it is acknowledged that the simplistic, but valid, AUC method and the presented regression analysis within this document do not sufficiently address the OCTAVE Sustain re-randomisation design issue, and as demonstrated in ERG question B2 of this document do not alter the overall conclusion to the economic analysis as presented in the CS. As mentioned in response to ERG question B1 (a); to further address any uncertainty in interpretation of the OCTAVE EQ-5D utility values and to increase comparability with previous NICE technology appraisal in moderately to severely ulcerative colitis, Pfizer used Woehl et al 2008 utilities as the base-case health state values for the economic analyses in the CS.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

B2. **Priority question**: A more appropriate method for analysis of the longitudinal EQ-5D data would be to use a mixed model to estimate utility scores at the end of the induction and maintenance periods (when clinical response and clinical remission were measured) with adjustment for previous utility scores, treatment group, patient subgroup (biologic-naive or prior exposure) and possibly other baseline covariates. Please explain why this type of analysis was not used?

Pfizer response:

Please see response to B1 for the methods, and Appendix F (Table 265a.7.1 and Table 265a7.6 for induction and Table 265a.7.2 and Table 265a.7.7 for maintenance) for the results of the regression analysis. Alternative methods for the OCTAVE Sustain EQ-5D analysis were considered but it was concluded that based on the OCTAVE Sustain trial design and the available data, alternative methods were unlikely to lead to different conclusions than presented in the CS.

Table 11 presents the results of the regression analysis based on methods the ERG referred to in B1 (without the treatment stratification (placebo or tofacitinib)). The EQ-5D scores follow a logical direction with the non-clinical response values showing the lowest scores and clinical remission showing the highest scores. Overall the OCTAVE Sustain EQ-5D scores are higher than the OCTAVE Induction 1 and 2 scores, within each efficacy category.

Table 11 EQ-5D HRQoL in pooled OCTAVE 1 and 2 and OCTAVE Sustain trials

Efficacy endpoint[a]	OCTAVE Induction 1 and 2 Values at week 8		OCTAVE Sustain Values at week 52	
	N[b]	Adjusted mean (95% CI)[c]	N[d]	Adjusted mean (95% CI)[e]
Non-clinical response	[redacted]	[redacted]	[redacted]	[redacted]
Clinical response (but not clinical remission)	[redacted]	[redacted]	[redacted]	[redacted]
Clinical remission	[redacted]	[redacted]	[redacted]	[redacted]

[a]Efficacy endpoints are based on NRI and Local Read of Endoscopy.
[b] N = number of subjects with non-missing EQ-5D data at week 8
[c] Adjusted mean derived from the linear mixed effects model: Score = Treatment + Prior treatment with TNFi therapy + Corticosteroid use at baseline + Geographic region + Week + Treatment*Week + Baseline EQ-5D with subjects as random effect
[d] N = number of subjects with non-missing EQ-5D data at week 52
[e] Adjusted mean derived from the linear mixed effects model: Score = Treatment + Induction Treatment + Baseline Remission Status + Week + Treatment*Week + Baseline EQ-5D with subjects as random effect.
Abbreviations: CI, confidence interval

The non-clinical response EQ-5D score from OCTAVE Induction 1 and 2 reflects the HRQoL of patients who did not respond to treatment at 8 weeks. The non-clinical response EQ-5D score from OCTAVE Sustain reflects the HRQoL of patients who responded to treatment at 8 weeks, but have lost response at some point in the following 52 weeks.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Using the results of this regression analysis, Pfizer have conducted two additional scenarios in the economic model. For the induction health state utility value (HSUV) and the maintenance HSUV we used the response and remission scores at 8 weeks and 52 weeks, respectively. Scenarios assumed for the Active UC HSUV:

1. Pfizer used the non-clinical response EQ-5D score from OCTAVE Induction 1 and 2
2. Pfizer used the non-clinical response EQ-5D score from OCTAVE Induction 1 and 2 for induction phase non-responders, and from OCTAVE Sustain for patients who responded initially but lost response (discontinued treatment).

In all but one scenario the ICER of tofacitinib against conventional treatment remained at a level below £20,000 per QALY. In one instance (Scenario 2; biologic-exposed), the ICER was £22,000 per QALY. Based on the OCTAVE trial design, the results should be interpreted with caution.

The results of the first scenario are presented in **Error! Reference source not found.** and **Error! Reference source not found.** for the biologic-naïve and biologic-exposed populations, respectively. The ICER of tofacitinib versus conventional therapy was £16,599 per QALY for the biologic-naïve population and £19,953 per QALY in the biologic-exposed population. When compared to tofacitinib, vedolizumab generated an additional QALY of [redacted] with an ICER of £1.16 million per QALY in the biologic-naïve population and an additional QALY of [redacted] with an ICER of £12.8 million per QALY in the biologic-exposed population.

Table 12 Scenario 1 - Biologic-naïve patients: full incremental cost-effectiveness results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental	ICER (£/QALY) TOF vs comparator
	QALYs	Costs (£)	QALYs	Costs (£)		
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£16,599.21
Adalimumab	[redacted]	[redacted]			Dominated	Dominated
Golimumab	[redacted]	[redacted]			Dominated	Dominated
Infliximab	[redacted]	[redacted]			Dominated	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£16,599.21	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£1,161,372.77	£1,161,372.77

Abbreviations: ICER, incremental cost-effectiveness ratio; N/A, not applicable; QALY, quality-adjusted life year; TOF: tofacitinib.

Table 13 Scenario 1 - Biologic-exposed patients: full incremental cost-effectiveness results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental	ICER (£/QALY) TOF vs comparator
	QALYs	Costs (£)	QALYs	Costs (£)		
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£19,953.22
Adalimumab	[redacted]	[redacted]			Dominated	Dominated
Golimumab	[redacted]	[redacted]			Dominated	Dominated
Infliximab	[redacted]	[redacted]			Dominated	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£19,953.22	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£12,802,191.16	£12,802,191.16

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Abbreviations: ICER, incremental cost-effectiveness ratio; N/A, not applicable; QALY, quality-adjusted life year; TOF: tofacitinib.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

In the second scenario, for biologic naïve patients (Table 14), the ICER for tofacitinib compared with conventional treatment was £18,494 per QALY. Compared with tofacitinib, vedolizumab generated a marginal additional [redacted] QALYs, with an ICER of £1.27 million per QALY.

For biologic-experienced patients (Table 15), the ICER for tofacitinib compared with conventional treatment was £22,440 per QALY. Compared with tofacitinib, vedolizumab generated a marginal additional [redacted] QALYs, with an ICER of £20.8 million per QALY.

Table 14 Scenario 2 - Biologic-naïve patients: full incremental cost-effectiveness results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental	ICER (£/QALY) TOF vs comparator
	QALYs	Costs (£)	QALYs	Costs (£)		
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£18,493.79
Adalimumab	[redacted]	[redacted]			Dominated	Dominated
Golimumab	[redacted]	[redacted]			Dominated	Dominated
Infliximab	[redacted]	[redacted]			Dominated	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£18,493.79	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£1,266,535.84	£1,266,535.84

Abbreviations: ICER, incremental cost-effectiveness ratio; N/A, not applicable; QALY, quality-adjusted life year; TOF: tofacitinib.

Table 15 Scenario 2 - Biologic-exposed patients: full incremental cost-effectiveness results

Strategy	Total		Incremental		ICER (£/QALY) fully incremental	ICER (£/QALY) TOF vs comparator
	QALYs	Costs (£)	QALYs	Costs (£)		
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£22,440.03
Adalimumab	[redacted]	[redacted]			Dominated	Dominated
Golimumab	[redacted]	[redacted]			Dominated	Dominated
Infliximab	[redacted]	[redacted]			Dominated	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£22,440.03	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£20,825,984.97	£20,825,984.97

Abbreviations: ICER, incremental cost-effectiveness ratio; N/A, not applicable; QALY, quality-adjusted life year; TOF: tofacitinib.

B3. Table 38 [redacted].

Pfizer response:

All patients received 10 mg twice daily for 8 weeks in both OCTAVE Induction studies. If patients achieved clinical response in the Induction studies they were enrolled into the OCTAVE Sustain maintenance study. If patients had not achieved clinical response at 8 weeks they had the opportunity to enrol in the open label OCTAVE Open study. Those patients entering OCTAVE Open received tofacitinib 10 mg twice daily for a further 8 weeks

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

and were assessed at months 1, 2, 4, 6, 9 and 12, with a four week follow up after last dose. Those patients in OCTAVE Open who did not achieve clinical response at 16 weeks were withdrawn from the study. Therefore, for all patients continuing beyond 8 weeks in the OCTAVE programme, a clinical assessment of response was made at 16 weeks.

It should be noted that current biological treatments also include assessment recommendations within their respective Summary of Product Characteristics (SmPC) for assessing response to treatment at specific intervals, ranging from 2 weeks to 14 weeks. (18, 19, 20, 21). The impact of the different assessment intervals was explored with clinicians which stated that patients with moderately to-severely disease would usually be seen 4 weeks after starting therapy, and again at 8 weeks and at endoscopic assessment between weeks 12-16. However, local variability was acknowledged. Experts emphasised that these typical assessment times also apply to conventional therapy and as such are not limited to biological therapies. No additional service implications are expected for the inclusion of an extended 8 week induction treatment period for tofacitinib within the licence.

B4. It is stated in CS (section B.3.4.1 p.139) and CS Appendices (M.4 p. 422) that an analysis of patient level EQ-5D data was conducted to estimate change in EQ-5D utility over time, based on the destination health state: “For instance, for remitters at week 8 in OCTAVE Induction 1 and 2 trials, the analysis looked back at week 4 and at baseline”.

(a) Is the reference to ‘week 4’ in this sentence an error (elsewhere it is stated that EQ-5D was measured at 0, 2 and 8 weeks in the induction trials)?

(b) Was the analysis conducted with the OCTAVE Sustain trial as well as the OCTAVE Induction 1 and 2 trials?

(c) Please explain the statistical methods used for the analysis and present the results. Did the analysis include any adjustment for co-variates? If so, how were they chosen? What do you mean by the ‘non-responder imputation method’? How did you test for homogeneity in mean EQ-5D results at the end of the trial?

(d) Adjusting for baseline utility as a covariate is more efficient than using change from baseline scores (Manca et al. Health Economics 14(5): 487-496). Was this approach considered? If so, please present the results.

Pfizer response:

(a) Apologies, Pfizer can confirm this is a typographical error and should read week 2 instead of week 4.

(b) The analysis was conducted for the OCTAVE Sustain trial only.

(c) Details on the methods and results of the regression analysis as conducted on the OCTAVE EQ-5D index utilities is presented in A13 of this document. A subsequent analysis using summary statistics on EQ-5D scores at each visit was used to inform the methods in Appendix M4 of the CS. The analysis did not include adjustment for covariates.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

The non-responder imputation approach assumed that subjects with missing response outcomes at the end of the trial were presumed non-responders. Therefore, their EQ-5D was allocated to the no-response category.
By homogeneity Pfizer referred to similarity of EQ-5D values, concluded by visual inspection of the trends. No formal statistical tests of homogeneity were conducted.

(d) See response to B1 and B2

B5. Table 49 (CS p.140) reports utility estimates by health state from the OCTAVE trials used for scenario analysis in the economic model. Please report a measure of variance (standard error or confidence interval) for the active US, response no remission, and remission estimates.

Pfizer response:

Table 16 presents the mean and SD observed at baseline in OCTAVE 1, 2 and Sustain trials. The original submission estimated precision across the health state utilities using the minimum and maximum values of the averages across the trial arms from OCTAVE Sustain. Additional estimates of mean values and variance for each health state are now provided in Table 16, which are based on the meta-regression performed in response to question B2.

Table 16 Mean baseline utility and standard deviation from OCTAVE 1, 2 and Sustain trials

	OCTAVE 1	OCTAVE 2	OCTAVE Sustain
Tofacitinib 5 mg BID	[redacted]	[redacted]	[redacted]
Tofacitinib 10 mg BID	[redacted]	[redacted]	[redacted]
Placebo	[redacted]	[redacted]	[redacted]

B6. **Priority question**: The ERG is unable to replicate the following scenarios. Please provide further clarification on how you conducted these analyses.

a. ITT population analysis (described in CS section B.3.7.2 p.156)
b. Tofacitinib maintenance dose (described in CS section B.3.7.2 p.157)
c. Central read NMA results (CS section B.3.8.4, scenario 7 p.165)
d. Adalimumab maintenance dose (CS section B.3.8.4, scenario 9 p.165)
e. Golimumab maintenance dose (CS section B.3.8.4, scenario 10 p.165)
f. Vedolizumab maintenance dose (CS section B.3.8.4, scenario 11 p.165)

][.Pfizer response:

a. The overall ITT population analysis uses the results of the NMA of clinical response and clinical remission performed on the overall ITT populations from included studies. Results as treatment effects for both induction and maintenance phase analyses are presented in Appendix D, Table 106 of the CS. To implement these data in the economic model, the biologic-exposed population efficacy parameters ('EfficacySafety' tab) were modified, along with the relevant anchor and cut-offs (presented in the Table 17).

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Patients characteristics were manually updated from the 'PatientsChrs' tab, to match the overall ITT population baseline characteristics from OCTAVE Induction 1 and 2 (Age: 41.21 years; Male: 59.22%; Patient weight: 73.614 kg).

Table 17 treatment effect assumptions for the ITT population economic analysis

Model parameter	Induction	Maintenance
	Median (95% Credible Interval)	
Anchor (PBO)	[redacted]	[redacted]
Cut-off (z) Remission	[redacted]	[redacted]

b. The tofacitinib maintenance dose scenario uses a weighted average of the total efficacy and total costs associated with tofacitinib 5 mg ([redacted]) and tofacitinib 10 mg ([redacted]). The model was run separately for the tofacitinib 5 mg and the 10 mg maintenance dose and the weighted average was calculated from the results of those two analyses. Inputs for other treatments remained unchanged between model runs.
Please note that Table 65 (p.151) in the company submission contains an error in the Incremental QALY and Costs columns and Table 18 presents the correct values. The conclusions however remain unchanged.

Table 18 Biologic naïve population: tofacitinib [redacted] maintenance dose mix

Strategy	Total		Incremental		ICER (£/QALY) fully incremental
	QALYs	Costs (£)	QALYs	Costs (£)	
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£12,627.81

c. The central read NMA results analysis uses the efficacy parameters from the NMA sensitivity analyses using central endoscopic sub-score readings from OCTAVE Induction 1, OCTAVE Induction 2 and OCTAVE Sustain. Results as treatment effects for both induction and maintenance phase analyses are presented in Appendix D, Table 101 of the CS. To implement these data in the economic model, the efficacy parameters were updated in the 'EfficacySafety' tabs, along with the relevant anchor and cut-off (Table 19), and the model was run for both populations.

Table 19 Model set up - Central read analysis

	Model parameter	Biologic-naïve	Biologic-exposed
		Median (95% Credible Interval)	
Induction	Anchor (PBO)	[redacted]	[redacted]
	Cut-off (z) Remission	[redacted]	[redacted]
Maintenance	Anchor (PBO)	[redacted]	[redacted]
	Cut-off	[redacted]	[redacted]

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

d. The scenario exploring adalimumab mixed maintenance dose can be selected from the 'ModelSettings' tab by selecting "Ada_40mgEOW+EW" in the "Determine maintenance treatment" section. This results in the calculation of the cost of adalimumab as the weighted average of adalimumab 40 mg EOW (73%) and EW (27%).

e. The maintenance dose of golimumab 100 mg can be selected from the 'ModelSettings' tab by choosing "GOL_100mg" in the "Determine maintenance treatment" section. This results in the selection of the 100 mg maintenance dose for the efficacy calculations in 'TransProb' tab and the drug cost calculations in the 'Cost_Drug' tab.

f. The maintenance dose of vedolizumab Q4W can be selected from the 'ModelSettings' tab by choosing "VED_Q4W" in the "Determine maintenance treatment" section. This results in the selection of the 300 mg Q4W maintenance dose for the the efficacy calculations in the 'TransProb' tab and the drug cost calculations in the 'Cost_Drug' tab.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

References

1. Feagan BGR, D. T. Danese, S. Vermeire, S. Abhyankar, B. Sankoh, S. James, A. Smyth, M. Efficacy of Vedolizumab Induction and Maintenance Therapy in Patients With Ulcerative Colitis, Regardless of Prior Exposure to Tumor Necrosis Factor Antagonists. Clinical Gastroenterology and Hepatology. 2017;15(2):229-39.e5
2. Feagan BGR, S Vermeire, WJ Sandborn, W Reinsch, J Panes, D Tarabar, C Su, W Niezychowski, H Zhang, G Friedman, D Woodworth, BE Sands; Tofacitinib for maintenance therapy in patients with active ulcerative colitis in the Phase 3 OCTAVE Sustain trial: results by local and central endoscopic assessments; United European Gastroenterology Journal 2017; 5 (Supplement 1) P416
3. OCTAVE UC Clinical Research Organization Project Charter
4. Irvine EJ, Feagan B, Rochon J, Archambault A, Fedorak RN, Groll A, Kinnear D, Saibil F, McDonald JW. Quality of life: a valid and reliable measure of therapeutic efficacy in the treatment of inflammatory bowel disease. Gastroenterology. 1994 Feb 1;106(2):287-96.
5. Irvine EJ. Development and subsequent refinement of the inflammatory bowel disease questionnaire: a quality-of-life instrument for adult patients with inflammatory bowel disease. Journal of pediatric gastroenterology and nutrition. 1999 Apr 1;28(4):S23-7.
6. Han SW, McColl E, Steen N, Barton JR, Welfare MR. The inflammatory bowel disease questionnaire: a valid and reliable measure in ulcerative colitis patients in the North East of England. Scandinavian journal of gastroenterology. 1998 Jan 1;33(9):961-6.
7. Feagan BG, Patel H, Colombel JF, Rubin DT, James A, Mody R, Lasch K. Effects of vedolizumab on health-related quality of life in patients with ulcerative colitis: results from the randomised GEMINI 1 trial. Alimentary pharmacology & therapeutics. 2017 Jan 1;45(2):264-75.
8. Travis S, Feagan BG, Peyrin-Biroulet L, Panaccione R, Danese S, Lazar A, Robinson AM, Petersson J, Pappalardo BL, Bereswill M, Chen N. Effect of adalimumab on clinical outcomes and health-related quality of life among patients with ulcerative colitis in a clinical practice setting: Results from InspirADA. Journal of Crohn's and Colitis. 2017 Oct 27;11(11):1317-25.
9. Binion DG, Louis E, Oldenburg B, Mulani P, Bensimon AG, Yang M, Chao J. Effect of adalimumab on work productivity and indirect costs in moderate to severe Crohn's disease: a meta-analysis. Canadian Journal of Gastroenterology and Hepatology. 2011;25(9):492-6.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

10. Tracey D, Klareskog L, Sasso EH, Salfeld JG, Tak PP. Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. Pharmacology & therapeutics. 2008 Feb 1;117(2):244-79.

11. Mukherjee Arnab et al, 1.Mo1781 Tofacitinib Plasma Concentration Monitoring Is Not Needed for Optimization of Induction Therapy in Moderate to Severe Ulcerative Colitis: Results of Pooled Exposure-Response Analyses of Phase 3 Induction Studies; Gastroenterology , Volume 150 , Issue 4 , S774

12. L.C.S. De Vries, M.E. Wildenberg, W.J. De Jonge, G.R. D'Haens; The Future of Janus Kinase Inhibitors in Inflammatory Bowel Disease, Journal of Crohn's and Colitis, Volume 11, Issue 7, 1 July 2017, Pages 885–893, https://doi.org/10.1093/ecco-jcc/jjx003

13. Danese S, Grisham M, Hodge J, Telliez JB. JAK inhibition using tofacitinib for inflammatory bowel disease treatment: a hub for multiple inflammatory cytokines. American Journal of Physiology-Gastrointestinal and Liver Physiology. 2015 Nov 25;310(3):G155-62.

14. Pouillon L, Bossuyt P, Peyrin-Biroulet L. Tofacitinib Is the Right OCTAVE for Ulcerative Colitis. Gastroenterology. 2017 Sep 1;153(3):862-4.

15. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials.

16. Archer R, Tappenden P, Ren S, James MMS, Harvey R, Basarir H, et al. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (Including a review of TA140 and TA262): Clinical effectiveness systematic review and economic model. Health Technology Assessment. 2016;20(39).

17. Essat M, Tappenden P, Ren S, Bessey A, Archer R, Wong R, et al. Vedolizumab for the Treatment of Adults with Moderate-to-Severe Active Ulcerative Colitis: An Evidence Review Group Perspective of a NICE Single Technology Appraisal. PharmacoEconomics. 2016;34(3):245-57.

18. Infliximab SmPC; https://www.medicines.org.uk/emc/product/3831/smpc, accessed 20th June 2018

19. Vedolizumab SmPC; https://www.medicines.org.uk/emc/product/5442; accessed 20th June 2018

20. Golimumab SmPC; https://www.medicines.org.uk/emc/product/5133/smpc; accessed 20th June 2018

21. Adalimmumab SmPC; https://www.medicines.org.uk/emc/product/7986/smpc, accessed 20th June 2018

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Appendix A – Tofacitinib Phase 2 Inclusion and Exclusion Criteria

Linked to ERG Clarification Question A2

Full inclusion and exclusion criteria as listed on pages 32 to 36 in the Phase II CSR;

List of inclusion criteria:

Subjects were required to meet all of the following inclusion criteria to be eligible for enrolment into the study:

1. [redacted]

List of exclusion criteria:

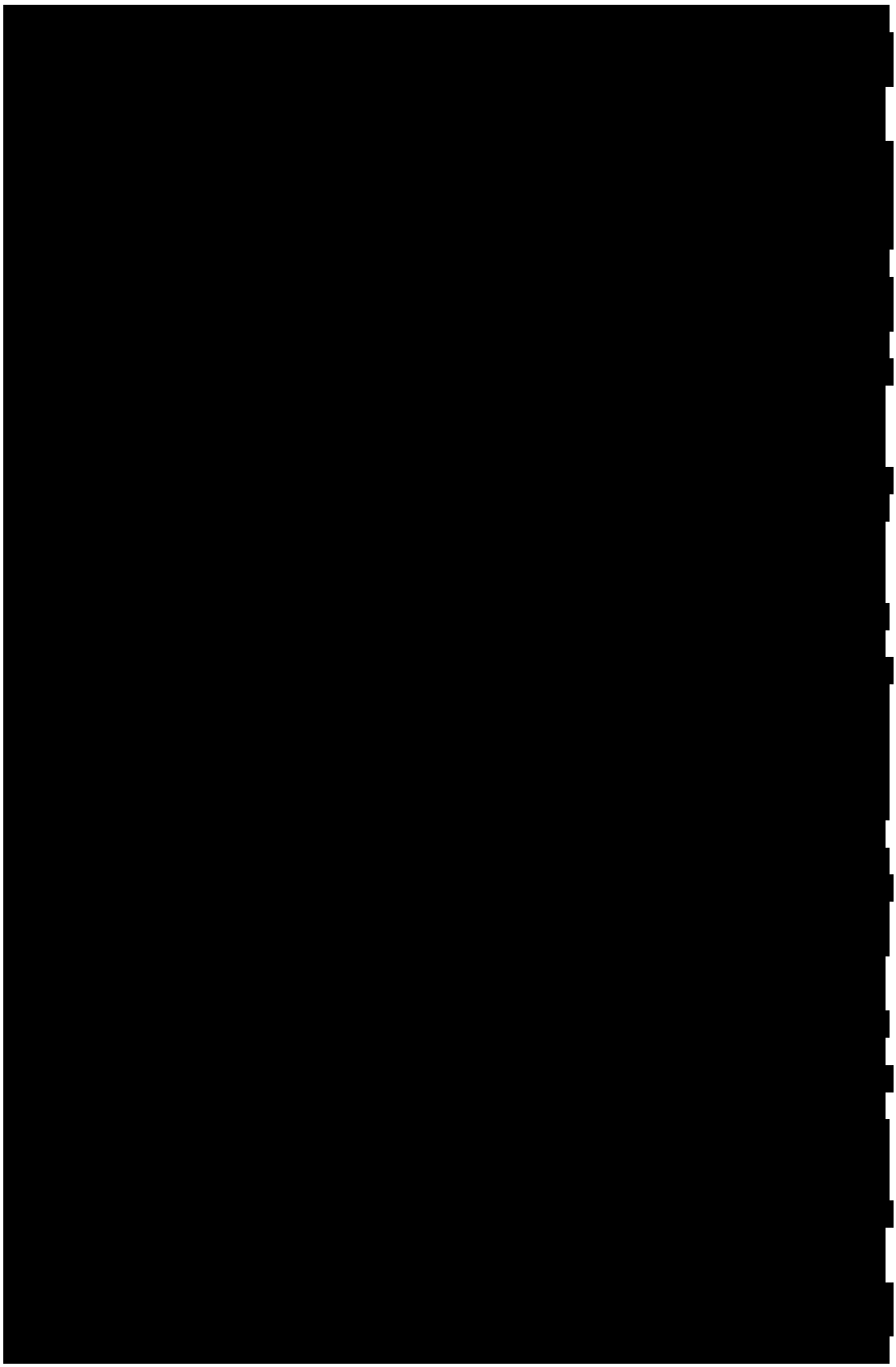
Subjects presenting with any of the following were not included in the study:

1. [redacted]

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140



NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Appendix B – List of excluded studies and reason for exclusion

Linked to ERG Clarification Question A5

	Reference	Reason for exclusion
1	W. F. Sandborn, B. Marano, C. Strauss, R. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Gibson, P. Collins, J. Jarnerot, G. Rutgeerts, P; A phase 3 randomized, placebo-controlled, double-blind study to evaluate the safety and efficacy of subcutaneous golimumab maintenance therapy in patients with moderately to severely active ulcerative colitis: pursuit-maintenance. American journal of gastroenterology. 2012.	Duplicate of included study
2	.G. R. R. Lichtenstein, P. Sandborn, W. J. Sands, B. E. Diamond, R. H. Blank, M. Montello, J. Tang, L. Cornillie, F. Colombel, J. F.; A pooled analysis of infections, malignancy, and mortality in infliximab-and immunomodulator-treated adult patients with inflammatory bowel disease. American Journal of Gastroenterology. 2012	Study design; narrative review
3	G. J. S. Mantzaris, M. Archavlis, E. Petraki, K. Christidou, A. Karagiannidis, A. Triadaphyllou, G; A prospective randomized observer-blind 2-year trial of azathioprine monotherapy versus azathioprine and olsalazine for the maintenance of remission of steroid-dependent ulcerative colitis. American Journal of Gastroenterology. 2004.	Incorrect interventions; not in scope
4	S. M. M. Wilhelm, K. A. Rivait, K. N. Kale-Pradhan, P. B.; A review of infliximab use in ulcerative colitis. Clinical Therapeutics. 2008.	Study design; non systematic review
5	W. J. V. A. Sandborn, G. Reinisch, W. Colombel, J. D'Haens, G. Wolf, D. C. Kron, M. Tighe, M. B. Lazar, A. Thakkar, R. B.; Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2012	Duplicate of included study
6	B. G. S. Feagan, W. J. Yang, M. Lomax, K. G. Lazar, A. Thakkar, R. B. Mulani, P. M. Chao, J; Adalimumab therapy reduces hospitalization and Colectomy rates in patientswith Ulcerative Colitis: Data from controlled trials. Inflammatory Bowel Diseases. 2011.	Duplicate study with no additional data
7	S. D. Dayanand, P. Manickam, S. Williams, F.; An indirect comparisons analysis of TNF inhibitors vs. selective adhesion molecule inhibitors in the induction of response in ulcerative colitis. Gastroenterology. 2015.	Disease severity incorrect
8	E. D. S. Shah, C. A. Chong, K. Melmed, G.; Anti-TNF agents may pose a higher risk of infection than anti-integrin agents in treating inflammatory bowel disease. Gastroenterology. 2015	Study design; non systematic review
9	G. D. Fiorino, S. Peyrin-Biroulet, L.; Anti-TNF therapy in inflammatory bowel diseases. Clinical Update on Inflammatory Disorders of the Gastrointestinal Tract. 2010.	Study design; narrative review

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
10	Z. D. Zhou, C. Liu, W. X.; Anti-TNF-A therapy about infliximab and adalimamab for the effectiveness in ulcerative colitis compared with conventional therapy: A meta-Analysis. Hepato-Gastroenterology. 2015.	Unable to Access
11	D. P. T. Jewell, S. C; Azathioprine in ulcerative colitis. Gut. 1972.	Disease severity incorrect
12	K. K. O. Joergensen, I. C. Goll, G. L. Lorentzen, M. Bolstad, N. Berset, I. P. Haavardsholm, E. A. Lundin, K. E. Mork, C. Kvien, T. K. Jahnsen, J; Biosimilar infliximab (CT-P13) is not inferior to originator infliximab: Explorative ibd subgroup-analyses in Crohn's disease and ulcerative colitis from the nor-switch trial. Gastroenterology. 2017.	Disease severity incorrect
13	G. L. O. Goll, I. C. Jorgensen, K. K. Lorentzen, M. Bolstad, N. Haavardsholm, E. A. Lundin, K. E. A. Mork, C. Jahnsen, J. Kvien, T. K; Biosimilar infliximab (CT-P13) is not inferior to originator infliximab: Results from a 52-week randomized switch trial in Norway. Arthritis and Rheumatology. 2016.	Disease severity incorrect
14	D. B. Laharie, A. Branche, J. Allez, M. Bouhnik, Y. Filippi, J. Zerbib, F. Nachury, M. Savoye, G. Moreau, J. Delchier, J. Ricart, E. Cosnes, J. Lopez-Sanroman, A. Dewit, O. Carbonnel, F. Bommelaer, G. Coffin, B. Van Assche, G. Esteve, M. Faarkila, M. Perez, A. Mary, J. Colombel, J. Lemann, M; Ciclosporin versus infliximab in acute severe ulcerative colitis refractory to intravenous steroids: A randomized study. Journal of Crohn's and Colitis. 2011.	Outcomes out of scope
15	D. B. Laharie, A. Branche, J. Allez, M. Bouhnik, Y. Filippi, J. Zerbib, F. Savoye, G. Nachury, M. Moreau, J. Delchier, J. C. Cosnes, J. Ricart, E. Dewit, O. Lopez-Sanroman, A. Dupas, J. L. Carbonnel, F. Bommelaer, G. Coffin, B. Roblin, X. Van Assche, G. Esteve, M. Farkkila, M. Gisbert, J. P. Marteau, P. Nahon, S. De Vos, M. Franchimont, D. Mary, J. Y. Colombel, J. F. Lemann, M.; Ciclosporin versus infliximab in patients with severe ulcerative colitis refractory to intravenous steroids: A parallel, open-label randomised controlled trial. The Lancet. 2012.	Outcomes out of scope
16	F. C. Balzola, G. Ho, G. T. Hoentjen, F. Russell, R. K; Ciclosporin versus infliximab in patients with severe ulcerative colitis refractory to intravenous steroids: A parallel, open-label randomised controlled trial. Inflammatory Bowel Disease Monitor. 2013	Study design; summary of published trial report
17	D. L. A. B. J. Branche; Ciclosporine versus infliximab in patients with severe ulcerative colitis refractory to intravenous steroids: a parallel, open-label randomised controlled trial. The Lancet. 2012.	Duplicate of excluded study
18	W. F. Sandborn, B. Marano, C. Strauss, R. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Gibson, P. Collins, J. Jarnerot, G. Rutgeerts, P; Clinical response is a meaningful endpoint in ulcerative colitis clinical studies. Inflammatory Bowel Diseases. 2012.	Study design; non RCT
19	J. G. A. Williams, M. F. Alrubaiy, L. Clement, C. Cohen, D. Croft, G. P. Grey, M. Hutchings, H. A. Morgan, J. M. Rapport, F. Russell, I. T. Seagrove, A. C. Watkins, A.; Comparative clinical effectiveness of infliximab and ciclosporin for acute severe ulcerative colitis: Early results from the construct trial.United European Gastroenterology Journal. 2014.	Disease severity incorrect, acute severe colitis

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
20	A. M. Yoshida, T. Ueno, F. Kanoshima, K. Shirai, M. Morikawa, Y. Endo, Y. Comparative effectiveness of calcineurin inhibitors and biologics for inducing mucosal healing in patients with ulcerative colitis: A network meta-analysis. American journal of gastroenterology. 2015.	Incorrect interventions; not in scope
21	J. G. A. Williams, M. F. Alrubaiy, L. Clement, C. Cohen, D. Grey, M. Hilton, M. Hutchings, H. A. Longo, M. Morgan, J. M. Rapport, F. L. Seagrove, A. C. Watkins, A. Comparison Of iNfliximab and ciclosporin in STeroid Resistant Ulcerative Colitis: Pragmatic randomised trial and economic evaluation (CONSTRUCT). 2016. Health Technology Assessment	Incorrect interventions; not in scope
22	E. V. S. Loftus, S. Ramachandran, P. Yang, Z. Guo, C. Y. Gasink, C. Comparison of rates of active tuberculosis infection in the phase 2 and 3 clinical trial programs for anti-IL12/23 and anti-TNFs. 2017. Gastroenterology	Mixed population with no usable data
23	B. S. Feagan, M. Thakkar, R. Lazar, A. Yang, M. Macaulay, D. Chao, J. Sandborn, W. J. Effect of adalimumab dose escalation on hospitalization risk in patients with moderately to severely active ulcerative colitis. 2015. American Journal of Gastroenterology	Study design; single arm
24	J. F. J. Colombel, B. Sandborn, W. J. Feagan, B. Peyrin-Biroulet, L. Eichner, S. F. Robinson, A. M. Mostafa, N. M. Zhou, Q. Thakkar, R. B. Effects of concomitant immunomodulators on the pharmacokinetics, efficacy and safety of adalimumab in patients with Crohn's disease or ulcerative colitis who had failed conventional therapy. 2017. Alimentary Pharmacology and Therapeutics	Study design; narrative review
25	B. G. S. Feagan, W. Smyth, M. D. Sankoh, S. Parikh, A. Fox, I. Effects of continued vedolizumab therapy for ulcerative colitis in week 6 induction therapy nonresponders. 2014. Gastroenterology	Study design; non-RCT
26	B. S. Feagan, W. J. Smyth, M. Sankoh, S. Parikh, A. Fox, I. Effects of continued vedolizumab therapy for ulcerative colitis in week 6 induction therapy nonresponders. 2014. Journal of Crohn's and Colitis	Duplicate of excluded study
27	O. X. Adedokun, Z. Marano, C. Strauss, R. Zhang, H. Johanns, J. Ford, J. Zhou, H. Davis, H. Colombel, J. F. Reinisch, W. Feagan, B. Rutgeerts, P. Sandborn, W. Effects of immunomodulators on the pharmacokinetics and efficacy of golimumab in patients with moderately to severely active ulcerative colitis: results from phase 2/3 pursuit-SC induction and maintenance studies. 2013. American journal of gastroenterology.	Outcomes out of scope
28	E. K. Gavalas, J. Stergiopoulos, C. Zavos, C. Gisakis, D. Nikolaidis, N. Giouleme, O. Chatzopoulos, D. Kapetanakis, N. Efficacy and safety of infliximab in steroid-dependent ulcerative colitis patients. 2007. Hepato-Gastroenterology	Study design; non randomised
29	W. J. S. Sandborn, B. E. D'Haens, G. Vermeire, S. Schreiber, S. Danese, S. Panes, J. Feagan, B. G. Reinisch, W. Niezychowski, W. Friedman, G. Lawendy, N. Yu, D. Woodworth, D. Mukherjee, A. Healey, P. Zhang, H. Su, C. Efficacy and safety of oral tofacitinib as induction therapy in patients with moderate-to-severe ulcerative colitis: Results from 2 phase 3 randomised controlled trials. 2016. Journal of Crohn's and Colitis	Duplicate study with no additional data

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
30	Sandborn W.J., Sands B.E., Danese S., D'Haens G.R., Vermeire S., Schreiber S., Feagan B., Reinisch W., Friedman G., Woodworth D., Zhang H., Lawendy N., Niezychowski W., Su C., Panés J. Efficacy and safety of oral tofacitinib as maintenance therapy in patients with moderate to severe ulcerative colitis: results from a phase 3 randomised controlled trial. 2017.	Duplicate study with no additional data
31	J. Lofland. Efficacy of adalimumab and infliximab for the treatment of moderate to severe ulcerative colitis: Number needed to treat analysis of randomized controlled trials. 2012. American Journal of Gastroenterology	Study design; narrative review
32	A. Parikh. Efficacy of vedolizumab in ulcerative colitis by prior treatment failure in gemini i, a randomized, placebo-controlled, double-blind, multicenter trial. 2012. Inflammatory Bowel Diseases	Duplicate of included study
33	B. G. S. Feagan, C. A. Melmed, G. Y. Isaacs, K. Lasch, K. Rosario, M. Green, A. Abhyankar, B. Efficacy of vedolizumab with and without continued immunosuppressant use in GEMINI 1 and GEMINI 2. 2015. United European Gastroenterology Journal	Duplicate study with no additional data
34	J. D. R. Lewis, W. Bressler, B. Parikh, A. Yang, H. Rosario, M. Roseth, A. Danese, S. Feagan, B. G. Sands, B. E. Ginsburg, P. Dassopoulos, T. Xu, J. Wyant, T. Faecal calprotectin reductions in patients achieving mucosal healing with vedolizumab induction therapy in GEMINI 1. 2016. Journal of Crohn's and Colitis	Outcomes out of scope
35	J. R. Lewis, W. Bressler, B. Parikh, A. Yang, H. Rosario, M. Roseth, A. Danese, S. Feagan, B. G. Sands, B. E. Ginsburg, P. Dassopoulos, T. Xu, J. Wyant, T. Fecal calprotectin reductions in patients with mucosal healing during vedolizumab induction therapy in GEMINI 1. 2016. Gastroenterology	Outcomes out of scope
36	S. A. Kedia, V. Makharia, G. K. Golimumab for moderately to severely active ulcerative colitis. 2016. Expert Review of Clinical Pharmacology	Study design out of scope
37	W. V. A. Sandborn, G. Reinisch, W. Colombel, J. F. D'Haens, G. Wolf, D. Kron, M. Tighe, M. Lazar, A. Thakkar, R. Induction and maintenance of clinical remission by adalimumab in patientswith moderate-to-severe Ulcerative Colitis. 2011. Inflammatory Bowel Diseases	Duplicate study with no additional data
38	G. H. Jarnerot, E. Friis-Liby, I. Blomquist, L. Karlen, P. Granno, C. Vilien, M. Strom, M. Danielsson, A. Verbaan, H. Hellstrom, P. M. Magnuson, A. Curman, B. Infliximab as rescue therapy in severe to moderately severe ulcerative colitis: A randomized, placebo-controlled study. 2005. Gastroenterology	Acute disease out of scope
39	M. S. Radin, S. Roccatello, D. Cuadrado, M. J. Infliximab Biosimilars in the Treatment of Inflammatory Bowel Diseases: A Systematic Review. 2017. BioDrugs	Disease severity incorrect
40	T. S. Ochsenkuhn, M. Goke, B. Infliximab for acute, not steroid-refractory ulcerative colitis: A randomized pilot study. 2004. European Journal of Gastroenterology and Hepatology	Disease severity incorrect; acute

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
41	A. C. F. Moss, R. J. Infliximab for induction and maintenance therapy for ulcerative colitis. 2006. Gastroenterology	Study design out of scope
42	A. K. Akobeng. Infliximab for induction and maintenance therapy for ulcerative colitis. 2006. Journal of Pediatric Gastroenterology and Nutrition	Study design; summary of published trial report
43	P. S. Rutgeerts, W. J. Feagan, B. G. Reinisch, W. Olson, A. Johanns, J. Travers, S. Rachmilewitz, D. Hanauer, S. B. Lichtenstein, G. R. de Villiers, W. J. Present, D. Sands, B. E. Colombel, J. F. Infliximab for induction and maintenance therapy for ulcerative colitis.[Erratum appears in N Engl J Med. 2006 May 18;354(20):2200]. 2005. New England Journal of Medicine	Duplicate of included study
44	F. C. Balzola, G. Hoentjen, F. Ho, G. T. Russell, R. Infliximab in steroid-dependent ulcerative colitis: Effectiveness and predictors of clinical and endoscopic remission. 2013. Inflammatory Bowel Disease Monitor	Study design out of scope
45	B. E. T. Sands, W. J. Sandborn, W. J. Rutgeerts, P. J. Hanauer, S. B. Mayer, L. Targan, S. R. Podolsky, D. K. Infliximab in the treatment of severe, steroid-refractory ulcerative colitis: a pilot study. 2001. Inflammatory Bowel Diseases	Intervention out of scope; single infusion IFX treatment
46	A. D. P. Armuzzi, B. Lupascu, A. Fedeli, P. Leo, D. Mentella, M. C. Vincenti, F. Melina, D. Gasbarrini, G. Pola, P. Gasbarrini, A. Infliximab in the treatment of steroid-dependent ulcerative colitis. 2004. European Review for Medical & Pharmacological Sciences	Duplicate of included study
47	A. P. Armuzzi, B. Lupascu, A. Fedeli, P. Leo, D. Mentella, M. C. Vincenti, F. Melina, D. Gasbarrini, G. Pola, P. Gasbarrini, A. Infliximab in the treatment of steroid-dependent ulcerative colitis. 2004. European review for medical and pharmacological sciences	Duplicate of included study
48	N. M. Narula, J. Colombel, J. F. Leontiadis, G. Muqtadir, Z. Reinisch, W. Infliximab or cyclosporine as rescue therapy in patients with severe ulcerative colitis refractory to steroids: A systematic review and meta-analysis. 2014. American Journal of Gastroenterology	Study design out of scope
49	J. G. A. Williams, M. F. Alrubaiy, L. Arnott, I. Clement, C. Cohen, D. Gordon, J. N. Hawthorne, A. B. Hilton, M. Hutchings, H. A. Jawhari, A. U. Longo, M. Mansfield, J. Morgan, J. M. Rapport, F. Seagrove, A. C. Sebastian, S. Shaw, I. Travis, S. P. L. Watkins, A. Infliximab versus ciclosporin for steroid-resistant acute severe ulcerative colitis (CONSTRUCT): a mixed methods, open-label, pragmatic randomised trial. 2016 The Lancet Gastroenterology and Hepatology	incorrect comparator; not in scope
50	G. R. Lichtenstein Is infliximab effective for induction of remission in patients with ulcerative colitis? 2001. Inflammatory Bowel Diseases	Study design out of scope
51	A. F. Parikh, I. Leach, T. Xu, J. Scholz, C. Patella, M. Feagan, B. G. Long-term clinical experience with vedolizumab in patients with inflammatory bowel disease. 2013. Inflammatory Bowel Diseases	Incorrect interventions; unlicensed dose

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
52	A. L. Parikh, T. Xu, J. Feagan, B. Long-Term clinical experience with vedolizumab in patients with mild to moderate ulcerative colitis. 2011. American Journal of Gastroenterology	Incorrect interventions; unlicensed dose
53	E. V. C. Loftus, J. F. Feagan, B. G. Vermeire, S. Sandborn, W. J. Sands, B. E. Danese, S. D'Haens, G. R. Kaser, A. Panaccione, R. Rubin, D. T. Shafran, I. McAuliffe, M. Kaviya, A. Sankoh, S. Mody, R. Abhyankar, B. Smythm, M. Long-term efficacy of vedolizumab for ulcerative colitis. 2017. Journal of Crohn's and Colitis	Study design; single arm
54	E. V. Loftus, Jr. Colombel, J. F. Feagan, B. G. Vermeire, S. Sandborn, W. J. Sands, B. E. Danese, S. D'Haens, G. R. Kaser, A. Panaccione, R. Rubin, D. T. Shafran, I. McAuliffe, M. Kaviya, A. Sankoh, S. Mody, R. Abhyankar, B. Smyth, M. Long-term Efficacy of Vedolizumab for Ulcerative Colitis. 2017. Journal of Crohn's & colitis	Duplicate of excluded study
55	B. K. Feagan, A. Smyth, M. Panaccione, R. Sankoh, S. Abhyankar, B. Long-term efficacy of vedolizumab therapy for ulcerative colitis. 2014. United European Gastroenterology Journal	Study design; single arm
56	G. L. J. Goll, K. K. Sexton, J. Olsen, I. C. Bolstad, N. Lorentzen, M. Haavardsholm, E. A. Mork, C. Jahnsen, J. Kvien, T. K. Long-term safety and efficacy of biosimilar infliximab (CT-P13) after switching from originator infliximab: Results from the 26-week open label extension of a randomized Norwegian trial. 2017 Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP	Population out of scope
57	P. R. R. Gibson, W. Sandborn, W. Feagan, B. G. Marano, C. W. Strauss, R. Johanns, J. Zhang, H. Padgett, L. Colombel, J. F. Collins, J. Rutgeerts, P. J. Long-term safety and efficacy of golimumab in patients with moderately to severely active ulcerative colitis: results from the pursuit-SC maintenance study extension. 2014. Gastroenterology.	Study design; single arm
58	J. F. P. Colombel, R. Ghosh, S. Sandborn, W. J. Rutgeerts, P. Hanauer, S. Van Assche, G. Reinisch, W. Peyrin-Biroulet, L. Robinson, A. M. Lau, W. Huang, B. Pappalardo, B. Read, H. A. Long-term safety of adalimumab in clinical trials in adult patients with Crohn's disease or ulcerative colitis. 2015. United European Gastroenterology Journal	Study design out of scope
59	J. J. Gao, X. L. Low-dose infliximab for corticosteroid-refractory ulcerative colitis: impact of number of infusions on efficacy and safety. 2013. World chinese journal of digestology	Intervention out of scope; Unlicensed dose of IFX
60	P. R. F. Gibson, B. G. Sandborn, W. J. Marano, C. Strauss, R. Johanns, J. Padgett, L. Collins, J. Tarabar, D. Hebzda, Z. Rutgeerts, P. Reinisch, W.. Maintenance of Efficacy and Continuing Safety of Golimumab for Active Ulcerative Colitis: PURSUIT-SC Maintenance Study Extension Through 1 Year. 2016. Clinical and Translational Gastroenterology	Study design; single arm
61	Z. W. Yang, Q. Wu, K. Fan, D. Meta-analysis: pre-operative infliximab treatment and short-term post-operative complications in patients with ulcerative colitis (Structured abstract). 2010. Alimentary Pharmacology and Therapeutics	Outcomes out of scope

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
62	G. V. A. D'Haens, G. Wolf, D. Sandborn, W. Colombel, J. F. Lazar, A. Kron, M. Robinson, A. Thakkar, R.. Mucosal healing in ulcerative colitis patients with week 8 response to adalimumab: Subanalysis of ultra 2. 2012. Inflammatory Bowel Diseases	Duplicate of included study
63	W. J. Sandborn. Mucosal healing with infliximab: Results from the active ulcerative colitis trials. 2012. Gastroenterology and Hepatology	Study design out of scope
64	A. G. Amiot, C. Serrero, M. Grimaud, J. C. Peyrin-Biroulet, L. Zallot, C. Bigard, M. A. Filippi, J. Hebuterne, X. Pariente, B. Nachury, M. Desreumaux, P. Roblin, X. del Tedesco, E. Buisson, A. Bommelaer, G. Stefanescu, C. Bouhnik, Y. Boureille, A. Trang-Poisson, C. Altwegg, R. Marteau, P. Dray, X. Carbonnel, F. Vaysse, T. Seksik, P. Beaugerie, L. Cosnes, J. Sokol, H. Landman, C. Bourrier, A. Nancey, S. Boschetti, G. Laharie, D. Poullenot, F. Allez, M. Gornet, J. M. Baudry, C. Savoye, G. Moreau, J. Vuitton, L. Koch, S. Viennot, S. Aubourg, A. Picon, L. Pelletier, A. L. Sickersen, G. Bouguen, G. Abitbol, V. Chaussade, S. Nahon, S. Winkfield, B. Brixi-benmansour, H. Gincul, R. Barberis, J. C. Bonaz, B. Michiels, C. Zerbib, F. de Beauregard, M. B. Locher, C. Davin-Couve, S. Poirette, A. Guillem, L. Stetiu-Mocanu, M. Philippe, B. Beorchia, S. Al Qaddi, J. One-year effectiveness and safety of vedolizumab therapy for inflammatory bowel disease: a prospective multicentre cohort study. 2017. Alimentary Pharmacology and Therapeutics	Study design out of scope
65	A. W. Croft, A. Doecke, J. Cooley, R. Howlett, M. Radford-Smith, G. Outcomes of salvage therapy for steroid-refractory acute severe ulcerative colitis: Ciclosporin vs. infliximab. 2013. Alimentary Pharmacology and Therapeutics	Study design out of scope
66	W. C. Reinisch, J. F. Feagan, B. G. Han, C. Marano, C. Strauss, R. Gibson, P. Sandborn, W. J. Huyck, S. Cornillie, F. Rutgeerts, P. Patient-reported outcomes can be used to monitor continuous clinical response in patients with moderately to severely active ulcerative colitis treated with golimumab: Results from the pursuit maintenance study. 2015. United European Gastroenterology Journal	Study design out of scope
67	J. F. S. Colombel, W. Reinisch, W. Robinson, A. Wang, W. Huang, B. Lazar, A. Thakkar, R. Patient-reported symptom measures differ in their association with mucosal healing in adults with moderately to severely active ulcerative colitis: Results from ultra 1 and 2. 2014. Gastroenterology	Study design out of scope
68	M. W. Rosario, T. Milch, C. Parikh, A. Feagan, B. Sandborn, W. J. Yang, H. Fox, I. Pharmacokinetic and pharmacodynamic relationship and immunogenicity of vedolizumab in adults with inflammatory bowel disease: Additional results from the GEMINI 1 and 2 studies. 2014. Journal of Crohn's and Colitis	Study design out of scope
69	J. F. H. Colombel, C. Reinisch, W. Feagan, B. Marano, C. Strauss, R. Johanns, J. Zhang, H. Gibson, P. Collins, J. Rutgeerts, P. Sandborn, W. Predictive value of patient-reported outcomes to mucosal healing in patients with moderately to severely active ulcerative colitis. 2014. Value in Health	Study design out of scope; non RCT

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
70	B. S. Feagan, W. Reinisch, W. Ghosh, S. Robinson, A. Lazar, A. Zhou, Q. Skup, M. Thakkar, R. Predictors of hospitalization in patients with moderately to severely active ulcerative colitis from ultra 1 and ultra 2. 2014. American Journal of Gastroenterology	Outcomes out of scope
71	Randomised clinical study: discrepancies between patient-reported outcomes and endoscopic appearance in moderate to severe ulcerative colitis. 2015. Alimentary pharmacology & therapeutics	Duplicate of excluded study
72	B. S. Jharap, W. J. Reinisch, W. D'Haens, G. Robinson, A. M. Wang, W. Huang, B. Lazar, A. Thakkar, R. B. Colombel, J. F. Randomised clinical study: Discrepancies between patient-reported outcomes and endoscopic appearance in moderate to severe ulcerative colitis. 2015. Alimentary Pharmacology and Therapeutics	Study design; single arm
73	D. D. H. Wolf, G. Sandborn, W. Colombel, J. F. Assche, G. Lazar, A. Zhou, Q. Robinson, A. Chao, J. Thakkar, R.. Rate of and response to dose escalation in patients treated with adalimumab for moderately-to-severely active ulcerative colitis: ultra 2 subanalysis. 2012. Inflammatory bowel diseases.	Study design; single arm
74	J. G. Lewis, S. Sandborn, W. Van Assche, G. D'Haens, G. Lazar, A. Eichner, S. Huang, B. Robinson, A. Thakkar, R. Rates of "patient-defined" remission with adalimumab in patients with ulcerative colitis: Subanalysis of ULTRA 1 and ULTRA 2. 2013. Journal of Crohn's and Colitis	Duplicate study with no additional data
75	A. T. Armuzzi, C. Panes, J. Lakatos, P. Fisseha, N. Pappalardo, B.. Real-world effectiveness of adalimumab in patients with ulcerative colitis. 2016. Journal of Crohn's and Colitis	Outcomes out of scope
76	E. V. C. Loftus, J. F. Previtali, A. Smyth, M. Response and remission rates with up to 3 years of vedolizumab treatment in patients with ulcerative colitis. 2016. Journal of Crohn's and Colitis	LTE including non-randomised patients, therefore a single ar, study and out of scope
77	Anonymous. Safety and Efficacy of Subcutaneous Golimumab Induction Therapy in Patients with Moderately to Severely Active Ulcerative Colitis. 2012. Clinical Advances in Hematology and Oncology	Duplicate of included study
78	E. C. Loftus, J. F. Siegel, C. Lewis, J. Abhyankar, B. Sankoh, S. Smyth, M. Milch, C. Safety of vedolizumab alone or with concomitant corticosteroids and/or immunosuppressants in patients with ulcerative colitis or Crohn's disease. 2014. American Journal of Gastroenterology	Outcomes out of scope
79	J. O. Florholmen, G. Olsen, T. Rismo, R. Cui, G. Christiansen, I. Short-and long-term clinical outcomes of infliximab in fulminant ulcerative colitis. Ulcers	Disease severity incorrect; acute
80	F. C. Balzola, G. Ho, G. T. Russell, R. K.. Subcutaneous golimumab induces clinical response and remission in patients with moderate to severe ulcerative colitis. 2013. Inflammatory Bowel Disease Monitor	Based on study design, Summary of published trial report;

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
81	W. C. Sandborn, J. F. Yang, M. Thakkar, R. Mulani, P. Chao, J.. Sustained clinical remission of ulcerative colitis is associated with greater improvements in quality of life, work productivity and activity. 2011. American Journal of Gastroenterology	Study design; non-RCT
82	K. K. O. Jørgensen, I. C. Goll, G. L. Lorentzen, M. Bolstad, N. Haavardsholm, E. A. Lundin, K. E. A. Mørk, C. Jahnsen, J. Kvien, T. K. Switching from originator infliximab to biosimilar CT-P13 compared with maintained treatment with originator infliximab (NOR-SWITCH): a 52-week, randomised, double-blind, non-inferiority trial. 2017. Lancet (london, england)	disease severity incorrect
83	S. V. Lula, S. Gils, A. Accossato, P. Marren, A. Systematic literature review on the immunogenicity of biologics in inflammatory bowel disease. 2016 American Journal of Gastroenterology	Outcomes out of scope
84	N. M. Narula, J. Colombel, J. F. Leontiadis, G. I. Maqtadir, Z. Reinisch, W. Systematic review and meta-analysis: Infliximab or ciclosporin as rescue therapy in patients with severe ulcerative colitis refractory to steroids. 2015. Journal of Crohn's and Colitis	Disease severity incorrect; acute
85	E. D. S. Shah, C. A. Chong, K. Melmed, G. Y.. The comparative effectiveness of biologics and immunomodulators for the treatment of ulcerative colitis. 2014. Gastroenterology	Study design; abstract of SLR
86	J. F. S. Colombel, B. E. Rutgeerts, P. Sandborn, W. Danese, S. D'Haens, G. Panaccione, R. Loftus, E. V. Sankoh, S. Fox, I. Parikh, A. Milch, C. Abhyankar, B. Feagan, B. G. The safety of vedolizumab for ulcerative colitis and Crohn's disease. 2017. Gut	Study design; narrative review
87	F. C. Balzola, G. Ho, G. T. Russell, R. K. Wehkamp, J. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. 2012. Inflammatory Bowel Disease Monitor	Study design out of scope
88	S. J. B. Bickston, B. W. Tsoulis, D. J. Cheng, J. MacDonald, J. K. Khanna, R. Feagan, B. G. Vedolizumab for induction and maintenance of remission in ulcerative colitis. 2014. The Cochrane database of systematic reviews	Duplicate of included study
89	A. L. Parikh, T. Wyant, T. Scholz, C. Sankoh, S. Mould, D. R. Ponich, T. Fox, I. Feagan, B. G. Vedolizumab for the treatment of active ulcerative colitis: A randomized controlled phase 2 dose-ranging study. 2012. Inflammatory Bowel Diseases	Unlicensed doses and dosing schedules therefore out of scope
90	N. B. Shahidi, B. Panaccione, R.. Vedolizumab for the treatment of ulcerative colitis. 2016. Expert Opinion on Biological Therapy	Study design; narrative review
91	S. G. Chaplin, R. Patel, K. Irving, P.. Vedolizumab for ulcerative colitis and crohn's disease. 2015. Prescriber	Study design out of scope
92	E. P. P. Juanes Calabuig, P. Juarez Gimenez, J. C.. Vedolizumab in the treatment of Ulcerative Colitis and Crohn's Disease. 2016. European Journal of Clinical Pharmacy	Study design; narrative review

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
93	T. L. Wyant, J. Reinisch, W. Dassopoulos, T. Ginsburg, P. Sands, B. Feagan, B. Danese, S. Roseth, A. Rosario, M. Yang, H. Parikh, A. Bressler, B. Vedolizumab induces clinical response and remission across a range of baseline fecal calprotectin levels in patients with ulcerative colitis: Results from GEMINI 1. 2014. American Journal of Gastroenterology	Outcomes out of scope
94	Vedolizumab Induction Therapy for Ulcerative Colitis. 2012. Clinical advances in hematology & oncology	Disease severity incorrect
95	A. Rezaie. Vedolizumab, a gut-specific monoclonal antibody, renews hope for an alternative to anti-TNF therapy in inflammatory bowel diseases. 2014. Annals of Gastroenterology	Study design out of scope
	Systematic reviews and meta-analyses	
96	S. E.-A. Nikfar, S. Abdollahi, M.. A systematic review and meta-analysis of the efficacy and adverse events of infliximab in comparison to corticosteroids and placebo in active ulcerative colitis. 2011. International Journal of Pharmacology	SLR
97	X. H. Chen, J. Yuan, Y. Huang, C. Liu, T. Mo, C. Li, H. Chen, B. Xu, Q. Hou, Z. He, W. Liu, F. Adalimumab for Moderately to Severely Active Ulcerative Colitis: A Systematic Review and Meta-Analysis. 2016. BioDrugs	SLR
98	K. D. Thorlund, E. Mills, E. J. Fedorak, R. N. Marshall, J. K.. Adalimumab versus infliximab for the treatment of moderate to severe ulcerative colitis in adult patients naive to anti-TNF therapy: An indirect treatment comparison meta-analysis. 2014. Journal of Crohn's and Colitis	Indirect treatment comparison
99	K. D. Thorlund, E. Mills, E. Fedorak, R. Marshal, J.. Adalimumab versus infliximab for the treatment of moderate-to-severe ulcerative colitis in adult patients with no prior anti-TNF experience: An indirect comparison meta-analysis. 2013. American Journal of Gastroenterology	Indirect treatment comparison
100	P. P. Kawalec, A. An indirect comparison of in?iximab versus adalimumab or golimumab for active ulcerative colitis. 2016. Archives of Medical Science	Indirect treatment comparison
101	S. F. Danese, G. Peyrin-Biroulet, L. Lucenteforte, E. Virgili, G. Moja, L. Bonovas, S. Biological agents for moderately to severely active ulcerative colitis. 2014. Annals of Internal Medicine	SLR
102	G. R. D. Lichtenstein, R. H. Wagner, C. L. Fasanmade, A. A. Olson, A. D. Marano, C. W. Johanns, J. Lang, Y. Sandborn, W. J.. Clinical trial: Benefits and risks of immunomodulators and maintenance infliximab for IBD-subgroup analyses across four randomized trials. 2009. Alimentary Pharmacology and Therapeutics	SLR; ACT-1 & ACT-2, ACCENT 1 & 2

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
103	K. A. Thorne, L. Akbari, A. Samuel, D. G. Morrison-Rees, S. Roberts, S. E. Colectomy rates in patients with ulcerative colitis following treatment with infliximab or ciclosporin: A systematic literature review. 2016. European Journal of Gastroenterology and Hepatology	SLR
104	K. D. Thorlund, E. Eapen, S. Mills, E. Comparative efficacy and safety of golimumab, infliximab and adalimumab for the treatment of moderate to severe ulcerative colitis: A bayesian indirect treatment comparison meta-analysis. 2014. Value in Health	Indirect treatment comparison
105	D. T. A. Rubin, A. O. Zhang, Y. Xu, Y. Fahrbach, K. Chen, L. A. Manuchehri, A. Kayhan, C. Woolcott, J. C. Cappelleri, J. C. Healey, P. Comparative efficacy and safety of tofacitinib and biologics as induction therapy for moderately-to-severely active ulcerative colitis: A systematic review and network meta-analysis. 2017. Value in Health	SLR
106	S. G. Singh, S. K. Wang, Z. Murad, M. H. Loftus, E. V. Comparative efficacy of biologic therapy in the management of biologic-naive patients with ulcerative colitis: An indirect treatment comparison meta-analysis. 2014. Gastroenterology	Indirect treatment comparison
107	A. D. M. Vickers, R. Bergman, A. Ling, C. S. Ainsworth, C. Medjedovic, J. Smyth, M. D. Comparative efficacy of biologics in the treatment of moderately to severely active ulcerative colitis (UC): A systematic review and network meta-analysis. 2015 Gastroenterology	CA of SLR
108	K. D. Thorlund, E. Toor, K. Mills, E. J.. Comparative efficacy of golimumab, infliximab, and adalimumab for moderately to severely active ulcerative colitis: A network meta-analysis accounting for differences in trial designs. 2015. Expert Review of Gastroenterology and Hepatology	NMA
109	Z. M. L. Zhang, W. Jiang, X. L.. Efficacy and safety of adalimumab in moderately to severely active cases of ulcerative colitis: a meta-analysis of published placebo-controlled trials. 2016. Gut and Liver	NMA
110	Y. N. Z. Song, P. Efficacy and safety of tumor necrosis factor-alpha blockers for ulcerative colitis: A systematic review and meta-analysis of published randomized controlled trials. 2015. Journal of Food and Drug Analysis	SLR
111	A. C. S. Ford, W. J. Khan, K. J. Hanauer, S. B. Talley, N. J. Moayyedi, P.. Efficacy of biological therapies in inflammatory bowel disease: Systematic review and meta-analysis. 2011 American Journal of Gastroenterology	SLR
112	M. G. Chen, S. Black, C. M. Fan, T. Chaudhary, M. A. Jansen, J. P. Efficacy of infliximab and adalimumab for the treatment of ulcerative colitis-an indirect comparison of RCT evidence. 2013. United European Gastroenterology Journal	CA of SLR
113	A. F. Lopez, A. C. Colombel, J. F. Reinisch, W. Sandborn, W. J. Peyrin-Biroulet, L. Efficacy of tumour necrosis factor antagonists on remission, colectomy and hospitalisations in ulcerative colitis: Meta-analysis of placebo-controlled trials. 2015. Digestive and Liver Disease	SLR

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
114	E. D. S. Shah, C. A. Chong, K. Melmed, G. Y. Evaluating study withdrawal among biologics and immunomodulators in treating ulcerative colitis: A meta-analysis of controlled clinical trials. 2016. Inflammatory Bowel Diseases	SLR
115	M. V.-C. Galvan-Banqueri, M. D. Castillo-Munoz, M. A. Beltran Calvo, C. Molina Lopez, T. Indirect comparison for Anti-TNF drugs in moderate to severe ulcerative colitis. 2015. Farmacia hospitalaria : organo oficial de expresion cientifica de la Sociedad Espanola de Farmacia Hospitalaria	SLR
116	C. L. K. Wheat, C. W. Clark-Snustad, K. Grembowski, D. Thornton, T. A. Devine, B. Inflammatory Bowel Disease (IBD) pharmacotherapy and the risk of serious infection: A systematic review and network meta-analysis. 2017. BMC Gastroenterology	SLR
117	W. Q. H. Mei, H. Z. Liu, Y. Li, Z. C. Wang, W. G. Infliximab is superior to other biological agents for treatment of active ulcerative colitis: A meta-analysis. 2015. World Journal of Gastroenterology	NMA
118	R. T. Archer, P. Ren, S. James, M. M. S. Harvey, R. Basarir, H. Stevens, J. Carroll, C. Cantrel, A. Lobo, A. Hoque, S. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (Including a review of TA140 and TA262): Clinical effectiveness systematic review and economic model. 2016. Health Technology Assessment	SLR
119	Y. L. Jin, Y. Lin, L. J. Zheng, C. Q. Meta-analysis of the effectiveness and safety of vedolizumab for ulcerative colitis. 2015. World Journal of Gastroenterology	NMA
120	M. B. Chen, C. M. Gurunath, S. Jansen, J. P. Chaudhary, M. A. Fan, T. Network meta-analysis of approved biologic interventions for the maintenance of response in ulcerative colitis. 2013. Value in Health	NMA
121	M. C. Z. Wang, L. Y. Han, W. Shao, Y. Chen, M. Ni, R. Wang, G. N. Wei, F. X. Zhang, Y. W. Xu, X. D. Zhang, Y. C. PRISMA - Efficacy and safety of vedolizumab for inflammatory bowel diseases: A systematic review and meta-analysis of randomized controlled trials. 2014. Medicine (United States)	SLR
122	E. D. F. Shah, J. P. Siegel, C. A. Chong, K. Melmed, G. Y..Risk for Overall Infection with Anti-TNF and Anti-integrin Agents Used in IBD: A Systematic Review and Meta-analysis. 2017. Inflammatory Bowel Diseases	SLR
123	P. K. Mocko, P. Pilc, A.. Safety Profile of Biologic Drugs in the Therapy of Ulcerative Colitis: A Systematic Review and Network Meta-Analysis. 2016. Pharmacotherapy	SLR
124	Z. Y. Yang, X. Q. Zhu, Y. Z. Liu, Z. Zou, Y. Deng, Y. Guo, C. C. Garg, S. K. Feng, J. S. Short-term effect and adverse events of adalimumab versus placebo in inducing remission for moderate-to-severe ulcerative colitis: A meta-analysis. 2015. International Journal of Clinical and Experimental Medicine	NMA

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
125	P. M. Kawalec, A. lopuch, S. Systematic review of the effectiveness of biological therapy for active moderate to severe ulcerative colitis. 2014. Journal of Gastroenterology and Hepatology (Australia)	SLR
126	A. H. Cholapranee, G. S. Kaplan, G. G. Peyrin-Biroulet, L. Ananthakrishnan, A. N. Systematic review with meta-analysis: comparative efficacy of biologics for induction and maintenance of mucosal healing in Crohn's disease and ulcerative colitis controlled trials. 2017. Alimentary Pharmacology and Therapeutics	SLR
127	E. J. H. Mao, G. S. Kaplan, G. G. Peyrin-Biroulet, L. Ananthakrishnan, A. N.. Systematic review with meta-analysis: comparative efficacy of immunosuppressants and biologics for reducing hospitalisation and surgery in Crohn's disease and ulcerative colitis. 2017. Alimentary Pharmacology and Therapeutics	SLR
128	D. F. Christophorou, N. Duny, Y. Valats, J. C. Bismuth, M. Pineton De Chambrun, G. Daures, J. P. Blanc, P. Systematic review with meta-analysis: Infliximab and immunosuppressant therapy vs. infliximab alone for active ulcerative colitis. 2015. Alimentary Pharmacology and Therapeutics	SLR
129	C. J. P.-B. Williams, L. Ford, A. C.. Systematic review with meta-analysis: malignancies with anti-tumour necrosis factor-alpha therapy in inflammatory bowel disease. 2014. Alimentary Pharmacology & Therapeutics	SLR
130	A. D. A. Vickers, C. Mody, R. Bergman, A. Ling, C. S. Medjedovic, J. Smyth, M. Systematic review with network meta-analysis: Comparative efficacy of biologics in the treatment of moderately to severely active ulcerative colitis. 2016. PLoS ONE	SLR
131	R. W. L. Stidham, T. C. Higgins, P. D. Deshpande, A. R. Sussman, D. A. Singal, A. G. Elmunzer, B. J. Saini, S. D. Vijan, S. Waljee, A. K. Systematic review with network meta-analysis: the efficacy of anti-tumour necrosis factor-alpha agents for the treatment of ulcerative colitis. 2014. Alimentary Pharmacology & Therapeutics	SLR
132	J. P. G.-L. Gisbert, Y. Mate, J..Systematic review: Infliximab therapy in ulcerative colitis. 2007. Alimentary Pharmacology and Therapeutics	SLR
133	K. M. M. LeBlanc, H. Parker Claire, E. MacDonald John, K. The impact of biological interventions for ulcerative colitis on health-related quality of life. 2015. Cochrane Database of Systematic Reviews	SLR
134	N. B. Shahidi, B. Panaccione, R. The role of vedolizumab in patients with moderate-to-severe Crohn's disease and ulcerative colitis. 2015. Therapeutic Advances in Gastroenterology	SLR
135	M. M. T. Lawson, A. G. Akobeng, A. K. Tumour necrosis factor alpha blocking agents for induction of remission in ulcerative colitis. 2006. Cochrane database of systematic reviews (Online)	SLR

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

	Reference	Reason for exclusion
136	J. B. B. Bickston Stephen, W. Tsoulis David, J. Cheng, Jianfeng MacDonald John, K. Khanna, Reena Feagan Brian, G. Vedolizumab for induction and maintenance of remission in ulcerative colitis. 2014 Cochrane Database of Systematic Reviews	SLR
137	M. H. M. Mosli, J. K. Bickston, S. J. Behm, B. W. Tsoulis, D. J. Cheng, J. Khanna, R. Feagan, B. G. Vedolizumab for induction and maintenance of remission in ulcerative colitis: A cochrane systematic review and meta-analysis. 2015. Inflammatory Bowel Diseases	SLR

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Appendix C – List of included studies

Linked to ERG Clarification Question A6

Ref. count	Ref number in table 84 of CS	Reference
1	22	Colombel JFR, P. Reinisch, W. Esser, D. Wang, Y. Lang, Y. Marano, C. W. Strauss, R. Oddens, B. J. Feagan, B. G. Hanauer, S. B. Lichtenstein, G. R. Present, D. Sands, B. E. Sandborn, W. J. Early mucosal healing with infliximab is associated with improved long-term clinical outcomes in ulcerative colitis. Gastroenterology. 2011;141(4):1194-201.
2	73	Reinisch W, Sandborn WJ, Hommes DW, D'Haens G, Hanauer S, Schreiber S, et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. Gut. 2011;60(6):780-7
3	74	Rutgeerts P, Sandborn WJ, Feagan BG, Reinisch W, Olson A, Johanns J, et al. Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis. New England Journal of Medicine. 2005;353(23):2462-76
4	75	Sandborn WJ, Feagan BG, Marano C, Zhang H, Strauss R, Johanns J, et al. Subcutaneous golimumab induces clinical response and remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2014;146(1):85-95; quiz e14-5
5	83	Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel J-F, Sandborn WJ, et al. Vedolizumab as Induction and Maintenance Therapy for Ulcerative Colitis. New England Journal of Medicine. 2013;369(8):699-710
6	96	Panaccione RG, S. Middleton, S. Márquez, J. R. Scott, B. B. Flint, L. Hoogstraten, H. J. Chen, A. C. Zheng, H. Danese, S. Rutgeerts, P. Combination therapy with infliximab and azathioprine is superior to monotherapy with either agent in ulcerative colitis. Gastroenterology [Internet]. 2014; 146(2):[392-400.e3 pp.]. Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/777/CN-00958777/frame.html http://www.gastrojournal.org/article/S0016-5085(13)01526-6/pdf.
7	98	Sandborn WJ, Su C, Sands BE, D'Haens GR, Vermeire S, Schreiber S, et al. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2017;376(18):1723-36.
8	99	Sandborn WJ, Ghosh S, Panes J, Vranic I, Su C, Rousell S, et al. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. N Engl J Med. 2012;367(7):616-24
9	109	Sandborn WJ, van Assche G, Reinisch W, Colombel JF, D'Haens G, Wolf DC, et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2012;142(2):257-65 e1-3
10	110	Hibi T, Imai Y, Senoo A, Ohta K, Ukyo Y. Efficacy and safety of golimumab 52-week maintenance therapy in Japanese patients with moderate to severely active ulcerative colitis: a phase 3, double-blind, randomized, placebo-controlled study-(PURSUIT-J study). Journal of gastroenterology. 2017;52(10):1101-11
11	111	Jiang XL, Cui HF, Gao J, Fan H. Low-dose Infliximab for Induction and Maintenance Treatment in Chinese Patients With Moderate to Severe Active Ulcerative Colitis. J Clin Gastroenterol. 2015;49(7):582-8.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Ref. count	Ref number in table 84 of CS	Reference
12	113	Mshimesh BAR. Efficacy and safety of adalimumab versus infliximab in patients suffered from moderate to severe active ulcerative colitis. Asian Journal of Pharmaceutical and Clinical Research. 2017;10(3):300-7.
13	114	Suzuki Y, Motoya S, Hanai H, Matsumoto T, Hibi T, Robinson AM, et al. Efficacy and safety of adalimumab in Japanese patients with moderately to severely active ulcerative colitis. Journal of gastroenterology. 2014;49(2):283-94
14	194	Sandborn WJR, P. Feagan, B. G. Reinisch, W. Olson, A. Johanns, J. Lu, J. Horgan, K. Rachmilewitz, D. Hanauer, S. B. Lichtenstein, G. R. de Villiers, W. J. S. Present, D. Sands, B. E. Colombel, J. F. Colectomy Rate Comparison After Treatment of Ulcerative Colitis With Placebo or Infliximab. Gastroenterology. 2009;137(4):1250-60.
15	195	Feagan BGR, W. Rutgeerts, P. Sandborn, W. J. Yan, S. Eisenberg, D. Bala, M. Johanns, J. Olson, A. Hanauer, S. B. The effects of infliximab therapy on health-related quality of life in ulcerative colitis patients. American Journal of Gastroenterology. 2007;102(4):794-802
16	196	Armuzzi ADP, B. Lupascu, A. Fedeli, P. Leo, D. Mentella, M. C. Vincenti, F. Melina, D. Gasbarrini, G. Pola, P. Gasbarinni, A. Infliximab in the treatment of steroid-dependent ulcerative colitis. European Review for Medical and Pharmacological Sciences. 2004;8(5):231-3.
17	197	Feagan BG GG, Wild G, et al. . Treatment of ulcerative colitis with a humanized antibody to the alpha4beta7 integrin. New England Journal of Medicine. 2005;352(24):2499-507
18	198	Loftus ES, C. Panaccione, R. Sandborn, W. Smyth, M. James, A. Xu, J. Abhyankar, B. Corticosteroid dose reduction in ulcerative colitis patients treated with vedolizumab. Gut. 2016;65:A254-A5
19	199	Loftus Jr EVS, C. A. Panaccione, R. Sandborn, W. J. Smyth, M. Green, A. Xu, J. Abhyankar, B. Corticosteroid dose reduction in ulcerative colitis patients treated with vedolizumab. United European Gastroenterology Journal. 2015;1):A255-A6
20	200	Sandborn WC, J. F. Panaccione, R. Lasch, K. Sankoh, S. Abhyankar, B. Deep clinical remission in patients with Ulcerative Colitis: Evaluating the effects of vedolizumab on various combinations of endoscopic and patient-reported outcomes. Journal of Crohn's and Colitis. 2015;9:S237-S8
21	201	Sandborn WC, J. F. Panaccione, R. Lasch, K. Mody, R. Green, A. Abhyankar, B. Deep remission as a predictor of clinical outcomes in vedolizumab-treated patients with ulcerative colitis. Journal of Crohn's and Colitis. 2015;9:S299-S300.
22	202	Sandborn WC, J. F. Panaccione, R. Lasch, K. Sankoh, S. Abhyankar, B. Deep remission with vedolizumab in patients with ulcerative colitis: Evaluating various combinations of endoscopic and patient-reported outcomes. American Journal of Gastroenterology. 2014;109:S483
23	203	Feagan BGP, H. Colombel, J. F. Rubin, D. T. James, A. Mody, R. Lasch, K. Effects of vedolizumab on health-related quality of life in patients with ulcerative colitis: results from the randomised GEMINI 1 trial. Alimentary Pharmacology and Therapeutics. 2017;45(2):264-75
24	204	Sands BES, I. Farraye, F. A. Cheifetz, A. S. Abhyankar, B. Sankoh, S. Smyth, M. Efficacy and safety of retreatment with vedolizumab in patients with ulcerative colitis. Journal of Crohn's and Colitis. 2015;9:S37

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Ref. count	Ref number in table 84 of CS	Reference	
25	205	Yajnik VK, N. Dubinsky, M. Axler, J. James, A. Abhyankar, B. Lasch, K. Efficacy and Safety of Vedolizumab in Ulcerative Colitis and Crohn's Disease Patients Stratified by Age. Advances in Therapy. 2017;34(2):542-59	
26	206	Yajnik VK, N. Dubinsky, M. Axler, J. Green, A. Abhyankar, B. Lasch, K. Efficacy and safety of vedolizumab with advancing age in patients with ulcerative colitis: Results from the GEMINI 1 study. Journal of Crohn's and Colitis. 2015;9:S363-S4	
27	207	Feagan BS, B. Sankoh, S. Milch, C. Fox, I. Efficacy of vedolizumab in ulcerative colitis by prior treatment failure in gemini i, a randomized, placebo-controlled, double-blind, multicenter trial. Inflammatory bowel diseases. 2012;18:S1-S2.	
28	208	Feagan BGR, D. T. Danese, S. Vermeire, S. Abhyankar, B. Sankoh, S. James, A. Smyth, M. Efficacy of Vedolizumab Induction and Maintenance Therapy in Patients With Ulcerative Colitis, Regardless of Prior Exposure to Tumor Necrosis Factor Antagonists. Clinical Gastroenterology and Hepatology. 2017;15(2):229-39.e5	
29	209	Feagan BS, C. Melmed, G. Isaacs, K. Lasch, K. Rosario, M. James, A. Abhyankar, B. Efficacy of vedolizumab with and without continued immunosuppressant use in gemini 1 and gemini 2. Gut. 2016;65:A244-A5.	
30	210	Colombel JFLJ, E. V. Siegel, C. A. Lewis, J. D. Smyth, M. Sankoh, S. Abhyankar, B. Efficacy of vedolizumab with concomitant corticosteroid or immunomodulator use in patients with ulcerative colitis from GEMINI 1. Journal of Crohn's and Colitis. 2015;9:S296-S7.	
31	211	Cited incorrectly in the CS: Feagan BGS, W. Smyth, M. D. Sankoh, S. Parikh, A. Fox, I. Effects of continued vedolizumab therapy for ulcerative colitis in week 6 induction therapy nonresponders. Gastroenterology. 2014;1:S-590	Should have been cited in the CS: Brian G. Feagan, Jean-Frederic Colombel, David T. Rubin, Reema Mody, Serap Sankoh,Karen Lasch. Health-Related Quality of Life in Patients With Ulcerative Colitis After Treatment With Vedolizumab: Results From the Gemini 1 Study. Gastroenterology. 2014; 1: S-590
32	212	Feagan BC, J. F. Rubin, D. Mody, R. Sankoh, S. Lasch, K. Improvements in health-related quality of life in patients with ulcerative colitis treated with vedolizumab. Journal of Crohn's and Colitis. 2014;8:S51-S2	
33	213	Feagan BG RP, Sands BE, et al; GEMINI 1 Study Group,. Induction and maintenance therapy with vedolizumab, a novel biologic therapy for ulcerative colitis. Gastroenterology and Hepatology. 2014;10(1):64-6	
34	214	Feagan BGR, P. J. Sands, B. E. Colombel, J. Sandborn, W. J. Colombel, J. Hanauer, S. B. Van Assche, G. A. Axler, J. Kim, H. Danese, S. Fox, I. Milch, C. Sankoh, S. Wyant, T. Xu, J. Parikh, A. Induction therapy for ulcerative colitis: Results of gemini i, a randomized, placebo-controlled, double-blind, multicenter phase 3 trial. Gastroenterology. 2012;1):S160	
35	215	Sands BEC, R. Isaacs, K. Fedorak, R. N. Abhyankar, B. Sankoh, S. Smyth, M. Infusion-related reactions with vedolizumab treatment in patients with UC or CD during the GEMINI 1 and GEMINI 2 clinical trials. Journal of Crohn's and Colitis. 2015;9:S392-S3	
36	216	Colombel JFS, B. E. Feagan, B. G. Loftus, E. V. Sankoh, S. Fox, I. Parikh, A. Milch, C. Integrated safety analysis of vedolizumab for the treatment of ulcerative colitis or crohn's disease. Gastroenterology. 2013;1):S113	
37	217	Sands BH, S. Colombel, J. F. Danese, S. Abreu, M. Ahuja, V. Ponich, T. Hilmi, I. Sankoh, S. Smyth, M. Abhyankar, B. Fox, I. Feagan, B. Reductions in corticosteroid use in patients with ulcerative colitis or crohn's disease treated with vedolizumab.	

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Ref. count	Ref number in table 84 of CS	Reference
		American Journal of Gastroenterology. 2013;108:S503.
38	218	Colombel JFS, C. A. Abhyankar, B. Loftus Jr, E. V. Lewis, J. D. Sankoh, S. Smyth, M. Milch, C. Safety of vedolizumab alone or with concomitant corticosteroids and/or immunosuppressants in patients with ulcerative colitis or crohn's disease. United European Gastroenterology Journal. 2014;1):A82
39	219	Bokemeyer BS, A. Axler, J. L. Curtis, R. I. Ehehalt, R. Schreiber, S. Geransar, P. James, A. Kaviya, A. Khalid, J. M. Wolf, D. C. Feagan, B. G. Sustained remission with vedolizumab in patients with moderately to severely active ulcerative colitis: A GEMINI 1 post hoc analysis of week 14 remitters. Gastroenterology. 2017;152 (5 Supplement 1):S604
40	220	Sandborn WS, B. Rutgeerts, P. Sankoh, S. Rosario, M. Milch, C. Fox, I. Sustained therapeutic benefit of vedolizumab throughout 1 year in ulcerative colitis in GEMINI I, a randomized, placebo-controlled, double-blind, multicenter trial. Journal of Crohn's and Colitis. 2013;7:S138-S9
41	221	D'Haens GRC, J. F. Dubinsky, M. Abhyankar, B. James, A. Lasch, K. The efficacy of vedolizumab by disease activity and prior tumor necrosis factor a antagonist failure in patients with ulcerative colitis or Crohn's disease: Post HOC analyses from the gemini 1 and gemini 2 studies. Gastroenterology. 2016;1):S804-S5.
42	222	D'Haens GC, J. F. Dubinsky, M. Abhyankar, B. James, A. Lasch, K. The efficacy of vedolizumab by disease activity and prior tumour necrosis factor-alpha antagonist failure in patients with ulcerative colitis or Crohn's disease: Post-hoc analyses from the GEMINI 1 and GEMINI 2 studies. Journal of Crohn's and Colitis. 2016;10:S58.
43	223	Lam MCWB, B. Vedolizumab for ulcerative colitis and Crohn's disease: Results and implications of GEMINI studies. Immunotherapy. 2014;6(9):963-71.
44	224	Feagan BR, P. Sands, B. Sandborn, W. Colombel, J. F. Hanauer, S. Van Assche, G. Axler, J. Kim, H. J. Danese, S. Fox, I. Milch, C. Sankoh, S. Wyant, T. Xu, J. Parikh, A. Vedolizumab maintenance therapy for ulcerative colitis: Results of gemini I, a randomized, placebo-controlled, double-blind, multicenter phase 3 trial. American Journal of Gastroenterology. 2012;107:S609-S10
45	225	Kobayashi TS, Y. Motoya, S. Hirai, F. Ogata, H. Ito, H. Sato, N. Ozaki, K. Watanabe, M. Hibi, T. First trough level of infliximab at week 2 predicts future outcomes of induction therapy in ulcerative colitis: A post-hoc analysis of a multicenter prospective randomized controlled trial. American Journal of Gastroenterology. 2015;110:S787
46	226	Suzuki YM, S. Hirai, F. Ogata, H. Ito, H. Sato, N. Ozaki, K. Watanabe, M. Hibi, T. Infliximab therapy for Japanese patients with ulcerative colitis: Efficacy, safety, and association between serum infliximab levels and early response in a randomized, double-blind, placebo-controlled study. Journal of Crohn's and Colitis. 2015;9:S372-S3
47	227	Sandborn WS, B. E. D'Haens, G. R. Vermeire, S. Schreiber, S. Danese, S. Panes, J. Feagan, B. G. Reinisch, W. Niezychowski, W. Friedman, G. Lawendy, N. Yu, D. Woodworth, D. A. Mukherjee, A. Healey, P. J. Zhang, H. Su, C. Efficacy and safety of oral tofacitinib as induction therapy in patients with moderate to severe ulcerative colitis: Results from two phase 3

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Ref. count	Ref number in table 84 of CS	Reference
		randomized controlled trials. Gastroenterology. 2016;1):S157
48	228	Panes JS, C. Marren, A. Yu, D. Woodworth, D. Zhang, H. Healey, P. Improvement in patient-reported outcomes in 2 Phase 3 studies of tofacitinib in patients with moderately to severely active ulcerative colitis. Journal of Crohn's and Colitis. 2016;10:S283-S4
49	229	Danese SS, W. J. Panes, J. Zhang, H. Woodworth, D. Marren, A. Su, C. Onset of efficacy of tofacitinib for induction therapy in patients with active ulcerative colitis in two multinational, phase 3 clinical trials. Digestive and Liver Disease. 2017;49 (Supplement 2):e90-e1
50	230	Sandborn WJD, S. Panes, J. Zhang, H. Woodworth, D. Marren, A. Su, C. Onset of efficacy of tofacitinib for induction therapy in patients with active ulcerative colitis in two multinational, phase 3 clinical trials. American Journal of Gastroenterology. 2016;111:S260-S1
51	231	Feagan BV, S. Sandborn, W. J. Reinisch, W. Tarabar, D. Su, C. Niezychowski, W. Zhang, H. Woodworth, D. Yu, D. Sands, B. Tofacitinib for induction therapy in patients with active ulcerative colitis in two phase 3 clinical trials: Results by local and central endoscopic assessments. American Journal of Gastroenterology. 2016;111:S319-S20
52	232	Sands BDH, G. Sandborn, W. J. Hibi, T. Su, C. Niezychowski, W. Ghosh, S. Zhang, H. Yu, D. Woodworth, D. Healey, P. Marren, A. Panes, J. Tofacitinib has induction efficacy in moderately to severely active ulcerative colitis, regardless of prior TNF inhibitor therapy. American Journal of Gastroenterology. 2016;111:S261
53	233	Sandborn WJS, B. E. Danese, S. D'Haens, G. R. Vermeire, S. Schreiber, S. Feagan, B. G. Reinisch, W. Friedman, G. Woodworth, D. A. Zhang, H. Lawendy, N. Niezychowski, W. Su, C. Panes, J. Efficacy and safety of oral tofacitinib as maintenance therapy in patients with moderate to severe ulcerative colitis: Results from a phase 3 randomized controlled trial. Gastroenterology. 2017;152 (5 Supplement 1):S199
54	234	Panes JR, D. T. Vermeire, S. Lindsay, J. O. Sands, B. E. Su, C. Friedman, G. Zhang, H. Kayhan, C. Manuchehri, A. Healey, P. J. Maintenance of quality of life improvement in a phase 3 study of tofacitinib for patients with moderately to severely active ulcerative colitis. Gastroenterology. 2017;152 (5 Supplement 1):S601-S2
55	235	Probert CSH, S. D. Schreiber, S. Kuhbacher, T. Ghosh, S. Arnott, I. D. Forbes, A. Infliximab in moderately severe glucocorticoid resistant ulcerative colitis: a randomised controlled trial. Gut. 2003;52(7):998-1002.
56	236	Sandborn WJF, B. G. Marano, C. W. Strauss, R. Johanns, J. Zhang, H. Colombel, J. Reinisch, W. Gibson, P. Collins, J. Jarnerot, G. A. Rutgeerts, P. J. A phase 2/3 randomized, placebo-controlled, double-blind study to evaluate the safety and efficacy of subcutaneous golimumab induction therapy in patients with moderately to severely active ulcerative colitis: Pursuit sc. Gastroenterology. 2012;1):S161.
57	237	Sandborn WF, B. Marano, C. Strauss, R. Johanns, J. Zhang, H. Colombel, J. F. Reinisch, W. Gibson, P. Collins, J. Tarabar, D. Hebzda, Z. Rutgeerts, P. Assessment of fecal markers and clinical outcomes in patients with moderately to severely active ulcerative colitis: Results from pursuit-SC induction. American Journal of Gastroenterology. 2014;109:S502-S3
58	238	Feagan BG, P. Marano, C. Strauss, R. Han, C. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Collins, J. Jarnerot, G. Rutgeerts, P. Sandborn, W. Impact of golimumab sc on disease specific and generic hrqol in patients with moderately

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Ref. count	Ref number in table 84 of CS	Reference
		to severely active UC: Pursuit-SC induction. Inflammatory bowel diseases. 2012;18:S43
59	239	Feagan BG, P. Marano, C. Strauss, R. Han, C. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Collins, J. Jarnerot, G. Rutgeerts, P. Sandborn, W. Impact of subcutaneously administered golimumab on disease specific and generic health-related quality of life in patients with moderately to severely active ulcerative colitis: Results from PURSUIT-SC induction. Journal of Crohn's and Colitis. 2013;7:S99-S100
60	240	Rutgeerts PF, B. Marano, C. Strauss, R. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Gibson, P. R. Collins, J. Jarnerot, G. Sandborn, W. J. Phase 2/3 randomized, placebo-controlled, double-blind study of SC golimumab induction in moderate to Severe UC. Journal of Gastroenterology and Hepatology. 2013;28:591.
61	241	Sandborn WF, B. Colombel, J. F. Reinisch, W. Gibson, P. Rutgeerts, P. Weng, H. Yao, R. Marano, C. Zhang, H. Strauss, R. Relationship between clinical outcomes and disease duration, extent, and severity in patients with ulcerative colitis who received 6 weeks of treatment with golimumab. United European Gastroenterology Journal. 2014;1):A538-A9.
62	242	Anonymous. Safety and Efficacy of Subcutaneous Golimumab Induction Therapy in Patients with Moderately to Severely Active Ulcerative Colitis. Clinical Advances in Hematology and Oncology. 2012;10 (10):25
63	243	Gibson PF, B. Marano, C. Strauss, R. Han, C. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Collins, J. Jarnerot, G. Rutgeerts, P. Sandborn, W. Subcutaneously administered golimumab induction therapy improves health states measured by EQ-5D in patients with moderately to severely active ulcerative colitis: Results from PURSUIT-SC. Journal of Crohn's and Colitis. 2013;7:S65
64	244	Sandborn W, Colombel JF, Feagan B, Marano C, Strauss R, Han C, et al. Early and sustained remission after treatment with subcutaneously administered golimumab is associated with normalized health-related quality of life in patients with moderate to severe ulcerative colitis: Post-HOC analysis from pursuit induction and maintenance trials. American Journal of Gastroenterology. 2013;108:S519.
65	245	Sandborn WR, W. Feagan, B. Marano, C. Strauss, R. Han, C. Johanns, J. Zhang, H. Gibson, P. Collins, J. Jarnerot, G. Rutgeerts, P. Colombel, J. F. Early responses in patient-reported outcomes are predictors for mucosal healing in patients with moderately to severely active ulcerative colitis treated with golimumab: Results from pursuit SC induction. American Journal of Gastroenterology. 2013;108:S519
66	246	Rutgeerts PF, B. G. Marano, C. W. Padgett, L. Strauss, R. Johanns, J. Adedokun, O. J. Guzzo, C. Zhang, H. Colombel, J. F. Reinisch, W. Gibson, P. R. Sandborn, W. J. Randomised clinical trial: A placebo-controlled study of intravenous golimumab induction therapy for ulcerative colitis. Alimentary Pharmacology and Therapeutics. 2015;42(5):504-14
67	247	Sandborn WJF, B. G. Marano, C. Zhang, H. Strauss, R. Johanns, J. Adedokun, O. J. Guzzo, C. Colombel, J. Reinisch, W. Gibson, P. R. Collins, J. Jarnerot, G. Rutgeerts, P. Subcutaneous Golimumab Maintains Clinical Response in Patients With Moderate-to-Severe Ulcerative Colitis. Gastroenterology
68	248	Sandborn WF, B. Marano, C. Strauss, R. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Gibson, P.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Ref. count	Ref number in table 84 of CS	Reference
		Collins, J. Jarnerot, G. Rutgeerts, P. A phase 3 randomized, placebo-controlled, double-blind study to evaluate the safety and efficacy of subcutaneous golimumab maintenance therapy in patients with moderately to severely active ulcerative colitis: Pursuit-maintenance. American Journal of Gastroenterology. 2012;107:S673-S4
69	249	Rutgeerts PF, B. G. Marano, C. Strauss, R. Johanns, J. Zhang, H. Guzzo, C. Colombel, J. F. Reinisch, W. Gibson, P. Collins, J. Jarnerot, G. Sandborn, W. Phase 3 study to evaluate the SC golimumab maintenance therapy in moderate to severe UC: Pursuit-maintenance. Journal of Gastroenterology and Hepatology. 2013;28:593.
70	250	Rutgeerts PR, W. Feagan, B. Sandborn, W. Tarabar, D. Habzda, Z. Weng, H. Yao, R. Zhang, H. Adedokun, O. Marano, C. Strauss, R. Weight-based maintenance dosing of golimumab in patients with moderate to severe ulcerative colitis. United European Gastroenterology Journal. 2014;1):A529-A30
71	251	Panes JS, C. Bushmakin, A. G. Cappelleri, J. C. Mamolo, C. Healey, P. Tofacitinib in active ulcerative colitis: Analysis of efficacy based on patient-reported outcomes. Journal of Crohn's and Colitis. 2014;8:S51
72	252	Panes JS, C. Bushmakin, A. G. Cappelleri, J. C. Mamolo, C. Healey, P. Randomized trial of tofacitinib in active ulcerative colitis: Analysis of efficacy based on patientreported outcomes. BMC Gastroenterology. 2015;15 (1) (no pagination)(9).
73	253	Panes JS, W. J. Zhang, H. Yu, D. Niezychowski, W. Su, C. Evaluation of the relationship between fecal calprotectin concentrations and clinical and endoscopic outcome measures in a phase 2 study of tofacitinib, an oral janus kinase inhibitor, in active ulcerative colitis. Journal of Crohn's and Colitis. 2013;7:S107S-8.
74	254	Panes JS, C. Bushmakin, A. Cappelleri, J. Healey, P. Direct and indirect effects of tofacitinib on treatment satisfaction in patients with ulcerative colitis. Journal of Crohn's and Colitis. 2015;9:S380-S1.
75	255	Panaccione RG, S. Middleton, S. Marquez, J. Flint, L. J. van Hoogstraten H Chen, A. Zheng, H. Danese, S. Rutgeerts, P. Improvement in patient quality of life during treatment with infliximab, azathioprine, or combination infliximab+azathioprine for moderate-to-severe ulcerative colitis. American Journal of Gastroenterology. 2013;108:S544-S5.
76	256	Panaccione RG, S. Middleton, S. Velazquez, J. R. M. Khalif, I. Flint, L. Van Hoogstraten, H. J. Zheng, H. Danese, S. Rutgeerts, P. Infliximab, azathioprine, or infliximab + azathioprine for treatment of moderate to severe ulcerative colitis: The UC SUCCESS trial. Journal of Crohn's and Colitis. 2011;5 (1):S8.
77	257	Feagan BS, W. Lazar, A. Thakkar, R. Skup, M. Yang, M. Chao, J. Mulani, P. Adalimumab induction dose reduces the risk of hospitalizations and colectomies in patients with ulcerative colitis during the first 8 weeks of therapy. American Journal of Gastroenterology. 2012;107:S647
78	258	Feagan BGS, W. J. Lazar, A. Thakkar, R. B. Huang, B. Reilly, N. Chen, N. Yang, M. Skup, M. Mulani, P. Chao, J. Adalimumab therapy is associated with reduced risk of hospitalization in patients with ulcerative colitis. Gastroenterology. 2014;146(1):110-8.e3.
79	259	Feagan BS, W. Yang, M. Lomax, K. Lazar, A. Thakkar, R. Mulani, P. Chao, J. Adalimumab therapy reduces hospitalization and colectomy rates in patients with ulcerative colitis: Data from controlled trials. American Journal of Gastroenterology. 2011;106:S452

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Ref. count	Ref number in table 84 of CS	Reference
80	260	Feagan BS, M. Ozbay, A. B. Lazar, A. Thakkar, R. Chen, N. Chao, J. Sandborn, W. Adalimumab therapy reduces hospitalization rates in patients with moderate ulcerative colitis. American Journal of Gastroenterology. 2014;109:S480.
81	261	Sandborn WF, B. Yang, M. Skup, M. Thakkar, R. Lazar, A. Mulani, P. Chao, J. Adalimumab therapy reduces hospitalization rates in patients with ulcerative colitis among initial responders. Inflammatory bowel diseases. 2012;18:S40.
82	262	Panaccione RC, J. F. Sandborn, W. Reinisch, W. D'Haens, G. R. Hommes, D. W. Ghosh, S. Pappalardo, B. Petersson, J. Robinson, A. M. Lazar, A. Zhou, Q. Thakkar, R. Durable clinical remission and response in adalimumab-treated patients with ulcerative colitis. Gastroenterology. 2015;1):S264
83	263	Colombel JFS, W. J. Ghosh, S. Wolf, D. C. Panaccione, R. Feagan, B. Reinisch, W. Robinson, A. M. Lazar, A. Kron, M. Huang, B. Skup, M. Thakkar, R. B. Four-year maintenance treatment with adalimumab in patients with moderately to severely active ulcerative Colitis: Data from ULTRA 1, 2, and 3. American Journal of Gastroenterology. 2014;109(11):1771-80
84	264	Sandborn WJVA, G. Reinisch, W. Colombel, J. D'Haens, G. Wolf, D. C. Kron, M. Tighe, M. Lazar, A. Thakkar, R. B. Induction and maintenance of clinical remission by adalimumab in patients with moderate-to-severe ulcerative colitis. Journal of Crohn's and Colitis. 2011;5 (1):S68.
85	265	Sandborn WVA, G. Reinisch, W. Colombel, J. F. D'Haens, G. Wolf, D. Kron, M. Tighe, M. Lazar, A. Thakkar, R. Induction and maintenance of clinical remission by adalimumab in patientswith moderate-to-severe Ulcerative Colitis. Inflammatory bowel diseases. 2011;17:S3-S4
86	266	Hanauer SS, W. J. Colombel, J. F. Reinisch, W. Vermeire, S. Petersson, J. Pappalardo, B. Robinson, A. M. Lazar, A. Zhou, Q. Thakkar, R. B. Rapid changes in laboratory parameters and early response to adalimumab: Pooled analysis from ULTRA 1 and ULTRA 2. Journal of Crohn's and Colitis. 2015;9:S289-S90
87	267	Sandborn WJW, D. C. Van Assche, G. Yang, M. Mulani, P. M. Reilly, N. Lazar, A. Thakkar, R. B. Chao, J. Rapid onset of adalimumab and long-term efficacy among week-8responders in adults with moderate to severe active Ulcerative Colitis. Inflammatory bowel diseases. 2011;17:S4
88	268	Lewis JDG, S. Sandborn, W. Van Assche, G. A. D'Haens, G. R. Lazar, A. Eichner, S. Huang, B. Robinson, A. Thakkar, R. Rates of "patient-defined" remission with adalimumab in patients with ulcerative colitis: Subanalysis of ultra 1 and ultra 2. Gastroenterology. 2013;1):S232
89	269	Colombel JFG, S. Sandborn, W. Van Assche, G. Reinisch, W. Lazar, A. Eichner, S. Huang, B. Robinson, A. Thakkar, R. Safety of adalimumab in global clinical trials of ulcerative colitis patients. Inflammatory bowel diseases. 2013;19:S78-S9
90	270	Reinisch WS, W. J. Hommes, D. W. D'Haens, G. R. Hanauer, S. B. Schreiber, S. Panaccione, R. Fedorak, R. N. Tighe, M. B. Huang, B. Lazar, A. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis. Gastroenterology. 2010;1):S114-S5
91	271	Reinisch WS, W. Thakkar, R. Lazar, A. Huang, B. Mulani, P. Chao, J. Adalimumab induction therapy improves health-related quality of life in patients with moderately to severely active ulcerative colitis. American Journal of Gastroenterology. 2010;105:S441-S2.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Ref. count	Ref number in table 84 of CS	Reference
92	272	Feagan BS, W. Yang, M. Skup, M. Thakkar, R. Lazar, A. Mulani, P. Chao, J. Adalimumab therapy reduces hospitalization and colectomy rates in patients with ulcerative colitis among initial responders. American Journal of Gastroenterology. 2012;107:S642-S3
93	273	Sandborn WJVA, G. Thakkar, R. B. Lazar, A. Kron, M. Yang, M. Chao, J. Mulani, P. M. Adalimumab improves health-related quality of life for 52 weeks in patients with ulcerative colitis. Journal of Crohn's and Colitis. 2011;5 (1):S87-S8.
94	274	Sandborn WT, R. Lazar, A. Skup, M. Yang, M. Chao, J. Mulani, P. Van Assche, G. Time-in-remission analysis of adalimumab vs. placebo for the treatment of ulcerative colitis. American Journal of Gastroenterology. 2012;107:S641-S2.
95	275	Colombel JFS, W. Van Assche, G. D'Haens, G. Wolf, D. Kron, M. Lazar, A. Robinson, A. Chao, J. Thakkar, R. Benefit-risk assessment of adalimumab as maintenance treatment for ulcerative colitis: Near analysis of week 8 responders in ultra 2. American Journal of Gastroenterology. 2012;107:S643-S4
96	276	D'Haens GVA, G. Wolf, D. Sandborn, W. Colombel, J. F. Lazar, A. Kron, M. Robinson, A. Thakkar, R. Mucosal healing in ulcerative colitis patients with week 8 response to adalimumab: Subanalysis of ultra 2. Inflammatory bowel diseases. 2012;18:S21
97	277	Sandborn WJC, J. F. D'Haens, G. Van Assche, G. Wolf, D. Kron, M. Lazar, A. Robinson, A. M. Yang, M. Chao, J. D. Thakkar, R. One-year maintenance outcomes among patients with moderately-to-severely active ulcerative colitis who responded to induction therapy with adalimumab: Subgroup analyses from ULTRA 2. Alimentary Pharmacology and Therapeutics. 2013;37(2):204-13
98	278	Sandborn WJDH, G. R. Colombel, J. F. Van Assche, G. A. Wolf, D. C. Kron, M. Lazar, A. Robinson, A. Chao, J. Thakkar, R. One-year response and remission rates in ulcerative colitis patients with week 8 response to adalimumab: subanalysis of ultra 2. Gastroenterology. 2012;1):S565
99	279	Van Assche GW, D. D'Haens, G. Sandborn, W. Colombel, J. F. Lazar, A. Kron, M. Robinson, A. Thakkar, R. Reduced steroid usage in ulcerative colitis patients with week 8 response to adalimumab: Subanalysis of ultra 2. Inflammatory bowel diseases. 2012;18:S21-S2
100	280	Ghosh SW, D. Sandborn, W. Colombel, J. F. Lazar, A. Eichner, S. Robinson, A. Thakkar, R. Sustained efficacy in patients with ulcerative colitis treated with adalimumab: Results from ULTRA 2. Journal of Crohn's and Colitis. 2013;7:S238
101	281	Colombel JFP, S. Sandborn, W. D'Haens, G. Robinson, A. Petersson, J. Huang, B. Lazar, A. Thakkar, R. Time to remission and response in adalimumabtreated patients with moderately to severely active ulcerative colitis from ultra 2. United European Gastroenterology Journal. 2013;1):A219-A20
102	282	Sandborn WDH, G. Colombel, J. F. Van Assche, G. Wolf, D. Kron, M. Lazar, A. Robinson, A. Chao, J. Thakkar, R. One-year response and remission rates in ulcerative colitis patients with week 8 response to adalimumab: Subanalysis of ULTRA 2. Journal of Crohn's and Colitis. 2012;6:S92.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Appendix D – Tofacitinib Phase II study summary

A summary of the Phase II study (published as Sandborn *et al*. *NEJM*. 2012;367(7):616-24) is provided below;

Study design

The phase II trial was a randomised, double-blind, phase II trial of tofacitinib versus placebo for adults with moderately to severely active ulcerative colitis. Eligible patients were those aged at least 18 years who had a confirmed diagnosis of ulcerative colitis for at least 3 months, a Mayo score of ≥ 6 and moderately or severely active disease on sigmoidoscopy (i.e., a Mayo endoscopic subscore of 2 or 3); these eligibility criteria are similar to those in the Phase III trials.

Patients (N = 194) were randomly assigned in a 2:2:2:3:3 ratio to receive tofacitinib at a dose of 0.5 mg (n = 31), 3 mg (n = 33), 10 mg (n = 33), or 15 mg (N = 49) twice daily, or placebo (n = 48). Randomisation was performed centrally with the use of permuted blocks balanced within each randomisation stratum (e.g., previous exposure to anti-TNF therapy: yes or no). Patients were treated for 8 weeks and followed for a further 4 weeks. The Mayo score was determined at baseline and at 8 weeks.

The primary outcome was clinical response at 8 weeks, defined as an absolute decrease from baseline in Mayo score of ≥ 3 and a
decrease from baseline of 30% or more with an accompanying decrease in the rectal bleeding subscore of ≥ 1 point or an absolute rectal bleeding subscore of ≤ 1 – this definition is identical to that used in the Phase III trials.

One difference between the Phase II and Phase III studies is the use of centrally read endoscopic subscores in the latter. The tofacitinib OCTAVE development programme was the first in ulcerative colitis to use central reads to assess primary efficacy endpoints, and the benefit of using central reads has not yet been established. Results based on local reading, as in the Phase II trial, may be closer to real-world data than central reading and remain relevant for prescribers (see section B.2.3.1.2.4). Locally read endoscopic subscores are used in the base case of the NMA and economic model (see sections B.2.9 and B.3); centrally read data are used in sensitivity analysis.

Baseline characteristics

The demographic and baseline disease characteristics were generally similar across the five treatment groups (see Appendix L1.6, Table 237). Across all treatment groups, 131 (67.5%) patients received concomitant aminosalicylates and 85 (43.8%) received concomitant

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

glucocorticoids at some point during the study. Patient characteristics were generally similar to those in the Phase III trials.

Results

The primary outcome, clinical response at 8 weeks, occurred in 32%, 48%, 61%, and 78% of patients receiving tofacitinib at a dose of 0.5 mg ($p = 0.39$), 3 mg ($p = 0.55$), 10 mg ($p = 0.10$), and 15 mg ($p < 0.001$), respectively, compared with 42% of patients receiving placebo (see section B.2.6.4, Figure 18).

At week 8, clinical remission (measured with the same definition used in the Phase III trials: a Mayo score ≤2, with no subscore >1) was achieved by 13%, 33%, 48%, and 41% of patients receiving tofacitinib at a dose of 0.5 mg ($p = 0.76$), 3 mg ($p = 0.01$), 10 mg ($p < 0.001$), and 15 mg ($p < 0.001$), respectively, compared with 10% of patients receiving placebo.

The most commonly reported adverse events related to infection were influenza and nasopharyngitis (in six patients each) (see Appendix F, Table 166). Two patients receiving tofacitinib 10 mg twice daily had serious adverse events from infection (postoperative abscess in one and anal abscess in the other).

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Induction base-case NMA results (Table 25 from page 92 of the CS)

Table 25 Induction phase base-case NMA results – comparative effects and probabilities of achieving response and remission

Comparator	Comparator vs PBO			TOF vs comparator				
	Treatment effect, median (95% CrI)	Odds ratio, median (95%CrI)		Odds ratio, median (95%CrI)		Absolute probability		SUCRA [a]
	Probit scale	Clinical response	Clinical remission	Clinical response	Clinical remission	Clinical response	Clinical remission	
TNFi-naïve subgroup								
PBO								
TOF 10 mg								
INF 10 mg/kg								
ADA 160/80/40 mg [b]								
GOL 200/100 mg [c]								
VED 300 mg [d]								
TNFi-exposed subgroup								
PBO								
TOF 10 mg								
ADA 160/80/40 mg [b]								
VED 300 mg [d]								

[a] based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2.
Abbreviations: ADA, adalimumab; CrI, credible interval; GOL, golimumab; INF, infliximab; PBO, placebo; SUCRA, surface under cumulative ranking curve; TNFi, tumour necrosis factor inhibitor; TOF, tofacitinib, VED, vedolizumab.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Appendix E – NMA Winbug code

Linked to ERG Clarification Question A17

DATA AND INITIAL VALUES FOR BASE CASE MULTINOMIAL PROBIT MODELS

Induction - TNF naive base case analysis

Baseline model

Data

[redacted]

END

Initial values

list(mu=c(0,0,0,0,0, 0,0,0,0,0, 0,0), sd.m=1, m=0) #chain 1
list(mu= c(-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1), sd.m=2, m= -1) #chain 2
list(mu= c(1,1,1,1,1,1,1,1,1,1,1,1), sd.m=0.5, m= 1) #chain 3

Induction - TNF exposed base case analysis

Baseline model

Data

list(NObs=5)

[redacted]

END

Initial values

list(mu=c(0,0,0,0,0,0), sd.m=1, m=0) #chain 1
list(mu= c(-1,-1,-1,-1,-1,-1), sd.m=2, m= -1) #chain 2
list(mu= c(1,1,1,1,1,1), sd.m=0.5, m= 1) #chain 3

Maintenance - TNF naive base case analysis

Baseline model

Data

list(NObs=7)

[redacted]

END

Initial values

list(mu=c(0,0,0,0,0,0,0,0), sd.m=1, m=0) #chain 1
list(mu= c(-1,-1,-1,-1,-1,-1,-1,-1), sd.m=2, m= -1) #chain 2
list(mu= c(1,1,1,1,1,1,1,1), sd.m=0.5, m= 1) #chain 3

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Maintenance - TNF exposed base case analysis

Baseline model

Data

list(NObs=3)

[redacted]

END

Initial values

list(mu=c(0,0,0), sd.m=1, m=0) #chain 1
list(mu= c(-1,-1,-1), sd.m=2, m= -1) #chain 2
list(mu= c(1,1,1), sd.m=0.5, m= 1) #chain 3

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

NICE National Institute for Health and Care Excellence

Professional organisation submission

Tofacitinib for moderately to severely active ulcerative colitis [ID1218]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**British Society of Gastroenterology**

3. Job title or position	**Consultant gastroenterologist , BSG IBD committee chairman, and committee member**
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	**British Society of Gastroenterology is the professional body representing UK gastroenterologists. Funding is through annual subscription from members and from income from the annulal conference, which is in turn funded by attendance fees and company sponsorship**
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**no**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	To initiate healing of colonic mucosal inflammation, and prevent recurrence of inflammation in future. This will result in symptoms (bleeding, diarrhoea, urgency of defaecation, abdominal pain, and fatigue) resolving and prevent their recurrence.

or prevent progression or disability.)	
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Clinical and endoscopic remission: Complete resolution of symptoms (as above) with evidence of mucosal healing of all affected areas of the colon at colonoscopy. This can be quantified by Disease Activity Score (Mayo score), and the UCEIS score to measure mucosal healing. Patient Reported Outcome Measures (PROMs) such as the IBD Control PROM is increasingly used as an overall measure of patient well-being Clinical response: improvement in symptoms (as measured by Disease Activity Score) Maintenance of remission. Measured by same scores at intervals such as 1year after start of therapy
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. 10-20% of patients currently fail medical therapy for UC, either as a result of acute severe UC, or due to failure of medical therapy resulting in chronic treatment refractory, or corticosteroid-dependent disease
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	1st line therapy: mesalazine, oral and topical for mild to moderate disease. More severe presentations require oral (or intravenous) corticosteroids to induce remission, followed by maintenance therapy with mesalazine 2nd line therapy is required a) for in-patients with acute severe UC failing to respond to IV corticosteroids, who are treated with either infliximab iv, or ciclosporin iv; b) out-patients with relapses requiring 2 or more courses of corticosteroids in a year despite high-dose mesalazine. They are generally treated with thiopurines, and may need early escalation to biological therapy (infliximab, adalimumab, or vedolizumab). These drugs are generally continued as maintenance therapy. Colectomy for failure to respond to medical therapy, or for life-threatening complications (perforation, toxic

	megacolon, haemorrhage)
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	NICE guidelines ECCO guidelines 2016
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Yes. Most steps in treatment pathway clear. There is some debate in specific areas eg a) starting dose of mesalazine in moderate active UC; b) who should receive early escalation to biologics therapy after failure to respond to mesalazine maintenance therapy; c) duration of maintenance therapy after achieving remission with biologics; d) treatment choices in patients with medically refractory therapy
• What impact would the technology have on the current pathway of care?	Offers an alternative treatment to UC patients failing to respond to mesalazine therapy, either as an alternative to conventional treatment (thiopurines/anti-TNF drugs/vedolizumab) or as a further treatment option for patients who fail to respond to one or more of these drugs. It offers a different class of treatment and therefore may have value in patients who are non-responsive to drugs with diferrent mode of action
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes. Oral treatment however is far more convenient for patients in comparison to both iv and subcutaneous treatments
• How does healthcare	Oral medication, so does not require infusion facilities (nursing staff, and space to administer iv therapy)

resource use differ between the technology and current care?	nor training for giving subcutaneous injections.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Hospital outpatients
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	No. Treatment would not require additional resources, apart from information and monitoring which would be coordinated by specialist IBD nurses already involved with UC patients treatment
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. Particularly likely to benefit a group of patients who have failed current treatment options, and increase their chance of avoiding colectomy
• Do you expect the technology to increase length of life more than current care?	No. Mortality is not increased overall in UC, and deaths due to severe disease or complications are rare in the UK
• Do you expect the technology to increase	Yes. In those patients who have failed current treatment options, as it will reduce their need for long-term corticosteroid use, and reduce their chance of requiring colectomy which is highly likely to require formation

health-related quality of life more than current care?	of ileostomy, for 6-12 months, or permanently
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Treatment-refractory or corticosteroid-dependent UC in an out-patient setting
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional	Yes Less need for infusions Orally administered -better for patient's and healthcare providers

tests or monitoring needed.)	
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Yes. Annual assessment of response to treatment as for current biologics therapy
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	Yes. For many young people, colectomy and ileostomy formation is a devastating consequence of failure to manage medically. It often occurs at a time when they are completing secondary education, or during university, before they have formed a long-term relationship, or at a time before completing their family and has a significant effect on education attainment, formation of relationships, sexual activity and pregnancy/delivery. Many of these factors are not picked up by QALY scores, and current literature is poor on QOL scores in different states (remission, mild/moderate/severe disease, ileostomy after colectomy, ileoanal pouch after colectomy)
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it	Yes Orally administered Small molecule so reduced chance of immunogenicity compared to monoclonal antibody therapies Good side-effect profile

improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Yes. For reasons stated above
• Does the use of the technology address any particular unmet need of the patient population?	Yes. See above
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Phase 2 and 3 data has minor and acceptable side effects – 10-15% rise in lipids (significance unclear), fall in neutrophil counts, slight increase in infections such as shingles (risk approximately 4 in 100 patient years). Vaccination is recommended before treatment is started.
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes, with the exception that patients with proctitis (disease extent less than 15cm from anal verge) were excluded from the OCTAVE trials. This is common practice in clinical trials as proctitis may have different clinical symptom patterns (less diarrhoea for instance). There is no reason to believe that proctitis would not respond in a similar way to tofacitinib, and represents a group who often have refractory disease, and

	for whom colectomy is often inappropriate.
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	Clinical remission and mucosal healing after induction therapy (8 weeks). Clinical remission and mucosal healing at 1 year (and longer). Normalised quality of life. Colectomy avoidance. Corticosteroid-free remission over one or more years. In the OCTAVE studies all these were measured, but only 1 year follow-up data, and no information on colectomy rates (which would be low). Patients studied were not necessarily refractory to multiple treatments, but over half had had prior treatment with anti-TNF, and nearly half were on corticosteroids at start of induction trials.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	Mucosal healing at one year, is associated with lower subsequent relapse rate
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
19. Are you aware of any relevant evidence that might	No

not be found by a systematic review of the trial evidence?	
20. Are you aware of any new evidence for the comparator treatments since the publication of NICE technology appraisal guidance TA329 and TA342?	No
21. How do data on real-world experience compare with the trial data?	Very limited at present
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	This is only an issue if high cost prevents access to the treatment in some areas of UK
22b. Consider whether these issues are different from issues	Same issues regarding cost, and less variability in treatment provision for an oral agent in comparison to IV

with current care and why.	therapies

Key messages

23. In up to 5 bullet points, please summarise the key messages of your submission.

- Treatment option for patients failing standard therapies
- Huge unmet need in young patients where active disease, and colectomy has impact on education, relationships and pregnancy
- Oral treatment has advantages over iv or subcutaneous alternatives
- Small molecule, so likelihood of immunogenicity and loss of response over time is less

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Professional organisation submission

Tofacitinib for moderately to severely active ulcerative colitis [ID1218]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**UKCPA**

3. Job title or position	**Consultant pharmacists – Gastroenterology; UKCPA committee members**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ✓☐ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	The UK Clinical Pharmacy Association (UKCPA) is a member association providing education and training for clinical pharmacy practitioners. The UKCPA actively develops clinical pharmacy practice and individual practitioners. Activities include establishing professional curricula, developing professional recognition (credentialing) processes, and developing professional tools and frameworks for practitioners. The UKCPA Gastroenterology Interest Group provides a network for information exchange and training for any pharmacist working within the speciality of gastroenterology. In 2016 the UKCPA was awarded Royal Pharmaceutical Society accreditation as a Foundation Training Provider and a Faculty Training Provider.
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
The aim of treatment for this condition	

6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	Induce & maintain steroid-free remission (clinical/ endoscopic) for 12 months or longer & prevent need for surgical intervention. Improve quality of life for patient suffering with UC having failed conventional therapy without introducing additional risk factors (ie cancer risk, ADRs)
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Mayo score ≤ 2 Sustained remission for 12 months Improvement in Quality of life to near equal of healthy individuals Steroid free periods for ≥ 12 months
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes: 1. Other biologics formulated as injection – IV (requires hospital admission), subcutaneous. a. This may be an issue if needle phobic b. Capacity issues in secondary care for infusions 2. First drug of this class offering alternative mode of action to other biologics which may have failed or not tolerated/contra-indicated 3. Patient convenience, acceptability and potentially increased compliance 4. Low immunogenicity: Tofacitinib is not considered immunogenic unlike other licenced biologics which therefore are not effective for all and loss of response over time likely due to immunogenicity

What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	PO/PR 5 ASA, steroids, immunomodulators, biologics (anti-TNF infliximab/adalimumab/golimumab), vedolizumab, (dietary interventions unlikely to be effective, surgery
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	• NICE CG166 (currently being updated) • NICE TAs for anti-TNF agents, vedolizumab – various • European Crohn's & colitis Consensus Guidelines. J. Crohns & Colitis 2017
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Choice of first line biologics may vary nationally Locally defined treatment pathways - commissioners /secondary care Interpretation of NICE guidance varies by commissioners resulting in variability of access in England
• What impact would the technology have on the current pathway of care?	• Alternative to injectable biologics • Higher Patient acceptability as oral • Alternative mode of action if other treatment targets fail • Less chance of immunogenicity developing • Short half life

10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes
• How does healthcare resource use differ between the technology and current care?	Likely to be more expensive than current conventional therapy and biologics
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary care – under the supervision of specialist gastroenterologist with interest in IBD
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Staff Education Homecare contracts management
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes and higher patient acceptability Rapid onset of action, may be used instead of steroid in acute management of UC

• Do you expect the technology to increase length of life more than current care?	Potentially yes as low immunogenicity, but longterm effectiveness needs to be evaluated
• Do you expect the technology to increase health-related quality of life more than current care?	Potentially yes; delay need for surgery, increased steroid free periods and quality of life improvement Replacing steroids in management of acute flares
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Appropriate: Needle phobic Failed/intolerant to other biologics Less appropriate: Compliance concerns
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for	Higher patient acceptability Easier to use lower hospital resource requirements (oral vs injections), less nursing time required similar monitoring requirements expected in view of efficacy

example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	may need access to VZIG (UK limited resources) and VZ antibody testing
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Expect as per TA approval criteria for other biologics
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	Potentially: no need to attend infusion clinics (time) higher acceptability of treatment, lower psychological barriers to treatment

16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	yes
• Is the technology a 'step-change' in the management of the condition?	Yes – alternative mode of action and route of administration
• Does the use of the technology address any particular unmet need of the patient population?	Yes – those who have failed /intolerant other biologics
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Similar profile to other licenced biologics Drug Drug Interactions (DDI): Caution drug interactions as metabolised via CYP3A4 systems (exposure is increased when coadministered with potent inhibitors of CYP3A4 (e.g., ketoconazole) or when

	administration of one or more concomitant medications results in both moderate inhibition of CYP3A4 and potent inhibition of CYP2C19 (e.g., fluconazole) Contraindication in pregnancy: This is a young patient group and potential teratogenicity and pregnancy whilst on treatment need to be taken into account. Infection risk will need extensive pre-assessment similar to biologics, VZ may be an issue as VZIG resources are limited Considerable effect on FBC (lymphopenia, neutropenia and anaemia potential) Needs dose adjustment in liver disease. Hyperlipidaemia may be an issue
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	yes
• If not, how could the results be extrapolated to the UK setting?	yes

• What, in your view, are the most important outcomes, and were they measured in the trials?	Just as effective in biologics naïve vs experienced,
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not aware, cancer risk seems to be similar to biologics particularly when used in in patients with thiopurines exposure, increased ADR
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	no
20. Are you aware of any new evidence for the comparator treatments since the publication of NICE technology	no

appraisal guidance TA329 and TA342?	
21. How do data on real-world experience compare with the trial data?	
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	Not aware of any
22b. Consider whether these issues are different from issues with current care and why.	
Key messages	

23. In up to 5 bullet points, please summarise the key messages of your submission.

- Alternative mode of action for a biologic
- Oral formulation
- Concerns about pregnancy and family planning
- Agranulocytosis, anaemia and increased risk with old age
- DDIs

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Tofacitinib for moderately to severely active ulcerative colitis [ID1218]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	

2. Name of organisation	Crohn's and Colitis UK
3. Job title or position	Health Service Development Manager
4a. Brief description of the organisation (including who funds it). How many members does it have?	Crohn's and Colitis UK is the leading charity dedicated to improving the lives of everyone affected by Crohn's Disease, Ulcerative Colitis and other forms of Inflammatory Bowel Disease (IBD). Founded as a patient's organisation in 1979, we now have over 35,000 members and 50 local networks. Working together, we provide information and support, campaign to improve services and healthcare, and fund vital research. Funding is through membership subscriptions and a wide range of fundraising activities, including events, grants, legacies and corporate partnerships. Full details are available in our annual accounts. https://www.crohnsandcolitis.org.uk/about-us/annual-accounts
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	Through our day-to-day work providing information and support, we are in contact with thousands of people affected by IBD, who share their experiences through our helplines, online forum and at events. We also regularly conduct surveys and hold focus groups exploring issues that are relevant to people with Ulcerative Colitis and Crohn's Disease and fund and support qualitative and quantitative research. In relation to this submission, we have heard directly from patients who have been treated with tofacitinib, including those we have nominated as patient experts for this appraisal.

Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	The symptoms of Ulcerative Colitis, and their unpredictable nature, can have a profound and devastating impact on all aspects of a person's life, especially given that 25% are diagnosed in the first two decades of life. Frequent diarrhoea, abdominal pain and fatigue, anaemia, extra-intestinal manifestations such as joint, skin and eye problems, and the side effects of medications, all affect an individual's ability to work, study, socialise, participate in leisure activities or have intimate relationships. (IBD Quality of Life Survey, 2018; IBD Standards, 2013). Emotional wellbeing can be significantly affected by difficulty in coping with personal lives and feelings of anger, embarrassment, frustration, sadness and fears of needing surgery or developing cancer (Cosnes et al, 2011). Stigma and lack of wider understanding of the condition exacerbate the impact. Anxiety and depressive illness is higher in people with IBD, with mood disorders at least in part a consequence of the IBD itself (Graff, 2009) and its medical treatment (e.g. corticosteroid therapy), surgery, including specifically colectomy and stoma formation. Additionally, most reports indicate that stress may be involved in triggering relapse. *"The isolation I have felt has been overwhelming. I can't take my children to the park, for a walk or play date or any of the other simple things that I used to take for granted. I do not have any kind of social life myself as it is simply not possible for me to go out when I may need to open my bowels with no warning. My relationship with my husband is also under a lot of pressure for obvious reasons. I am 35 years old and feel like an old woman. Instead of being in the prime of my life, my world has become so small due to this horrific disease and so have the lives of my children. The constant tiredness and the mood swings caused by all the steroids I take have made me a pretty poor mum. Recently my 2 year old son had to have an operation and I was too ill to stay with him in hospital overnight. It broke my heart. I want to get a job, enjoy my children and just start living again."* Quote from a person living with Ulcerative Colitis.

	"When I am unwell the constant anaemia make everyday life feel like wading through treacle, the pain can be crippling. The very real concern of faecal incontinence gives me physical symptoms of stress as well as affecting me emotionally and mentally." Quote from a person living with Ulcerative Colitis. The experience of caring for someone with IBD can be especially difficult given that it is to some degree an invisible condition and due to the unpredictable nature of the symptoms, which many also find extremely uncomfortable to talk about, and the effects of treatment. For parents of young people, there are challenges around providing support, while enabling independence and seeing lives and aspiration affected by the son or daughter's condition. *"He was struggling to maintain a healthy weight, was constantly feeling sick, rushing to the toilet and in pain and missing a great deal of his work at a stage in his career that was very important to him. He was unable to continue his sport and his social life was negligible. He had to take medication to digest his food."* Quote from the parent of a person living with Ulcerative Colitis.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	While a range of options are available for treating Ulcerative Colitis, these remain far from optimal for patients, a substantial number of whom experience lack of response (primary or secondary) and/or adverse reactions. The effects of steroids, in particular, are extremely unpleasant and long-term safety profile of other treatments, including biologics, of some concern. *"During the majority of my time living with UC and the ever-changing drugs, I had no quality of life. I was off sick from work for 8 months. I was unable to drive my children to or from school or make them their breakfast as this was the time, usually until about midday, that I could not leave the toilet. There was no fun time with my 3 wonderful children or my husband, I was always in bed, in pain or on the toilet. We did not cuddle or play, because if any of them touched my tummy, it would be so sore. This period of illness really affected my confidence. My friends gave up coming around as I was so poorly. My quality of work really dropped. I continuously made mistakes because of the side effects from all the drugs."* Quote from a person living with Ulcerative Colitis.

	"My 'moon face' from the constant use of prednisolone was depressing and because of my ill health my hair became really thin. Prednisolone also affected my mood. I was so angry and unhappy. This also kept me awake at night, so I took sleeping pills. Quote from a person living with Ulcerative Colitis. For many patients with Ulcerative Colitis, the prospect of surgery is one they face with considerable anxiety and it can bring with it a range of potential complications, which may require further treatment and ongoing management. There can also be an associated profound psychological and social impact, for example, in terms of body image and self-esteem. For those who are facing this at an age when they have just begun to form relationships and do not yet have a family, this can be especially difficult, as it can for those of some religious faiths and cultures. Clinical outcomes after pouch surgery remain variable and fertility in women can be significantly affected. *"Surgery would have been a massive emotional and psychological barrier for our son at this stage in his life."* Quote from a person living with Ulcerative Colitis. *"Personally I'm not prepared for the drastic surgery of having my colon removed."* Quote from a person living with Ulcerative Colitis.
8. Is there an unmet need for patients with this condition?	For some patients, acceptable treatment options run out and the impact of this can be profound. For others, a more limited quality of life can be normalised due to inadequate response. *"I was steroid dependent and all conventional UC therapies failed – including anti TNF (Infliximab). Long term steroid use resulted in osteoporosis at age 28. I was housebound for many years due to UC and was unable to work. Quality of life was zero."* Quote from a person living with Ulcerative Colitis. *"Despite corticosteroids, azathioprine and other drugs, nothing was working. The steroids caused him great distress and made him very difficult company. Trying to work was a nightmare and he experienced crushing fatigue."* Quote from a person living with Ulcerative Colitis.

Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	There is a clear value to an oral treatment in terms of convenience and acceptability to patients, in addition to important factors such as reduced likelihood of loss of response (immunogenicity) due to it being a small molecule and good side effect profile. The management of treatment such as infusion therapy and drug monitoring for immunosuppressants such as azathioprine can impact significantly on patient's lives and work requiring regular attendance at hospital and additional travel and parking costs. Some patients find the prospect of subcutaneous injection unacceptable, or at least unpleasant, home delivery has to be managed and the need to store these drugs at an appropriate temperature can impact on travel plans.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	The main potential disadvantage to patients would be in terms of prescription cost (in England) which can impact on adherence and lead to complications and increased cancer risk and cost to the NHS (York Health Economics Consortium, 2018). However, this is outweighed by the value of an additional treatment option, which has a different mode of action, reduced likelihood of loss of response and a convenient delivery method. There should also be an associated reduction in NHS costs due to reduced infusions.

Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Patients who have had little or no success with currently available medical treatment options, and wish to avoid or delay surgery, are likely to benefit in particular. This would include young people wishing to complete studies or start a family and those for whom surgery would be considered unacceptable due to cultural or religious factors. *"I was diagnosed with UC in 2012 and had multiple failed attempts at other medication over 2 years which affected my studies, work and social life before joining the trial for tofacitinib. I was on this treatment successfully for 2½ years during which time I completed my studies and qualified as an adult nurse. I stopped only to get pregnant in Oct 2016 and gave birth to my daughter in September 2017."* Person living with Ulcerative Colitis who has had personal experience of treatment with tofacitinib. *"This treatment has completely changed my life. Diagnosed in 2004 I had been in and out of remission for many years, housebound at the least, hospitalized at the worst. I have tried numerous drugs and trials but nothing that compare. I am now in my 4th year of taking Tofacitinib and it is like I am a new person, and feel so privileged to have been able to trial it...it has been a true success for me!"* Person living with Ulcerative Colitis who has had personal experience of treatment with tofacitinib.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	As referred to above, potential equality issues that should be considered in relation to this condition and technology would include the impact for women who have not yet completed their family and those who, consider surgery to be unacceptable due to cultural or religious factors. Cost may also be a factor associated with lower income.

Other issues	
13. Are there any other issues that you would like the committee to consider?	None

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission:

- The symptoms of Ulcerative Colitis, and their unpredictable nature, together with the side effects of medications, can have a profound and devastating impact on all aspects of a person's life.
- Current treatments remain far from optimal for patients, a substantial number of whom experience lack of response (primary or secondary) and/or adverse reactions to medical treatments and may face the prospect of surgery with considerable anxiety.
- For some patients, acceptable treatment options run out and the impact of this can be profound. For others, a more limited quality of life can be normalised due to inadequate response.
- There is a clear value to an additional treatment option, which through being orally administered, offers greater convenience and acceptability to patients, as well as a good side effect profile and reduced likelihood of loss of response (immunogenicity).
- Due to its different mode of action, those who have had little or no success with currently available medical treatment options and wish to avoid or delay surgery will benefit in particular.

NICE National Institute for Health and Care Excellence

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Clinical expert statement

Tofacitinib for previously treated active ulcerative colitis [ID1218]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Dr Patrick Allen**
2. Name of organisation	**Ulster Hospital**

3. Job title or position	**Consultant Gastroenterologist**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ☐ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	**x**☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ **yes**

NICE National Institute for Health and Care Excellence

Patient expert statement

Tofacitinib for previously treated active ulcerative colitis [ID1218]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Shirley Samantha Leather**
2. Are you (please tick all that apply):	☒ a patient with the condition? ☐ a carer of a patient with the condition? ☐ a patient organisation employee or volunteer?

	☐ other (please specify):
3. Name of your nominating organisation	Crohn's and Colitis UK
4. Did your nominating organisation submit a submission?	☒ yes, they did ☐ no, they didn't ☐ I don't know
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)

6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes
7. How did you gather the information included in your statement? (please tick all that apply)	☒ I have personal experience of the condition ☐ I have personal experience of the technology being appraised ☐ I have other relevant personal experience. Please specify what other experience: ☐ I am drawing on others' experiences. Please specify how this information was gathered:
Living with the condition	
8. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	I was diagnosed with Ulcerative Colitis in 2004. Having just had a new baby boy my first 6 weeks of motherhood were more than demanding. I would wake up in the night to breast feed, get him attached only to have to go to the loo urgently, passing diarrhoea with what seemed like so much blood. In honesty I just thought it was all part of the process of having a baby…how wrong I was. Over the last 14 years I have been hospitalized four times, with stays from the shortest being 10 days to the longest 4 weeks, and that doesn't include all the day visits and appointments, much of my time being hooked up to machines pumping weird and wonderful drugs through my body. I am one of the lucky ones in the fact that I did not permanently suffer with the awful abdominal cramps associated with the disease, however the tiredness and fatigue were overwhelming, and most days I would have to sleep in the afternoon, trying to make my two young children understand that mummy needed to sleep and that they must play quietly.

	My husband, who I would never have got through without this, has been my rock and when our first son was only 6months old would work all day and then come visit me in hospital with him, as babies were not allowed on the cancer ward I had been placed on…It was heart breaking for me and him, and my time there was life changing. Life has been a rollercoaster of a ride with Ulcerative Colitis, going in and out of remission for so many years, trying so many different drugs I was not able to even think about going back to work. The steroids were probably the worst and each time I took them I would balloon up which was so depressing, and other drugs had so much monitoring so it seem that every week I was in a hospital for one thing or another. At my worst I would go to the toilet 27 times a day, and in a house with four of us and only one toilet that was challenging. My kids, too, have been amazing and when I could no longer make the short walk to school with them they would happily go with another parent….I was housebound and in some ways felt I had failed as a wife and a mum. Holidays were a no when I was ill though once we did do a road trip but can't tell you how many times we had to stop for me to go to the toilet 'just in case'. The feeling of constant apprehension as to whether you will make it to the next one is indescribable and my social life depleted as I could not bear the thought of being out and having an accident. Huge thanks go to Dr Bloom who got me onto this trial. In my head this was my last chance as I had tried and failed with so many drugs and trials previously, and though surgery was my absolute last resort I did not have any standard of life left.
Current treatment of the condition in the NHS	
9. What do patients or carers think of current treatments and care available on the NHS?	Firstly, I would like it noted that the care I have received from the NHS throughout has been outstanding. I have tried so many treatments to help keep my condition under control with the hope of long term remission. Oral steroids in the first few years were always the back stop and usually would help initially, though once I had cut down then the bleeding and diarrhoea would start again, to the point it just became a vicious circle, and the way I ballooned meant my self-esteem was so low. Cyclosporin was another, rather toxic, treatment though with my initial use gave me excruciating pains down my shins which I had to take Tramadol for. Azathioprine, seemed to work well for a few years, but then I got Neutropenia and had

	to be taken off that. I have used numerous enemas and suppositories some which helped, some that didn't and the pain, let alone the dignity of inserting them, sometimes didn't seem worth the result. Methotrexate was the last oral drug I took before this trial, though that only lasted a few months before I relapsed yet again. With the opportunity of Infliximab I think we were all hopeful that this would be the one for me, and after the three dose infusions there was no improvement, my disappointment could not be measured. By the time I was offered this trial, I was at my very lowest and desperate to find anything that could give me back some sort of quality of life. I can honestly say that Tofacitinib has totally and utterly changed my life, it is like I am a completely new person, and amazingly I have now been in remission for over three years. I am honoured and privileged to be able to have trialled this drug!!
10. Is there an unmet need for patients with this condition?	
Advantages of the technology	
11. What do patients or carers think are the advantages of the technology?	Tofacitinib is amazing in the fact that it is just a pill and can be taken in the comfort of your own home. Other drugs that I have trialled have meant going in to hospital to have infusions, others were to be self-injected which would be no good for me as I still, after all these years, cannot stand the sight of needles!!
Disadvantages of the technology	
12. What do patients or carers think are the disadvantages of the technology?	I can't see that there are any disadvantages.

Patient population	
13. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Surely everyone can benefit from this drug, and I hope that they may, in the future, be able to.
Equality	
14. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	
Other issues	
15. Are there any other issues that you would like the committee to consider?	

Key messages
16. In up to 5 bullet points, please summarise the key messages of your statement: • • • • •

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Patient expert statement

Tofacitinib for previously treated active ulcerative colitis [ID1218]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Charlotte Hughes**
2. Are you (please tick all that apply):	☒ a patient with the condition? ☐ a carer of a patient with the condition? ☐ a patient organisation employee or volunteer?

	☐ other (please specify):
3. Name of your nominating organisation	Chrones & Colitis UK
4. Did your nominating organisation submit a submission?	☒ yes, they did ☐ no, they didn't ☐ I don't know
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)

6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes
7. How did you gather the information included in your statement? (please tick all that apply)	☒ I have personal experience of the condition ☐ I have personal experience of the technology being appraised ☐ I have other relevant personal experience. Please specify what other experience: ☐ I am drawing on others' experiences. Please specify how this information was gathered:
Living with the condition	
8. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	I was diagnosed with ulcerative colitis when I was 23 (UC) in March 2012 during a week's hospital admission with severe symptoms such as pain when passing stools, blood in my stools and increased frequency up to 20 times a day. This was a life altering diagnosis. I found work and socialising much more difficult, feeling constantly anxious about where I would be able to find a toilet, embarrassment when accidents happened in public. I made massive adjustments to my life without even really realising at the time how much impact it had. I could no longer walk home from the bus without an overwhelming fear of needing to go to the toilet. Food, which had always been a great pleasure of mine, also became an issue I couldn't finish a meal without needing to leave to go to the toilet, I had constant low energy and fatigue. I would have to plan my journeys to and from work/university with a stop off to go the toilet. Although I have always been a positive person and try to make the most of situations I definitely found myself feeling depressed and very alone. Trying to go on a date and develop new relationships was really hard as UC had such an overwhelming impact on my daily life but it wasn't something I wanted to discuss on a first

	date! I had endless hospital appointments and endoscopy's which took up a huge amount of time of lasting for hours. This meant I had to organise lots of time off work, or spent my free time at appointments sometimes my days off before a nightshift so I wouldn't miss work. Having each treatment take months of monitoring, appointments and bloods tests and then fail was emotionally distressing as I felt my options were running out and I didn't want to have to surgery.
Current treatment of the condition in the NHS	
9. What do patients or carers think of current treatments and care available on the NHS?	Since I was diagnosed I went through various different treatments all of which didn't not work for me. I had Azathioprine and Methotrexate both of which required daily, then weekly then monthly blood tests, which was very time consuming. Both of these medications had no effect on my UC but seemed to make my general health almost worse, I often was ill catching colds that would last for ages or vomiting bugs. I then was offered a trial medication called Humeria in the summer of 2013, which was a subcutaneous injection, this had to be kept cold so was difficult to organise taking it away with me and I had to learn to inject myself. Taking part in a trail was very time consuming; there were a lot of very lengthy appointments, scans and paperwork. This medication also had no effect on my UC. In between the different treatments I had various courses of steroids, these did had some positive effects on my UC but only temporarily and also came with side effects such as insomnia and mood swings.
10. Is there an unmet need for patients with this condition?	Yes, I think there are patients like me living with UC who have been unresponsive to all other treatments and feel their have run out of options.
Advantages of the technology	
11. What do patients or carers think are the advantages of the	This medication worked very well and very quickly for me when I had failed to respond to all other treatments. I felt better than I had in years; I was able to be out of the house without constant anxiety of where to find a toilet for the first time in years. My energy slowly returned and I only then realised fully how ill I had been. I completed my studies to qualify as an adult nurse and took on my first job on the wards, one that would have struggled with in my previous state. I have had no sign effects from this medication;

technology?	in fact I think it also improved my eczema. Taking an oral tablet is a major advance as supposed to injections or infusions as you can travel easily and don't need to worry about equipment or keeping medication cold. I was successfully on Tofacitinib for over 2 years I ended the trial only to become pregnant and I now have a little girl who is 10months old, I have continued to be in remission having stopped the medication in October 2017.
Disadvantages of the technology	
12. What do patients or carers think are the disadvantages of the technology?	The only very minor disadvantage would be the prescription cost, as I am on various other prescription medications although for me this would be massively worth the money. That you currently cannot take the medication whilst being pregnant or breastfeeding.
Patient population	
13. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	People like me who have tried all other options available and are running out of options. Patient's for whom injections or infusions is not a viable option.
Equality	
14. Are there any potential equality issues that should be	**No.**

taken into account when considering this condition and the technology?	
Other issues	
15. Are there any other issues that you would like the committee to consider?	No.
Key messages	
16. In up to 5 bullet points, please summarise the key messages of your statement: • Living with active UC is a lonely and life altering, making work and social life very difficult • The current NHS treatments had all failed to work for me • Having another oral option for patients who have failed all current treatment is vital • I responded quickly and very well to Tofacitnib and I'm now in remission • Tofacitinib changed my life	

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

CONFIDENTIAL UNTIL PUBLISHED

Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE

Tofacitinib for moderately to severely active ulcerative colitis

Produced by Southampton Health Technology Assessments Centre

Authors Neelam Kalita, Research Fellow (Health Economics), SHTAC
Petra Harris, Research Fellow, SHTAC
Joanne Lord, Professorial Fellow, SHTAC
Geoff Frampton, Senior Research Fellow, SHTAC
David A. Scott, Director, Diligent Agile Synthesis Limited
Joanna Picot, Senior Research Fellow, SHTAC

Correspondence to Dr. J. Picot
Southampton Health Technology Assessments Centre (SHTAC)
Wessex Institute
Alpha House
Enterprise Road, University of Southampton Science Park
Southampton SO16 7NS
www.southampton.ac.uk/shtac

Date completed 19/07/2018

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 17/109/06.

Declared competing interests of the authors

None from the authors. The department Dr Parkes works within has been involved in the Phase III trials of tofacitinib (and its potential future competitor filgotonib) – for which the department receives payment. Dr Parkes does not receive personal payment. Dr Parkes is also on the research advisory board for Crohn's and Colitis UK. Dr Brown currently prescribes all of the comparator drugs in ulcerative colitis and he is currently a principal investigator for a phase IIb/III clinical trial of an investigational product (upadacitinib, which is a JAK1 inhibitor) being developed by Abbvie for treatment of ulcerative colitis.

Acknowledgements

We are very grateful to Dr Miles Parkes, Directory of Gastroenterology and Inflammatory Bowel Disease Lead, Addenbrooke's Hospital Cambridge and Dr. Matthew Brown, Consultant Gastroenterologist, Basingstoke and North Hampshire Hospital who offered clinical advice with Dr. Brown also offering comment on the draft report. We would also like to thank: Karen Welch, Information Scientist, SHTAC, for appraising the literature search strategies in the company's submission, and for running searches where necessary; and Dr Jonathan Shepherd, Principal Research Fellow, SHTAC, for providing a quality assurance review of the draft ERG report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Kalita, N.; Harris, P.; Lord, J.; Frampton, G.; Scott D.A.; Picot, J. Tofacitinib for moderately to severely active ulcerative colitis: A Single Technology Appraisal. Southampton Health Technology Assessments Centre, 2018.

Contributions of authors

Neelam Kalita critically appraised the health economic systematic review, critically appraised the economic evaluation, and drafted the report; Petra Harris critically appraised the clinical effectiveness systematic review and drafted the report; Joanne Lord critically appraised the health economic systematic review, critically appraised the economic evaluation, and drafted the report; David Scott critically apprised the network meta-analyses and drafted the report; Geoff Frampton critically appraised the clinical effectiveness systematic review and drafted the report; Joanna Picot critically appraised the clinical effectiveness systematic review, drafted the report, project managed the review and is the project guarantor.

Word count: 67948

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADA	Adalimumab
AE	Adverse event
BID	Twice daily
BSG	British Society of Gastroenterology
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CrI	Credible interval
CSR	Clinical study report
DIC	Deviance information criterion
EMA	European Medicines Agency
ERG	Evidence review group
EQ-5D	5-dimensions EuroQol questionnaire
FAS	Full analysis set
GOL	Golimumab
HDAS	NICE Healthcare Databases Advanced Search
HRQoL	Health-related quality of life
IBDQ	Inflammatory Bowel Disease Questionnaire
ICER	Incremental cost-effectiveness ratio
INF	Infliximab
IQR	Inter quartile range
LLN	Lower limit of normal
MCID	Minimal clinically important difference
MCS	Mental health component summary (of SF-36)
mFAS	Modified full analysis set
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NRI	Non-responder imputation
PAS	Patient access scheme
PBO	Placebo
PCS	Physical component summary (of SF-36)
QALY	Quality-adjusted life year
Q4W	Once every four weeks

Q8W	Once every eight weeks
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SF-36	36-Item Short Form survey
SLR	Systematic literature review
SmPC	Summary of Product Characteristics
STA	Single technology appraisal
SUCRA	Surface under cumulative ranking curve
TA	Technology appraisal
TEAE	Treatment-emergent adverse event
TNF	Tumour necrosis factor
TNFi	Tumour necrosis factor inhibitor
TOF	Tofacitinib
UC	Ulcerative colitis
ULN	Upper limit of normal
VAS	Visual analogue scale
VED	Vedolizumab
WHO	World Health Organisation
WPAI-UC	Work Productivity and Activity Impairment-Ulcerative Colitis

SUMMARY

Scope of the company submission

The company's submission (CS) presents evidence for the clinical effectiveness and cost-effectiveness of tofacitinib (Xeljanz®) for the treatment of adult patients with moderately to severely active ulcerative colitis who have had an inadequate response, lost response, or were intolerant to either conventional therapy or a biologic agent. Tofacitinib is an orally administered small-molecule selective inhibitor of the Janus kinase (JAK) family of tyrosine kinases. The inhibition of JAKs by tofacitinib attenuates the signalling of several interleukins and type I and II interferons, which leads to modulation of the immune and inflammatory response in ulcerative colitis. The recommended dose is 10mg twice daily for induction for eight weeks and 5mg given twice daily for maintenance.

Summary of submitted clinical effectiveness evidence

The company's systematic review of clinical effectiveness identified four relevant placebo controlled randomised controlled trials (RCT) of tofacitinib.

- one phase II RCT [treatment arms: TOF 0.5 mg twice a day (BID), 3 mg BID, 10 mg BID, 15 mg BID and placebo]
- two identical phase III induction RCTs (OCTAVE Induction 1 and OCTAVE Induction 2; treatment arms: TOF 10 mg BID and placebo)
- one maintenance RCT (OCTAVE Sustain; treatment arms: 5 mg BID, 10 mg BID and placebo)

The ERG believes the company has identified all the relevant RCTs of tofacitinib. In addition to the RCTs, an open-label uncontrolled long-term extension study of tofacitinib, OCTAVE Open, is ongoing.

The CS focusses on the three large phase III trials OCTAVE Induction 1 and OCTAVE Induction 2 and the OCTAVE Sustain maintenance study. The small phase II trial is included in network meta-analyses (NMAs) but it not reported on in detail in the CS.

The OCTAVE 1 and OCTAVE 2 induction trials followed identical methods and both were multicentre, worldwide RCTs. To be enrolled patients had to have moderately to severely active ulcerative colitis. Eligible patients were randomised on a 4:1 ratio to 10 mg twice a day (BID) of oral tofacitinib or placebo for eight weeks (a third 15 mg BID tofacitinib arm was discontinued prior to full recruitment based on feedback from regulatory authorities). Randomisation was

stratified by previous treatment with TNFi therapies, glucocorticoid use at baseline, and geographic region.

People who had participated in the OCTAVE 1 and 2 induction trials completed 8 weeks of induction therapy and met the criteria for a clinical response were eligible to be re-randomised into the OCTAVE Sustain maintenance study. Eligible patients were randomised in a 1:1:1 ratio to 5 mg BID tofacitinib, 10mg BID tofacitinib or placebo. Randomisation was stratified by induction-trial group assignment and remission status at maintenance-trial entry. The duration of treatment was 52 weeks but any patient who met treatment failure criteria was required to withdraw from the study.

The CS reports the effects of tofacitinib treatment across a range of outcomes relevant to the NICE scope and the company decision problem, which are summarised below.

Remission is the primary outcome of the OCTAVE Induction trials and the OCTAVE Sustain maintenance trial. In both the OCTAVE 1 and OCTAVE 2 Induction trials, a statistically significant difference in remission at week 8 in comparison to placebo was observed in participants who received tofacitinib 10 mg twice daily. The same results were obtained regardless of whether centrally read or locally read endoscopic data were used (albeit the mean differences between the tofacitinib and placebo group were higher in both trials when using locally read endoscopic data). In the OCTAVE Sustain maintenance trial there was a statistically significant difference in remission at week 52 in comparison to placebo for participants who received tofacitinib 10 mg twice daily and those who received tofacitinib 5 mg twice daily. Locally read endoscopic data again produced less conservative results than centrally read endoscopic data. Sustained remission (remission at both week 24 and week 52) results were also in favour of tofacitinib.

Mucosal healing is a key secondary outcome of the OCTAVE trials. A statistically significant difference in the proportion of participants with mucosal healing in favour of tofacitinib was observed both at week 8 in the OCTAVE 1 and OCTAVE 2 induction trials as well as at week 52 in the OCTAVE Sustain trial.

The OCTAVE Sustain maintenance trial reported the outcome of sustained corticosteroid-free remission among those in remission at baseline in this trial. This outcome also favoured the

tofacitinib groups with statistically significant differences between the 5mg and the 10mg tofacitinib arms versus placebo.

Remission, mucosal healing and sustained corticosteroid-free remission did not contribute data to the economic model.

Clinical remission is an outcome with an almost identical definition to the primary outcome of remission. The difference being that the rectal bleeding sub-score of the Mayo score does not have to be zero to achieve clinical remission. The outcomes of clinical remission and clinical response contribute data to the economic model.

Using locally read data (which were used in the base case economic evaluation) in OCTAVE 1, the mean difference between the tofacitinib group and the placebo group was 13.3 percentage points (95% CI 6.5 to 20.2, p=0.0017). The corresponding data for OCTAVE 2 were a mean difference from placebo of 15.6 percentage points (95% CI 9.9 to 21.3, p=0.0002). At week 52 in the OCTAVE Sustain maintenance trial the results for clinical remission also favoured tofacitinib (difference versus placebo 35.1%, 95% CI 26.7 to 43.5, p<0.0001 using locally read data).

Clinical response at both week 8 (OCTAVE Induction trials) and week 52 (OCTAVE Sustain trial) was also statistically significantly higher among participants who received tofacitinib.

Subgroup analyses according to prior TNFi-exposure status were conducted for the main clinical effectiveness outcomes. The results were consistent regardless of prior TNFi-exposure status.

HRQoL was reported using generic (EQ-5D and SF-36) and disease specific (IBDQ and WPAI-UC) instruments. HRQoL was typically improved by tofacitinib treatment however for some HRQoL measures the ERG was uncertain about the impact of missing data. Data from the EQ-5D-3L did not inform the base-case economic model but were included in a scenario analysis.

Safety data for tofacitinib in patients with moderate to severely active ulcerative colitis comes from the Phase II trial, the three Phase III OCTAVE trials and the ongoing OCTAVE Open extension study. Rates of adverse events of any type were broadly similar for the tofacitinib and

placebo arms within each trial with serious adverse events affecting fewer than 10% of patients. Ulcerative colitis was the most frequent serious adverse event and most other serious adverse events were related to ulcerative colitis. Serious infections were uncommon (data on serious infections were included in the economic model). Overall, and in comparison with evidence from the use of tofacitinib in patients with rheumatoid arthritis, no new safety signals were identified.

There are no head-to-head RCTs of tofacitinib versus the comparators defined in the company's decision problem. Therefore the company used NMA to estimate the relative effectiveness and safety of tofacitinib in both the induction and maintenance phases of treatment in comparison to TNF-alpha inhibitors (infliximab, adalimumab and golimumab), vedolizumab and conventional therapies. The company's systematic review identified 21 RCTs that were considered for inclusion in the NMA. Four of these were the tofacitinib RCTs listed above, a further 14 were included in one or more NMA networks and three studies could not be included in any of the NMA networks.

Table 1 NMAs conducted by the company

	TNFi-naïve population subgroup	**TNFi-exposed population subgroup**
Induction phase	Clinical response and clinical remission Mucosal healing	Clinical response and clinical remission Mucosal healing
	Safety outcomes (discontinuation due to AEs, SAEs, serious infections)	
Maintenance phase	Clinical response and clinical remission Mucosal healing	Clinical response and clinical remission Mucosal healing

The ERG judged the NMAs to be generally well conducted but identified nine issues:

- Use of the probit scale to model clinical response/clinical remission is an improvement on a previous approach in NICE guidance TA342 but a multinomial logit model could have been considered.
- Potential inconsistency in a closed loop of the maintenance TNFi-naïve network was not examined

- The ERG would have made different choices regarding model fit, in general for the efficacy outcomes the ERG would have chosen the random effects model as the more conservative approach given the known between study heterogeneity. For the safety outcome of serious infections, the absence of any events in the placebo arms of the tofacitinib trials causes very wide credible intervals.
- The ERG was unable to replicate the same baseline (placebo) credible intervals used in the probit or logit models to estimate absolute probabilities. The company's estimates may be conservative.
- The phase II trial may have had a disproportionate effect on the random effect safety NMA because of the relatively high serious infection rate in the tofacitinib arm of this study.
- No safety NMA was conducted for the maintenance period.
- The company did not attempt to adjust for differences in lengths of induction and maintenance treatment and the ERG is concerned that this could have introduced potential bias against those treatments where studies had shorter induction phase and benefit those treatments with a shorter maintenance phase.
- There are differences between patient populations in the re-randomised design maintenance trials. OCTAVE Sustain re-randomised all responders from the OCTAVE induction trials to either placebo or tofacitinib treatment. In contrast, the other re-randomised studies, only re-randomised patients who had received and responded to active treatment into the maintenance phase of the study.
- Adjustments to treat-through trials were made, and although the ERG does not believe these introduce additional bias, it is nevertheless the case that non-responders at the end of the induction phase are ignored (and these participants potentially could have become responders by the end of the maintenance phase).

For the three outcomes synthesized by NMA which contribute data to the economic model the results were as follows.

The induction phase NMA for the TNFi-naïve population provided strong evidence of benefit for all treatments over placebo with infliximab having the largest treatment effect for both clinical response and clinical remission. In the TNFi-exposed population, tofacitinib had the largest treatment effect on clinical response and clinical remission compared to placebo. Only tofacitinib and vedolizumab showed strong evidence of benefit.

In the maintenance phase NMA for the TNFi-naive population all treatments showed strong evidence of benefit over placebo with tofacitinib 10mg having the largest treatment effect on clinical response and clinical remission. In the TNFi-exposed population, tofacitinib 10mg had the largest treatment effect on clinical response and clinical remission compared to placebo. Tofacitinib 5mg, 10mg and vedolizumab 300mg Q4W and Q8W all showed a strong evidence of benefit over placebo.

[redacted]

Summary of submitted cost effectiveness evidence

The company's submission includes a review of published cost-effectiveness evidence and a new economic model developed for this appraisal. The model compares the cost-effectiveness of Tofacitinib for treating people with moderately to severely active ulcerative colitis who are either intolerant of, or whose disease has had an inadequate response or loss of response to conventional therapy (oral corticosteroids and/or immunosuppressants) or a TNF-alpha inhibitor.

The company adheres the NICE scope; but excludes adalimumab as a comparator in TNfi-exposed sub group analysis

Broadly, the company model adheres with the NICE scope. We present a top-line view of the ERG's observations on patient characteristics, sub-groups and comparators included within the company model.

For their base case analyses, patient characteristics (including initial age, weight and gender mix) for the two sub-groups of TNFi-naïve and TNFi-exposed are based on means from the tofacitinib arms in the OCTAVE Induction trials. We view that these baseline characteristics should be assumed similar for people with and without prior exposure to TNFi drugs.

In line with the NICE scope, the company conducts sub-group analysis according to previous treatment with one or more biologics. However, as the NMA results used in the model are

defined by prior exposure to TNF-alpha inhibitors alone, we view it appropriate to label the sub groups based on status of patients' exposure to TNF-alpha inhibitors- i.e. TNFi- naïve and TNFI-exposed. The company presents cost-effectiveness analyses for these two sub-groups. In addition, they also present cost-effectiveness results based on analysis of the whole ITT population. We view that the company's 'ITT' cost-effectiveness scenario is highly uncertain and that it omits relevant comparators (the TNFi drugs), so does not address the specified decision problem. Hence, we focus on separate analyses for the two TNFi exposure subgroups in our discussion and additional analysis.

For patients in TNFi-exposed sub group, the company excludes adalimumab, infliximab and golimumab as comparators. Whilst clinical response and remission rates are not available for infliximab or golimumab in this sub group, but they are available for adalimumab. Hence, we consider adalimumab as a relevant comparator for at least some patients with prior exposure to a TNFi agent, although we understand that further treatment with a TNFi may not be appropriate for all patients in this subgroup.

The structure and assumptions of the submitted model are mostly reasonable, albeit a few issues

The company submitted a Markov cohort model consisting of 9 health states, with a cycle length of 8 weeks and patient lifetime horizon. Costs and QALYs were discounted at 3.5% annually. The model uses 3 sets of input parameters: clinical inputs (governing the rates of response and remission and adverse event rates for comparator treatments, as well as the incidence and complication/mortality rates for surgery), health state utilities; and resource use and costs.

The company assumes treatment effect to be maintained with ongoing treatment and non-responders are given conventional therapy as second-line. We agree with this approach which follows the independent economic analysis in TA329. However, the model does not reflect NICE recommendations for annual assessment of benefit and need for continued treatment in previous appraisals TA329 and TA342. Clinical advice suggests that withdrawal of treatment for patients in remission is unlikely in practice, and the effects of this are difficult to quantify given the model structure and limited evidence over long-term maintenance of remission.

The company model applies a constant risk of relapse across each 8-week cycle of maintenance, with treatment stopping immediately when patients lose response. We consider

this assumption to reflect UK practice. However, we have concerns that the costs of monitoring and follow-up in the company's model do not reflect the full cost of ensuring that treatment can be withdrawn within 8 weeks of a relapse. We address this in our additional analyses.

The company conducted NMA to inform clinical inputs within the model. To populate clinical remission and response, the company used a simple fixed effect approach. The ERG has a general preference for the random effect NMA models, as we believe that the fixed effect models may underestimate uncertainty due to heterogeneity between the studies.

In their base case NMA, the company combined outcomes for subgroups defined as TNFi-failed for vedolizumab with TNFi-exposed subgroups for tofacitinib and adalimumab. We consider that combining results for TNFi-failed and TNFi-exposed subgroups is a potential source of bias in favour of tofacitinib and view that using a more like-for-like comparison between tofacitinib and vedolizumab by using data for the TNFi-failed subgroups from the OCTAVE and GEMINI trials, is reasonable.

The company transformed the results of the clinical response/remission NMAs from the probit scale to the natural scale and converted to absolute probabilities to inform the economic model. For simplicity, they assume a constant ratio of patients in remission and response throughout maintenance phase and beyond in extrapolation. This is inconsistent with the clinical advice to the ERG as experience indicates the risk is greatest in the first 6-12 months; and falls thereafter. However, due to absence of evidence we were unable to adapt the model to reflect clinical evidence. Extrapolation of relapse and discontinuation rates from the maintenance trials is likely to underestimate the average duration of treatment and hence both the costs and QALYs of active treatments. However, it is not possible to estimate the net direction of bias in ICERs between comparators, because trends in long-term risks may vary between TNFi drugs, vedolizumab and tofacitinib.

Adverse events, except serious infections, were excluded from the economic analysis. We agree with the company's approach, but acknowledge that the omission of non-infection SAEs does introduce a risk of bias. However, given the frequency of these events, this is unlikely to influence the cost-effectiveness results. The company estimated risk of serious infections using a binomial logit NMA model in the induction trails and chose random effects model for their base case. However, there was considerable uncertainty around the model estimates. The ERG had

concerns as our verification checks indicated an even higher level of uncertainty around tofacitinib estimates, and we were unable to replicate the company's base case NMA values. We therefore applied a frequentist NMA approach to estimate the risk of serious infection, which we use as a scenario in ERG analysis.

We agree with the company approaches to modelling surgery risks, perioperative- and post-operative complications and mortality.

Health state utilities are estimated from published literature for the base case.

We agree with the company that the utility estimates by a published study by Woehl et al. provide an appropriate source for base case parameters. For scenario analysis, the company also conducted simple and regression-based analyses of EQ-5D data from the OCTAVE trials. However, these estimates are problematic as sources of utility parameters for the economic model due to the re-randomisation design and lack of intermediate assessments of clinical response and remission between week 8 and week 52 which complicate the interpretation of results.

In general, company's approach to costing is appropriate and consistent with related NICE guidance, albeit with a few errors in estimation

Costs and resources associated with drug acquisition, drug administration, monitoring and follow up and treatment of serious infections were included in the company's cost-effectiveness analyses. Overall, the costs inputs and sources used were appropriate although the ERG identified a few inconsistencies:

- We identified an error in the estimation of cost associated with elective surgery with complications which we corrected in the ERG corrected company's base case model.
- The company made an error in estimating weight wastage. Correction of this error had no influence on the base case results as they used 'fitting distribution' approach for wastage calculation.
- We noted a few minor changes in NHS prices for included drugs: sulfasalazine, prednisolone and azathioprine. The price changes lead to a very small decrement in the estimated cost of CT alone, with biologic drugs and with tofacitinib.
- No cost was assumed for administering adalimumab and golimumab which are administered by subcutaneous injection. We address this by assuming an initiation of self-administration of subcutaneous injections by adding the cost of a non-consultant led

clinic attendance to the cost of induction for adalimumab and golimumab in our additional analyses

- Health care usage assumptions were made based on the study by Tsai et al. (2008). Whilst we agree with the company's approach for the base case, we conduct scenarios testing alternative resource use based on expert advice.
- We question company's assumption that maintenance treatment will always stop within 8 weeks of a loss of response which is consistent with the number of outpatient appointments. We test this assumption in our additional analyses
- The company excludes cost of stoma care and the estimated cost of surgery is low compared with previous appraisals. We address these in our additional analysis.

Company's base case results

The company's base case results are presented in

Table 2 and Table 3

Table 2 Cost effectiveness: Company base case, no prior TNFi (with tofacitinib PAS)

Strategy	Total		Incremental analysis			Pairwise ICERs tofacitinib vs. comparator (£/QALY)
	QALYs	Costs (£)	QALYs	Costs (£)	ICER (£/QALY)	
Conventional	[redacted]	[redacted]	-	-	-	£8,554
Adalimumab	[redacted]	[redacted]	-	-	Dominated	Dominated
Golimumab	[redacted]	[redacted]	-	-	Dominated	Dominated
Infliximab	[redacted]	[redacted]	-	-	Dominated	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£8,554	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£615,057	£615,057

Table 3 Cost effectiveness: Company base case, with prior TNFi (with tofacitinib PAS)

Strategy	Total		Incremental analysis			Pairwise ICERs tofacitinib vs. comparator (£/QALY)
	QALYs	Costs (£)	QALYs	Costs (£)	ICER (£/QALY)	
Conventional	[redacted]	[redacted]	-	-	-	£10,302
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£10,302	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£7,838,238	£7,838,238

A range of uncertainty analyses were conducted by the company, but they have been selective in the scenarios they present

The company performed a range of deterministic-, probabilitistic- and scenario analyses to assess the methodological as well as parameter uncertainty of their base case analyses. The ERG agrees with their assumptions for DSA and PSA and their results, in general. However, we identified errors in the scenarios relating the use of central read NMA results and tofacitinib maintenance using [redacted] split. The company corrected the error in the latter scenario in their response to clarification question. For the scenario analyses, we view that the company has been selective in the scenarios they present.

Commentary on the robustness of submitted evidence

Strengths

- The model structure is consistent and follows the conventional design for ulcerative colitis appraisals.
- The model generally adheres to the NICE scope for this appraisal.
- The perspective of the analysis aligns with the NICE guide to the methods of Technology Appraisal.
- The model uses a lifetime time horizon to allow estimation of all relevant costs and quantity of life impairment.
- The model uses appropriate sources for costs and resource use and in line with other technology appraisals
- The model allows the flexibility to incorporate treatment sequencing which provides a closer reflection of clinical practice.
- The ERG agrees with the company's approach to modelling surgery and its related risks, source of costs and utilities for the base case and mortality.
- The economic model was of good quality, with very few errors in input parameters, logic or coding.
- In the TNFi-naïve arm, the model results were comparable with the clinical data for the tofacitinib arm.

Weaknesses and Areas of uncertainty

- For their base case analyses, patient characteristics (including initial age, weight and gender mix) for the two sub-groups of TNFi-naïve and TNFI-exposed are based on means from the

tofacitinib arms in the OCTAVE Induction trials. We view that these baseline characteristics should be assumed similar for people with and without prior exposure to TNFi drugs.

- We do not consider the company's cost effectiveness analysis with the ITT population scenario to be reliable due to the high level of uncertainty in the underlying NMA. The scenario also omits relevant comparators (the TNFi drugs), so does not address the specified decision problem.
- The company excludes adalimumab as a comparator for patients with prior exposure to a TNFi, despite available evidence to support this.
- The company assumes equal use of 4 drugs in aminosalicylate class (balsalazide, mesalazine, olsalazine & sulfasalazine). However, clinical advice to ERG suggests most patients receive mesalazine in UK and the doses for active ulcerative colitis are potentially higher than specified in company base case.
- The company assumes treatment effect to be maintained with ongoing treatment and non-responders are given conventional therapy as second-line. However, the economic model does not reflect NICE recommendations for annual assessment of benefit and need for continued treatment in previous appraisals TA329 and TA342.
- The company model applies a constant risk of relapse across each 8-week cycle of maintenance, with treatment stopping immediately when patients lose response. Although this is reflective of UK practice, the costs of monitoring and follow-up in the company's model do not reflect the full cost of ensuring that treatment can be withdrawn within 8 weeks of a relapse.
- The company use fixed effects NMA models to inform the economic model . The ERG has a general preference for the random effect NMA models, as we believe that the fixed effect models may underestimate uncertainty due to heterogeneity between the studies.
- The base case NMAs combine outcomes for subgroups defined as TNFi-failed for vedolizumab with TNFi-exposed subgroups for tofacitinib and adalimumab. Combining results for TNFi-failed and TNFi-exposed subgroups introduces a potential source of bias in favour of tofacitinib.
- The company assume constant ratio of patients in remission and response throughout maintenance phase and beyond in extrapolation. These assumptions might not be realistic as clinical -experience indicates the risk is greatest in the first 6-12 months; and falls thereafter.
- There is considerable uncertainty in the NMA estimates for risks of serious infections. We have reservations about the company's approach to estimating this parameter as our

verification checks indicated an even higher level of uncertainty around tofacitinib estimates, and we were unable to replicate the company's base case NMA values.

- The company's simple and regression-based analyses of EQ-5D data from the OCTAVE trials are problematic as sources of utility parameters for the economic model. They are relevant to the decision problem and clinical evidence, but the re-randomisation design and lack of intermediate assessments of clinical response and remission between week 8 and week 52 complicate the interpretation of results.
- The company did not include any costs associated with an initiation of self-administration of subcutaneous injections for adalimumab and golimumab
- The company excludes cost of stoma care and the estimated cost of surgery is low compared with previous appraisals.

Summary of additional work undertaken by the ERG

The ERG conducted a number of scenario analyses. Our preferred assumptions, alongside the scenarios are presented in Table 4.

Table 4 ERG's preferred assumptions and scenarios

Aspect of the model		ERG preferred	ERG scenarios	
Patients	Age (yrs)	Average of all patients in OCTAVE 1 and 2: 41	Range: 28-52	
	Weight (kgs)	Average for all patients in OCTAVE 1 and 2: 73.5	Range: 70-80 kg	
Comparator	TNFi-exposed	Include adalimumab		
	Treatment sequencing	No change	INF-ADA-CT INF-VED-CT INF-TOF-CT VED-ADA-CT	GOL-ADA-CT GOL-VED-CT GOL-TOF-CT ADA-VED-CT

			TOF-ADA-CT	ADA-TOF-CT
NMA models	Remission and response rates	Use RE except for TNFi-experienced maintenance (RE would not run)	FE for both subgroups, induction and maintenance	
		No change	Use TNFi-failed for both vedolizumab and tofacitinib with TNFi-experienced for adalimumab	
	Serious infections	Frequentist random effects NMA model	Bayesian random effect model	
Utilities	Sources for pre and post-surgery health states	Same as company	• Swinburn et al. • OCTAVE 8 weeks • OCTAVE 52 weeks	
Resource use and costs	Drug stopping rule	Same as company	Additional OP visits to assess response within 8 weeks	
	Conventional drug usage	Same as company	Patient use of mesalazine: 50.3% (CT), 46.2% (concurrent). No other aminoslicylates	
	Health state resource use	Same as company	Reduced admissions, outpatient follow up and endoscopy	
	Drug administration costs	Same as company	Assume 1 OP visit at start of treatment for training on subcutaneous injections	
	Hospitalisation and surgery costs	NHS Reference costs + cost of stoma care post surgery (Buchanan et al. uprated for inflation)	Buchannan et al. estimate of surgery cost (uprated to 2016/17 prices) – includes repeat procedures	
Surgery	Incidence rate	Same as company	Chhaya et al.	
	Complications	Same as company	Tappenden et al.: Probabiity of perioperative complications (elective 0.2386; emergency	

			0.2614), probability of post surgery complications (0.173)

The results of the ERG's preferred assumptions are presented in Table 5. Collectively, our preferred assumptions give very similar results to the company's model. TNF-inhibitors remain dominated (with higher costs and fewer QALYs) than tofacitinib in both the sub-groups. While the pairwise ICER for tofacitinib compared with vedolizumab fall in the south-west quadrant (meaning tofacitinib is less effective but also less costly than vedolizumab) in the TNFi-naïve subgroup; in patients with prior exposure to TNFi, vedolizumab is dominated by tofacitinib under our preferred set of assumptions.

Table 5 Cost effectiveness: ERG preferred assumptions (with Tofacitinib PAS)

TNFi- naïve			
Conventional	[redacted]	[redacted]	£7,815
Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
Golimumab	[redacted]	[redacted]	Tofacitinib dominant
Infliximab	[redacted]	[redacted]	Tofacitinib dominant
Tofacitinib	[redacted]	[redacted]	--
Vedolizumab	[redacted]	[redacted]	£607,571 (SW)
TNFi-exposed			
Conventional	[redacted]	[redacted]	£9,389
Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
Tofacitinib	[redacted]	[redacted]	--
Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

1 Introduction to ERG Report

This report is a critique of the company's submission (CS) to NICE from Pfizer on the clinical effectiveness and cost effectiveness of tofacitinib for moderately to severely active ulcerative colitis. It identifies the strengths and weakness of the CS. Clinical experts were consulted to advise the ERG and to help inform this review.

Clarification on some aspects of the CS was requested from the company by the ERG via NICE on 12 June 2018. A response from the company via NICE was received by the ERG on 27 June 2018 and this can be seen in the NICE committee papers for this appraisal.

2 BACKGROUND

2.1 Critique of the company's description of the underlying health problem

The CS provides a generally clear and accurate overview of ulcerative colitis (CS section B.1.3).

Aetiology

Ulcerative colitis is a chronic inflammatory disease that is characterised by relapsing and remitting mucosal inflammation which typically affects the rectum and extends proximally to affect either a variable area of the colon, or its entire mucosal surface.[1,2] Ulcerative colitis is classified as proctitis, left-sided colitis, or extensive colitis, according to its maximal extent seen on colonoscopy. The CS (citing the British Society of Gastroenterology (BSG) guidelines for the management of inflammatory bowel disease in adults[3]) states that about 50% of patients with ulcerative colitis have a relapse in any year (CS section B.1.3.1). The NICE scope concurs that an estimated 30–60% of people with ulcerative colitis will have at least one relapse per year, of which about 80% are mild to moderate and about 20% are severe. However, the BSG guidelines[3] and the NICE scope do not specify the sources of these data. Patients with more extensive disease are at greater risk of developing dysplasia or colorectal cancer, and are generally advised to have surveillance colonoscopy.[4]

The pathogenesis of ulcerative colitis is complex and multifactorial, involving genetic predisposition, defects of the intestinal epithelial barrier, dysfunction of immune responses, and environmental factors.[2] The CS emphasises the importance of understanding the role of the

immune system and inflammatory cascade for understanding the disease and the role of current and future treatment options.

Risk factors

Risk factors for ulcerative colitis are not specified in the CS, but include: a family history of inflammatory bowel disease; Jewish ethnicity; the use of oral contraceptives, hormone replacement therapy, or non-steroidal anti-inflammatory drugs (NSAIDs); and former cigarette smoking. Conversely, in active smokers and some people who have had an appendectomy the risk of developing ulcerative colitis is reduced. Males and females do not differ in their risk of developing ulcerative colitis.[2,4]

Symptoms

Patients typically present with bloody diarrhoea, and some may also have rectal bleeding, urgency, faecal incontinence, nocturnal defecation and fatigue. Greater severity and extent of disease are associated with worsening bloody diarrhoea and the development of systemic signs, but any extent of colitis can be associated with constitutional symptoms, including fatigue and fever.[2,4]

Diagnosis

A gold standard for diagnosing ulcerative colitis is not available. Diagnosis is based on the history of symptoms, endoscopic findings on colonoscopy, histology, and excluding other causes of colonic inflammation (e.g. infection).[1] The key feature of ulcerative colitis on endoscopy is a diffuse continuous mucosal inflammation of the rectum and a variable extent of the colon. Other typical findings include erythema, loss of the normal vascular pattern, bleeding, erosions and ulcerations. The extent of inflammation observed on colonoscopy is related to the risk of disease complications.[1,2] Ulcerative colitis may be diagnosed at any age, but most commonly affects adults aged in their 20s to 40s (CS section B.1.3.1).

Severity of ulcerative colitis is classified as mild, moderate or severe based on a combination of factors which include, among others, the number of bowel movements per day and presence or absence of blood in the stool.[5,6]

Incidence and prevalence

The UK has among the highest incidence and prevalence rates of ulcerative colitis in the world.[7] The incidence of ulcerative colitis in the UK has been estimated at around 13.9 per 100,000 people, with a prevalence around 243 per 100,000 people. The most recent estimate available, for 2011, suggests that there were approximately 146,000 people in the UK who had ulcerative colitis.[3,7] The CS acknowledges that this may be a substantial underestimate, given the broad age of onset and lifelong duration of the condition (CS section B.1.3.1).

2.2 Critique of the company's overview of current service provision

Current treatments for moderately to severely active ulcerative colitis may be pharmacological or surgical, with all patients managed pharmacologically initially, before surgery in some cases. Clinical advice to the ERG is that surgery is reserved for patients who are non-responsive to the available drug treatments. Rarely, surgery may be carried out earlier if absolutely necessary, e.g. if a patient has a high risk of colorectal cancer.

Patients with moderately to severely active ulcerative colitis are typically managed according to a step-up approach based on the patient's history, treatment response and tolerance of individual therapies. Patients who have an inadequate response to conventional therapies (aminosacylates, corticosteroids or thiopurines) may be offered a biological therapy (a tumour necrosis factor (TNF) inhibitor or the anti-integrin agent vedolizumab).[8,9]

The CS briefly describes the clinical pathway of care (CS section B.1.3.3; discussed further below in section 2.3) but does not mention the staff, infrastructure or other resources associated with current service provision for patients with moderately to severely active ulcerative colitis, or whether these would change if tofacitinib is recommended for patients in the NHS. Clinical advice to the ERG is that a nurse-led service is used for intravenous therapies. This would not be applicable for tofacitinib which is administered orally. Tofacitinib can be taken with or without food, and the tablet can be crushed if patients have swallowing difficulties, so we assume that treatment self-administration by patients would be straightforward.

2.3 Critique of the company's definition of the decision problem

Population

The population stated in the NICE scope is "people with moderately to severely active ulcerative colitis who are intolerant of, or whose disease has had an inadequate response or loss of response to conventional therapy (oral corticosteroids and/or immunosuppressants) or a TNF-alpha inhibitor". This is consistent with the indication as specified in the Summary of Product Characteristics (SmPC),[10] as acknowledged by the company in CS Table 2.

In their decision problem table (CS Table 1) the company gives a broader description of the population, as "people with moderately to severely active ulcerative colitis". This description is applied to the NICE scope column within the decision problem table, thereby inaccurately reflecting the NICE scope. The company confirmed that this is a semantic error in CS Table 1 and that the decision problem does reflect the NICE scope (clarification response A1).

We note that the pivotal Phase III trial populations (in OCTAVE 1, OCTAVE 2 and OCTAVE Sustain) are consistent with the population definition as given in the NICE scope. However, according to the trial publication and protocol,[11] patients in the tofacitinib Phase II trial did not have to be intolerant of, or have had an inadequate response or loss of response to conventional therapy. The indication for tofacitinib in the Phase II trial therefore does not appear to be consistent with the NICE scope. In response to a clarification question from the ERG, the company stated that patients were only included in the Phase II trial if they continued to have moderate to severe disease despite previous treatment, and the company provided supporting data on the baseline characteristics of the Phase II trial participants listing the proportions who had failed prior treatments (clarification response A2). Although it is not clear from the CS or trial publication, the Phase II trial therefore does appear to meet the NICE scope.

The population in the pivotal OCTAVE Induction trials had a mean age of around 40-42 years (CS Table 15) and age ranged from 18 to 81 years [Table 13 in each clinical study report (CSR)]. Expert advice from one advisor to the ERG is that patients presenting in NHS clinical practice would typically be younger than the mean age in the trials, with the peak age at presentation being nearer 20 years on average. Although younger patients tend to have more severe ulcerative colitis;[4] the ERG's clinical advisor suggested that the age difference between the trials and clinical practice would be unlikely to affect patients' disease characteristics or their treatment.

Intervention

The intervention is tofacitinib citrate, a 5 mg or 10 mg oral tablet, brand name Xeljanz®. The description of the intervention in CS Table 2 is consistent with the SmPC.[10] The positive CHMP opinion, which is consistent with the NICE scope, was adopted by the EMA on 31st May 2018 for tofacitinib 10 mg to be used in the treatment of adult patients with moderately to severely active ulcerative colitis who have had an inadequate response, lost response, or were intolerant to either conventional therapy or a biologic agent.

Tofacitinib is a small-molecule selective inhibitor of the Janus kinase (JAK) family of tyrosine kinases. The inhibition of JAKs by tofacitinib attenuates the signalling of several interleukins and type I and II interferons, which leads to modulation of the immune and inflammatory response in ulcerative colitis.[10] The mode of action of tofacitinib, including its role in inhibition of the JAK-STAT pathway, is summarised in CS section B.1.3.4.

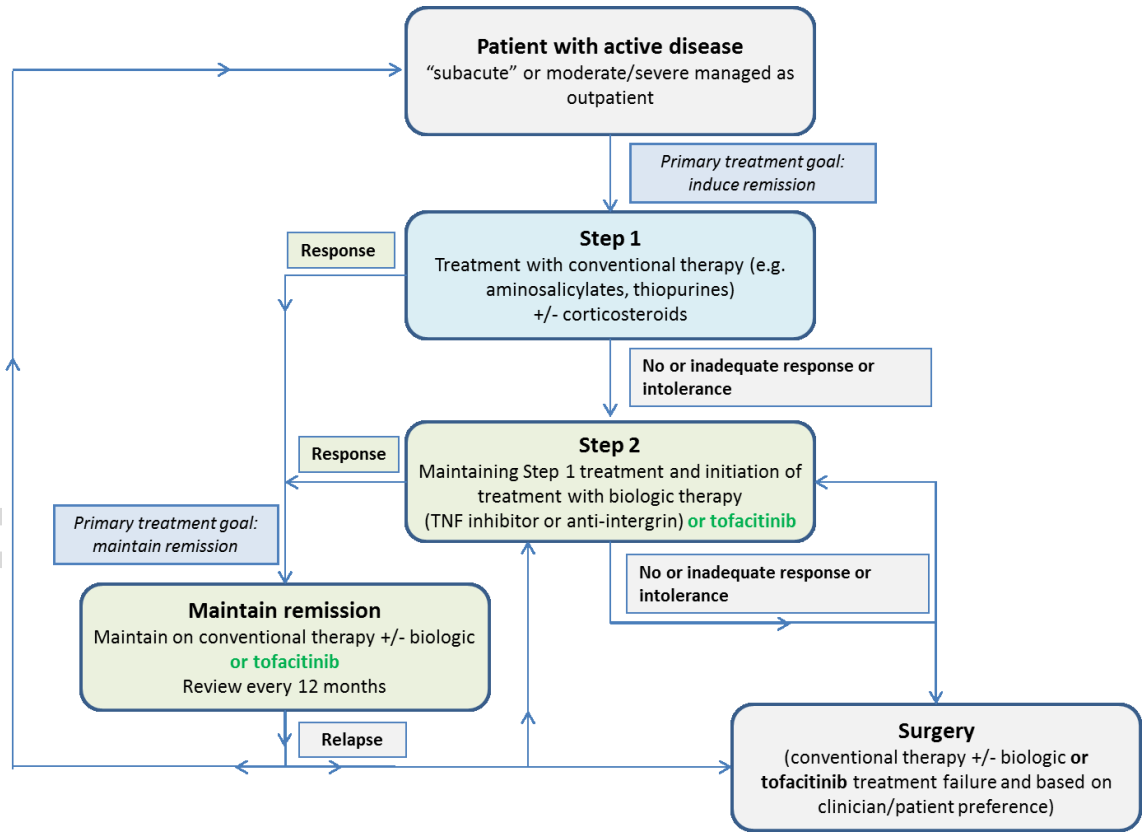
Comparators

The comparators in the company's decision problem (CS Table 1) are as specified in the NICE scope. These are:

- Conventional therapy, which may include a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclometasone, budesonide, hydrocortisone or prednisolone), thiopurines (mercaptopurine or azathioprine);
- TNF inhibitors (infliximab, adalimumab and golimumab[8]);
- Vedolizumab[9] (an anti-integrin agent).

These comparators are all used in the NHS (although experts advising the ERG commented that there has been a decline in the usage of thiopurines). CS Figure 1 illustrates the stepwise manner in which these therapies would be used in clinical practice in the NHS, showing that the conventional therapies are the standard first-line approach (step 1) whilst biological therapies (step 2) would not be employed without first trying a conventional therapy. Experts advising the ERG agreed that CS Figure 1 does broadly reflect current NHS practice. In response to a clarification request from the ERG, the company stated that CS Figure 1 is based on current NICE guidelines and clinical practice, but it is "*a simplification of the clinical pathway as the treatment of ulcerative colitis is dependent on multiple factors, the including patient's medical*

history and clinical decision making on the appropriateness of therapies, and therefore may not adequately capture the nuances of clinical practice when comparing to the NICE scope" (clarification response A3). In their clarification response the company provided a simplified version of CS Figure 1 in order to better represent the position of tofacitinib in the treatment pathway in relation to the NICE scope (reproduced in Figure 1).



Source: company's clarification response A3

Figure 1 Proposed position of tofacitinib within the treatment pathway

Outcomes

The outcomes included in the CS are clinically meaningful and are consistent with the NICE scope and EMA guidance on methods for clinical trials in ulcerative colitis.[12] The primary outcome in the phase 3 OCTAVE trials was clinical remission whilst the primary outcome in the phase 2 trial was clinical response. HRQoL was a secondary outcome in all the tofacitinib trials, and mucosal healing was a secondary outcome in the phase 3 trials. Details of the outcome selection are discussed further below in section 3.1.4. In summary, the key issues noted by the ERG are:

- Time to surgical intervention, listed as an outcome in the NICE scope, is not reported in the CS as it was not assessed in the pivotal trials (CS Table 1)
- The CS only provides brief results from the Phase II trial, for the primary outcome only (further results were requested by the ERG)
- The CS does not report all of the patient-reported outcomes that were measured in the pivotal trials (although as noted below in section 3.1.3 this does not appear likely to have resulted in bias)

Other relevant factors

The NICE scope indicates that, if evidence allows, subgroups of people who have been previously treated with one or more biologics and people who have not received prior biologics should be considered. Although the company presents subgroup analyses in their submission (CS Appendix E) their focus is on subgroups of people by TNFi-exposure status. There is no subgroup analysis for subgroups of people by prior biologic therapy (biologic therapy would include not only the TNF inhibitors but also vedolizumab). Nevertheless the ERG is mindful that subgroups by TNFi-exposure status are important, particularly because the existing evidence base for comparator treatments has demonstrated that primary non-response and secondary non-response to TNFi agents are limitations of the existing therapies adalimumab, golimumab and infliximab.

The CS does not identify any inequities that could be associated with the provision (or non-provision) of tofacitinib (CS section B.1.4) and the ERG is not aware of any equality issues with tofacitinib.

3 CLINICAL EFFECTIVENESS

3.1 Critique of company's approach to systematic review

3.1.1 Description of company's search strategy

The company conducted six systematic literature searches (five of which are standard for an STA, plus an additional surgery review):

- Clinical effectiveness: start year unspecified ("no limits") to 16/11/2017
- Non RCT evidence: start year unspecified ("no limits") to 15/11/2017

- Cost effectiveness: start year unspecified ("no limits") to October 2017
- Health Related Quality of Life: start year unspecified ("no limits") to 15/11/2017
- Cost and healthcare resource identification: start year unspecified ("no limits") to 20/10/2017
- Surgery Literature Review - dates not given

The key literature searches were systematic, transparent, well documented and reproducible. A typographical error was found in line 18 of the cost effectiveness searches in Medline and Embase ("mdel*" instead of model*) however correct spelling elsewhere in both of these strategies coupled with accurate spelling in the Cochrane search, should have counteracted this error. The additional surgery review was undertaken to inform the economic analysis on the probability of colectomy and ensuing complications. This search is not fully documented, although a synopsis of the terms used are recorded which is acceptable. Key conferences were adequately searched and ongoing trials were sought via clinicaltrials.gov and the WHO International Clinical Trials Registry Platform.

Overall the searches are deemed fit for purpose. However, all searches in the CS are between six to eight months out of date. Due to time constraints the ERG has prioritised updating the cost effectiveness, HRQoL, and cost & healthcare resource searches, replicating the documented strategies. Two additional cost-effectiveness papers were identified by the ERG's updated search (see Section 4.2) but no additional relevant references were identified by the updated HRQoL or healthcare resource searches.

To identify any new clinical effectiveness evidence, we conducted a rapid search using HDAS (NICE Healthcare Databases Advanced Search) and Delphis (a broad-scope University of Southampton search engine powered by Ebsco). Four additional full-text publications on tofacitinib clinical effectiveness and/or safety which are not listed in the CS were identified. These reported on: the phase II trial;[13] a NMA comparing tofacitinib against biologic therapies;[14] an analysis of HRQoL in the OCTAVE trials using the Inflammatory Bowel Disease Questionnaire (IBDQ) and SF-36;[15] and a subgroup analysis of effectiveness and safety outcomes from the OCTAVE trials in East Asian participants.[16] Two new conference abstracts reporting results from the OCTAVE trials were also identified.[17,18] These new publications largely duplicate information already present in the CS, or are not directly relevant to the current scope.

3.1.2 Statement of the inclusion/exclusion criteria used in the study selection.

The company provides a clear description of the eligibility criteria (both inclusion and exclusion criteria) for the systematic literature review (SLR) (Appendix D.1.1.3, Table 83). The SLR aimed to identify clinical effectiveness and safety evidence not only for tofacitinib but also for relevant comparators, which could potentially be used in network meta-analysis (NMA).

The population eligibility criteria are specified as 'adult patients with moderately and/or severely active UC' (either treatment-naïve or treatment-experienced). The criteria exclude patients with non-specific inflammatory bowel disease (IBD) and those with acute severe ulcerative colitis or ulcerative colitis exacerbation/flare requiring hospitalisation as well as paediatric patients and animal/in vitro studies. This population is reflective of the decision problem (CS Table 1), although, as noted in section 3.2 above, the population in the company's decision problem is broader than the population specified in the NICE final scope.

The company confirmed (clarification response A1) that the population had been inaccurately described in their decision problem and that their interpretation of the population is in fact consistent with that described in the NICE scope. Whilst the population eligibility criteria as explicitly stated in CS Table 83 are wider than the NICE scope, the ERG is satisfied that the populations in the studies finally included in the company's SLR are consistent with the NICE scope.

The interventions and comparators for the company's SLR generally reflect the NICE scope, the anticipated licensed indication for tofacitinib and current NHS practice. Calcineurin inhibitors and surgical intervention were excluded as comparators in the SLR, whereas placebo was included as a comparator (not specified in the NICE final scope). Clinical expert advice to the ERG suggests that the exclusion of calcineurin inhibitors and surgical intervention is reasonable. It is presumed that the company included placebo as a comparator because it is the comparator in the OCTAVE trials, on which the clinical evidence for tofacitinib is based.

To be included studies had to RCTs (both blinded and open-label RCTs were eligible) and had to report at least one of the following outcome measures: response, remission, mucosal healing, relapse or loss of response/remission, discontinuation, treatment duration, rates of surgical

intervention, time to surgical intervention, Mayo score/Disease activity index, hospitalisation, mortality, adverse events (AEs), serious AEs (SAEs), treatment-related AEs, injection or infusion site reaction and HRQoL (EQ-5D, SF-36, IBDQ).

Setting did not form part of the company's eligibility criteria for the SLR. The company placed no limits on the quality of the included RCTs in their eligibility criteria. The ERG agrees this is appropriate.

The CS provides a flow diagram illustrating the number of records identified in the SLR and reasons for the exclusion of studies at the full text screening stage (CS Appendix D.1.2.1, Figure 39). References linked to the included studies are listed in CS Appendix D.1.2.1, Table 84. It was difficult for the ERG to equate these to the 102 references listed in the CS. In response to a request from the ERG, the company provided the information more clearly, including a correction to a referencing error (clarification response A6 and clarification response Appendix C). A list of the 137 references excluded at the full text stage of the reference screening process is not included in the CS. This was subsequently provided by the company (clarification response A5 and clarification response Appendix B).

The evidence was limited to studies published in the English language, which the ERG considers appropriate for a submission to NICE. However, the company did not discuss any potential bias that may have arisen from the restrictions of the eligibility criteria specified for the SLR. The ERG notes that RCTs are, by design, potentially at a lower risk of bias than other study design and that all the included RCTs were subject to quality assessment using the concise critical appraisal checklists provided by NICE in the STA user guide (CS D.1.2.2.2 Table 86).

3.1.3 Identified studies

The company's SLR included 21 RCTs. In four of these the intervention was tofacitinib:

- one Phase II RCT (treatment arms: tofacitinib 0.5 mg, 3 mg, 10 mg, 15 mg and placebo)[11]
- two identical Phase III induction RCTs (OCTAVE Induction 1 and OCTAVE Induction 2; treatment arms: tofacitinib 10 mg and placebo) (also a 15 mg arm which was discontinued)[19]
- one maintenance RCT (OCTAVE Sustain; treatment arms: 5 mg, 10 mg and placebo)[19]

In all four RCTs the comparator was placebo. All these RCTs were used in support of the company's application for a marketing authorisation and were sponsored by Pfizer, the manufacturer of tofacitinib.

The Phase II trial is not described in detail in the CS but it is included in the company's NMA (CS section B.2.9) and data from this trial are also included in the adverse events section (CS Appendix F Table 166). As the Phase II trial was a small dose-finding study with 194 patients, of whom only 33 received the licensed 10 mg BID dose (company clarification response A16), the CS focuses on the Phase III trials. The ERG agrees that this is reasonable and accordingly the current ERG report also focuses primarily on the Phase III trials.

It was unclear to the ERG from the description of the Phase II trial population reported both in the CS and in the trial publication whether this matched the NICE scope. The company confirmed that it does match the scope, as "*patients were only included if they continued to have moderate to severe disease despite previous treatment*" (clarification response A2). In addition, the company provided a table detailing the failed drug treatments at baseline (clarification response Table 1) and full details of the inclusion and exclusion criteria (clarification response Appendix A).

The number of centres in the studies ranged from 51 (Phase II trial) to 297 (OCTAVE Sustain), but it should be noted that a number of centres in the Phase III trials randomised just one patient (16 centres in OCTAVE 1; 25 centres in OCTAVE 2; and 66 centres in OCTAVE Sustain[19]). While each study included some patients from the UK, this number was low [redacted].

OCTAVE 1 and 2 were double-blind, randomised placebo-controlled tofacitinib induction trials with an 8 week treatment phase, and used identical methods (see Table 6).

In addition to the criteria listed above, patients had to have moderately to severely active disease (6 to 12 on the Mayo score, with a rectal bleeding sub-score of 1 to 3 and an endoscopic sub-score of 2 or 3). Prohibited therapies included TNFi therapies within 8 weeks of baseline; azathioprine, methotrexate, and 6-mercaptopurine within 2 weeks; and ciclosporin and intravenous corticosteroids (CS Tables 9 and 10). Permitted concomitant medications for ulcerative colitis included oral aminosalicylates (stable dose ≥4 weeks prior to baseline and

during study); oral glucocorticoids (maximum dose 25 mg per day of prednisone or a prednisone equivalent; stable dose ≥2 weeks prior to baseline and during study); and antibiotics used for chronic ulcerative colitis (e.g., metronidazole and rifaximin; stable dose ≥2 weeks prior to baseline and during study). Eligible patients were randomised on a 4:1 ratio to 10 mg twice a day (BID) of oral tofacitinib or placebo. The trials initially included a third treatment arm of 15 mg BID oral tofacitinib, but this was discontinued prior to full recruitment based on feedback from regulatory authorities. The company clarified that patients assigned to the tofacitinib 15 mg BID arm continued to receive blinded treatment for the remainder of the induction trial period and, of these, 19 patients were eligible to enter the OCTAVE Sustain trial (clarification response A10).

Patients were eligible to join the OCTAVE Sustain trial if they: met the eligibility criteria of the OCTAVE Induction trials; completed the 8 weeks of induction therapy; and met the clinical response criteria for the induction trials (see Figure 2). This was a randomised, double-blind, placebo-controlled trial lasting 52 weeks. Eligible patients from OCTAVE 1 and 2 were randomised in a 1:1:1 ratio to receive either 5 mg or 10 mg BID oral tofacitinib, or placebo.

The OCTAVE induction and maintenance trials conform to a re-randomisation design. That is, participants are first randomised to tofacitinib or placebo groups of the OCTAVE Induction study. Following 8-weeks of induction therapy, those participants who have met clinical response criteria are re-randomised into one of the three arms of the OCTAVE Sustain maintenance study. An alternative, utilised by some of the other clinical trials that have taken place in this disease area, is a treat-through design. In a treat-through trial participants are randomised to induction therapy and outcomes are measured at the end of the induction phase. Participants then continue in their original randomised group into the maintenance phase and outcomes are measured again at the end of the maintenance phase.

In addition to the four RCTs the OCTAVE study programme also includes the OCTAVE Open extension study which is ongoing.

Table 6 Summary characteristics of tofacitinib RCTs

Phase II trial[11] (efficacy/dose RCT)		**OCTAVE 1[19] (induction RCT)**		**OCTAVE 2[19] (induction RCT)**		**OCTAVE Sustain[19] (maintenance RCT)**		**OCTAVE Open[20] (extension study)**
Tofacitinib 0.5 mg (n=31) 3 mg BID (n=33) 10 mg BID (n=33) 15 mg BID (n=49)	Placebo (n=48)	Tofacitinib 10 mg BID (n=476)[a]	Placebo (n=122)	Tofacitinib 10 mg BID (n=429)[a]	Placebo (n=112)	Tofacitinib 10 mg BID (n=197) 5 mg BID (n=198)	Placebo (n=198)	Tofacitinib[b] 10 mg BID ([redacted]) 5 mg BID ([redacted])
Design: randomised, double- blind, placebo-controlled trial (2:2:2:3:3 ratio tofacitinib 0.5 mg: 3mg: 10 mg: 15 mg: placebo)		*Design:* identical randomised, double-blind, placebo-controlled trials (4:1 ratio tofacitinib: placebo, stratified according to previous treatment with TNFi therapies, glucocorticoid use at baseline, and geographic region)				*Design*: randomised, double-blind, placebo-controlled trial (1:1:1 ratio tofacitinib 5 mg: tofacitinib 10 mg; placebo)		*Design:* open-label extension
Location: 51 sites worldwide (UK = 2, [redacted][d])		*Location:* 144 sites worldwide (UK = 2, [redacted])		*Location:* 169 sites worldwide (UK = 3, [redacted])		*Location:* 297 sites worldwide (UK = 5, [redacted])		*Location:* 215 sites worldwide (UK = 5)
Inclusion: • age ≥18 years • confirmed diagnosis of UC for ≥3 months • score of 6 to 12 on the Mayo scale and		*Inclusion:* • age ≥18 years • confirmed diagnosis of UC for ≥4 months • moderately to severely active disease (6 to 12 on the Mayo score, with a rectal bleeding sub score of 1 to 3 and an endoscopic sub-score of 2 or 3)				*Inclusion:* • entry criteria for the Induction trials • completed 8 weeks induction therapy		*Inclusion:* • completed or demonstrated treatment failure in the OCTAVE Sustain maintenance study or

• moderately or severely active disease (i.e. Mayo-endoscopic findings sub-score of 2 or 3, respectively)	• treatment failure with/to or unacceptable side effects from treatment with ≥1 of: ○ oral or intravenous glucocorticoids ○ azathioprine ○ mercaptopurine ○ infliximab ○ adalimumab	• met clinical response criteria in OCTAVE Induction 1 and 2	• non-responders after completed 8 weeks of treatment in the OCTAVE 1 & 2 induction studies
Background therapy: Oral mesalamine or oral prednisone at a stable dose of ≤ 30 mg per day	*Background therapy:* oral aminosalicylates at a stable dose for ≥4 weeks prior to baseline and during study and oral glucocorticoids (at a maximum dose of 25 mg per day of prednisone or a prednisone equivalent) at a stable dose for ≥2 weeks prior to baseline and during study. Patients on chronic treatment for UC with antibiotics (e.g. metronidazole and rifaximin) were eligible if dose was stable for ≥2 weeks prior to baseline and during study.	*Background therapy:* oral amino-salicylates (stable dose) and chronic treatment for UC with antibiotics (e.g., metronidazole, rifaximin). Oral glucocorticoids at study entry were tapered mandatory starting 1st week at specified rate depending on starting dose (daily dose of prednisone or equivalent was decreased at a rate of 5 mg per week until dose reached 20 mg/day, then 2.5 to 5.0 mg per week until dose reached 10 mg/day, then by 2.5 mg per week until the dose was 0 mg).	*Background therapy:* oral aminosalicylates (stable dose) and chronic treatment for UC with antibiotics (e.g., metronidazole, rifaximin). Oral glucocorticoids at study entry were tapered mandatory as per the OCTAVE Sustain schedule.

Length of follow-up: 8 weeks of treatment and 4 weeks follow-up	*Length of follow-up:* 9 weeks (primary efficacy endpoint at 8 weeks)	*Length of follow-up:* 53 weeks (primary efficacy endpoint at 52 weeks)	*Length of follow-up:* up to 6 years (12-month interim results reported)

Sources Sandborn et al.[11], CS Table 7, 8, 9, 10, 13 and 14, and B.2.6.3.1

BID, twice daily; NR, not reported; UC, Ulcerative colitis.

[a] 15 mg BID tofacitinib treatment was discontinued based on feedback from regulatory authorities (OCTAVE 1: n=38, OCTAVE 2: n=18)

[b] Three subpopulations received tofacitinib 10 mg ([redacted]) in the open label extension study: Induction non-responders tofacitinib 10 mg [redacted]; maintenance completers tofacitinib 10 mg [redacted], maintenance treatment failures 10 mg ([redacted] comprising participants from OCTAVE Induction 1 and 2 who withdrew from OCTAVE Sustain due to treatment failure on tofacitinib (5 mg, [redacted]; 10 mg, [redacted]) or placebo ([redacted]). One subpopulation received tofacitinib 5 mg in the open label extension study: Maintenance: remission tofacitinib 5 mg [redacted]. Note that there appears to be a typographical error in CS Table 8 where the number of patients receiving tofacitinib 10 mg is given as [redacted].

[d] [redacted]

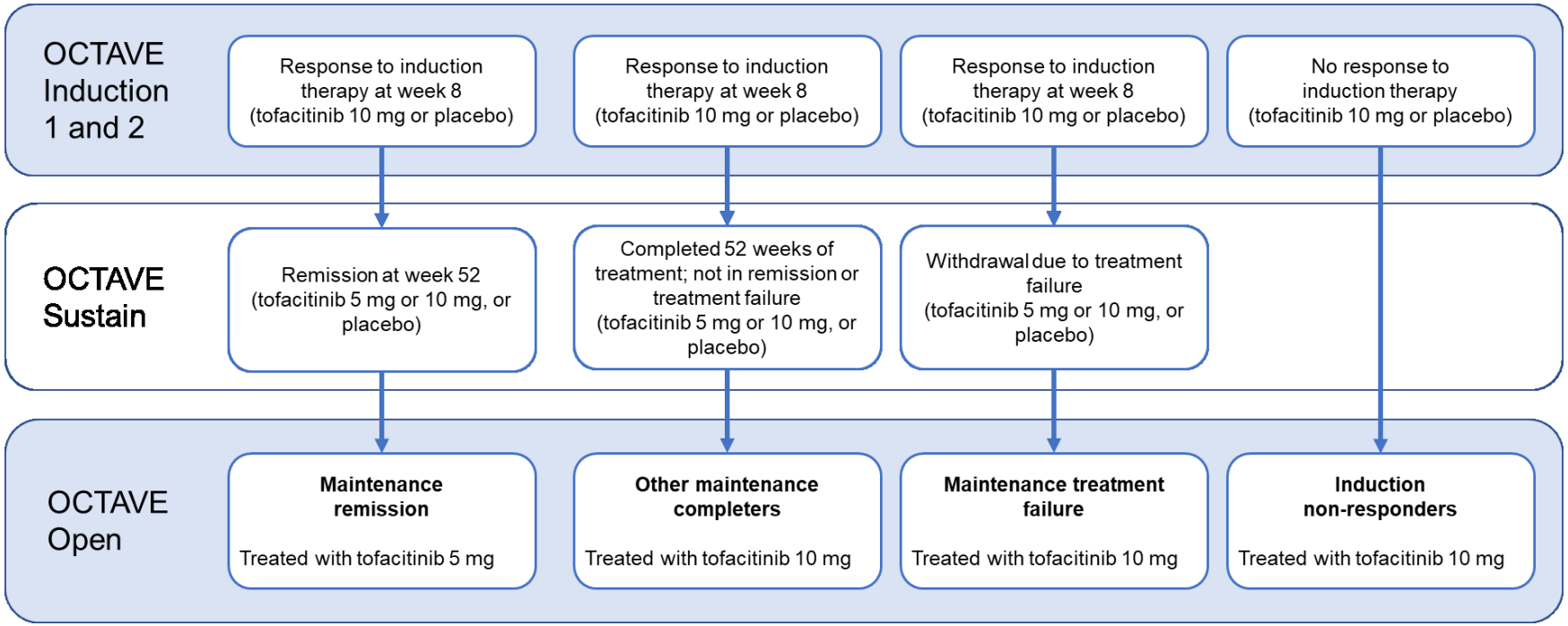


Figure 2 Participant flow in the OCTAVE trials (Source CS Figure 5)

3.1.3.1 OCTAVE RCTs baseline characteristics

The CS states that there were no significant differences in baseline characteristics between groups within each trial, apart from two exceptions. In OCTAVE 2, there was a statistically significant higher proportion of male patients in the tofacitinib group compared with the placebo group (tofacitinib 60.4% versus placebo 49.1%, p = 0.03). However, we note that sex is not considered to be a prognostic factor for ulcerative colitis[2] and a similar difference was not evident in OCTAVE 1. In OCTAVE Sustain, there was a significant difference in smoking status among the tofacitinib and placebo groups (p = 0.03), with a higher proportions of people who had never smoked and lower proportions of current smokers in the two tofacitinib groups than in the placebo group. Smoking is known to be a modifying factor in ulcerative colitis, with former cigarette smoking being a strong risk factor, yet active smokers are less likely to develop ulcerative colitis than former and non-smokers,[2] and active smoking is associated with milder disease.[21] If this imbalance were to affect the results it could disadvantage the tofacitinib groups, since these had a lower proportion of current smokers, yet it might also disadvantage the placebo group since this had a higher proportion of former smokers. We note that the difference between the tofacitinib and placebo arms in the number of patients who were current smokers amounted to only six patients (3.1 percentage points) whilst the difference in the number who were former smokers was only 10 patients for the tofacitinib 10 mg comparison (4.9 percentage points), although it was 24 patients (12.2 percentage points) for the tofacitinib 5 mg comparison. On balance, the risk of selection bias being introduced as a result of these imbalances in smoking status within OCTAVE Sustain appears to be low. Table 7 provides a summary of the trial characteristics of all the tofacitinib trials, including the ongoing extension trial Open.

Generally, patient characteristics appear to be balanced across the different OCTAVE trials, although patients enrolled in OCTAVE Sustain had lower Mayo scores and C-reactive protein levels than in either OCTAVE 1 or 2 (CS Table 15). This may be reflective of patients having had to achieve a response in order to be eligible to join OCTAVE Sustain.

Table 7 Summary of baseline patient characteristics of the OCTAVE 1 and 2 and Sustain

Characteristic	OCTAVE Induction 1		OCTAVE Induction 2		OCTAVE Sustain		
	TOF 10 mg (N=476)	Placebo (N=122)	TOF 10 mg (N=429)	Placebo (N=112)	TOF 5 mg (N=198)	TOF 10 mg (N=197)	Placebo (N=198)
Male sex, n (%) [a]	277 (58.2)	77 (63.1)	259 (60.4)	55 (49.1)	103 (52.0)	110 (55.8)	116 (58.6)
Age, years [b]	41.3±14.1	41.8±15.3	41.1±13.5	40.4±13.2	41.9±13.7	42.9±14.4	43.4±14.0
Induction trial group assignment, n (%)							
Placebo	—	—	—	—	22 (11.1)	24 (12.2)	24 (12.1)
Tofacitinib, 10 mg BID	—	—	—	—	170 (85.9)	167 (84.8)	167 (84.3)
Tofacitinib, 15 mg BID	—	—	—	—	6 (3.0)	6 (3.0)	7 (3.5)
Remission at maintenance trial entry, n (%)	—	—	—	—	65 (32.8)	55 (27.9)	59 (29.8)
Duration of disease, median yrs [b] (range)	6.5 (0.3–42.5)	6.0 (05–36.2)	6.0 (0.4–39.4)	6.2 (0.4–27.9)	6.5 (0.6–40.3)	6.8 (0.6–35.7)	7.2 (0.6–42.7)
Extent of disease, n/total n (%) [c,d]							
Proctosigmoiditis	65/475 (13.7)	19/122 (15.6)	67/428 (15.7)	16/111 (14.4)	28/196 (14.3)	33/196 (16.8)	21/198 (10.6)
Left-sided colitis	158/475 (33.3)	37/122 (30.3)	149/428 (34.8)	39/111 (35.1)	66/196 (33.7)	60/196 (30.6)	68/198 (34.3)
Extensive colitis or pancolitis	252/475 (53.1)	66/122 (54.1)	211/428 (49.3)	56/111 (50.5)	102/196 (52.0)	103/196 (52.6)	108/198 (54.5)
Total Mayo score [b,e]	9.0±1.4	9.1±1.4	9.0±1.5	8.9±1.5	3.3±1.8	3.4±1.8	3.3±1.8
Partial Mayo score [b,e]	6.3±1.2	6.5±1.2	6.4±1.3	6.4±1.2	1.8±1.3	1.8±1.3	1.8±1.4
C-reactive protein, median mg/litre [b] (range)	4.4 (0.1–208.4)	4.7 (0.1–82.5)	4.6 (0.2–156.0)	5.0 (0.2–205.1)	0.7 (0.1–33.7)	0.9 (0.1–74.3)	1.0 (0.1–45.0)
Oral glucocorticoid use at baseline, n (%) [b]	214 (45.0)	58 (47.5)	198 (46.2)	55 (49.1)	101 (51.0)	87 (44.2)	100 (50.5)
Previous treatment with TNFi, n (%) [c]	254 (53.4)	65 (53.3)	234 (54.5)	65 (58.0)	90 (45.5)	101 (51.3)	92 (46.5)

Characteristic	OCTAVE Induction 1		OCTAVE Induction 2		OCTAVE Sustain		
	TOF 10 mg (N=476)	Placebo (N=122)	TOF 10 mg (N=429)	Placebo (N=112)	TOF 5 mg (N=198)	TOF 10 mg (N=197)	Placebo (N=198)
Previous treatment failure, n (%) [c,f]							
TNF antagonist	243 (51.1)	64 (52.5)	222 (51.7)	60 (53.6)	83 (41.9)	93 (47.2)	89 (44.9)
Glucocorticoid	350 (73.5)	98 (80.3)	303 (70.6)	83 (74.1)	145 (73.2)	149 (75.6)	151 (76.3)
Immunosuppressant [g]	360 (75.6)	83 (68.0)	301 (70.2)	75 (67.0)	143 (72.2)	141 (71.6)	129 (65.2)
White race, n (%) [h]	395 (84.6)	98 (83.1)	331 (80.3)	88 (83.0)	164 (84.5)	153 (81.8)	155 (80.3)
Weight, kg	72.9 (16.8)	72.7 (16.7)	74.4 (16.8)	73.2 (16.2)	73.4 (17.8)	74.6 (15.1)	76.2 (16.7)
Smoking status, n (%) [c,i]							
Never smoked	301 (63.2)	80 (65.6)	268 (62.5)	81 (72.3)	142 (71.7)	128 (65.0)	113 (57.1)
Current smoker	22 (4.6)	4 (3.3)	25 (5.8)	5 (4.5)	7 (3.5)	6 (3.0)	12 (6.1)
Former smoker	153 (32.1)	38 (31.1)	136 (31.7)	26 (23.2)	49 (24.7)	63 (32.0)	73 (36.9)

Source: CS Table 15

Footnotes: see next page

Footnotes for Table 2

[a] In the OCTAVE Induction 2 trial, there was a significant difference between groups in the proportion of male patients (p = 0.03).

[b] For the OCTAVE Sustain trial, the baseline values were obtained at the time of entry in the OCTAVE Sustain trial.

[c] For the OCTAVE Sustain trial, the baseline values were obtained at the time of entry into one of the induction trials (OCTAVE Induction 1 or 2).

[d] Data on extent of disease are missing for three patients.

[e] The total Mayo score ranges from 0 to 12 and the partial Mayo score (i.e., the total Mayo score excluding the endoscopic subscore) ranges from 0 to 9, with higher scores indicating more severe disease.

[f] Previous treatment failure was determined by the investigator.

[g] Immunosuppressants included agents such as azathioprine and mercaptopurine and did not include biologic agents (e.g., TNF antagonists) or glucocorticoids.

[h] Unspecified race was treated as missing data.

[i] In OCTAVE Sustain, there was a significant difference for smoking status among placebo and tofacitinib groups (*p* = 0.03).

In summary, the CS appears to have identified all relevant RCTs and has provided all relevant study publications electronically, although CSR for the phase II trial had to be requested by the ERG and NICE (Clarification question A4).

3.1.3.2 Non-randomised trials

The company conducted a SLR to identify non-RCT evidence (CS Appendix D.1.4.2.), in order to provide long-term evidence (over 12 weeks for induction and over 52 weeks for maintenance therapy) regarding the efficacy and safety of tofacitinib for the treatment of moderately to severely active ulcerative colitis, hence relevant to the decision problem.

The company included one open-label, ongoing tofacitinib extension trial of up to 6 years duration (OCTAVE Open - NCT01470612).[20] Patients could enter OCTAVE Open from the OCTAVE 1 and 2 Induction trials if they did not have a response or enter from the OCTAVE Sustain trial once they completed 52 weeks of follow-up or if they withdrew due to treatment failure. Consequently, OCTAVE Open has four distinct patient groups (as depicted in Figure 2):

- Induction non-responders: patients from OCTAVE 1 AND 2 who did not have a response to induction therapy and did not enter OCTAVE Sustain (all allocated to 10 mg BID tofacitinib in OCTAVE Open)
- Maintenance remission: patients with a response to induction therapy in OCTAVE 1 and 2 who were in remission at week 52 in OCTAVE Sustain (all allocated to 5 mg BID tofacitinib in OCTAVE Open)
- Maintenance completers: patients who at the end of 52 weeks of maintenance therapy in Sustain were not in remission but did not meet the definition of treatment failure (all allocated to 10 mg BID tofacitinib in OCTAVE Open)
- Maintenance treatment failure: patients with a response in OCTAVE 1 and 2 who withdrew from OCTAVE Sustain due to treatment failure on tofacitinib (all allocated to 10 mg BID tofacitinib in OCTAVE Open)

Patient disposition for OCTAVE Open is presented in a confidential table (CS Appendix D.1.4, Table 119), with demographic and baseline characteristics in Appendix L.1.5 Table 231 and baseline disease characteristics in Appendix L.1.5 Table 232.

Evidence from this trial (which is still ongoing) presented in the CS is predominantly for patients with 12-month data because 24-month data are currently only available for a small number of patients. The CS presents a summary of results in sections B.2.6.3.2 to B.2.6.3.5 with full endpoint results shown in CS Appendix L (Tables 233 to 236). Additionally, a table of 12-month interim data for treatment emergent adverse events (CS Appendix F, Table 167) is presented.

3.1.3.3 Ongoing studies

Apart from the OCTAVE Open trial reported above, which may provide more data within the next 12 months, the CS states that preliminary results from a phase IIIb/IV study of tofacitinib in patients with ulcerative colitis in stable remission (NCT03281304) may also be available within the next 12 months (CS B.2.11). Apart from the ClinicalTrials.gov Identifier, no other information is provided in the CS. Details on the clinical trials website for the trial are shown in Table 8.

Table 8 Ongoing phase IIIb/IV study of tofacitinib

Title:	A Phase 3b/4, Multi-center, Double-blind, Randomized, Parallel Group Study Of Tofacitinib (Cp-690,550) In Subjects With Ulcerative Colitis In Stable Remission
Aim:	To evaluate flexible dosing in patients with ulcerative colitis
Start date:	Nov 2017
Estimated completion date:	Nov 2019
Number randomised:	130
Intervention:	5 mg BID tablet 10 mg BID tablet

3.1.4 Description and critique of the approach to validity assessment

The ERG has assessed the methodological quality of the four tofacitinib RCTs using NICE's recommended criteria (Table 9). The seven questions in Table 9 relate to risks of different types of bias that could arise within the trials. The company has phrased some of their quality assessment questions slightly differently to those recommended by NICE (indicated where appropriate in the table) for the quality assessment based on only the three OCTAVE trials (CS Table 19) and the quality assessment used for all the trials included in the NMA (CS Table 86). For question 5, about imbalances in dropouts, two versions of the question are given in the CS. For clarity we have labelled these as 5a and 5b, since the risk of bias interpretation differs according to how the question is phrased.

The OCTAVE Induction and Sustain RCTs appear to have a low risk of selection bias (questions 1 to 3), as the populations were generally well-balanced across the trial arms.

Participants and investigators in all four tofacitinib RCTs were blinded to the treatment allocations and so the risk of performance or detection bias that could arise through knowledge of treatment allocations appears to be low (question 4). Details of the blinding method of endoscopy readers are provided by the company in clarification response A11b.

The risk of attrition bias as a result of any treatment-related imbalances in dropouts between trial arms appears to be low in the OCTAVE Induction trials (question 5). However, there were some imbalances in dropouts in both the Phase II trial and the OCTAVE Sustain trial which

might have introduced bias. The ERG is uncertain about the direction and magnitude of any bias since there were several different reasons why patients withdrew. The CS acknowledges these imbalances in OCTAVE Sustain but not in the Phase II trial, and does not comment on whether they would have introduced bias.

Not all protocol-specified clinical effectiveness outcomes that were measured in the OCTAVE trials and Phase II trial are reported in the trial publications (question 6). However, the key outcomes are reported and the risk of reporting bias in these trials appears to be low.

The company used the "full analysis set" (FAS) as the primary analysis population but this was defined differently in the Phase II trial and the OCTAVE trials (question 7). The OCTAVE trials conducted an appropriate analysis in which the FAS was consistent with the ITT principle and accounted for missing remission data appropriately. Therefore, the risk of bias in the primary outcome, and all other outcomes analysed according to the FAS, appears to be low in the OCTAVE trials. In contrast, the Phase II trial conducted analyses in which the FAS was defined as being equivalent to a modified ITT population that did not include all randomised patients and not all missing data were included in analyses. As such, there is a risk of attrition bias in the Phase II trial, but with unclear direction and magnitude.

In summary, the OCTAVE Induction trials appear to be generally at low risk of the five types of bias assessed. The OCTAVE Sustain trial and the Phase II trial also appear to be at low risk of selection, performance, detection and reporting biases but could be at risk of attrition bias as a result of unbalanced dropouts between the tofacitinib and placebo arms.

Table 9 Company and ERG assessments of trial quality

Quality assessment question	**Judge-ments**	**Phase II trial**	**OCTAVE 1 & 2**	**OCTAVE Sustain**
1. Was randomisation carried out appropriately? ("yes" indicates low risk of selection bias)	**CS:**	Yes	Yes	Yes
	ERG:	Yes	Yes	Yes
ERG comments: A central randomisation method was employed (CS Tables 9, 13, 86) (not reported in the CS for the phase 2 trial, but stated in the trial publication).				
2. Was the concealment of treatment allocation adequate? ("yes" indicates low risk of selection bias)	**CS:**	Yes	Yes	Yes
	ERG:	Yes	Yes	Yes

ERG comments: Allocation concealment is not explicitly reported in the CS, CSRs or trial publications. However, central randomisation was telephone-based so the ERG assumes that the allocation sequence could not have been known to, foreseen, or influenced by the study investigators prior to them dialling in to receive each patient's random allocation to TOF or PBO.				
3. Were the groups similar at the outset of the study in terms of prognostic factors? ("yes" indicates low risk of selection bias)	**CS:**	Yes	Yes	Yes
	ERG:	Yes	Yes	Yes
ERG comments: Phase II trial (reported in the trial publication): The only statistically significant difference at baseline was in glucocorticoid use (placebo 58%, tofacitinib 10 mg 27%; p=0.03), although due to the small overall sample size this reflects a difference of only six patients. OCTAVE 1 and 2 (CS Table 15): The CS states that the only statistically significant difference between groups was in the proportion of male patients in OCTAVE 2. Where imbalances of >5% between arms occurred in the induction trials these did not systematically affect both trials. OCTAVE Sustain (CS Table 15): The CS states that the only statistically significant difference between groups was in smoking status. The proportion who never smoked differed between all three arms: 71.7% in the TOF 5 mg arm, 65.0% in the TOF 10 mg arm, and 57.1% in the PBO arm, but the difference was relatively small for TOF 10 mg vs PBO.				
4. Were the care providers, participants and outcome assessors blind to treatment allocation? ("yes" indicates low risk of performance and detection bias)	**CS:**	Yes	Yes	Yes
	ERG:	Yes	Yes	Yes
ERG comments: The CS states that the OCTAVE trials were patient-, investigator-, and sponsor-blinded (CS Tables 9 & 13) and the phase 2 trial was double blind (CS Table 86). The ERG assumes that "investigators" and "double blind" cover both the care providers and the outcome assessors, although this is not explicit in the CS. The method of blinding was to use a matching placebo tablet. NB this question is worded slightly differently in CS Table 86 compared to CS Tables 9 and 13, but in both cases a "yes" answer would suggest a low risk of bias.				
5a. Were there any unexpected imbalances in drop-outs between groups? (question as phrased in CS Tables 9 and 13; "no" indicates low risk of attrition bias)	**CS:**	Not reported	No	No
	ERG:	Yes	No	Yes
5b. Were discontinuations similar between groups? (question as phrased in CS Table 86; "yes" indicates low risk of attrition bias)	**CS:**	No	Yes	No
	ERG:	No	Yes	No
ERG comments: Phase II trial: Lower discontinuation rate in the TOF 10 mg group (6%) than the PBO group (27%) (reported in the publication appendix). The TOF discontinuations (n=2) were both due to lack of efficacy. The PBO discontinuations were due to lack of efficacy (n=5), AE (n=3), protocol violation (n=2), consent withdrawn (n=2) and loss to follow up (n=1). OCTAVE 1 & 2: Slight imbalances in discontinuations but these were not consistent in direction across both induction trials (OCTAVE 1: PBO 3.3%, TOF 6.5%; OCTAVE 2: PBO 13.4%, TOF 7.5%) (CS Table 17). OCTAVE Sustain: As noted in CS Table 18, discontinuation rates differed between PBO (73.2%), TOF 5 mg (43.9%) and TOF 10 mg (35.7%). The main reason for discontinuation was lack of clinical response (66.7%, 35.4%, 27.0% respectively); relatively few patients discontinued due to AE (<5% in each arm).				
	CS:	No	No	No

6. Is there any evidence to suggest that the authors measured more outcomes than they reported? (question phrased in CS Table 86 as "unreported outcomes suspected?") ("no" indicates low risk of reporting bias)	**ERG:**	No	Yes, but low bias risk	Yes, but low bias risk
ERG comments: The OCTAVE trial publication (as acknowledged in section 4 of the supplementary appendix)[19] does not report all outcomes that were measured in the OCTAVE Induction and Sustain trials. The publication does not explicitly state reasons why some outcomes were not reported. However, the most important clinical effectiveness outcomes are reported. Where outcomes were measured but not reported in the trial publication (e.g. several patient-reported outcome measures), these appear to favour TOF 10 mg over PBO, according to results in the CSRs. As such, the non-reporting of some outcome measures in the trial publication would appear unlikely to have introduced bias.				
7. Did the analysis (1) include an intention-to-treat (ITT) analysis? (2) If so, was this appropriate and (3) were appropriate methods used to account for missing data? [sub-questions numbered by ERG] ("yes" indicates low risk of attrition bias)	**CS:**	Stated ITT (CS Table 86) but see ERG comment below	Yes	Yes
	ERG:	1. No 2. NA 3. NA	1. Yes 2. Yes 3. Yes	1. Yes 2. Yes 3. Yes
ERG comments: Phase II trial: The trial protocol states that the Full analysis set was the main analysis population, defined as all randomised subjects, who have either withdrawn as a treatment failure or have completed at least one week of dosing and had at least one valid Mayo score during the active double-blind phase of the study (trial protocol section 5). This is a modified ITT rather than a true ITT population. The trial publication describes both a pre-specified analysis and a post-hoc analysis of the primary outcome, neither of which was based on all randomised participants. There is a possible risk of bias but the direction and magnitude are unclear since dropouts from the PBO arm occurred for several different reasons. OCTAVE 1 and 2 and Sustain: Full analysis set was the main analysis population, defined as all subjects as randomly assigned, which is consistent with the ITT principle (CS section B.2.4.1 and section 5 in the trial protocols). Crossovers are not mentioned in the CS, trial publications, protocol, or OCTAVE CSRs and the participant flow in CS Tables 17 and 18 do not mention that any crossovers occurred. Missing values for the primary outcome were analysed by non-responder imputation.				

NA: not applicable

3.1.5 Description and critique of company's outcome selection

The outcomes included in the CS match those in the NICE final scope and appear appropriate. However, time to surgical intervention, although specified in the NICE final scope, was not included, as this was not assessed in the OCTAVE trials.

In clinical trials of therapies for ulcerative colitis the Mayo Score is widely used and was used within the OCTAVE trials (CS Section B1.3.1 and CS Table 3). There are four components to the Mayo score, one of which is 'Endoscopic findings'. In the OCTAVE trials the Mayo endoscopic sub-score was assessed both locally (by the study site investigator) and centrally (from a video recording). Consequently the outcomes in the CS that utilise the endoscopic sub-score were reported separately using the local or the central read of the endoscopic data. The ERG notes that the FDA[22] state that central reading is the preferred approach and the OCTAVE clinical trial programme is the first in ulcerative colitis to use central reads (CS Section B.2.3.1.2.4).

The primary outcome in OCTAVE 1 and 2 was remission at week 8 based on centrally read endoscopic Mayo sub-scores, and at week 52 in OCTAVE Sustain (for definition of remission see Table 10). Higher Mayo scores indicate more severe disease. The company also defined key secondary outcomes: mucosal healing (OCTAVE 1 and 2: week 8; OCTAVE Sustain: week 52), and for OCTAVE Sustain only, sustained corticosteroid-free remission among patients in remission at baseline (week 52). Mucosal healing is associated with lower rates of hospitalisation and surgery,[23] while the use of corticosteroids long-term is not suitable due to side effects so a corticosteroid-free remission is important.[24]

Clinical response and clinical remission based on Mayo sores (for definitions see Table 10) were reported for all three trials (OCTAVE 1 and 2: week 8; OCTAVE Sustain: week 52). As can be observed from Table 10 the difference between the primary outcome of remission and the secondary outcome of clinical remission is that for the former the rectal bleeding sub-score must be zero whereas this is not necessary for the outcome of clinical remission. Clinical response and clinical remission were the only clinical effectiveness outcomes included in the economic model (the primary outcome did not contribute to the economic model), as they were thought to ensure comparability with trials of biological therapies for ulcerative colitis.

The remaining outcomes of disease activity were all based on a Mayo score (for definitions see Table 10):

- Endoscopic remission
- Symptomatic remission
- Deep remission
- Partial Mayo score (range 0-9)
- Total Mayo score (range 0-12)

Health-related quality of life (HRQoL) measures included in the CS were the disease-specific IBDQ, and the Work Productivity and Activity Impairment – Ulcerative Colitis (WPAI-UC) version 2 questionnaire. Generic measures were the 5-dimension EuroQol questionnaire (EQ-5D) and the 36-Item Short Form survey (SF-36). All four HRQoL measures are validated and have been used in other clinical trials in patients with ulcerative colitis.[25,26] However, where different versions of a measure exist (e.g. country specific versions of the IBDQ), the CS did not state which versions were used across the different countries and centres in which the OCTAVE trials took place.

- For the 32-item, disease-specific IBDQ, remission (defined in Table 10) and treatment response were reported for all three OCTAVE RCTs (OCTAVE 1 and 2 at weeks 4 and 8; OCTAVE Sustain at weeks 8, 24 and 52). IBDQ remission scores range from 32 to 224, with higher scores indicating better HRQoL. In HRQoL terms, a total IBDQ score ≥170 points is deemed to constitute clinical remission and a change of ≥16 points has previously been used as a minimal clinically important difference threshold in patients with ulcerative colitis.[27,28]
- For the EQ-5D-3L, both the utility score (based on five dimensions of health status: mobility, self-care, usual activities, pain/discompfort, and anxiety/depression) and visual analogue scale (VAS) outcomes were reported based on EQ-5D-3L version of the instrument, with UK preference weights. This is the only HRQoL measure included in the economic model and it is reported by all three OCTAVE RCTs (OCTAVE 1 and 2 at weeks 2, 8 and change from week 0-8; OCTAVE Sustain at weeks 8, 24, 52 and change from week 0-52).. The CS reports minimal clinically important differences (MCIDs) for UK patients with inflammatory bowel disease of 0.076 for the utility index and 10.9 for the VAS.[29]
    - All three OCTAVE RCTs reported Physical Component Summary (PCS) and Mental Component Summary (MCS) scores from the SF-36 version 2, using the acute form,

which has a recall period of 1 week (in OCTAVE 1 and 2 assessed at baseline and week 8; in OCTAVE Sustain assessed at baseline and weeks 24 and 52). Higher scores indicate better HRQoL. A systematic review[30] of the SF-36 in patients with ulcerative colitis suggests that a group-level clinically important difference threshold of 3 points for both summary scores and responder-level thresholds of 3.1 for PCS and 3.8 for MCS based on the SF-36v2 manual.[31]

- The WPAI-UC score, based on a 6-item questionnaire (version 2) assessing work productivity, is also reported by all three OCTAVE RCTs (OCTAVE 1 and 2 at baseline and week 8; OCTAVE Sustain at baseline and week 52). The questionnaire yields four scores expressed as impairment percentages: absenteeism; presenteeism; work productivity loss; non-work activity impairment. A higher score indicates greater impairment.[32] As part of the response to NICE and the ERG's clarification question A12, the company states that it is not aware of any validated MCID for this outcome in patients with ulcerative colitis. However the company also state that extrapolating from Crohn's Disease suggests a 7% decrease is the MCID for the WPAI.[33,34]

Table 10 Clinical effectiveness outcomes and outcome definitions of the OCTAVE RCTs

Outcome	Definition	When assessed, week		Used in Model
		OCTAVE 1 & 2	OCTAVE Sustain	
Primary: Remission based on centrally-read endoscopic sub-scores	Mayo score ≤2, no individual sub-score >1, rectal bleeding sub-score = 0	8	52	No
Key secondary: Mucosal healing	Mayo endoscopic sub-score ≤1	8	52	No
Key secondary: Sustained corticosteroid-free remission among patients in remission at baseline	Remission (as defined above for the primary outcome) plus no treatment with steroids for ≥4 weeks before the 24-week and 52-week visits	Not assessed	52	No
Clinical response	Mayo score decrease from baseline ≥ 3, and ≥ 30%, with a decrease in rectal	Week 8	52	Yes

	bleeding sub-score of ≥1 or absolute rectal bleeding sub-score of ≤ 1			
Clinical remission	Mayo score ≤2, no individual sub-score >1	8	52	Yes
Endoscopic remissio	Mayo endoscopic sub-score = 0	8	52	No
Symptomatic remission	Mayo score ≤2, no individual sub-score >1, rectal bleeding sub-score and stool frequency sub-score = 0	8	52	No
Deep remission	Mayo score ≤2, no individual sub-score >1, rectal bleeding sub-score and endoscopic sub-score = 0	8	52	No
Partial Mayo score (range 0-9)	Total Mayo score excluding the endoscopic sub-score	2, 4 & 8; change 0-8	Not assessed	No
Total Mayo score (range 0-12)	Sum of 4 sub-scores (stool frequency, rectal bleeding, endoscopic findings, physician's global assessment), each 0-3 with higher scores indicating more severe disease (details in CS Table 3)	Change 0-8	Not assessed	No
HRQoL	**Details/definition**	**OCTAVE 1 & 2**	**OCTAVE Sustain**	**Used in Model**
IBDQ remission	IBDQ score ≥170	4 & 8	8, 24 & 52	No
IBDQ treatment response	IBDQ score increase ≥16 from induction trial baseline	4 & 8	8, 24 & 52	No
EQ-5D score (utility and visual analogue scale versions)	Based on EuroQol-5D 3 level version (no problems, some problems and extreme problems) with UK preference weights	2 & 8; change week 0-8	8, 24 & 52; change 0-52	Yes
SF-36 (PCS and MCS score)	Acute Physical Component Summary & Mental Component Summary scores based on Short-Form 36-item survey (v2)	8; change 0-8	24 & 52; change 0-52	No
WPAI-UC score (assesses work productivity)	6-item Work Productivity and Activity Impairment-Ulcerative Colitis questionnaire (version 2)	8; change 0-8	52; change 0-52	No

3.1.6 Description and critique of the company's approach to trial statistics

The ERG has assessed the approach to trial statistics for the Phase II trial, the OCTAVE 1 and 2 Induction trials and the OCTAVE Sustain trial. The OCTAVE Open study is ongoing and only

summary statistics have been generated (CS Table 16) therefore only the sources of information for this study have been indicated below.

The CS focusses on outcomes from the OCTAVE 1 and 2 Induction trials and the OCTAVE Sustain trial. A brief summary of results from the Phase II trial (which contributes data to the NMAs) is included (CS Figure 18). Interim data from the OCTAVE Open study are summarised in CS Sections B.2.6.3.2 to B.2.6.3.5.

Outcomes and their units of measurement are defined in CS Tables 11 and 12 (OCTAVE 1 and 2), CS B.2.3.1.3.3 (OCTAVE Sustain) and CS B.2.3.1.4.2 (Open study). Outcomes for the Phase II trial are not defined in the CS. Outcomes were defined in the same way in OCTAVE 1, 2 Sustain and Open. The two OCTAVE Induction trials and the OCTAVE Sustain trial are both complete. The only interim data presented in the CS come from the OCTAVE Open study but these do not contribute data to the economic model. The CS has presented appropriate measures of effects (proportions or mean differences with p-values for comparisons between placebo and tofacitinib groups) with uncertainty for continuous outcomes indicated by confidence intervals.

Statistical power

The primary outcome in both the OCTAVE 1 and OCTAVE 2 trials was remission at week 8, based on centrally read Mayo endoscopic subscores. The power calculation is reported in CS Table 16. For each of these trials the company calculated that approximately 545 participants per trial (randomised 4:1, i.e. 436 patients to the 10 mg tofacitinib group and 109 patients to the placebo group) would provide 90% power to detect a difference of 17.5 percentage points between tofacitinib and placebo in the primary and key secondary outcomes. The CS does not justify or explain the rationale for being able to detect a 17.5 percentage point difference between the tofacitinib and placebo groups. One of the ERG's clinical experts thought this was a modest difference but similar to comparator drugs which are used in clinical practice. This power calculation assumed remission rates in the placebo groups of 15% for the primary outcome (remission at week 8) and 35% for the key secondary outcome of mucosal healing. These assumptions are not justified or explained in the CS. The required sample size was achieved for OCTAVE 1 (tofacitinib 10 mg N=476; placebo N=122) and was only narrowly missed for the tofacitinib arm of OCTAVE 2 (tofactinib 10 mg N=429; placebo N=112). Not all of the assumptions made for the power calculation are justified or explained in the CS.

Additionally, the actual sample size was slightly smaller than that calculated and actual rates of remission and mucosal healing were lower than assumed for the power calculation. Nevertheless, the ERG believes that the power calculation was conducted appropriately and the ERG considers that the trials were probably adequately powered.

The primary outcome in the OCTAVE Sustain trial was remission at week 52 based on centrally read Mayo endoscopic subscores. The company calculated that a total of 654 participants (randomised 1:1:1, so 218 in each group) would provide 90% power to detect a 17.5 percentage point difference in remission between the tofacitinib groups (5 mg; 10 mg) and the placebo group, assuming a remission rate in the placebo group of 30% (CS Table 16). The CS does not justify or explain the rationale for being able to detect a 17.5 percentage point difference between the tofacitinib and placebo groups or the assumption of a remission rate in the placebo group of 30%. The required sample size was not achieved, as 593 patients were randomised, which is 61 short of the 654 target (20-22 short per trial arm; CS Table 18). Although the sample size fell short by around 10% per arm, the power calculation was done at a fairly strict level (90% power). On balance the ERG believes that, although there is uncertainty in the statistical power achieved, it is likely to have been adequate.

Statistical power for the Phase II trial and for OCTAVE Open is not reported in the CS.

Analysis populations

The CS defines five main analysis sets (CS B.2.4.1) for the OCTAVE Induction and OCTAVE Sustain trials:

- Full Analysis Set (FAS)
- OCTAVE Induction modified Full Analysis Set (mFAS)
- OCTAVE Sustain mFAS
- Per-Protocol Analysis Set (PPAS)
- Safety Analysis Set (SAS)

FAS – this is the primary analysis population for effectiveness endpoints and is defined as all subjects randomly assigned to either placebo, tofacitinib 10 mg twice daily, or (for OCTAVE Sustain only) tofacitinib 5 mg twice daily. NB this is equivalent to an intention to treat analysis population.

mFAS – this is a subset of the OCTAVE 1 and 2 FAS from which 3 patients were excluded (all from a site in Japan) due to potential unblinding during the study.

OCTAVE Sustain mFAS – this is a subset of the OCTAVE Sustain FAS that included only those patients who had received tofacitinib in the induction trials (i.e. it excluded those patients from the OCTAVE Induction trials who received placebo and met the entry criteria for OCTAVE Sustain).

PPAS – this is a subset of the FAS population who had no major protocol violations that could have potentially had a significant impact on outcomes (this subset was determined by the sponsor prior to database lock).

SAS - included all randomised participants who received at least 1 dose of study medication.

Results from the mFAS and PPAS are not described in full detail in the CS but primary endpoint results are summarised in Appendix L Table 206 to 208.

Analysis populations are not defined in the CS for the Phase II trial or the OCTAVE Open study (NB the trial publication and CSR indicate that FAS in the Phase II trial was defined differently to the Phase III trials and did not include all randomised patients).

Analysis methods

In both the OCTAVE Induction trials, binary outcomes were analysed using a Cochran-Mantel-Haenszel (CMH) Chi-square test, stratified by prior treatment with TNFi therapy, corticosteroid use at baseline, and geographic region. This analysis was applied to the primary outcome (proportion of participants with remission), the key secondary outcome (proportion with mucosal healing), and other binary secondary outcomes (proportions with outcomes derived from the Mayo score, proportion with IBDQ remission, and proportion with IBDQ treatment response) (CS Table 16; OCTAVE Induction trials CSRs sections 9.7.4.2 to 9.7.4.4).

Binary outcomes in the OCTAVE Sustain trial were also analysed using a CMH Chi-square test, but stratification was by treatment received in the induction trials and remission status at baseline (OCTAVE Sustain CSR 9.7.5.2).

In both the OCTAVE Induction trials, continuous outcomes measured only at baseline and week 8 (e.g. the secondary endpoint of change from baseline to week 8 in the total Mayo score) were analysed with an analysis of covariance (ANCOVA) model with observed-cases data. Factors in the ANCOVA were prior treatment with TNFi therapy, corticosteroid use at baseline, and geographic region, whilst baseline score was a covariate. For continuous outcomes measured repeatedly over time (e.g. partial Mayo score at baseline and weeks 2, 4 and 8) data were analysed using a linear mixed-effects model with baseline, treatment group, prior treatment with TNFi therapy, corticosteroid use at baseline, geographic region, visit, and treatment group by visit interaction as fixed effects and subject as a random effect (CS Table 16 and OCTAVE 1 and 2 CSRs section 9.7.4.4)

Continuous outcomes in the OCTAVE Sustain trial (e.g. Mayo scores at baseline, weeks 24 and 52) were analysed using a linear mixed-effects model with induction study treatment assignment included as a baseline stratification factor. (CS Table 16 and OCTAVE Sustain CSR 9.7.5.4).

In OCTAVE 1 and 2 the type 1 error rate was controlled at 0.05 by a fixed-sequence testing procedure for the primary outcome and the key secondary outcome. In OCTAVE Sustain the type 1 error rate was controlled at the 0.05 level for the primary outcome and both of the key secondary outcomes by using a sequentially rejective Bonferroni-based iterative multiple test procedure.

Analysis methods are not reported in the CS for the Phase II trial. The CS states that summary statistics for the OCTAVE Open study have been produced for the interim analysis of the available data (CS Table 16).

In summary, the ERG is satisfied that the analysis methods for the OCTAVE trials were pre-specified and appear appropriate for binary and continuous outcome data. However, the type 1 error rate was controlled only for the primary and key secondary outcomes of the OCTAVE trials, with no adjustments made for multiple comparisons among the other secondary outcomes and therefore caution is needed in interpreting these analyses.

Missing data

Missing data for binary outcomes derived from the total or partial Mayo score were managed in the same way across OCTAVE 1, 2, OCTAVE Sustain and OCTAVE Open (CS Table 16). Patients with missing data for these outcomes were considered as not having had a response (i.e. a non-responder imputation was applied). The ERG agrees that for the binary outcomes based on the Mayo score this is a conservative approach.

For continuous secondary effectiveness outcomes only measured at two timepoints (e.g. at baseline and week 8, or at baseline and week 52) and for continuous effectiveness outcomes (e.g. partial Mayo score) measured repeatedly over time, missing values were not imputed. In the case of continuous effectiveness outcomes measured repeatedly, a linear mixed-effects model was used for the analyses where the missing data were assumed to be missing at random. No justification for the choice of methods to manage data missing from continuous outcomes is provided in the CS.

The CSRs for the OCTAVE 1, 2 and Sustain trials indicate that sensitivity analyses with different approaches for handling missing data (last observation carried forward and observed-cases analyses) were performed for the primary and the key secondary outcomes (OCTAVE 1 and 2 CSRs section 9.7.4.2, OCTAVE Sustain CSR section 9.7.3.1.1) but this is not commented on in the CS. According to the CSRs the results of the sensitivity analyses were consistent with the primary analyses using non-responder imputation.

Missing data for the patient reported outcomes were initially handled using the rules suggested by the developers of the questionnaires (OCTAVE 1 and 2 CSRs section 9.7.2.1; SUSTAIN CSR section 9.7.3.1), but the CS does not state what these rules were. For IBDQ binary outcomes if missing data could not be imputed using the tool developers' rules then they were treated as non-responders. The CS does not state how many of the missing data were accounted for using the developers' rules and how many were imputed as non-responders.

Missing data for the other patient reported outcome measures in the OCTAVE trials were handled differently between the outcomes and between the OCTAVE Induction and OCTAVE Sustain trials (CS Table 16):

- OCTAVE Induction trials: Missing data for EQ-5D continuous outcomes were assumed to be missing at random whilst missing values for SF-36 and WPAI were not imputed.

- OCTAVE Sustain trial: The missing at random assumption was applied to both EQ-5D and SF-36 continuous outcomes, whilst missing WPAI values were not imputed.

The company does not explain these methodological differences and no alternative methods to account for missing data are reported in the CS. The ERG notes that the proportions of data missing from the OCTAVE 1 and 2 trials (as calculated by the ERG from data presented in CS Appendix L Tables 95, 218 and 219) vary among the different patient-reported outcomes. These were lowest for the EQ-5D (0.8% to 8.0% missing data per arm at week 8) and highest for the SF-36 (4.9% to 12.5% missing data per arm at week 8). Furthermore there appear to be imbalances in missing data between trial arms but the company does not comment on this.

Methods for handling missing data are not described for the Phase II trial.

In summary, the ERG would have preferred the company to have provided a justification of the different approaches to handling missing data. For the primary outcomes and key secondary outcomes the company conducted appropriate sensitivity analyses which gave results consistent with the primary analysis. Different methods for accounting for missing data were not explored for patient-reported outcomes. The ERG would therefore interpret the patient reported outcome measures more cautiously than the primary outcome and key secondary outcomes where the exploration of the impact of missing data has been more thorough.

Subgroups

Subgroup analyses are reported in CS section B.2.7. Subgroups based on prior biologic therapy (people previously treated with one or more biologics and people who have not received prior biologic therapy) are listed in the NICE scope under 'Other considerations'. The CS does not report on subgroups based on prior biologic therapy but instead focuses on results according to the subgroups of patients who are TNFi-naïve and those who are TNFi-exposed (i.e. not a wider group of people who have received prior biologic therapy which could include vedolizumab which had been received by some participants in the OCTAVE Induction trials).

The CS highlights (CS Table 4) that the limitations to existing therapy with TNFi agents include that some patients will fail to respond to induction therapy (primary non-response to TNFi-agents) and up to 50% of initial responders will lose response over time (secondary non-

response). Consequently prior TNFi-therapy is an important factor in decisions regarding treatment options.

Pre-planned subgroup analyses were conducted for outcomes according to four factors in OCTAVE 1 and 2 (CS Table 9): prior TNFi exposure (yes vs no); prior TNFi failure (yes vs no), baseline corticosteroid use (yes vs no), and geographic region. However, results of subgroup analyses are not presented in the CS for geographic region. Two of the four factors were pre-specified subgroup analysis factors in the OCTAVE Sustain trial [prior TNFi exposure (yes vs no); prior TNFi failure (yes vs no)]. The OCTAVE Sustain trial included additional pre-planned subgroups, six of which are listed alongside the two noted above in CS Table 13: duration of disease (<6 years vs ≥ 6 years); prior corticosteroid failure at induction study baseline (yes vs no); induction study treatment assignment (tofacitinib 10 mg vs tofacitinib 10 mg or 15 mg vs placebo); remission at maintenance study baseline (yes vs no); mucosal healing at maintenance study baseline (yes vs no); corticosteroid use at maintenance study baseline (yes vs no). Results of these subgroup analyses in OCTAVE Sustain are presented in CS Appendix E.

In OCTAVE 1 and 2 two of the factors assessed by subgroup analyses, prior TNFi exposure and corticosteroid use, were stratification factors at randomisation. Similarly two of the factors assessed by subgroup analyses in OCTAVE Sustain, induction-trial group assignment and remission status at maintenance-trial entry, were stratification factors at randomisation. This would help to ensure that the patient characteristics in these subgroups were well-balanced between the trial arms (confirmed for OCTAVE 1 and 2 by the baseline characteristics of the TNFi exposure subgroups provided by the company in clarification response A7).

The ERG presumes that type 1 error (a false positive, identifying an effect that isn't real) was not controlled for in the subgroup analyses as no statement relating to this has been identified in the CS.

The CS points out (CS section B.2.7.2) that the OCTAVE trials were not powered to test the statistical significance of subgroup analyses due to the limited patient numbers in the subgroups. To increase statistical power, subgroup analyses were also conducted for the pooled OCTAVE 1 and 2 trial population, although the CS does not comment on the statistical power that would have been achieved in these analyses.

In summary, the company has pre-specified the factors for which subgroup analyses were conducted, which is good practice. The CS focuses on the analyses by prior TNFi exposure which, being a randomisation stratification factor in the OCTAVE Induction trials, should improve the balance in patient characteristics between the tofacitinib and placebo arms (i.e. reduce the risk of selection bias) for these subgroups. The ERG agrees that pooling subgroups for the OCTAVE 1 and 2 trials was appropriate for maximising the available statistical power for the TNFi exposure subgroups to increase confidence in the subgroup analyses of OCTAVE 1 and 2. However, the subgroup analyses for OCTAVE Sustain were not powered to test the statistical significance of effects and thus should be interpreted cautiously.

3.1.7 Description and critique of the company's approach to the evidence synthesis

The ERG describes and critiques the company's approach to evidence synthesis by NMA. The ERG identified a number of issues which are discussed in section 3.1.7.1 to section 3.1.7.9.

In an absence of direct head-to-head comparisons between active treatments, the company conducted a network meta-analysis (NMA). NMA is an extension of pairwise network meta-analysis which combines direct and indirect evidence through a connected network of comparators. The NMA compared the relative effects of tofacitinib (5 mg and 10 mg) with adalimumab (40/80/160mg), golimumab (200/100mg and 100mg), infliximab (5 mg/kg), vedolizumab (300mg Q4W and Q8W), and placebo. EMA-licensed doses were included and treated as separate treatments in the NMA. All studies in a moderate to severely active ulcerative colitis population who had failed to tolerate conventional therapy were included.

Effectiveness outcomes included in the NMA consisted of clinical response, clinical remission, and mucosal healing. Safety outcomes included discontinuations due to adverse events, serious adverse events, and serious infections. We have focused our critique on those outcomes included in the economic model: clinical response, clinical remission, and serious infections.

Baseline characteristics of included studies are presented in CS Table 87. The company noted heterogeneity between studies in terms of certain patient characteristics (including prior TNFi exposure, disease duration, and studies in Asian patients) and study design (treat-through or re-randomisation for the maintenance period).

To reduce heterogeneity the company undertook separate NMAs for the TNFi-naïve and TNFi-experienced/failure populations. This choice was informed by subgroup analysis from the OCTAVE programme, a similar assumption in NICE TA342, and a "single integrated induction phase NMA" conducted by the company which showed a statistically significant effect for the interaction between treatment and prior TNFi exposure.

Separate analyses were conducted for the induction (6 to 8 weeks) and maintenance periods (up to one year). Evidence networks and included studies are shown in Figure 3 to Figure 5 below. Most treatments were compared to placebo apart from the Mshimesh 2017 trial[35] which compared adalumimab to infliximab (Induction TNFi-naïve and safety networks), and the UC-SUCCESS study[36] which compared azathioprine to infliximab (safety network only).

Safety outcomes were analysed independently of TNFi exposure status to maximise statistical power for rare events and assumed that prior TNFi-exposure has no effect on safety outcomes. The company stated no NMA was conducted for the safety outcomes in the maintenance period due to the differences in study design.

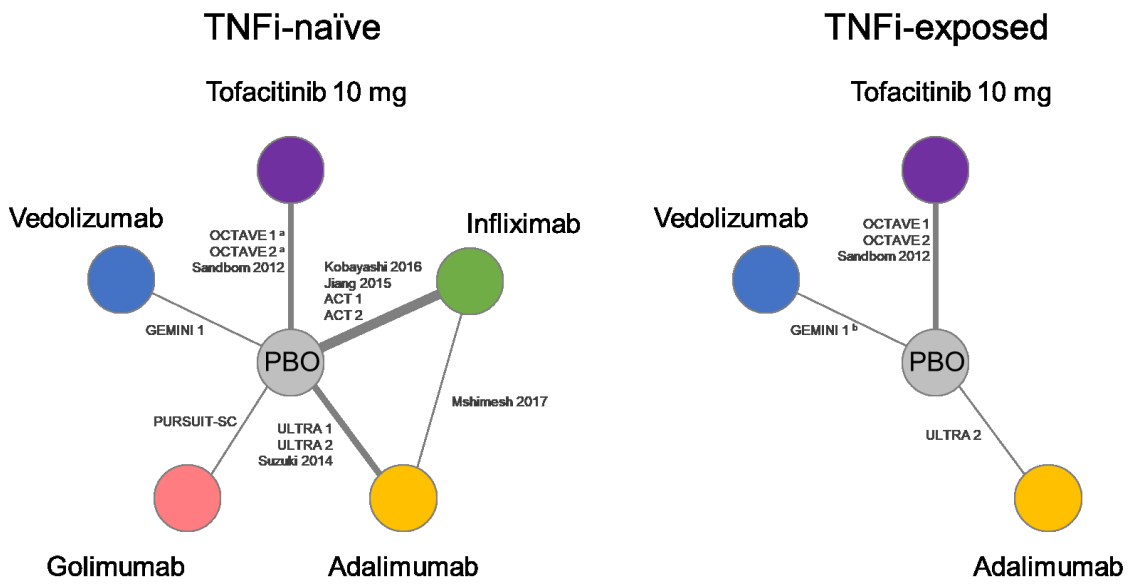


Figure 3 Base-case network of evidence for induction phase clinical response and clinical remission by TNFi-exposure subgroup (taken from CS Figure 28)

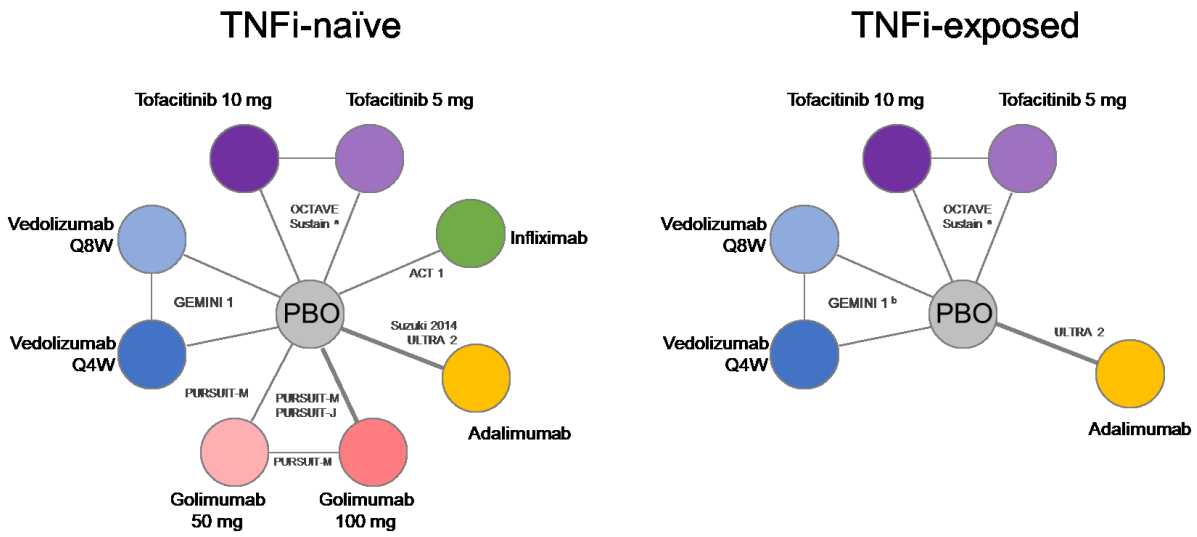


Figure 4 Base-case network of evidence for maintenance phase clinical response and clinical remission by TNFi-exposure subgroup (taken from CS Figure 29)

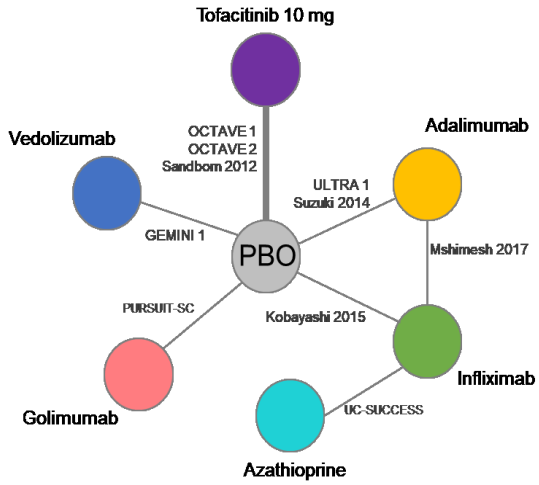


Figure 5 Base-case network of evidence for induction phase safety outcomes (discontinuation due to AEs, serious AEs and serious infections) (taken from CS Figure 30)

Fixed and random effects models were conducted. Where there was a difference in the deviance information criterion (DIC) of less than 3, the company favoured the fixed effects model.

Table 11 summarises the outcomes and comparators included in the analyses undertaken by the company.

Table 11 Outcomes and comparators included in the NMA analyses reported in the CS

Treatment	Clinical response/clinical remission, mucosal healing				Safety (discontinuations due to AEs, serious AEs, serious infections)
	Induction phase, TNFi-naive	Induction phase, TNFi-exposed	Maintenance phase, TNFi-naive	Maintenance phase, TNFi-exposed	
Tofacitinib	X	X	X	X	X
Adalimumab	X	X	X	X	X
Golimumab	X		X		X
Infliximab	X		X		X
Placebo	X	X	X	X	X
Vedolizumab	X	X	X	X	X
Azathioprine[a]					

[a] azathioprine was included in the safety evidence network but not in the NMA results

The company used a multinomial probit model for clinical response and clinical remission. Essentially this modelled clinical response and clinical remission jointly, treating them as ordered categorical data, thus maintaining the correlation between outcomes. This also assumed a common relative treatment effect across response categories. A binomial logit model was used for the safety endpoints.

As noted above, two alternative study designs were used in the maintenance phase. The tofacitinib, golimumab, and vedolizumab studies used a "re-randomised" design, whilst the adalimumab and infliximab studies used a "treat-through" design. Whilst the "treat-through" studies followed a traditional parallel design, randomising patients at baseline, "re-randomised" studies only included induction phase responders in the maintenance phase and re-randomised them to the active treatment or placebo.

The company adjusted for the differences in maintenance study design by adjusting the treat-through study results (ULTRA 2, Suzuki 2014[37] [Adalimumab]; ACT 1[38] [infliximab]) to match those of the re-randomised studies (OCTAVE Sustain[19] [tofacitinib]; PURSUIT-M[39] and PURSUIT-J[40] [Golimumab]; GEMINI 1[41] [vedolizumab]) using similar methods to Takeda in TA342.[9]

Response and remission results are presented on the probit scale (where a negative coefficient indicates treatment is more effective than placebo), and as odds ratios and absolute probabilities. Safety outcomes are presented as log odds, odds ratios, and absolute probabilities.

The company conducted three sets of sensitivity analyses for the effectiveness outcomes: centrally read (as opposed to locally read) endoscopic sub-scores; excluding Asian studies;[35,37,40,42,43] and using prior TNFi-failure as opposed to prior TNFi-exposure data. A further sensitivity analysis in the response to clarification questions (question A16) excluded the Phase II tofacitinib study.

One sensitivity analysis was conducted on safety outcomes: excluding Asian studies[35,37,42,43] and the tofacitinib Phase II (non-Asian) study (Sandborn 2012[11]).

The company's approach to data synthesis by NMA was generally well conducted. A summary of the ERG's appraisal of the company's approach is presented in Table 12.

However, a number of issues were identified which are discussed in sections 3.1.7.1 to 3.1.7.9 which follow Table 12.

Table 12 ERG appraisal of the NMA approach

Checklist		**Response**
Does the MS present an NMA?		Yes
Are the NMA results used to support the evidence for the clinical effectiveness of the intervention		Yes
Are the NMA results used to support the evidence for the cost-effectiveness of the intervention		Yes, selected endpoints
Homogeneity		
	1. Is homogeneity considered?	Yes
	2. Are the studies homogenous in terms of patient characteristics and study design?	No, difference in TNFi exposure status and study design
	3. Is the method used to determine the presence of statistical heterogeneity adequate? (e.g. Chi-squared test, I-squared statistic)	Yes, meta-regression (interaction between treatment/TNFi exposure status)
	4. If the homogeneity assumption is not satisfied, is clinical or methodological homogeneity across trials in each set involved in the indirect comparison investigated by an adequate method? (e.g. subgroup analysis, sensitivity analysis, meta-regression)	Yes, separate analyses by TNFi-exposure status. Adjustments made for differences in study design.
Similarity		
	1. Is the assumption of similarity stated?	No
	2. Have they justified their assumption?	Yes, see above
Consistency		
	1. Does the analysis explicitly assess consistency?	Yes, partially

	2. Does the method described include a description of the analyses/ models/ handling of potential bias/ inconsistency/ analysis framework?	Yes
	3. Are patient or trial characteristics compared between direct and indirect evidence trials?	No
	4. If Q3 is yes, and inconsistency is reported, is this accounted for by not combining the direct and indirect evidence?	n/a no inconsistency reported ($p<0.05$)

3.1.7.1 Use of the probit scale to model clinical response/clinical remission

The company used the multinomial ordered probit scale for clinical response and clinical remission. By modelling clinical response and clinical remission jointly, the company avoided a situation where "*it would be possible to end up with a model that makes impossible predictions, for example that more patients experience clinical remission than experience clinical response*" (CS section B.1.1.1.1). The ERG agrees with this assessment.

In the previous NICE TA342,[9] Takeda used separate binomial logit models for clinical response and clinical remission which was criticised by the ERG:
The results for clinical response and remission should be interpreted with further caution because these were estimated without considering the dependence/correlation between response and remission (TA342, ERG report, p65).

We concur with the company on this point. Hence, the use of a multinomial probit model is an improvement as it takes account of this correlation between outcomes, which is fundamental for the economic model. It is also consistent with the Mayo score, which is essentially a continuous score divided into ordered categories. However, interpreting coefficients on the probit scale is difficult and non-intuitive. We suggest an alternative, the logit model, could have been considered which would have the advantage that the coefficients would be more interpretable.

We queried the company's use of the probit model in the clarification questions (question A18). We agree that separate binomial logit models for response and remission could have introduced inconsistent results across categories of response. However, a multinomial logit analysis could have been considered. The multinomial logit has been previously used in psoriatic arthritis for ordered categorical data.[44] We do not expect such a model would have

resulted in different results but would have aided the interpretability and readability of the company's submission.

Whilst the main analyses tables report the odds ratios and probabilities along with the probit or log odds (e.g. CS Tables 25 and 26), other tables in the sensitivity analysis report results on just the probit scale (e.g. CS Tables 27 and 28). Furthermore, some tables headings are labelled as median treatment effect without acknowledging the scale (e.g. CS Tables 43 and 45, CS section B.3.3.1.1, should read "probit scale", CS Table 48, B.3.3.3, should read "log odds"). This lack of clarity added to the difficulty interpreting the probit scale impedes the readability of the company's submission.

3.1.7.2 Assessment of inconsistency

We noted the presence of closed loops in some of the networks. The company provided details of inconsistency checking and results in their response to the clarification questions (A19). They found no statistically significant inconsistency in the TNFI-naïve subgroup induction network (CS Figure 28) nor safety network (CS Figure 30). However, inconsistency in the maintenance TNFi-naïve network between the two-arm and three-arm trial was not examined.

3.1.7.3 Validation of company results and assessment of model fit.

The ERG replicated selected results to validate the analysis. No errors were found in the company's code. Our validation prioritised the following outputs which contributed to the economic model but we also looked at serious adverse events given the rarity of the serious infections endpoint.

1. CS Table 26 – response/remission fixed effects model in TNFi-naive subgroup (Maintenance phase), using input data from CS Table 93. Probit.
2. CS Table 25 – response/remission fixed effects model in TNFi-exposed subgroup (Induction phase), using input data from CS Table 43. Probit.
3. CS Table 34 - Serious infections random effects model (Induction phase), using input data from CS Table 96. Log odds.
4. CS Table 33 - Serious adverse events fixed effects model (Induction phase), using input data from CS Table 96. Log odds.
5. CS Table 28 – response/remission fixed effects model in TNFi-failure subgroup (maintenance phase), using data from Table 99. Probit.

Furthermore, the ERG conducted a number of additional analyses based around best model fit.

The company states in Appendix D (CS section D.1.3.3) that "*where the difference in DIC suggested indifference [i.e. a difference of less than 3 points], the simpler fixed effects model was preferred*". The ERG would have chosen the random effects model as the more conservative approach in such circumstances to account for between-study heterogeneity.

The ERG believes there is potential heterogeneity between studies which would favour using a random effects model in the base case, model fit being equal. Selected baseline characteristics of studies included in the NMA are presented in CS Table 87. A visual inspection of this table shows disease duration varies from 4.3 to 10.9 years, and IBDQ score varies from 114 to 167. One of our experts identified disease extent (extensive/pan-colitis vs left sided disease) which is not well reported but varies between studies (CS Table 87) as well as albumin, haemoglobin, and baseline C-reactive protein as other potential effect modifiers. These are potentially unobserved sources of heterogeneity. Furthermore, baseline characteristics are not compared by TNFi-exposure status which further precludes an effective qualitative assessment of heterogeneity. In addition to the differences in maintenance design, there is also a difference in the inclusion criteria of re-randomised trials. GEMINI 1,[41] PURSUIT-M,[39] and PURSUIT-J[40] allowed only active treatment responders to enter the maintenance period, whereas OCTAVE Sustain allowed all responders, whether on active treatment or placebo, to enter the maintenance period. The ERG in NICE TA342 also noted that due to the presence of heterogeneity, the fixed effects model would underestimate uncertainty.

Model fit statistics are presented in CS Table 23 (response/remission - Induction phase; CS section B.2.9.2.1.1), Table 24 (response/remission - maintenance phase; CS section B.2.9.2.1.1), and Table 31 (safety outcomes– Induction period; CS section B.2.10.8.1). The first column of CS Table 24 is mislabelled as Induction whereas it is for the Maintenance phase. We have summarised the choice of company base-case model and the ERG preferred model for each of the analyses in Table 13.

Table 13 Company choice of base-case and ERG preference

	Company base-case model	ERG favoured model
Clinical response/clinical remission, Induction TNFi-naive	Random effects	Random effects
Clinical response/clinical remission, Induction TNFi-exposed	Fixed effects	Random effects
Clinical response/clinical remission, Maintenance TNFi-naive	Fixed effects	Random effects
Clinical response/clinical remission, Maintenance TNFi-exposed	Fixed effects	Fixed effects
Serious infections, Induction	Random effects	Fixed effects

In the induction phase TNFi-exposed subgroup, the fixed effects model was preferred despite similar DIC and similar total residual deviance. The ERG would have selected the random effects model as the more conservative analysis. Whilst the base case models are presented in the main NMA results (CS Table 25) the alternative model is not reported. We would prefer to have seen this explored as a sensitivity analysis.

Similarly, the company preferred the fixed effects model in the maintenance phase TNFi-naïve population for clinical response/remission. The ERG would have chosen the random effects model for both the lower DIC and total residual deviance. The ERG would prefer to have seen this explored as a sensitivity analysis.

Finally, the company chose the random effects model for serious infections. In response to a clarification request the company provided the random effect standard deviation (1.82, 95%CrI 0.15, 4.59) (clarification question A22). This wide CrI indicates weak support for the random effects model which has a similar DIC, thus we might have favoured the fixed effects model. The ERG would prefer to have seen the fixed effects model included in a sensitivity analysis.

Table 14 and Table 15 show the results of the ERG validation and exploratory analysis for the response and remission analyses. The ERG ran the same number of chains, burn-in and

simulations reported by the company (section D.1.3.3). Models converged and our results concur to two decimal places.

The alternative choice random effects models show wider credible intervals and some variation in the median estimates for adalimumab and golimumab in the maintenance analysis for the TNFi-naïve population as smaller studies are given more weight under the random effects than the fixed effects model.

Table 14 ERG replication and additional analysis on model choice - clinical response and clinical remission for TNFi-naïve subgroup

Comparator	Treatment effect vs placebo, median (95% CrI), probit scale[a]		
	Company base-case (fixed effects)	ERG replication of base-case (fixed effects)	ERG alternative model selection (random effects)
Maintenance phase			
Tofacitinib 5 mg	[redacted]	[redacted]	[redacted]
Tofacitinib 10 mg	[redacted]	[redacted]	[redacted]
Infliximab 5 mg/kg	[redacted]	[redacted]	[redacted]
Adalumimab 40 mg Q2W	[redacted]	[redacted]	[redacted]
Golimumab 50 mg	[redacted]	[redacted]	[redacted]

Source of company base-case (fixed effects) is CS Table 26
[a] On the probit scale, negative coefficients indicate improvement over placebo. Where the upper and lower CrI are both negative, treatments show strong evidence of benefit versus placebo.

Table 15 ERG replication and additional analysis on model choice - clinical response and clinical remission for TNFi-exposed subgroup

Comparator	Treatment effect vs placebo, median (95% CrI), probit scale[a]		
	Company base-case (fixed effects)	ERG replication of base-case (fixed effects)	ERG alternative model selection (random effects)
Induction phase			
Tofacitinib 10 mg	[redacted]	[redacted]	[redacted]
Adalumimab 160/80/40 mg	[redacted]	[redacted]	[redacted]
Vedolizumab 300 mg	[redacted]	[redacted]	[redacted]

Source of company base-case (fixed effects) is CS Table 25
[a] On the probit scale, negative coefficients indicate improvement over placebo. Where the upper and lower CrI are both negative, treatments show strong evidence of benefit versus placebo.

However, when we attempted to replicate the serious infections results there was a higher level of uncertainty around the coefficients particularly for tofacitinib (Table 16). The wider credible intervals persisted under the fixed effects model conducted by the ERG.

Table 16 ERG replication and additional analysis on model choice - serious infections

Comparator	Treatment effect vs placebo, median (95% CrI), logit scale		
	Company base-case (random effects)	ERG replication of base-case (random effects)	ERG alternative model selection (fixed effects)
Tofacitinib 10 mg	[redacted]	41.42 (4.66, 125.3)	38.72 (3.52, 96.9)
Infliximab 10 mg/kg	[redacted]	-0.56 (-6.82, 5.61)	-0.51 (-2.8, 1.52)
Adalumimab 160/80/40 mg	[redacted]	-0.21 (-5.86, 5.44)	-0.1 (-1.74, 1.49)
Golimumab 200/100 mg	[redacted]	-2.28 (-10.07, 5.28)	-2.12 (-5.50, -0.17)
Vedlizumab 300 mg	[redacted]	-1.90 (-9.71, 5.79)	-1.78 (-5.23, 0.47)
Azathioprine	[redacted]	-0.59 (-10.74, 9.6)	-0.55 (-4.8, 3.63)

Source of company base-case (fixed effects) is CS Table 34

The very wide credible intervals for tofacitinib are caused by the lack of any serious infections across placebo arms in the three tofacitinib studies, hence the difficulty to estimate a relative treatment effect compared to placebo (Table 17). There was also considerable autocorrelation in the tofacitinib coefficient despite thinning and running an extended number of simulations.

The reasons for the difference in our results are unclear, particularly how the company arrived at their estimate for tofacitinib.

Table 17 Tofacitinib induction phase serious infections used in NMA (data from CS Table 96)

Study name	Treatment arm	Serious Infections, n/N (%)
OCTAVE Induction 1	Placebo	0/122 (0%)
	Tofacitinib 10 mg	6/476 (1%)

Study name	Treatment arm	Serious Infections, n/N (%)
OCTAVE Induction 2	Placebo	0/112 (0%)
	Tofacitinib 10 mg	1/429 (0%)
Phase II trial	Placebo	0/48 (0%)
	Tofacitinib 10 mg	2/33 (6%)

As an alternative, we ran the induction phase serious infections analysis in a frequentist framework using the NMA web app developed by Owen and colleagues at the Complex Reviews Support Unit (CRSU) [https://crsu.shinyapps.io/metainsightc/]. The engine underneath this app is Netmeta, which being frequentist, adds 0.5 to zero cells, which results in better convergence and a smaller variance for tofacitinib (Figure 6).

We acknowledge the controversy over adding an arbitrary 0.5 to cells.[45] Nevertheless, we would argue this is a reasonable approximation under the circumstances. If we assume the placebo arms across studies are homogeneous then it seems unjust to encumber tofacitinib with a huge variance for not having a serious infection in any of their placebo arms (the OCTAVE Sustain placebo arm had two serious infections, akin to active treatment which had two in the 5 mg dose, and one in the 10 mg dose but no safety NMA was conducted for the maintenance phase). Of the other five studies with a placebo arm included in the safety analysis, only one had zero events (Suzuki 2014[37]), but similar treatment comparisons in other studies enabled relative treatment effects to be calculated. Random effects results (Figure 6) were generally consistent with the CS albeit all credible intervals were much smaller as was the mean effect for tofacitinib.

Although the UC-SUCCESS study[36] comparing azathioprine to infliximab was included in the safety network it is unclear why azathioprine was not included in the NMA results. We have retained azathioprine in our additional analysis as it appears to meet the inclusion criteria.

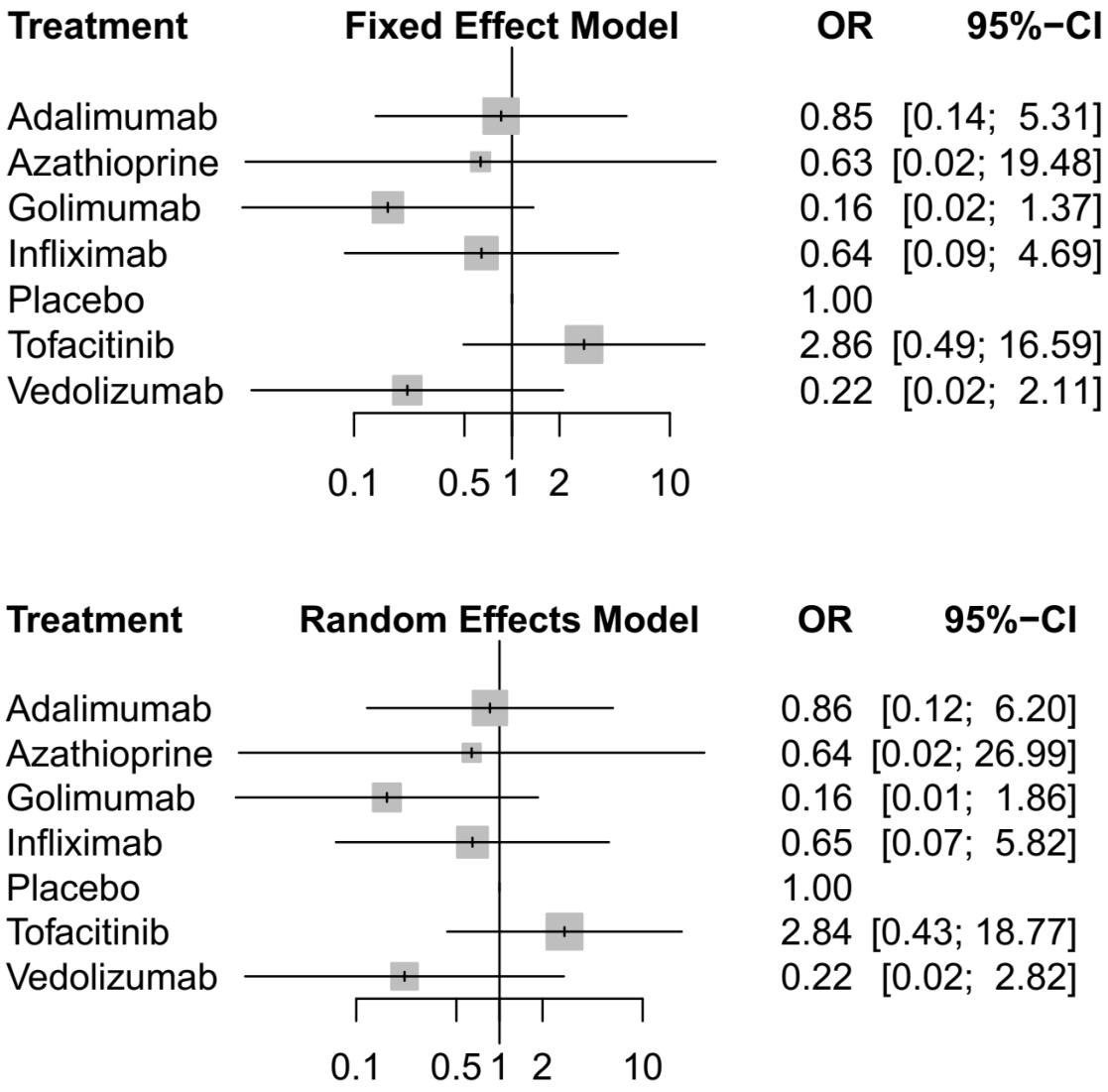


Figure 6 ERG additional analysis - frequentist models for serious infections, courtesy of CRSU web app

Finally, in the NMA base-case analysis for the TNFi-exposed subgroup in the maintenance phase, only TNFi-failed data were available from GEMINI 1 (vedolizumab) (CS Table 22). This may have introduced bias against vedolizumab.

The company conducted a sensitivity analysis using TNFi-failure data from both OCTAVE Sustain and GEMINI 1. However, the maintenance phase analysis "*could not be run because there were too few data points to estimate the multinomial probit model parameters*" (CS section B.2.9.3.2, CS Table 28), essentially because ULTRA 2 was dropped from this analysis. The ERG conducted the analysis using the TNFi-exposed data from ULTRA 2 (adalimumab). In our opinion, this introduced no more bias than the base

case which combined TNFi-exposed data for tofacitinib and adalimumab with TNFi-failure data for vedolizumab. Our scenario analysis at least included comparable data for tofacitinib and vedolizumab.

In the event, as Table 18 shows, use of TNFi-failure data makes little difference to the response/remission results for tofacitinib.

Table 18 ERG scenario analysis using TNFi-failure data from both OCTAVE Sustain and GEMINI 1

Comparator	Treatment effect vs placebo, median (95% CrI), probit scale[a]		
	Company base-case (fixed effects)	ERG replication of base-case (fixed effects)	ERG exploratory scenario analysis (fixed effects)
Maintenance phase			
Tofacitinib 5 mg	[redacted]	[redacted]	[redacted]
Tofacitinib 10 mg	[redacted]	[redacted]	[redacted]
Adalumimab 40 mg Q2W	[redacted]	[redacted]	[redacted]
Vedolizumab 300 mg Q8W	[redacted]	[redacted]	[redacted]
Vedolizumab 300 mg Q4W	[redacted]	[redacted]	[redacted]

Source of company base-case (fixed effects) is CS Table 28
[a] on the probit scale, negative coefficients indicate improvement over placebo. Where the upper and lower CrI are both negative, treatments show strong evidence of benefit versus placebo.

3.1.7.4 Baseline response models – uncertainty around absolute probabilities

To estimate absolute probabilities of each event, treatment effects from the NMA were combined with an estimate of the placebo (baseline) response from the placebo arms of included studies. In response to clarification request A17 the company provided the data, priors and output (meanA, precA) in WinBUGs code format for the probit baseline models. We were able to replicate selected median estimates for the baseline calculations. However, despite running the CS code [validated against NICE DSU Technical Support Document (TSD) 2[46]] and data we were unable to replicate the baseline credible intervals used in the

probit or logit models. The company models tended to lead to wider credible intervals compared to our calculations, thus would lead to conservative results. A summary of the differences in our findings is provided in Table 19 below.

Table 19 ERG replication of baseline (placebo) response results

Comparator	Treatment effect vs placebo, median (95% CrI)	
	Company baseline	ERG replication of company baseline
Induction TNFi-exposed, probit scale		
Response/remission	[redacted]	[redacted]
Maintenance TNFi-naïve, probit scale		
Response/remission	[redacted]	[redacted]
Induction, logit scale		
Serious Infections	[redacted]	[redacted]
Serious adverse events	[redacted]	[redacted]

3.1.7.5 Inclusion of the tofacitinib phase II trial

The Sandborn 2012 Phase II (induction) tofacitinib trial[11] is less well described in the CS despite being included in the NMAs. Furthermore, the company state:
All studies, except for one [Sandborn 2012], were conducted in patients with moderately to severely active ulcerative colitis who had an inadequate response to or had failed to tolerate one or more of the following conventional therapies: oral or intravenous corticosteroids, azathioprine, and/or 6-mercaptopurine (CS section B.2.9.1.1).

The ERG thus questioned the eligibility of this trial. The company confirmed that the Phase II trial met the inclusion criteria for the NMA and they also provided selected NMA results obtained with the Phase II trial excluded from the NMA (Table 7 in clarification response A16). These results for response and remission for the TNFi-naïve and TNFi-exposed populations in the induction period were similar to the base case (CS Table 25).

Base case results without the Phase II trial were not provided for the safety outcomes. However, given the relatively high serious infection rate in the tofacitinib arms of the Phase II trial compared to the OCTAVE trials (6% [2/33] patients had an event compared to 1% [6/476] in OCTAVE Induction 1 and none in OCTAVE Induction 2), the Phase II trial may

have had a disproportionate effect on the random effects NMA results, and we consider this to be a conservative analysis.

3.1.7.6 No safety NMA in the trials' maintenance period

The company said they were unable to perform an NMA for safety in the maintenance phase due to the aforementioned differences in study design, and carryover effects of active treatment on the placebo responders (OCTAVE trials only). These are the same reasons given for the need to adjust the treat-through trials for the response/remission outcomes. Of course, the latter bias could have been averted by using the mFAS population (i.e. excluding the placebo responders from OCTAVE Sustain thereby matching the GEMINI 1 population) of OCTAVE Sustain. Clinical experts advising the ERG suggested that adverse events are likely to increase with drug exposure over time, but it is unclear whether this would have introduced bias and, if so, in which direction.

3.1.7.7 Adjustment for differing lengths of the induction and maintenance periods across trials

The length of the induction phase ranged from six weeks for golimumab and vedolizumab to eight weeks for the other treatments. The maintenance phase ranged from 44 weeks to 54 weeks. Adalimumab had the shortest maintenance phase and golimumab the longest. These are summarised in Table 44 of the company's submission (CS section B.3.3.1.2) which is summarised here (Table 20).

Table 20 Duration of induction and maintenance phases of trials (CS Table 44)

	Induction phase (weeks)	Maintenance phase (weeks)	Total duration (weeks)	Maintenance design
Tofacitinib	8	52	60	Re-randomised
Adalimumab	8	44	52	Treat-through
Golimumab	6	54	60	Re-randomised
Infliximab	8	46	54	Treat-through
Vedolizumab	6	46	52	Re-randomised

In response to our request for clarification the company confirmed that they did not attempt to adjust for different lengths of the induction and maintenance phases across studies. The company noted:
it would have been impossible to properly estimate what difference was due to the treatment effect and what difference was the effect of an earlier measure (company's clarification response A20).

The meaning of this is unclear. However, one of our clinical experts suggested that a shorter induction phase may influence response and that it was entirely possible to see a higher response rate at week 8 than week 6. Our expert referred to the GEMINI 3 (vedolizumab) study in Crohn's disease where it became clear that the 6-week induction phase had failed to capture a majority of responders.

In our opinion, this could have introduced potential bias against studies with shorter induction phases, namely golimumab and vedolizumab. The company in TA342 performed a complementary log-log model (TSD2) to adjust for differences in follow-up in the induction phase (TA342 Company's submission, section 6.7.5). This assumes a Poisson process for each trial arm and a constant event rate and can be applied to binomial and multinomial models (TSD2[46]).

Furthermore, in the induction phase, CS Table 96 (Appendix D, p233) suggests 12-week induction data for the Phase II trial[11] and 14-week data for Kobayashi 2015[43] were used in the safety analysis. This appears to contradict CS Table 44.

With respect to the maintenance phase, the company referred to previous NICE appraisals, in particular that the ERG and appraisal committee for NICE TA342[9] did not believe differences in the length of the maintenance phase would impact results. However, this seems to refer to a difference of between 52 and 54 weeks in the maintenance period. (6.7.3, p125 TA342 company's submission) which are smaller than the differences in Table 1 above. In any case, we concur that the company's base case is likely to be a conservative assumption. Studies with a shorter maintenance phase would experience fewer responders losing response, given the assumption that response wanes slowly over time. Hence this could benefit those treatments with a shorter maintenance phase (i.e. golimumab and vedolizumab) but would be conservative for tofacitinib.

3.1.7.8 Differences between patient populations in the re-randomised maintenance trials.

As noted in section 3.1.7.3 above, unlike the other re-randomisation trials OCTAVE Sustain allowed all responders, whether on active treatment or placebo, to enter the maintenance period. This is a source of heterogeneity and might also be a potential source of bias if placebo responders in OCTAVE Sustain were less able to sustain their response or potentially more susceptible to active treatment, although the direction of any bias is unclear.

However, the company conducted an analysis using a modified Full Analysis Set (mFAS) population which explicitly excluded placebo responders (CS section B.2.4.1). This mFAS population is consistent with the GEMINI 1, PURSUIT-M, and PURSUIT-J maintenance populations and would also have ensured comparability across the placebo groups of the re-randomised trials. Selected results from the mFAS population for OCTAVE Sustain are presented in Appendix L, but only include centrally-read clinical remission. Consequently there are insufficient data to conduct this analysis for the NMA or economic model. As a proxy for the direction of effect of any bias, we compared centrally read remission at 52 weeks in the FAS (CS Figure 10, section B.2.6.2.1.1) and mFAS (CS Table 207, Appendix L.1.2) populations. Remission at 52 weeks was slightly lower in both tofacitinib arms using the mFAS population, suggesting that the base case NMA results may be slightly biased in favour of tofacitinib.

Hence, the ERG believes the mFAS population could have been made the base case or at least explored in a sensitivity analysis.

3.1.7.9 Adjustments to treat-through trials

The company considered that heterogeneity in the study design in the maintenance phase would have introduced bias had they used the reported clinical response and clinical remission data. Furthermore, some placebo patients in the maintenance phase had also received active treatment in the induction phase (OCTAVE Sustain only).

The company considered two methods to adjust for these differences in design. The first was to adjust the re-randomised trials to better match the treat-through design (following an approach used by Thorlund 2015a[47]) and the second was to adjust the threat-through studies to match the re-randomised (the approach used by Takeda[48] in NICE TA342).

The company favoured the latter approach similar to Takeda in NICE TA342 because it required "less data manipulation" and was more aligned with clinical practice and the economic model. The ERG concurs with this choice, which is also acknowledged by Thorlund 2015b[49] for whom "Published data did not allow us to adjust [treat-through] results to fit a re-randomised design" but recognised that the "Re-randomisation design … may mimic a more realistic clinical application of biologic therapy, wherein patients are given a trial of therapy for induction and those who respond are subsequently considered for maintenance dosing."

In NICE TA342, Takeda assumed that patients who responded at 12 months must also have responded at the end of induction, and they used inflation factors to adjust the event rates in both the active treatment and placebo arms for the treat-through trials. However, the exact calculations utilised are unknown since details are unavailable on the NICE website. Hence we cannot tell if the same methods were used in the CS.

The ERG in NICE TA342 criticised Takeda's approach since it "ignores the fact that non-responders at the end of induction could have become responders at the end of the maintenance phase" and
The ERG believes that the adjustment applied to the trials without re-randomisation at the end of the induction phase by the company did not adjust the bias sufficiently, rather, it is possible that their adjustment method actually introduced more bias into the analysis
(TA342, ERG report, p64)

In the CS, the company made the same assumption that the "*number of responders at end of induction period is a proxy for the total number of patients entering maintenance*" (CS Appendix D.1.3.2.1). This could potentially introduce bias against comparators in those studies which had a shorter induction phase as noted above.

The company made the following adjustments to the data in the treat-through trials to better match the re-randomised trials:

- the proportion of patients achieving "sustained clinical response" was used as the clinical response for the treat-through trials "*as this mitigates the risk of counting maintenance phase responders who were induction phase non-responders*" (CS Appendix D.1.3.2.1).
- For the Suzuki 2014 trial,[37] sustained clinical response was not reported for the placebo arm; instead, the ratio of sustained clinical responders to clinical responders

was estimated by the company from the ULTRA 2 (adalimumab) trial[50] and applied to Suzuki 2014.

- The average proportion of clinical remitters among clinical responders in the re-randomised placebo arms was applied to the placebo arms of the treat-through trials.
- For active treatments, the company used numbers of clinical remitters who were induction phase responders.

The ERG is unclear whether these calculations would have introduced any further bias beyond the original criticism in NICE TA342 that non-responders at the end of the induction phase are ignored. Whilst the re-randomisation design ignores non-responders, bias could have been introduced if relative treatment effects on non-responders differ or if the induction phase were of a different length. We believe the use of sustained clinical response has the potential to introduce additional bias against the "treat-through" studies albeit we are unclear whether it has done so.

Summary of the ERG's critique of the NMA approach

The company's NMAs were generally well conducted and made a number of efforts to minimise bias. Nevertheless, the ERG believes a number of potential biases remain.

- Our choice of random effects models for the induction TNFi-exposed and maintenance TNFi-naïve subgroups may have mitigated some concerns over heterogeneity
- Our choice of a frequentist model for serious infections may have mitigated bias from high uncertainty around rare events
- The differences in the uncertainty around our baseline response calculations compared to the company's may lead to conservative results
- There may be undetected inconsistency in the maintenance TNFi-naïve network which could lead to bias in the golimumab estimates. The direction of effect is unclear.
- The lack of safety analysis in the maintenance period may have introduced bias from longer drug exposure but the direction of effect is unclear.
- In the adjustment to the treat-through maintenance trials to match the re-randomised trials, the use of sustained clinical response has the potential to introduce bias albeit we are unclear whether it has done so.
- Potential bias remains with respect to the differences between the re-randomised populations (inclusion of placebo responders in OCTAVE Sustain). The direction of bias is uncertain but may favour tofacitinib.

- Bias may remain with respect to the different lengths of the induction and maintenance phases. The direction of bias may be in favour those studies with a shorter maintenance phase analysis and against those studies with a shorter induction phase.

3.2 Summary statement of company's approach

The ERG's assessment of the company's approach to the evidence synthesis is summarised in Table 21.

Table 21 Quality assessment (CRD criteria) of CS review

CRD Quality Item with ERG comments	
1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question?	1. Yes. Eligibility criteria are tabulated (CS Appendix D.1.1.3 Table 83) and generally appropriate. An exception is that the stated population eligibility criteria are broader than the NICE scope; however, the populations of the studies that were finally included in the company's SLR do match the NICE scope. Outcome measures did not form part of the eligibility criteria.
2. Is there evidence of a substantial effort to search for all relevant research? i.e. all studies identified	2. Yes. All literature searches were systematic and transparent, and are well-documented and reproducible, although over 6 months out of date. An adequate range of bibliographic databases was searched. Supplementary sources and key conferences were also searched. The ERG conducted a rapid update search, which identified four additional relevant full-text publications not listed in the CS and two new conference abstracts reporting results from the OCTAVE trials (see Section 3.1.1). However, these publications either duplicated information already present in the CS or are not directly relevant to the current scope.
3. Is the validity of included studies adequately assessed?	3. Yes. The company assessed the risk of bias in the OCTAVE RCTs (CS Table 19) and the RCTs included in the CS that form part of the NMA (CS Table 86), using the critical appraisal checklist provided by NICE in the Single Technology Appraisal (STA) user guide.[51] (for further details see Section 3.1.4).
4. Is sufficient detail of the individual studies presented?	4. Yes. The CS presents sufficient detail of OCTAVE 1, 2 and OCTAVE Sustain, including general methods (CS

	Tables 9 and 13), eligibility criteria (CS Table 10), participant baseline characteristics (CS Table 15), statistical methods (CS Table 16), outcomes (CS Tables 9, 11, 12, 13), subgroups (CS Tables 9 & 13) and participant flow (CS Tables 17 and 18; Figures 46 and 47). The CS reports limited information on the Phase II trial, but the company and ERG considered this trial to be of less importance than the Phase III trials and the ERG considers the brevity of reporting acceptable.
5. Are the primary studies summarised appropriately?	5. Yes. Clinical effectiveness results from the OCTAVE 1, 2 and OCTAVE Sustain trials are clearly summarised (CS sections B.2.6 and B.2.7, with results from the NMA trials summarised in CS section B.2.9 and Appendix D.1.2). Adverse events are summarised in CS section B.2.10 and CS Appendix F.

The company's evidence synthesis is generally well structured and uses standard methodology. The company's search for clinical effectiveness studies is over 6 months out of date. However, an ERG search update did not identify any missing tofacitinib trials.

The population eligibility criteria for the company's SLR as stated in the CS are broader than the NICE final scope, but the populations of the studies finally included in the SLR are consistent with the NICE scope. With the exception of time to surgical intervention (not reported in the OCTAVE trials), outcome measures reported in the CS match the outcome categories listed in the NICE final scope.

The CS does not include all of the patient-reported outcomes (PROs) that were measured in the OCTAVE Induction and OCTAVE trials, although the PROs that are reported appear adequate (EQ-5D informs the economic analysis) and the risk of reporting bias appears to be low (see Table 9).

Overall, there appears to be a low risk of systematic error in the systematic review of the CS based on the methods employed.

3.3 Summary of submitted evidence

A noted earlier (section 3.1.5) the Mayo endoscopic sub-score was assessed both locally and centrally in the OCTAVE trials. Consequently outcomes that utilise the endoscopic sub-

score were reported separately using the local or the central read of the endoscopic data in the CS. Locally read data were used in the base-case NMAs for the outcomes that contribute data to the economic model.

3.3.1 Summary of results for Remission (Primary endpoint in OCTAVE 1 and 2 and Sustain)

OCTAVE Induction trials 1 and 2:

In both the OCTAVE 1 and OCTAVE 2 trials, a statistically significant difference in remission at week 8 in comparison to placebo was observed in participants who received tofacitinib 10 mg twice daily (Table 22). When endoscopic sub-scores were centrally read in OCTAVE 1, 18.5% of those in receipt of tofacitinib 10 mg were in remission at week 8 in comparison to 8.2% of the placebo group (mean difference from placebo 10.3 percentage points, 95% CI 4.3 to 16.3, p-value 0.007). The corresponding data for OCTAVE 2 are 16.6% in the tofacitinib group in remission at week 8 versus 3.6% of the placebo group, mean difference from placebo 13.0 percentage points, 95% CI 8.1 to 17.9, p-value <0.001).

The locally read data produced mean differences that were 2-3 percentage points higher than those of the centrally read data, but were still statistically significant (Table 22),

This pattern was also observed for the pooled induction population (central read difference from placebo 11.6 versus 14.3 for the local read, p-values in both cases <0.0001).

As previously stated, these remission data are not used for economic modelling.

Table 22 Remission at week 8 in induction trials (FAS, NRI, central and local reads)

Parameter, n (%)	TOF 10 mg	PBO	Difference vs PBO, mean (95% CI)
OCTAVE 1	**N=476**	**N=122**	
Central read	88 (18.5)	10 (8.2)	10.3 (4.3–16.3); p=0.007
Local read	118 (24.8)	14 (11.5)	13.3 (6.5–20.2); p=0.0017
OCTAVE 2	**N=429**	**N=112**	
Central read	71 (16.6)	4 (3.6)	13.0 (8.1–17.9); p<0.001
Local read	89 (20.7)	6 (5.4)	15.4 (9.7–21.1); p=0.0002
OCTAVE 1 & 2 pooled data	**N=905**	**N=234**	
Central read	159 (17.6)	14 (6.0)	11.6 (7.7–15.5); p<0.0001
Local read	207 (22.9)	20 (8.5)	14.3 (9.8–18.8); p<0.0001

Source CS Figure 6 and Appendix L.1.4 Table 213

OCTAVE Sustain

In the OCTAVE Sustain maintenance trial, a statistically significant difference in remission at week 52 in comparison to placebo was observed both for participants who received tofacitinib 10 mg twice daily and those who received tofacitinib 5 mg twice daily (Table 23). When endoscopic sub-scores were centrally read, 34.3% of those in receipt of tofacitinib 5 mg were in remission at week 52 in comparison to 11.1% of the placebo group (mean difference from placebo 23.2 percentage points, 95% CI 15.3 to 31.2, p-value <0.001). A greater proportion of the tofacitinib 10 mg group were in remission at week 52 (40.6%), so consequently the difference in comparison to placebo was also greater (29.5 percentage points, 95% CI 21.4 to 37.6, p<0.001).

The local read data again produced less conservative results than the central read data, with the percentage difference between tofacitinib 5 mg and placebo approximately 3 percentage points higher at 26.3 (95% CI 19.0 to 34.5, p<0.0001) and that between tofacitinib 10 mg and placebo approximately 5 percentage points higher at 34.6 (95% CI 26.2 to 43.0, p<0.0001).

In addition to remission at week 52, data were also reported for participants with sustained remission (i.e. remission at both week 24 and week 52). For both the 5 mg and 10 mg tofacitinib doses and regardless of whether the central or local read data were used, the results were statistically significantly in favour of tofacitinib, with the greater percentage difference in comparison to placebo being obtained with the 10 mg dose (Table 23).

Lastly, remission and sustained remission were also reported for the subset of patients who entered the OCTAVE Sustain maintenance trial in remission. Among these patients, less than 12% of those in the placebo group maintained their remission, whereas in the 5 mg tofacitinib group there mean percentage difference in comparison to placebo was over 30 percentage points and was over 42 percentage points in the 10 mg tofacitinib group. Differences against placebo (at either tofacifinib dose and using both central and local read data) were all statistically significant (Table 23).

None of the remission data are used for economic modelling.

Table 23 Remission outcomes in maintenance trial (FAS, NRI, central and local reads)

Parameter, n (%)	TOF 5 mg N=198	PBO N=198	% difference vs PBO (95% CI)	TOF 10 mg N=197	% difference vs PBO (95% CI)
Remission at week 52					
Central read	68 (34.3)	22 (11.1)	23.2 (15.3–31.2); p<0.001	80 (40.6)	29.5 (21.4–37.6); p<0.001
Local read	78 (39.4)	26 (13.1)	26.3 (18.0–34.5); p<0.0001	94 (47.7)	34.6 (26.2–43.0); p<0.0001
Sustained remission at weeks 24 and 52					
Central read	44 (22.2)	10 (5.1)	17.2 (10.6–23.7); p<0.001	50 (25.4)	20.3 (13.5–27.1); p<0.001
Local read	62 (31.3)	19 (9.6)	21.7 (14.1–29.4) p<0.0001	73 (37.1)	27.5 (19.6–35.4); p<0.0001
Remission at week 52 among patients in remission at baseline, n/total n (%)					
Central read	30/65 (46.2)	6/59 (10.2)	36.0 (21.6–50.3); p<0.001	31/55 (56.4)	46.2 (31.0–61.4); p<0.001
Local read	32/65 (49.2)	7/59 (11.9)	37.4 (22.7–52.1) p<0.0001	32/55 (58.2)	46.3 (30.9–61.7); p<0.0001
Sustained remission at wks 24 & 52 among patients in remission at baseline, n/total n (%)					
Central read	24/65 (36.9)	3/59 (5.1)	31.8 (18.8–44.8); p<0.001	26/55 (47.3)	42.2 (27.9–56.5); p<0.001
Local read	32/65 (49.2)	7/59 (11.9)	37.4 (22.7–52.1) p<0.0001	32/55 (58.2)	46.3 (30.9–61.7); p<0.0001

Source CS Figure 10 and Appendix L.1.4 Table 221

3.3.2 Summary of results for mucosal healing (Key secondary endpoint in OCTAVE 1 and 2 and OCTAVE Sustain)

OCTAVE Induction trials 1 and 2

The proportion of participants with mucosal healing at week 8 was statistically significantly greater in the tofacitinib 10 mg group in both the OCTAVE 1 and OCTAVE 2 trials in comparison to the placebo group (Table 24). For centrally read data in OCTAVE 1, 31.3% of those in the tofacitinib group had mucosal healing at week 8 in comparison to 15.6% of the placebo group (mean difference from placebo 15.7 percentage points, 95% CI 8.1 to 23.4, p-value 0.001). The corresponding data for OCTAVE 2 are 28.4% in the tofacitinib group in

remission at week 8 versus 11.6% of the placebo group, mean difference from placebo 16.8 percentage points, 95% CI 9.5 to 24.1, p-value <0.001).

Greater differences between the two arms of the trials in favour of tofacitinib 10 mg twice daily were observed when using the local read data. For OCTAVE 1 the local read difference from placebo was almost four percentage points higher than the central read difference at 19.5 (95% CI 10.8 to 28.2) and the local read difference was over four percentage points higher than the central read difference for OCTAVE 2 (21.2, 95% CI 13.1 to 29.2). In both trials the local read difference from placebo was statistically significant (p<0.0001).

In the pooled induction population the central read difference from placebo was 16.3 versus 20.3 for the local read, with p-values in both cases <0.0001.

These mucosal healing data are not used for economic modelling, but an NMA was conducted for this outcome (Section 3.3.9.2 below).

Table 24 Mucosal healing week 8 in induction trials (FAS, NRI, central and local reads)

Parameter, n (%)	TOF 10 mg	PBO	Difference vs PBO, mean (95% CI)
OCTAVE 1	**N=476**	**N=122**	
Central read	149 (31.3)	19 (15.6)	15.7 (8.1–23.4); p<0.001
Local read	202 (42.4)	28 (23.0)	19.5 (10.8–28.2); p<0.0001
OCTAVE 2	**N=429**	**N=112**	
Central read	122 (28.4)	13 (11.6)	16.8 (9.5–24.1); p<0.001
Local read	156 (36.4)	17 (15.2)	21.2 (13.1–29.2); p<0.0001
OCTAVE 1 & 2 pooled data	**N=905**	**N=234**	
Central read	271 (29.9)	32 (13.7)	16.3 (11.0–21.6); p<0.0001
Local read	358 (39.6)	45 (19.2)	20.3 (14.4–26.3); p<0.0001

Source CS Figure 7 and Appendix L.1.4 Table 214

OCTAVE Sustain maintenance trial

At week 52 in the OCTAVE Sustain maintenance trial, the proportion of participants with mucosal healing was statistically significant better in comparison to placebo for participants who received tofacitinib 10 mg twice daily, and those who received tofacitinib 5 mg twice daily (Table 25). When endoscopic sub-scores were centrally read, 37.4% of those in

receipt of tofacitinib 5 mg were in remission at week 52 in comparison to 13.1% of the placebo group (mean difference from placebo 24.2 percentage points, 95% CI 16.0 to 32.5, p-value <0.001). In the tofacitinib 10 mg group, a greater proportion were in remission at week 52 (45.7%, percentage difference vs placebo 32.6 percentage points, 95% CI 24.2 to 41.0, p<0.001).

The local read data again produced less conservative results than the central read data (Table 25).

In addition to mucosal healing at week 52, data were also reported for participants with sustained mucosal healing (i.e. mucosal healing at both week 24 and week 52). For both the
5 mg and 10 mg tofacitinib doses and regardless of whether the central or local read data were used, the results were statistically significantly in favour of tofacitinib (Table 25).

Lastly, among the subset of patients who entered the OCTAVE Sustain maintenance trial with mucosal healing, mucosal healing at week 52 and sustained mucosal healing at weeks 24 and 52 were reported. Differences against placebo (at either tofacitinib dose and using both central and local read data) were statistically significant (Table 25).

These mucosal healing data are not used for economic modelling, but an NMA was conducted for this outcome (Section 3.3.9.2 below).

Table 25 Mucosal healing outcomes in OCTAVE Sustain (FAS, NRI, central and local reads)

Parameter, n (%)	TOF 5 mg N=198	PBO N=198	% difference vs PBO (95% CI)	TOF 10 mg N=197	% difference vs PBO (95% CI)
Mucosal healing week 52					
Central read	74 (37.4)	26 (13.1)	24.2 (16.0–32.5); p<0.001	90 (45.7)	32.6 (24.2–41.0); p<0.001
Local read	89 (44.9)	31 (15.7)	29.3 (20.7–37.9) p<0.0001	106 (53.8)	38.2 (29.5–46.8); p<0.0001
Sustained mucosal healing at weeks 24 and 52					
Central read	55 (27.8)	13 (6.6)	21.2 (14.1–28.3); p< 0.001	65 (33.0)	26.4 (19.0–33.8); p< 0.001

Local read	82 (41.4)	25 (12.6)	28.8 (20.5–37.1) p< 0.0001	98 (49.7)	37.1 (28.7–45.5); p< 0.0001
Mucosal healing at week 52 among patients with mucosal healing at baseline, n/total n (%)					
Central read	44/105 (41.9)	12/101 (11.9)	30.0 (18.7–41.4); p< 0.001	49/89 (55.1)	43.2 (31.1–55.3); p< 0.001
Local read	48/105 (45.7)	14/101 (13.9)	31.9 (20.2–43.5) p< 0.0001	56/89 (62.9)	49.1 (37.0–61.1); p< 0.0001
Sustained mucosal healing at weeks 24 and 52 among patients with mucosal healing at baseline, n/total n (%)					
Central read	35/105 (33.3)	9/101 (8.9)	24.4 (13.8–35.0); p< 0.001	44/89 (49.4)	40.5 (28.7–52.3); p< 0.001
Local read	48/105 (45.7)	13/101 (12.9)	32.8 (21.3–44.4) p< 0.0001	53/89 (59.6)	46.7 (34.6–58.8); p< 0.0001

Source CS Figure 11 and Appendix L.1.4 Table 222

3.3.3 Summary of results for sustained corticosteroid-free remission among those in remission at baseline (Key secondary endpoint in OCTAVE Sustain)

OCTAVE Sustain

Among the 593 participants who had a response in either the OCTAVE 1 or 2 induction trials and were randomised into the OCTAVE Sustain trial, 179 were in remission at OCTAVE Sustain baseline. Of these participants, 35.4% in the tofacitinib 5 mg arm and 47.3% of the tofacitinib 10 mg arm were in a sustained corticosteroid-free remission at weeks 24 and 52 in comparison to 5.1% of the placebo group (based on central read data) (Table 26). The differences between the tofacitinib arms and the placebo arm were statistically significant (p<0.001 for both doses of tofacitinib vs placebo). Results based on local read data gave slightly higher percentage differences between tofacitinib and placebo.

These results did not contribute to economic modelling.

Table 26 Sustained corticosteroid-free remission at weeks 24 and 52 among those in remission at baseline (FAS, NRI, central and local reads)

Parameter, n (%)	TOF 5 mg N=65	PBO N=59	% difference vs PBO (95% CI)	TOF 10 mg N=55	% difference vs PBO (95% CI)
Central read	23 (35.4)	3 (5.1)	30.3 (17.4–43.2); p<0.001	26 (47.3)	42.2 (27.9–56.5); p<0.001
Local read	31 (47.7)	7 (11.9)	35.8 (21.1–50.5); p<0.0001	32 (58.2)	46.3 (30.9–61.7); p<0.0001

Source: CS Figure 12 and Appendix L.1.4 Table 225

3.3.4 Summary of results for clinical remission

Note that the definition of clinical remission is almost identical to the definition of the primary outcome remission, except that the rectal bleeding sub-score does not have to be zero (Section 3.1.5).

OCTAVE 1 and 2

Due to the similarity of the definitions for clinical remission and remission, the proportion of participants achieving clinical remission in OCTAVE 1 (Table 27) were identical to those achieving remission reported above (Table 22). In OCTAVE 2 a single patient in the tofacitinib group, who met the criteria for clinical remission but who had not met the criteria for remission, accounted for the difference between the remission and clinical remission outcomes. Consequently, the results were statistically significantly in favour of the tofacitinib 10 mg group for this outcome.

As observed with other outcomes, use of the local read data led to a greater percentage difference between tofacitinib and placebo than with the central read data. The local read data were used in the NMA (Section 3.3.9.1 below), which then contributed to the economic model.

Table 27 Clinical remission week 8 in induction trials (FAS, NRI, central and local reads)

Parameter, n (%)	TOF 10 mg	PBO	Difference vs PBO, mean (95% CI)
OCTAVE 1	**N=476**	**N=122**	
Central read	88 (18.5)	10 (8.2)	10.3 (4.3–16.3); p=0.007
Local read	118 (24.8)	14 (11.5)	13.3 (6.5–20.2); p=0.0017

OCTAVE 2	N=429	N=112	
Central read	72 (16.8)	4 (3.6)	13.2 (8.3–18.1); p<0.001
Local read	90 (21.0)	6 (5.4)	15.6 (9.9–21.3); p=0.0002
OCTAVE 1 & 2 pooled data	**N=905**	**N=234**	
Central read	160 (17.7)	14 (6.0)	11.7 (7.8–15.6); p<0.0001
Local read	208 (23.0)	20 (8.5)	14.4 (9.9–18.9); p<0.0001

Source: CS Figure 8 and Appendix L.1.4 Table 215

OCTAVE Sustain

Clinical remission outcomes were very similar to remission outcomes and favoured the tofacitinib groups. For the central read data, one participant in the tofacitinib 10 mg group attained clinical remission who had not met the criteria for remission, but outcomes in the tofacitinib 5 mg and placebo groups were identical to those for the primary outcome (Table 28). When locally read data were used, two patients (one in the tofacitinib 5 mg and one in the 10 mg tofacitinib group) met the criteria for clinical remission.

The local read data were used in the NMA (Section 3.3.9.1 below), which then contributed to the economic model.

Table 28 Clinical remission week 52 in maintenance study (FAS, NRI, central and local reads)

Parameter, n (%)	TOF 5 mg N=198	PBO N=198	% difference vs PBO (95% CI)	TOF 10 mg N=197	% difference vs PBO (95% CI)
Central read	68 (34.3)	22 (11.1)	23.2 (15.3–31.2); p<0.001	81 (41.1)	30.0 (21.9–38.2); p<0.001
Local read	79 (39.9)	26 (13.1)	26.8 (18.5–35.1); p<0.0001	95 (48.2)	35.1 (26.7–43.5); p<0.0001

Source: CS Figure 13 and Appendix L.1.4 Table 226

3.3.5 Summary of results for clinical response

OCTAVE 1 and 2

Over half of the participants in OCTAVE 1 and 2 achieved a clinical response by week 8 of treatment with tofacitinib 10 mg twice daily. In contrast, just under a third of participants in the placebo group had a clinical response (Table 29). The percentage difference between the tofacitinib group and the placebo group was statistically significant in both trials and for

both the central and locally read data. The 593 participants with a clinical response (central read) were eligible to enter the OCTAVE Sustain maintenance study.

The local read data were used in an NMA (Section 3.3.9.1) which contributed data to the economic model.

Table 29 Clinical response week 8 induction trials (FAS, NRI, central and local reads)

Parameter, n (%)	TOF 10 mg	PBO	Difference vs PBO, mean (95% CI)
OCTAVE 1	**N=476**	**N=122**	
Central read	285 (59.9)	40 (32.8)	27.1 (17.7–36.5); p<0.001
Local read	289 (60.7)	42 (34.4)	26.3 (16.8–35.8); p<0.0001
OCTAVE 2	**N=429**	**N=112**	
Central read	236 (55.0)	32 (28.6)	26.4 (16.8–36.0); p<0.001
Local read	249 (58.0)	33 (29.5)	28.6 (18.9–38.2); p<0.0001
OCTAVE 1 & 2 pooled data	**N=905**	**N=234**	
Central read	521 (57.6)	72 (30.8)	26.8 (20.1–33.5); p<0.0001
Local read	538 (59.4)	75 (32.1)	27.4 (20.6–34.2); p<0.0001

Source: CS Figure 9 and Appendix L.1.4 Table 216

OCTAVE *Sustain*

Just over 60% of participants who had achieved a response to induction therapy and were re-randomised into the OCTAVE sustain study tofacitinib 10 mg twice daily group had a clinical response at week 52 (Table 30). In the tofacitinib 5 mg twice daily arm just over 50% of participants had clinical response at week 52. For both the tofacitinib 10 mg and 5 mg groups the percentage difference in comparison to the placebo group (in which approximately 20% had a clinical response) was statistically significant (central reads TOF 5 mg vs placebo a difference of 31.3 percentage points, 95% CI 22.4 to 40.2, p<0.001; TOF 10 mg vs placebo 41.7, 95% CI 32.9 to 50.5, p<0.001).

The local read data, which were very similar to the central read data, were used in an NMA (see Section 3.3.9.1 below) which contributed data to the economic model.

Table 30 Clinical response week 52 in maintenance study (FAS, NRI, central and local reads)

Parameter, n (%)	TOF 5 mg N=198	PBO N=198	% difference vs PBO (95% CI)	TOF 10 mg N=197	% difference vs PBO (95% CI)
Central read	102 (51.5)	40 (20.2)	31.3 (22.4–40.2); p<0.001	122 (61.9)	41.7 (32.9–50.5); p<0.001
Local read	101 (51.0)	41 (20.7)	30.3 (21.3–39.3); p<0.0001	121 (61.4)	40.7 (31.9–49.5); p<0.0001

Source: CS Figure 14 and Appendix L.1.4 Table 227

3.3.6 Summary of other clinical effectiveness endpoints

Other clinical effectiveness outcomes are not reported in detail in the CS and do not contribute data to the economic model. The majority of outcomes were statistically significantly in favour of tofacitinib, but they are based on Mayo scores and no adjustment has been made for multiple testing, therefore the statistical significance of these should be interpreted cautiously.

3.3.7 Summary of health related quality of life

3.3.7.1 EQ-5D

EQ-5D outcomes (utility index score and VAS score) were obtained at week 2 and week 8 of the OCTAVE induction trials and results are reported for each trial arm as adjusted mean change from baseline to week 2 and week 8 (Table 31). Missing values were assumed to be missing at random in the linear mixed-effects model used to analyse these data. The mean difference in the change from baseline between the tofacitinib and placebo arm favoured the tofacitinib arm for both the EQ-5D based outcomes and at both time points and was statistically significant except for the EQ-5D utility score difference from placebo at 8 weeks in OCTAVE 2. The CS notes in section B.2.6.1.2 that the benefits observed with tofacitinib exceeded the estimated M for patients with inflammatory bowel disease (utility index 0.076; VAS 10.9).

In the OCTAVE Sustain maintenance trial EQ-5D outcomes were obtained at weeks 4, 8, 16, 24, 32, 40 and 52 (Table 32). Missing values were assumed to be missing at random in the linear mixed-effect model used to analyse these data. In comparison to baseline values the EQ-5-D Utility Index values rose slightly over the 52-week analysis period in both the

tofacitinib 5 mg and 10 mg trial arms whereas in the placebo group values fell, indicating worsening quality of life. Differences between the tofacitinib arms and placebo favoured tofacitinib and statistically significant differences were obtained both for the 5 mg dose and the 10 mg dose from week 8 onwards. A similar pattern was observed for the EQ-5D VAS score and statistically significant differences between the tofacitinib arms and placebo were obtained at week 4.

Caution is advised in interpreting these results as the proportions of missing observations differed between arms, the reasons for the data being missing are not explained, and the appropriateness of the missing at random assumption is not discussed. EQ-5D data do not contribute to the company's economic base-case model but are included in a scenario analysis. Estimates of health state utilities obtained from OCTAVE EQ-5D data are discussed in this report in section 4.3.5.

Table 31 Change from baseline to week 8 in EQ-5D utility index and VAS scores. Summary for OCTAVE induction trials (FAS, without imputation)

Outcome	Time point	Adjusted mean ± SE		Difference vs PBO, mean (95% CI)
		TOF 10 mg	PBO	
OCTAVE 1		**N=476**	**N=122**	
EQ-5D utility index score	Week 2	0.13 ± 0.01 n=466	0.08 ± 0.02 n =122	0.04 ± 0.02 (0.00–0.08), p=0.0264
	Week 8	0.15 ± 0.01 n=452	0.08 ± 0.02 n=121	0.08 ± 0.02 (0.04–0.12), p<0.0001
EQ-5D VAS score	Week 2	13.11 ± 0.83 n=466	9.09 ± 1.52 n=122	4.02 ± 1.67 (0.75–7.29); p=0.0162
	Week 8	17.67 ± 0.84 n=451	9.49 ± 1.52 n=121	8.19 ± 1.67 (4.90–11.48); p<0.0001
OCTAVE 2		**N=429**	**N=112**	
EQ-5D utility index score	Week 2	0.12 ± 0.01 n=420	0.04 ± 0.02 n=109	0.08 ± 0.02 (0.04–0.12); p=0.0001
	Week 8	0.14 ± 0.01 n=414	0.11 ± 0.02 n=103	0.03 ± 0.02 (−0.02, 0.07); p=0.2201
EQ-5D VAS score	Week 2	13.32 ± 0.91 n=421	5.31 ± 1.67 n=110	8.01 ± 1.84 (4.39–11.62); p<0.0001
	Week 8	16.52 ± 0.91 n=414	8.29 ± 1.70 n=104	8.23 ± 1.87 (4.55–11.91); p<0.0001

Source: CS Appendix L, Table 218

Table 32 Change from baseline to week 8 in EQ-5D utility index and VAS scores. Summary for Octave Sustain (FAS, without imputation)

Time point	Change, adjusted mean ± SE			Change, adjusted mean ± SE	
	TOF 5 mg N=198	PBO N=198	Difference vs PBO (95% CI)	TOF 10 mg N=197	Difference vs PBO (95% CI)
EQ-5D Utility Index					
Wk 4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 16	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 40	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 52	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D VAS Score					
Wk 4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 16	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 40	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Wk 52	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: based on Table 46 in OCTAVE Sustain CSR

3.3.7.2 IBDQ

A statistically significant difference in the proportion of participants achieving IBDQ remission had emerged in favour of tofacitinib at the week 4 time point in both OCTAVE 1 and 2 (Table 33). The proportion of participants with IBDQ remission increased in all trial arms at week 8 with the difference between tofacitinib and placebo also increasing.

In the OCTAVE Sustain maintenance trial statistically significant differences in the proportions of participants with IBDQ remission and IBDQ response emerged by week 8 in favour of both tofacitinib 5 mg and 10 mg in comparison to placebo (Table 34 and CS Figure 15).

[redacted]

The ERG were uncertain about how much of the missing data were accounted for by IBDQ developers' rules and how much were treated as non-responders and consequently these data should be interpreted cautiously. IBDQ data are not included in the economic model.

Table 33 IBDQ results summary for OCTAVE induction trials (FAS, NRI)

Outcome, n (%)	Time point	TOF 10 mg	PBO	Difference vs PBO, mean (95% CI)
OCTAVE 1		**N=476**	**N=122**	
IBDQ remission (IBDQ score of ≥ 170)	Week 4	167 (35.1)	27 (22.1)	13.0 (4.4–21.5); p=0.008
	Week 8	206 (43.3)	32 (26.2)	17.0 (8.1–26.0); p<0.001
IBDQ treatment response (increase in IBDQ score of ≥ 16 points from induction trials baseline)	Week 4	299 (62.8)	55 (45.1)	17.7 (7.9–27.6); p<0.001
	Week 8	307 (64.5)	56 (45.9)	18.6 (8.8–28.4); p<0.001
OCTAVE 2		**N=429**	**N=112**	
IBDQ remission (IBDQ score of ≥ 170)	Week 4	124 (28.9)	9 (8.0)	20.9 (14.3–27.5); p<0.001
	Week 8	173 (40.3)	20 (17.9)	22.5 (14.0–30.9); p<0.001

Outcome, n (%)	Time point	TOF 10 mg	PBO	Difference vs PBO, mean (95% CI)
IBDQ treatment response (increase in IBDQ score of ≥ 16 points from induction trials baseline)	Week 4	266 (62.0)	44 (39.3)	22.7 (12.6–32.9); p<0.001
	Week 8	288 (67.1)	54 (48.2)	18.9 (8.7–29.2); p<0.001

Source: CS Appendix L, Table 217

Table 34 IBDQ results summary for OCTAVE Sustain trial (FAS, NRI)

Time point	TOF 5 mg N=198, n (%)	PBO N=198	% Difference vs PBO (95% CI)	TOF 10 mg N=197, n (%)	% Difference vs PBO (95% CI)
IBDQ Remission (IBDQ score of ≥ 170)					
Baseline	[redacted]	[redacted]		[redacted]	
Week 8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 16	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 40	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 52	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IBDQ Response (increase in IBDQ score of ≥ 16 points from induction trials baseline)					
Baseline	[redacted]	[redacted]		[redacted]	
Week 8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 16	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Week 40	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 52	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Table 42 in OCTAVE Sustain CSR

3.3.7.3 SF-36

In the OCTAVE 1 and 2 trials the PCS and MCS scores for SF-36 increased (i.e. improved) from baseline to week 8 in both the tofacitinib and placebo arms, but the improvement was statistically significantly greater in the tofacitinib arms (Table 35).

In OCTAVE Sustain the PCS and MCS outcomes were analysed as changes from baseline at week 24 and at week 52. At week 52 both the PCS and MCS scores had decreased (i.e. deteriorated) in the placebo and tofacitinib arms, with the decrease being largest for the placebo group, whilst the scores in the tofacitinib 10 mg arm had increased. The difference in change from baseline versus placebo was statistically significant for both the tofacitinib 5 mg and 10 mg arms at both time points (Table 36).

The company do not discuss the clinical significance of the SF-36 results and in the analysis of SF-36 PCS and MCS scores, missing data were not imputed. The ERG observes that the proportion of missing data in the OCTAVE Sustain trial was greater in the placebo arm than in either of the two tofacitinib arms (28% missing from tofacitinib 10 mg, 35% from tofacitinib 5 mg and 64% from placebo arms at 52 weeks). We assume that the patients who had not contributed data are most likely to be those who had failed treatment, although the CS does not state this or provide any further explanation for the missing data. The robustness of these SF-36 results is therefore unclear and they should be interpreted with caution.

Table 35 Change from baseline to week 8 in SF-36 component summary scores for OCTAVE induction trials (FAS, without imputation)

Outcome	Adjusted means ± SE		Difference vs PBO, mean (95% CI)
	TOF 10 mg	PBO	
OCTAVE 1	**N=476**	**N=122**	
PCS score change from baseline	6.8 ± 0.3 n=443	2.5 ± 0.6 n=116	4.2 ± 0.7 (2.9–5.5); p<0.0001

MCS score change from baseline	6.8 ± 0.5 n=443	3.5 ± 0.9 n=116	3.4 ± 1.0 (1.5–5.3); p =0.0005
OCTAVE 2	**N=429**	**N=112**	
PCS score change from baseline	6.8 ± 0.4 n=397	4.6 ± 0.7 n=98	2.2 ± 0.7 (0.7–3.6); p=0.0035
MCS score change from baseline	7.6 ± 0.5 n=397	4.4 ± 1.0 n=98	3.2 ± 1.1 (1.1–5.4); p=0.0037

Source: CS Appendix L, Table 219

Table 36 Change from baseline to week 52 in SF-36 component summary scores in OCTAVE Sustain (FAS, without imputation)

Outcome	Adjusted mean ± SE		% Difference vs PBO (95% CI)	Adjusted mean ± SE	% Difference vs PBO (95% CI)
	TOF 5 mg N=198	PBO N=198		TOF 10 mg N=197	
Change from baseline in PCS at wk 24	−0.3 ± 0.7 n=189	−5.0 ± 0.7 n=180	4.8 ± 0.8 (3.2–6.4); p<0.0001	0.4 ± 0.7 n=187	5.4 ± 0.8 (3.8–7.0); p<0.0001
Change from baseline in PCS at wk 52	−0.0 ± 0.8 n=129	−5.2 ± 0.9 n=71	5.1 ± 1.0 (3.1–7.2); p<0.0001	0.3 ± 0.7 n=141	5.5 ± 1.0 (3.4–7.5); p<0.0001
Change from baseline in MCS at wk 24	−1.1 ± 0.9 n=189	−7.3 ± 0.9 n=180	6.3 ± 1.0 (4.2–8.3); p<0.0001	−0.4 ± 0.9 n=187	6.9 ± 1.0 (4.8–9.0); p<0.0001
Change from baseline in MCS at wk 52	−1.0 ± 1.0 n=129	−6.7 ± 1.2 n=71	5.8 ± 1.3 (3.1–8.4); p<0.0001	0.1 ± 1.0 n=141	6.8 ± 1.3 (4.2–9.4); p<0.0001

Source: CS Appendix L, Table 228

3.3.7.4 WPAI-UC

For the analysis of the WPAI-UC missing data were not imputed. We assume that the high proportion of missing data for some elements of the WPAI-UC is likely to be because participants were not in employment, but the CS does not provide an explanation. For the 'non-work activity impairment' item, which is answered by all people whether or not in employment, the proportion of missing data is low.

After 8 weeks WPAI-UC scores in the OCTAVE Induction trials for all four elements had decreased (which indicates an improvement) but the effect was greater in the tofacitinib group (Table 37). Differences from placebo were in favour of tofacitinib and statistically significant for three of the four measures (presenteeism, work productivity loss and non-work activity impairment).

In the OCTAVE Sustain trial at week 52 WPAI-UC scores had increased (i.e. worsened) in the placebo group for all four elements but had decreased in the tofacitinib 5 mg and 10 mg trial arms. The difference versus placebo was statistically significant for presenteeism and non-work activity impairment (Table 38).

The company has not discussed the clinical significance of these observed changes in WPAI-UC scores, instead relying only on statistical significance for their interpretation.

The WPAI-UC scores are not included in the economic model.

Table 37 Summary of change from baseline in WPAI-UC scores in the OCTAVE induction trials (FAS, without imputation)

Outcome	Adjusted mean ± SE		Difference vs PBO, mean (95% CI)
	TOF 10 mg	PBO	
OCTAVE 1	**N=476**	**N=122**	
Absenteeism	−11.2 ± 1.3 n=270	−7.1 ± 2.7 n=55	−4.2 ± 2.9 (−9.9, 1.6); p=0.1565
Presenteeism	−22.1 ± 1.6 n=273	−9.2 ± 3.3 n=60	−12.9 ± 3.5 (−19.8, −6.0); p=0.0003
Work Productivity Loss	−19.1 ± 2.0 n=180	−8.5 ± 3.9 n=43	−10.6 ± 4.3 (−19.1, −2.1); p=0.0143
Non-Work Activity Impairment	−25.4 ± 1.3 n=442	−11.5 ± 2.3 n=119	−14.0 ± 2.6 (−19.0, −8.9); p<0.0001
OCTAVE 2	**N=429**	**N=112**	
Absenteeism	−7.3 ± 1.6 n=223	−9.3 ± 3.0 n=52	2.1 ± 3.3 (−4.4, 8.5); p=0.5295
Presenteeism	−18.6 ± 1.7 n=235	−13.7± 3.3 n=56	−4.9 ± 3.6 (−12.0, 2.2); p=0.1767
Work Productivity Loss	−14.7 ± 2.2 n=168	−11.2 ± 3.8 n=45	−3.5 ± 4.3 (−11.9, 4.9); p=0.4123

Non-Work Activity Impairment	−24.0 ± 1.3 n=398	−12.2 ± 2.5 n=98	−11.8 ± 2.7 (−17.2, −6.4); p<0.0001

Source: CS Appendix L, Table 212

Table 38 Summary of change from baseline to week 52 in WPAI-UC scores in the OCTAVE Sustain trial (FAS, without imputation).

Outcome	Adjusted mean ± SE		% Diff vs PBO (95% CI)	Adjusted mean ± SE	% Diff vs PBO (95% CI)
	TOF 5 mg N=198	PBO N=198		TOF 10 mg N=197	
Absenteeism	−4.5 ± 2.2 n=63	1.1± 2.8 n=33	−5.6 ± 3.4 (−12.2, 1.0); p=0.0953	−3.1 ± 2.2 n=68	−4.2 ± 3.3 (−10.7, 2.4); p=0.2131
Presenteeism	−3.6 ± 2.8 n=67	7.2 ± 3.5 n=34	−10.9 ± 4.1 (−18.9, −2.8); p=0.0081	−4.3 ± 2.9 n=70	−11.5 ± 4.1 (−19.5, −3.4); p=0.0052
Work Productivity Loss	−3.4 ± 4.9 n=22	1.0 ± 5.4 n=17	−4.4 ± 6.8 (−17.8, 9.0); p=0.5198	−6.6 ± 4.8 n=26	−7.6 ± 6.6 (−20.6, −5.4); p=0.2528
Non-Work Activity Impairment	−2.8 ± 2.2 n=112	11.3 ± 2.8 n=54	−14.1 ± 3.3 (−20.6, −7.5); p<0.0001	−3.1 ± 2.2 n=125	−14.4 ± 3.3 (−20.8, −7.9); p<0.0001

Source: CS Appendix L, Table 229

3.3.8 Sub-group analyses results

3.3.8.1 Prior TNFi exposure status

The CS focuses on subgroup results according to prior TNFi exposure (yes vs no) (CS sections B.2.7.3 to B.2.7.5). This is a more restricted subgroup than prior biologic therapy listed in the NICE scope; prior biologic therapy would include other biologics such as vedolizumab, in addition to the TNF inhibitors.

Detailed sub-group analyses by TNFi exposure status are presented here only for the outcomes that contribute data (from NMAs) to the economic model. Subgroup analyses by TNFi-exposure status for outcomes that do not contribute to data to the economic model (which include the primary outcome of the OCTAVE trials, remission) are presented in

Appendix 1. As stated in section 3.1.6 of this report the OCTAVE trials were not powered to test the statistical significance of subgroup analyses and although pooling the OCTAVE 1 and 2 trials maximises the available statistical power for the TNFi-exposure subgroups the results should nevertheless be interpreted cautiously.

Clinical remission

[redacted] (Table 39). The p-value from the Breslow-Day test did not indicate any significant heterogeneity in effect between the subgroups. When the data from OCTAVE 1 and 2 were pooled the differences between tofacitinib and placebo were [redacted] whilst the differences with locally-read outcomes were [redacted] for the TNFi-naive subgroup and [redacted] for TNFi-exposed subgroup.

At week 52 in OCTAVE Sustain the proportion of participants with clinical remission was higher in those who had received tofacitinib (either the 5 mg or 10 mg maintenance dose) than those who had received placebo in both the prior TNFi-exposed and TNFi-naïve subgroups, both for centrally-read and locally-read outcomes (Table 40). However, the tofacitinib versus placebo difference was greater in the TNFi-naive subgroup than the TNFi-exposed subgroup and this is particularly apparent for the tofacitinib 5 mg versus placebo comparison (e.g. for centrally-read outcomes the differences between tofacitinib 5 mg and placebo were [redacted] in the prior TNFi-naive subgroup and [redacted] in the prior TNFi-exposed subgroup). No test for heterogeneity of effects among the subgroups is reported for OCTAVE Sustain.

Table 39 Proportion of patients in clinical remission in OCTAVE Induction 1 and 2 according to prior TNFi treatment (FAS, NRI, central and local reads)

Subgroup: prior-TNFi treatment	TOF 10 mg n/N (%)	PBO n/N (%)	Difference (95% CI); p-value	p-value for hetero-geneity
OCTAVE 1, week 8				
TNFi-naïve Central read	[redacted]	[redacted]	[redacted]; p=[redacted]	[redacted]
TNFi-exposed Central read	[redacted]	[redacted]	[redacted]; p=[redacted]	
TNFi-naïve	[redacted]	[redacted]	[redacted]; p=[redacted]	[redacted]

Local read				
TNFi-exposed Local read	[redacted]	[redacted]	[redacted]; p=[redacted]	
OCTAVE 2, week 8				
TNFi-naïve Central read	[redacted]	[redacted]	[redacted]; p=[redacted]	[redacted]
TNFi-exposed Central read	[redacted]	[redacted]	[redacted]; p=[redacted]	
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]; p=[redacted]	[redacted]
TNFi-exposed Local read	[redacted]	[redacted]	[redacted]; p=[redacted]	
OCTAVE 1 & 2 pooled data, week 8				
TNFi-naïve Central read	[redacted]	[redacted]	[redacted]; p=[redacted]	Not reported
TNFi-exposed Central read	[redacted]	[redacted]	[redacted]; p[redacted]	
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]; p=[redacted]	Not reported
TNFi-exposed Local read	[redacted]	[redacted]	[redacted]; p=[redacted]	

Source: CS Appendix E Table 123

Table 40 Proportion of patients in clinical remission in OCTAVE Sustain at week 52 according to prior TNFi treatment (FAS, NRI, central and local reads)

Subgroup: prior-TNFi treatment	TOF 5 mg n/N (%)	PBO n/N (%)	Difference vs PBO (95% CI)	TOF 10 mg n/N (%)	Difference vs PBO (95% CI)
TNFi-naïve Central read	[redacted]	[redacted]	[redacted]; p[redacted]	[redacted]	[redacted] p[redacted]
TNFi-exposed Central read	[redacted]	[redacted]	[redacted]; p=[redacted]	[redacted]	[redacted] p[redacted]
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]; p[redacted]	[redacted]	[redacted] p[redacted]

TNFi-exposed Local read	[redacted]	[redacted]	[redacted]; p=[redacted]	[redacted]	[redacted] p[redacted]

Source: CS Appendix E Table 127

Clinical response

Results of the subgroup analyses of clinical response by prior TNFi treatment differed between OCTAVE 1 and OCTAVE 2 (Table 41). In OCTAVE 1 for both centrally-read and locally-read data the difference in clinical response at 8 weeks favouring tofacitinib was greater among TNFi-exposed participants than TNFi-naïve participants. The p-value from the Breslow-Day test suggests that there was significant heterogeneity in treatment effect between the subgroups. This was not the case for OCTAVE 2, in which clinical response results for the TNFi-exposed and TNFi-naïve subgroups were more similar (heterogeneity test not significant), and the difference favouring tofacitinib over placebo was slightly larger in the treatment-naïve subgroup for both centrally- and locally-read data. When the data from OCTAVE 1 and 2 were pooled the central read differences between tofacitinib and placebo were [redacted] in the TNFi-naïve subgroup and [redacted] in the TNFi-exposed subgroup.

At week 52 in OCTAVE Sustain the proportion of participants with a clinical response was higher among those who had received tofacitinib (either the 5 mg or 10 mg maintenance dose) than those who had received placebo in both the TNFi-exposed and TNFi-naïve subgroups, both for centrally-read and locally-read outcomes (Table 42). The proportions of participants with a clinical response were consistently higher in all three trial arms in the prior TNFi-naïve subgroup than in the TNFI-experienced subgroup; however, the relative treatment effect (difference versus placebo) was almost identical in the TNFi-naïve and TNFi-exposed subgroups for both the 5 mg versus placebo and the 10 mg versus placebo comparisons, for both the centrally-read and locally-read data.

Table 41 Proportion of patients with a clinical response in OCTAVE Induction 1 and 2 according to prior TNFi treatment (FAS, NRI, central and local reads)

Subgroup: prior-TNFi treatment	TOF 10 mg n/N (%)	PBO n/N (%)	Difference (95% CI); p-value	p-value for hetero-geneity
OCTAVE 1, week 8				
TNFi-naïve Central read	[redacted]	[redacted]	[redacted]; p=[redacted]	[redacted]

TNFi-exposed Central read	[redacted]	[redacted]	[redacted] p[redacted]	
TNFi-naïve Local read	[redacted]	[redacted]	[redacted] p=[redacted]	[redacted]
TNFi-exposed Local read	[redacted]	[redacted]	[redacted] p[redacted]	
OCTAVE 2, week 8				
TNFi-naïve Central read	[redacted]	[redacted]	[redacted] p=[redacted]	[redacted]
TNFi-exposed Central read	[redacted]	[redacted]	[redacted] p=[redacted]	
TNFi-naïve Local read	[redacted]	[redacted]	[redacted] p=[redacted]	[redacted]
TNFi-exposed Local read	[redacted]	[redacted]	[redacted] p=[redacted]	
OCTAVE 1 & 2 pooled data, week 8				
TNFi-naïve Central read	[redacted]	[redacted]	[redacted] p[redacted]	Not reported
TNFi-exposed Central read	[redacted]	[redacted]	[redacted] p[redacted]	
TNFi-naïve Local read	[redacted]	[redacted]	[redacted] p[redacted]	Not reported
TNFi-exposed Local read	[redacted]	[redacted]	[redacted] p[redacted]	

Source: CS Appendix E Table 124

Table 42 Proportion of patients with a clinical response in OCTAVE Sustain according to prior TNFi treatment (FAS, NRI, central and local reads)

Subgroup: prior TNFi treatment	TOF 5 mg n/N (%)	PBO n/N (%)	Difference vs PBO (95% CI)	TOF 10 mg n/N (%)	Difference vs PBO (95% CI)
TNFi-naïve Central read	[redacted]	[redacted]	[redacted] p[redacted]	[redacted]	[redacted] p[redacted]
TNFi-exposed Central read	[redacted]	[redacted]	[redacted] p[redacted]	[redacted]	[redacted] p[redacted]

TNFi-naïve Local read	[redacted]	[redacted]	[redacted] p[redacted]	[redacted]	[redacted] p[redacted]
TNFi-exposed Local read	[redacted]	[redacted]	[redacted] p[redacted]	[redacted]	[redacted] p[redacted]

Source: CS Appendix E Table 128

In addition to the subgroup analyses by TNFi-exposure status reported above for clinical remission and clinical response, subgroup analyses by TNFi-exposure status were also reported for remission (the primary outcome of both OCTAVE 1, 2 and Sustain) and for sustained corticosteroid-free remission among patients who were in remission at baseline (OCTAVE Sustain) (CS sections B.2.7.4 and B.2.7.5). Neither of these outcomes contribute to data to the economic model. These subgroup analyses are summarised in Appendix 1.

3.3.8.2 Other subgroups

Pre-planned subgroup analyses for the outcomes of remission, mucosal healing, clinical response and clinical remission are reported for OCTAVE 1 and OCTAVE 2 at 8 weeks and for OCTAVE Sustain at 52 weeks according to prior TNFi failure (yes vs no) and corticosteroid use at baseline (yes vs no) (CS Appendix E Tables 130 to 141). As noted above, the OCTAVE trials were not specifically powered statistically for subgroup analyses and adjustments to account for multiple testing were not performed, so these results should be interpreted with caution.

Prior TNFi failure

For all four of these outcomes tofacitinib was [redacted] in both the TNFi failure and no-failure subgroups, for both the centrally-read and locally-read outcomes. Overall, the data suggest that for these four outcomes the treatment effect was [redacted] in the no-failure subgroup than in the prior TNFi-failure subgroup, or there was [redacted] between the subgroups. However, this apparent [redacted] between the subgroups was generally [redacted] (OCTAVE 1 and 2) or the significance was not reported (OCTAVE Sustain).

Corticosteroid use at baseline

For all four outcomes tofacitinib was [redacted] in both the subgroup who had corticosteroid use at baseline and those without corticosteroids, for both the centrally-read and locally-read outcomes. The treatment effects for remission, clinical response and clinical

remission in OCTAVE 1 and OCTAVE 2 generally [redacted] between the two subgroups. The treatment effect for mucosal healing in OCTAVE 1 and OCTAVE 2 tended to be [redacted] in the no-corticosteroid subgroup, but only for centrally-read data (subgroup heterogeneity was [redacted] in OCTAVE 2 only for the centrally-read data). In OCTAVE Sustain the treatment effect for all four outcomes was [redacted] in the no-corticosteroid subgroup, although a test of subgroup statistical heterogeneity is not reported.

Other subgroup analyses in OCTAVE Sustain

For OCTAVE Sustain further pre-planned subgroup analyses are reported for the outcomes of remission, mucosal healing, clinical response, clinical remission, sustained corticosteroid-free remission at weeks 24 and 52 and sustained clinical response at weeks 24 and 52 (CS Appendix E, Tables 142 to 153). The majority of the subgroup analyses were conducted according to the following factors: treatment assignment during the induction study, in remission at maintenance study baseline (yes vs no), mucosal healing at maintenance study baseline (yes vs no), prior corticosteroid failure (yes vs no) and disease duration (<6 years vs ≥6 years). The ERG notes that in Appendix E Tables 149, 150, and 151 an additional subgroup of 'Gender' is listed for some comparisons in place of disease duration (e.g. Table 149 local read data for tofacitinib 10 mg and placebo) but the ERG believes this may be an error and that these data are likely to be disease duration data. Overall, across the different subgroups investigated, a higher proportion of participants in the tofacitinib groups (5 mg and 10 mg) consistently achieved the desired outcome than in the placebo group.

Results from one further potentially relevant subgroup, geographic region, are not reported in the CS.

3.3.9 Network meta-analysis results

In this section we present a summary of the base-case NMA results, with clinical remission and clinical response presented together because these were modelled jointly using the multinomial probit model described earlier in section 3.1.7. Results are presented on the probit scale (for clinical response and clinical remission) or the logit scale (for mucosal healing), as odds ratios and absolute probabilities. On the probit scale a negative coefficient indicates treatment is more effective than placebo whereas an odds ratio greater than one indicates that the comparator treatment had a greater treatment effect than placebo (for columns headed 'Comparator vs PBO) or that tofacitinib had a greater treatment effect than the comparator (for columns headed 'TOF vs comparator). The 95% credible interval indicates the lower and upper extremes in which the odds ratio is expected to lie with a

probability of 95%). The surface under cumulative ranking curve (SUCRA) value is used to rank treatments based on their probability of ranking first through to last among the treatment options. If the SUCRA probability is 0% the treatment always ranks last and if it is 100% the treatment always ranks first.

3.3.9.1 Summary of NMA results for clinical response and clinical remission

In the induction phase for the TNFi-naïve population analysis all treatments were included. Infliximab had the largest treatment effect on the secondary outcomes of clinical remission and clinical response compared to placebo, whilst adalimumab had the smallest effect. All treatments showed strong evidence of benefit over placebo. In the TNFi-exposed population, tofacitinib, adalimumab, and vedolizumab were included. Tofacitinib had the greatest treatment effect on clinical remission and clinical response compared to placebo. Both tofacitinib 10 mg and vedolizumab 300 mg showed strong evidence of benefit over placebo (Table 43).

Table 43 Induction Phase base-case NMA results – comparative effects and probabilities of achieving clinical response and clinical remission

Comparator	Comparator vs PBO			TOF vs comparator		Absolute probability		SUCRA [a]
	Treatment effect, median (95% CrI)	Odds ratio, median (95%CrI)		Odds ratio, median (95%CrI)				
	Probit scale	Clinical response	Clinical remission	Clinical response	Clinical remission	Clinical response	Clinical remission	
TNFi-naïve subgroup								
PBO				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed subgroup								
PBO				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 25

[a] based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2.

The maintenance phase NMA results for clinical response and clinical remission showed a similar pattern to those for the induction phase analyses (Table 44). In the TNFi-naive population, tofacitinib 10 mg had the largest treatment effect on clinical response and clinical remission compared to placebo, with all treatments showing strong evidence of benefit over placebo.

In the TNFi-exposed population, tofacitinib 10 mg also had the largest treatment effect on clinical response and clinical remission compared to placebo

The NMA results for clinical remission in both the induction and maintenance phases of treatment are included in the economic model, with the exception that adalimumab is not presented as a comparator for the TNFi-exposed subgroup.

Table 44 Maintenance phase base-case NMA results – comparative effects and probabilities of achieving clinical remission

Comparator	Comparator vs PBO			TOF 5 mg vs comparator		Absolute probability		SUCRA [a]
	Treatment effect, median (95% CrI)	Odds ratio, median (95%CrI)		Odds ratio, median (95%CrI)				
	Probit scale	Clinical response	Clinical remission	Clinical response	Clinical remission	Clinical response	Clinical remission	
TNFi-naïve subgroup								
PBO								
TOF 5 mg								
TOF 10 mg								
INF 5 mg/kg								
ADA 40 mg Q2W								
GOL 50 mg								
GOL 100 mg								
VED 300 mg Q8W								
VED 300 mg Q4W								

Comparator	Comparator vs PBO			TOF 5 mg vs comparator		Absolute probability		SUCRA [a]
	Treatment effect, median (95% CrI)	Odds ratio, median (95%CrI)		Odds ratio, median (95%CrI)				
	Probit scale	Clinical response	Clinical remission	Clinical response	Clinical remission	Clinical response	Clinical remission	
TNFi-exposed subgroup								
PBO				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF 5 mg	[redacted]	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 40 mg Q2W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg Q8W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg Q4W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 26

[a] based on treatment effect on probit scale

3.3.9.2 Summary of NMA results for Mucosal healing

Mucosal healing was a key secondary outcome for the OCTAVE Induction 1 and 2 trials and the OCTAVE Sustain trial. This outcome is not included in the economic model.

In the induction phase for the TNFi-naïve subgroup all treatments showed strong evidence of benefit over placebo at achieving mucosal healing. Infliximab had the largest treatment effect on mucosal healing compared to placebo and adalimumab had the lowest. In the TNFi-exposed subgroup, tofacitinib 10 mg had the largest treatment effect compared to placebo (Table 45).

Table 45 Induction phase base-case NMA results – comparative effects and probabilities of achieving mucosal healing

Comparator	Comparator vs PBO		TOF vs Comparator	Absolute probability, median (95% CrI)	SUCRA
	Treatment effect, median (95% CrI) Logit scale	Odds ratio, median (95% CrI)	Odds ratio, median (95% CrI)		
TNFi-naïve subgroup					
PBO			[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]		[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed subgroup					
PBO			[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]		[redacted]	[redacted]
ADA 160/80/40 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Appendix D.1.3.5.2.2. Table 109 (doses have been added by the ERG based on CS Table 25)

In the maintenance phase for the TNFi-naïve subgroup tofacitinib 10 mg had the greatest effect on mucosal healing in comparison to placebo. Infliximab and adalimumab had the smallest effects on mucosal healing in comparison to placebo. The remaining treatments, golimumab and vedolizumab, showed strong evidence of benefit over placebo in mucosal healing. In the TNFi-exposed subgroup vedolizumab had the greatest effect on mucosal

healing, with tofacitinib (5 mg and 10 mg) also providing greater benefit than placebo (Table 46).

Table 46 Maintenance phase base-case NMA results – comparative effects and probabilities of achieving mucosal healing

Comparator	Comparator vs PBO		TOF vs comparator	Absolute probability, median (95% CrI)	SUCRA
	Treatment effect, median (95% CrI) Logit scale	Odds ratio, median (95% CrI)	Odds ratio, median (95% CrI)		
TNFi-naïve subgroup					
PBO			[redacted]	[redacted]	[redacted]
TOF 5 mg	[redacted]	[redacted]		[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
INF	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 50 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 100 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED Q8W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED Q4W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed subgroup					
PBO			[redacted]	[redacted]	[redacted]
TOF 5 mg	[redacted]	[redacted]		[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED Q8W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED Q4W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Appendix D.1.3.5.2.2. Table 110

3.3.9.3 NMA sensitivity analyses

The company conducted sensitivity analyses to test the impact of the following factors on NMA outcomes:

- Studies in which the majority of participants were Asian were excluded. These studies were Suzuki 2014, Mshimesh 2017, Jiang 2015, Kobayashi 2015 and Pursuit J. The CS does not provide an explanation for excluding Asian studies, but states

that this "*sensitivity analysis is aligned with the base-case assumptions made in the NMA supporting TA329*".

- Centrally read endoscopic subscores (instead of locally-read subscores) were analysed for the clinical response, clinical remission and mucosal healing outcomes.
- TNFi-failure subgroup: this sensitivity analysis limited the data from the OCTAVE trials and the ULTRA 2 trial to patients who had prior TNFi failure (i.e. a subset of the base case data which included all patients with prior TNFi-exposure)
- Overall ITT analysis: data were not divided into two subgroups by TNFi-exposure status but instead an overall analysis was conducted regardless of prior TNFi-exposure status.

Condensed versions of results tables for these sensitivity analyses are presented in Appendix 2 of this report and are available in full in CS Appendix D.1.3.5.

On the whole the NMA results were relatively robust to the changes made in the sensitivity analyses described above.

[redacted]

The ERG notes that for the sensitivity analyses using data from the TNFi-failure population, [redacted] (CS Table 28).

3.3.10 Summary of adverse events

3.3.10.1 Adverse events in the OCTAVE research programme

The CS presents safety data in patients with moderate to severe ulcerative colitis from the Phase II trial, the two OCTAVE Induction trials, the OCTAVE Sustain trial and the OCTAVE Open extension study. In total, tofacitinib has been evaluated in 1157 patients with ulcerative

colitis, equivalent to 1986 patient-years of tofacitinib exposure with a maximum of 4.4 years of treatment (CS section B.2.10).

The CS classifies adverse events as: common adverse events; serious adverse events; adverse events leading to discontinuation; and adverse events of special interest (CS section B.2.10 and CS Table 29). In CS Appendix F, adverse events are also classified as being treatment-emergent, although the data presented for the overall frequencies of serious adverse events and treatment-emergent serious adverse events in the OCTAVE Induction and OCTAVE Sustain trials are identical (see Table 47). The CS lists adverse events of special interest as being infections (in general), herpes zoster infections, malignancies, gastrointestinal perforations and cardiovascular events, but does not give an explicit rationale. The company presents data on a wide range of adverse events (CS Appendix F), but the only adverse events that inform the economic analysis are serious infections (discussed further below). The CS presents less detailed information on adverse events for the Phase II trial and the OCTAVE Open extension study than for the OCTAVE Induction and Sustain trials. Where data are available, we have summarised the frequency of the main classes of adverse events for the Phase II and Phase III trials in Table 47 and for the OCTAVE Open study in Table 48.

[redacted]

Overall incidence of adverse events

The proportion of patients with of adverse events of any type ranged from 42% to 80% across the Phase II and Phase III trials, being highest in the OCTAVE Sustain trial 10 mg tofacitinib arm. Rates of any adverse event were broadly similar for the tofacitinib and placebo arms within each trial (Table 47 and Table 48). The most frequent specific adverse events were worsening ulcerative colitis, nasopharyngitis, arthralgia, and headache.

Serious adverse events

Serious adverse events affected fewer than 10% of patients in the tofacitinib trials, ranging from 3% in the 10 mg tofacitinib arm of OCTAVE 1to 8% in the placebo arm of OCTAVE 2 (Table 47), and [redacted] of the OCTAVE Open study (Table 48). The most frequent serious adverse event was ulcerative colitis and most serious adverse events were related to ulcerative colitis (CS section B.2.10.3).

Infections

The frequency of any infections ranged from 15% to 40% across the Phase II and Phase III trials, and was highest (24% to 40%) in the OCTAVE Sustain trial (Table 47). In addition to

nasopharyngitis, a range of other types of infection occurred but most of these each affected ≤2% of patients (CS Tables 156 to 158). The only type of infection (besides nasopharyngitis) that occurred in ≥5% of patients was Herpes zoster, which affected 5.1% of patients in the tofacitinib 10 mg arm of the Sustain trial (<5% in all other trial arms).

Most infections were mild or moderate in severity (CS section B.2.10.5). Serious infections were uncommon, affecting a maximum of only 2 patients in any trial arm (≤2%). Serious infections occurred only in the tofacitinib arm within each trial, with the exception of OCTAVE Sustain where 2 patients in the placebo arm had serious infections. The CS lists the specific serious infections that occurred in the OCTAVE Induction and Sustain trials but does not specify those which occurred in the Phase II trial (n=2) or the OCTAVE Open study (n not reported). The patients who had serious infections in OCTAVE 1 (n=6), OCTAVE 2 (n=1) and OCTAVE Sustain (n=5) are notable in that they each had a different type of infection, i.e. no individual type of serious infection occurred in more than one patient (CS Table 162).

CS Table 168 summarises the incidence of serious adverse events that have occurred in tofacitinib-treated patients across the company's clinical research programme on ulcerative colitis. The data show that [redacted].

Discontinuation due to adverse events

The frequency of adverse events leading to discontinuation ranged from 2% to 8% in the Phase II trial and OCTAVE Induction trials, but was higher in the OCTAVE Sustain trial (9% to 19%) (Table 47). The most common reason for discontinuation was worsening ulcerative colitis (CS section B.2.10.4).

Adverse events of special interest

The CS is slightly inconsistent in the reporting of adverse events of special interest, since infections and Herpes zoster are not listed under adverse events of special interest in CS Table 29, although they are reported elsewhere in the table. Where reported, adverse events of special interest affected a maximum of 3 patients in any trial arm.

Abnormal laboratory test results

The CS tabulates, but does not comment on, selected abnormal laboratory test results relating to cholesterol and triglyceride metabolism, and also reports the frequency of abnormal creatine kinase results. Monitoring cholesterol and other lipid parameters is recommended in the SmPC due to known short-term effects of tofacitinib on these.[10] The CS does not report whether any other laboratory test results (e.g. relating to liver or renal function) were abnormal, although the SmPC lists abnormal liver function tests as being a possible uncommon adverse event.[10] Overall, 5% to 27% of patients in the Phase II and Phase III trials had elevated total cholesterol (>1.3 x the upper limit of normal [ULN]) and 8% to 31% of patients had elevated low-density lipoprotein (>1.2 x ULN), with the rates being consistently higher in the tofactinib than placebo arms (Table 47). A smaller proportion of patients had abnormalities in high-density lipoprotein (1% to 9%) and triglycerides (0% to 8%) without a consistent within-trial difference between arms. The proportion of patients with elevated creatine kinase ranged from 2% to 28% and was higher in the tofacitinib than placebo arms in OCTAVE 1 and OCTAVE Sustain, but not in OCTAVE 2 (not reported for the Phase II trial).

Malignancies

The CS reports the frequency of malignancies across the company's tofacitinib ulcerative colitis research programme (total 1157 patients), divided into non-melanoma skin cancer and all other malignancies (CS Table 30). In total, 15 patients (1.3%) had non-melanoma skin cancer and 13 patients (1.2%) had a malignancy other than non-melanoma skin cancer. The company comments that a potential elevated risk of non-melanoma skin cancer was identified during the ulcerative colitis clinical trial programme compared to the company's rheumatoid arthritis trial programme, which likely reflects an increased malignancy risk in patients who have inflammatory bowel disease. However, the CS also states that the draft SmPC includes effective routine risk minimisation measures (CS section B.2.13.1). We note that the licensed indication for tofacitinib in rheumatoid arthritis is different to that in ulcerative colitis,[10] [e.g, tofacitinib is often administered with methotrexate, and at a different daily dose], so comparisons between the ulcerative colitis and rheumatoid arthritis trials programmes should be made with caution.

Mortality

The CS reports that there were 5 deaths across the OCTAVE programme (CS section B.2.10.6). These were: 1 death in the tofacitinib arm of OCTAVE 1, caused by dissecting aortic aneurysm, assessed as unrelated to the study drug; and 4 deaths in OCTAVE Open, all in the 10 mg tofacitinib group. Three of the deaths in the OCTAVE Open study occurred

>28 days after the last dose of tofacitinib and were due to malignancies. The remaining patient had died of hepatic angiosarcoma, in which tofacitinib was considered to have played a contributory role.

Table 47 Summary of adverse events in the tofacitinib Phase II and Phase III trials

Adverse event (AE)	Phase II trial		OCTAVE Induction 1		OCTAVE Induction	
	TOF 10 mg (N=33)	PBO (N=48)	TOF 10 mg (N=476)	PBO (N=122)	TOF 10 mg (N=429)	PBO (N=112
Any AE, n (%)	14 (42)	23 (48)	269 (57)	73 (60)	232 (54)	59 (
Serious AE, n (%)	2 (6)	4 (8)	16 (3) [a]	5 (4) [a]	18 (4) [a]	9 (
Most frequent AE, n (%) [b]						
Worsening ulcerative colitis	2 (6)	9 (19)	11 (2)	5 (4)	13 (3)	6
Nasopharyngitis	1 (3)	1 (2)	34 (7)	9 (7)	21 (5)	4
Arthralgia	2 (6)	0	14 (3)	6 (5)	11 (3)	6
Headache	3 (9)	2 (4)	37 (8)	8 (7)	33 (8)	9
Infections, n (%)						
Any infection [c]	9 (27)	7 (15)	111 (23)	19 (16)	78 (18)	17 (
Serious infection	2 (6)	0	6 (1)	0	1 (0.2)	
Herpes zoster	1 (3)	0	3 (1)	1 (1)	2 (1)	
AE leading to discontinuation, n (%) [d]	1 (3)	4 (8)	18 (4)	2 (2)	17 (4)	8
AE of special interest, n						
Intestinal perforation	Not reported		1	0	0	
Cancer other than non-melanoma skin cancer	Not reported		0	0	0	
Non-melanoma skin cancer	Not reported		1	0	1	
Cardiovascular events	Not reported		2	0	2	
Abnormal laboratory test results, n (%) [f]						
N for laboratory data	33	48	471	122	424	
Total cholesterol >1.3× ULN	8 (24)	5 (10)	80 (17)	11 (9)	73 (17)	
Low-density lipoprotein >1.2× ULN	9 (27)	4 (8)	91 (19)	11 (9)	92 (22)	12
High-density lipoprotein <0.8× LLN	3 (9)	2 (4)	6 (1)	2 (2)	7 (2)	
Triglycerides >1.3× ULN	2 (6)	0	15 (3)	1 (1)	12 (3)	
Creatine kinase >2× ULN	Not reported		45 (10) (n=474)	2 (2)	40 (9) (n=425)	10
Addition or increase in dose of lipid lowering agent, n (%)	Not reported		4 (1)	0	2 (1)	1

Source: CS Tables 29 ad 166 and the Phase II trial publication.

[a] The CS reports that these data are serious AE (CS Table 29; not marked as confidential) and also reports that these same data are treatment-emergent serious AE (CS Tables 159-161; marked as academic in confidence).

[b] The four most frequent adverse events in OCTAVE Sustain.

[c] Reported as "adverse effects from infection" (CS Table 166).

[d] Including patients who discontinued treatment because of worsening ulcerative colitis.

[e] Invasive ductal breast carcinoma.

[f] Laboratory data were missing for some patients.

Table 48 Summary of treatment-emergent adverse events in OCTAVE Open

Adverse event (AE)	Tofacitinib 5 mg	Tofacitnib 10 mg	Total
Number of AE			
Patients with AE, n (%)			
Patients with serious AE, n (%)			
Patients with severe AE, n (%)			
Patients discontinued due to AE, n (%)			
Patients with dose reduced or temporary discontinuation due to AE, n (%)			

Source: CS Table 167

Except for the number of AE, subjects were counted only once per treatment in each row.

3.3.10.2 Adverse events NMA

Three safety outcomes were analysed by NMA: discontinuations due to adverse events; serious adverse events; and serious infections. Only data from the serious infections NMA contribute to the economic model. As stated in section 3.1.7, safety outcomes were analysed only for the induction phase of the included studies and with data for TNFi-exposed and TNFi-naïve subgroups combined in one analysis.

Discontinuation due to adverse events does not contribute to the economic model.

(Table 49).

Table 49 Induction phase base-case NMA results – comparative effects and probabilities of discontinuing due to AEs

Comparator	Odds ratio, median (95%CrI)		Absolute probability median (95% CrI)	SUCRA[a]
	Comparator vs PBO	TOF vs comparator		
PBO		[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]		[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [c]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 32

[a] Based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2.

Serious adverse events do not contribute data to the economic model.

[redacted]

Table 50 Induction phase base-case NMA results – comparative effects and probabilities of serious AEs

Comparator	Odds ratio, median (95%CrI)		Absolute probability median (95% CrI)	SUCRA[a]
	Comparator vs PBO	TOF vs comparator		
PBO		[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]		[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [c]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 33

[a] Based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2.

Serious infections are included in the economic model.

[redacted] (Table 51).

Table 51 Induction phase base-case NMA results – comparative effects and probabilities of serious infections

Comparator	Odds ratio, median (95%CrI)		Absolute probability median (95% CrI)	SUCRA[a]
	Comparator vs PBO	TOF vs comparator		
PBO		[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]		[redacted]	[redacted]
INF 10 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]
ADA 160/80/40 mg [b]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 200/100 mg [c]	[redacted]	[redacted]	[redacted]	[redacted]
VED 300 mg [d]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 34
[a] Based on treatment effect on probit scale. [b] 160 mg at week 0, 80 mg at week 2, 40 mg at weeks 4 and 6. [c] 200 mg at week 0, 100 mg at week 2. [d] At weeks 0 and 2.

3.4 Overall summary of clinical effectiveness evidence

The systematic review of clinical effectiveness evidence in the CS identified four RCTs of tofacitinib as a treatment for people with moderately to severely active ulcerative colitis. One is a Phase II dose finding study (in which only one arm received the licensed 10 mg BID dose), two were identical Phase III RCTs of tofacitinib as an induction therapy (OCTAVE Induction 1 and OCTAVE Induction 2), and the fourth was an RCT of tofacitinib as a maintenance therapy (OCTAVE Sustain). In all four trials the comparator was placebo. The OCTAVE clinical trial programme followed a re-randomisation design in which participants who had been randomised into one of the two induction phase trials (OCTAVE 1 or OCTAVE 2) and who had achieved a response to 8-weeks of induction therapy were eligible to be re-randomised into the 52-week OCTAVE Sustain trial.

In addition to the four tofacitinib RCTs there is also an ongoing open label extension study (OCTAVE Open) which all participants in the OCTAVE trial programme are eligible to enter.

The focus in the CS is on the two Phase III induction therapy RCTs (OCTAVE 1 and OCTAVE 2) and the maintenance therapy RCT (OCTAVE Sustain) which provide the bulk of the evidence presented in the CS.

The ERG judged the two OCTAVE Induction trials to be at a low risk of the five types of bias assessed. The OCTAVE Sustain trial and the Phase II trial could be at risk of attrition bias as a result of unbalanced dropouts between the tofacitinib and placebo arms but appear to be at a low risk of bias for the other four types of bias assessed. Overall the studies appear to have been well conducted. The main clinical effectiveness outcomes reported in the CS are remission (primary outcome), mucosal healing, sustained corticosteroid-free remission (OCTAVE Sustain only), clinical remission, and clinical response. Health related quality of life outcomes (both generic and disease specific) and adverse events were also reported.

The company's systematic review had broad inclusion criteria to enable the identification of evidence not only for tofacitinib but also for relevant comparators. It identified 21 RCTs in total, the four tofacitinib RCTs plus 17 RCTs, most of which compared an active treatment to placebo. In the absence of direct head-to-head comparisons between active treatments, these studies could potentially be used in NMA.

Key clinical effectiveness outcomes within the OCTAVE trials were based on components of the Mayo Score. One of the four components to the Mayo score, 'Endoscopic findings', was assessed both locally (by the study site investigator) and centrally (from a video recording). Results including this component of the Mayo score were reported separately using the local or the central read of the endoscopic data.

Remission, as opposed to clinical remission, was the primary outcome of the OCTAVE induction and maintenance trials but this outcome did not contribute to economic modelling. At week 8 the mean differences between tofacitib and placebo in OCTAVE 1 and 2 were in favour of tofacitinib and statistically significant regardless of whether central read or local read data were used. Results using local read data were less conservative regarding the effectiveness of tofacitinib than those using central read data. A similar effect of tofacitinib was observed in the OCTAVE Sustain trial for both the 5 mg and 10 mg tofacitinib

maintenance doses, with the percentage difference in comparison to placebo being greater for the 10 mg tofacitinib dose at week 52.

Mucosal healing was designated a key secondary outcome for the OCTAVE trials. A greater proportion of participants in the tofacitinib group achieved mucosal healing at week 8 in comparison to the placebo group in both OCTAVE 1 and OCTAVE 2 and the differences versus placebo were statistically significant for both central and local read data. In the OCTAVE Sustain maintenance trial, statistically significant differences in both mucosal healing at week 52 and sustained mucosal healing at weeks 24 and 52 were reported for the 5 mg and 10 mg tofacitinib maintenance doses in comparison to the placebo arm of the trial.

Sustained corticosteroid-free remission among those in remission at baseline (a further key secondary outcome) in the OCTAVE Sustain trial, was statistically significantly greater in the tofacitinib 5 mg and 10 mg arms than in the placebo arm.

Clinical remission, which has a very similar definition to the primary outcome of remission, contributed data to the economic model via the NMA. Due to the similarity of outcome definition the results from the OCTAVE trials were almost identical to those reported above for remission, favouring tofacitinib.

The outcome of clinical response also contributes data to the economic analysis via NMA. The percentage difference between the tofacitinib group and the placebo group in favour of tofacitinib was statistically significant in both OCTAVE induction trials and the OCTAVE Sustain maintenance trial and for both the central and locally read data.

HRQoL was reported using both generic (EQ-5D and SF-36) and disease specific (IBDQ and WPAI-UC) instruments. Results showed HRQoL was typically improved by tofacitinib treatment; however, for some HRQoL measures we are uncertain about the impact of the missing data. Data from the EQ-5D-3L do not inform the base-case economic model but were included in a scenario analysis.

Subgroup analyses focused on results according to prior TNFi-exposure. Note that this is a more restricted subgroup than that of prior biologic therapy (which would also include other biological therapies such as vedolizumab) which is listed in the NICE scope. The OCTAVE trials were not powered to test the statistical significance of subgroup analyses so the results should be interpreted cautiously. Overall, the results were consistent regardless of prior TNFi-exposure status.

Safety data for tofacitinib in patients with moderate to severely active ulcerative colitis comes from the Phase II tofacitinib trial, the two OCTAVE Induction trials, the OCTAVE Sustain trial and the ongoing OCTAVE Open extension study. In total tofacitinib has been evaluated in 1157 patients with ulcerative colitis with a maximum exposure to tofacitinib of 4.4 years.

Rates of adverse events of any type were broadly similar for the tofacitinib and placebo arms within each trial. Serious adverse events affected fewer than 10% of patients in the tofacitinib trials. The most frequent serious adverse event was ulcerative colitis, and most serious adverse events were related to ulcerative colitis.

Serious infections were uncommon, affecting a maximum of only 2 patients (≤2%) in any trial arm and occurred only in the tofacitinib arm within each trial, with the exception of OCTAVE Sustain where two patients in the placebo arm had serious infections. As previously noted, data on serious infections was included in the economic model.

There were five deaths across the OCTAVE trials programme and tofacitinib was considered to have played a role in one of these (the death of a patient with hepatic angiosarcoma).

Overall, and in comparison with evidence from the use of tofacitinib in patients with rheumatoid arthritis, no new safety signals were identified.

NMA was used to compare tofacitinib to other potential treatment options where there was available evidence. Analyses were conducted for the outcomes of clinical response, clinical remission, mucosal healing and safety (discontinuations due to AEs, serious AEs, serious infections). Of these outcomes, clinical response, clinical remission and serious infections contributed data to the economic model and results for these outcomes are summarised below.

Heterogeneity was present among the studies available to include in NMA. There were differences in study design (re-randomised design as for the OCTAVE trial programme versus treat-through design in which participants entering the maintenance phase of a study remain in the arm they were allocated to during the induction phase of the study) and some patient characteristics (TNFi-exposure status, disease duration, studies in predominantly Asian patients).

Separate analyses by NMA were undertaken for the induction and maintenance phases of treatment. In addition, to reduce heterogeneity, NMAs were conducted separately for the TNFi-naïve and TNFi-exposed populations (Table 11). An adjustment was also made to the outcomes from studies with a treat-through design to try to better align these outcomes with those from studies with a re-randomised design.

Because clinical response and clinical remission are correlated outcomes (both are based on Mayo score data) the company used a multinomial probit model which maintained the correlation between these outcomes. A binomial logit model was used for the safety outcomes. Fixed and random effects models were conducted and sensitivity analyses were also undertaken (using centrally-read rather than locally-read endoscopic subscores; exclusion of studies with predominantly Asian patients; using TNFi-failure data instead of TNFi-exposed data; and (in response to a clarification question), excluding the Phase II tofacitinib trial).

The ERG judged the NMA to be generally well conducted but identified the following issues:

- Use of the probit scale to model clinical response/clinical remission. Whilst an improvement on a previous approach in NICE TA342, the use of a probit model did not aid interpretability and readability of the CS. A multinomial logit model could have been considered.
- Assessment of inconsistency did not examine any potential inconsistency in the maintenance TNFi-naïve network between the two-arm and three-arm trial.
- Exploration of best model fit. The ERG conducted additional analyses and would have made different choices regarding model fit. In general, for the effectiveness outcomes, the ERG would have chosen random effects models as the more conservative approach given the known between-study heterogeneity. In contrast, for the safety outcome of serious infections, the absence of any events in the placebo arms of the tofacitinib trials caused very wide credible intervals even when the ERG investigated a fixed effect model. The ERG therefore also ran an analysis using a frequentist framework for this outcome which allows a value of 0.5 to be added to cells when a zero value is present in the input data.
- Uncertainty around absolute probabilities from baseline models. To estimate absolute probabilities of each event, treatment effects from the NMA were combined with an estimate of the placebo (baseline) response from the placebo arms of included studies. The ERG was unable to replicate the placebo credible intervals used in the probit or logit models. The company models tended to lead to wider

credible intervals compared to our calculations, thus would lead to conservative results.

- Inclusion of the tofacitinib Phase II trial. Results from NMAs for response and remission for TNFi-naïve and TNFi-exposed subgroups are similar to the base case when the tofacitinib Phase II trial is excluded. However, results without the Phase II trial were not provided for safety outcomes and in this NMA the Phase II trial may have had a disproportionate effect on the random effect NMA because of the relatively high serious infection rate in the tofacitinib arm of this study.
- No safety analysis for the maintenance period. The company stated they were unable to conduct a NMA for safety outcomes in the maintenance phase. However, the ERG believe this could have been achieved by using the mFAS population of OCTAVE sustain. However, the issue of combining studies with treat-through and re-randomised designs would still remain
- Lack of adjustment for differing lengths of induction and maintenance periods across trials. The company did not attempt to adjust for differences in lengths of induction and maintenance treatment and the ERG is concerned that this could have introduced potential bias against those treatments where studies had shorter induction phase and benefit those treatments with a shorter maintenance phase.
- Differences between patient populations in the re-randomised design maintenance trials. OCTAVE sustain re-randomised all responders from the OCTAVE induction trials to either placebo or tofacitinib treatment. In contrast, in the three other studies with a re-randomised design, only patients who had received and responded to active treatment were eligible to be re-randomised into the maintenance phase of the study. In the ERG's view the base-case may be biased in favour of tofacitinib and it would have been useful to have explored the mFAS population for OCTAVE Sustain in a sensitivity analysis.
- Adjustments to treat-through trials. Although the ERG does not believe the adjustments made by the company introduce additional bias, it is nevertheless the case that non-responders at the end of the induction phase are ignored (and these participants potentially could have become responders by the end of the maintenance phase).

In the induction phase for the TNFi-naïve population all treatments showed strong evidence of benefit over placebo with infliximab having the largest treatment effect for both clinical response and clinical remission. In the TNFi-exposed population, tofacitinib had the largest treatment effect on clinical response and clinical remission compared to placebo. Only tofacitinib and vedolizumab showed strong evidence of benefit.

In the maintenance phase for TNFi-naive population all treatments showed strong evidence of benefit over placebo, with tofacitinib 10 mg having the largest treatment effect on clinical response and clinical remission. In the TNFi-exposed population, tofacitinib 10 mg had the largest treatment effect on clinical response and clinical remission compared to placebo. Tofacitinib 5 mg, 10 mg and vedolizumab 300mg Q4W and Q8W all showed a strong evidence of benefit over placebo.

In the safety analysis, which was only conducted for the induction phase [redacted].

In final summary, the ERG has identified the following key limitations of the evidence presented in the CS:

- All the direct evidence on the effectiveness of tofacitinib is from trials of tofacitinib versus placebo. The majority of evidence for other active treatments also comes from placebo controlled trials. In the absence of direct head-to-head comparisons of the available active treatments NMAs were undertaken.
- Heterogeneity was present among the studies included in the NMAs. Although the company took steps to try and reduce heterogeneity the ERG would have preferred random effects models for the effectiveness outcomes.
- The NMA for serious infections in the induction phase was potentially affected disproportionately by the Phase II tofacitinib trial. In addition, for this outcome the placebo arms of the tofacitinib trials experienced zero events. The ERG would therefore have preferred a fixed effects model as a sensitivity analysis for this outcome. However, very wide credible intervals persisted even with a fixed effects model and therefore the ERG has run an alternative frequentist analysis to investigate the impact of adding a value to cells in analyses where there are no events in the tofacitinib or placebo arms.
- No NMA for safety outcomes was conducted for the maintenance phase.
- Biases may exist due to differing lengths of induction and maintenance periods across trials (with may bias against treatments with shorter induction phases and

benefit treatment with shorter maintenance phases), and differences between studies with a re-randomisation design (the base-case NMA may be biased in favour of tofacitinib).

4 COST EFFECTIVENESS

4.1 Overview

The company submission includes:

- A systematic review of published economic evaluations of tofacitinib and other therapies for people with moderately to severely active ulcerative colitis (CS B.3.1 pages 118 to 120 and Appendix G);
- A description of the methods and results of their model developed to assess the cost-effectiveness of tofacitinib in relation to the comparators and population specified in the NICE scope for this appraisal (CS B.3.2 to B.3.11 pages 120 to 172 and Appendices H, I, J and M).

We summarise and critique these elements of the CS in sections 4.2 and 4.3 below and present additional work conducted by the ERG in section 4.4, including model validation, corrections to the company's analyses and additional analysis.

All of the results in this chapter include a confidential patient access scheme (PAS) price discount that has been agreed for tofacitinib but not an existing confidential PAS discount for the comparator vedolizumab. Results including both PAS discounts are presented in a confidential addendum to the ERG report.

4.2 Company's review of published economic evaluations

The company conducted a systematic review of the literature to identify economic evaluations of tofacitinib or any other therapy for moderately or severely active ulcerative colitis. The methods and results of this review are described in section B.3.1 and Appendix G of the CS. The ERG considers that the company's search strategy and inclusion/exclusion criteria were appropriate. However, as the search was conducted in October 2017, we updated it to identify any more recent relevant publications.

The company included 53 publications, described in Table 175 (CS Appendix G.1.2.2). The main submission focusses on 10 UK studies reported in six full papers[25,52-56] and seven abstracts[57-61] (see Table 35 CS page 120). Three of the full papers reported analyses conducted by the Evidence Review Groups for previous NICE technology appraisals: Archer et al. (2015)[25] and Tappenden et al. (2016)[54] relate to TA329 of infliximab, adalimumab and golimumab;[8] and Essat et al. (2016)[53] relates to TA342 of vedolizumab. A paper by Wilson

et al. (2017)[62] reported on the cost-effectiveness analysis of vedolizumab compared with TNF-alpha inhibitors from the Takeda submission for TA342. Tsai et al. (2008)[55] reported a cost-effectiveness analysis of maintenance treatment for infliximab compared with standard care based on the ACT I and ACT II RCTs. This analysis was used to inform resource utilisation and cost estimates in TA329, TA342 and in this current appraisal – see section 4.3.6.3 below (page 163). The final paper identified in the company's search - Buckland et al. (2008)[52] - compared high and low dose mesalazine, so is not relevant to the current appraisal.

The ERG update search identified two additional publications: a full paper by Wilson et al. (2018),[62] reporting cost-effectiveness of vedolizumab compared with conventional therapy from the Takeda TA342 analysis; and a paper by Wu et al. (2018)[63] reporting a cost-utility analysis comparing sequenced strategies including conventional therapies, tofacitinib, adalimumab, vedolizumab, golimumab and infliximab from a UK and Chinese perspective.

The analysis by Wu et al. indicated that one of the treatment sequences shown in Table 52 would be optimal in the UK context, depending on the incremental cost-effectiveness ration (ICER) threshold. Other sequences gave fewer QALYs for a higher cost than one or more alternatives (simple or extended dominance). At a cost-effectiveness threshold of £30,000 per QALY gained, the optimal treatment sequence would be adalimumab at first line, tofacitinib at second line and then conventional therapy.

Table 52 Non-dominated treatment sequences, UK perspective (Wu et al. 2018)[63]

Sequence	Cost (£)	QALYs	ICER (£ per QALY)	Comparator
CT	132,769	10.49	-	-
ADA-CT	134,598	10.71	8,438	CT
ADA-TOF-CT	153,333	11.67	19,407	ADA-CT
TOF-ADA-CT	154,216	11.70	30,989	ADA-TOF-CT
TOF-VED-CT	182,728	12.37	42,511	TOF-ADA-CT

CT: conventional therapy; ADA adalimumab; TOF tofacitinib; VED vedalimumab

We consider the cost-effectiveness of sequential treatment strategies in exploratory ERG analysis, see section 4.3.2.2 below.

4.3 Critical appraisal of the company's submitted economic evaluation

4.3.1 NICE reference case

In most regards the company's economic evaluation follows the NICE reference case and the NICE scope for this appraisal (see Table 53). The exception is the company's exclusion of adalimumab, infliximab and golimumab as comparators for patients with prior exposure to a TNFi. For this subgroup, clinical response and remission rates are not available for infliximab or golimumab, but they are available for adalimumab. Therefore, the company could have included adalimumab as a comparator for the TNFi-exposed subgroup. We discuss the appropriateness of this comparison in section 4.3.2.2 below.

Table 53 NICE reference case

Criteria	Included?	Comment
Decision problem as in scope	Yes	
Comparators as listed in scope	No	Adalimumab, infliximab and golimumab not included for people with prior TNFi exposure
Perspective on costs: NHS and PSS	Yes	
Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes	
Perspective on outcomes: All direct health effects, whether for patients or, when relevant, carers	Yes	
Cost utility analysis with fully incremental analysis	Yes	
Synthesis of evidence on outcomes based on a systematic review	Yes	
Time horizon: Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes	
Health effect expressed in QALYs. EQ-5D is preferred measure of health related quality of life	Yes	
Health related quality of life reported directly by patients and/or carers.	Yes	
Preference data from representative sample the UK population	Yes	
An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit.	Yes	
Discount rate: 3.5% pa for costs & health effects	Yes	

4.3.2 Modelled decision problem

4.3.2.1 Population

The population in the company model aligns with the NICE scope - people with moderately to severely active ulcerative colitis who are either intolerant of, or whose disease has had an inadequate response or loss of response to conventional therapy or a TNF-alpha inhibitor.

Subgroup analysis by TNFi exposure

The scope requests subgroup analysis according to previous treatment with one or more biologic drugs. The company base case is presented for two separate subgroups, labelled as biologic-naive and biologic-exposed (CS B.3.2.1). They argue that this division is appropriate as clinical evidence indicates that prior exposure to biologics is an important treatment effect modifier and that patients' treatment history is a deciding factor in the treatment pathway in clinical practice. *We agree with this approach but note that labelling the subgroups according to 'biologic' exposure is misleading, as the NMA results used in the model are defined by prior exposure to TNF-alpha inhibitors alone (not vedolizumab).*

Analysis for the whole population (ITT NMA)

The company also presents a scenario analysis using results from an NMA for all patients in the induction trials and the re-randomisation responder trials of maintenance therapy (CS B.3.7.2.1 and D.1.3.5.1.2). The CS notes that this 'ITT' scenario analysis is susceptible to heterogeneity in the proportion of patients with prior TNFi exposure in the trials. In particular, the TNFi trials only included TNFi-naïve patients, whereas the vedolizumab and tofacitinib trials included a mixture of patients with and without prior TNFi exposure. The company argues that the comparison between tofacitinib and vedolizumab represents the 'least confounded' results from the ITT scenario and they exclude the TNFi drugs from the table of cost-effectiveness analysis (Table 63 CS page 156).

We note that there is a high degree of uncertainty over the results of the ITT NMA. In particular, the odds ratios for vedolizumab compared with tofacitinib are very close to 1 with wide credible intervals: for example, for maintenance therapy with 8-weekly vedolizumab compared with daily 5 mg tofacitinib, the estimated odds ratios are [redacted] for clinical response and [redacted] for clinical remission (CS Table 106 D.1.3.5.1.2). *The ERG does not consider the company's 'ITT' cost-effectiveness scenario to be reliable because of the high level of uncertainty in the underlying NMA. The scenario also omits relevant comparators (the TNFi drugs), so does not address the specified decision problem. We therefore focus on separate analyses for the two TNFi exposure subgroups in our*

discussion and additional analysis. This approach is consistent with committee considerations in the NICE appraisal of vedolizumab (TA342).[9]

Baseline characteristics

In the company model, utility values and mortality rates depend on the age and gender mix of the cohort. Assumptions about the distribution of body weight are used to estimate dose and hence costs for some medications (infliximab and azathioprine). The company base case assumes the following baseline characteristics for the two subgroups:

- *TNFi-naïve*: age 41.1 years, 59.7% male and body weight 74.8 kg
- *TNFi-exposed*: age 41.3 years, 58.8% male and body weight 72.6 kg

These characteristics are based on means from the tofacitinib arms in the OCTAVE Induction trials, see Table 54.

Table 54 Patient baseline characteristics (OCTAVE Induction trials)

Subgroup	Treatment	N	Male, n (%)	Age, mean (95% CI)	Weight, kg mean (95% CI)
TNFi-naive	Tofacitinib	417	249 (59.7)	41.1 (39.8, 42.4)	74.8 (73.2, 76.4)
	Placebo	104	64 (61.5)	43.2 (40.5, 45.9)	73.7 (70.8, 76.6)
	Total	**521**	**313 (60.1)**	**41.5 (39.9, 43.2)**	**74.6 (72.6, 76.5)**
TNFi-exposed	Tofacitinib	488	287 (58.8)	41.3 (40.0, 42.6)	72.6 (71.1, 74.1)
	Placebo	130	68 (52.3)	39.4 (36.9, 41.9)	72.3 (69.3, 75.3)
	Total	**618**	**355 (57.4)**	**40.9 (39.3, 42.5)**	**72.5 (70.6, 74.4)**
All patients	Tofacitinib	905	536 (58.8)	41.2 (39.9, 42.5)	73.6 (72.1, 75.2)
	Placebo	234	132 (52.3)	41.1 (38.5, 43.7)	72.9 (70.0, 75.9)
	Total	**1139**	**668 (58.6)**	**41.2 (39.6, 42.8)**	**73.5 (71.5, 75.4)**

Source: CS Table 36, page 121. Subgroup and treatment totals estimated by ERG.

We consider it more appropriate to characterise the modelled population using all patients randomised in the OCTAVE induction trials, including patients in tofacitinib and placebo arms. Furthermore, we note that the small differences between the subgroups may well be due to chance – a suggestion that is supported by the observation that the mean age of randomised patients in the TNFi-exposed subgroup (40.9 years) is less than that for those in the TNFi-naïve subgroup (41.5). This appears counter-intuitive, although clinical advice to the ERG is that most exacerbations requiring drug change occur in the first year. Thus, the average of patients in the two subgroups may well be similar.

For comparison, the median age at diagnosis of ulcerative colitis in the 2016 RCP audit was 32 years (interquartile range (IQR) 24 to 45) and the median age at initiation of biologic treatment was 39 years (IQR 28 to 52).[64] The gender distribution in the audit was 59% (529/903), similar to that in the OCTAVE trials.

We consider that the gender mix, initial age and weight of the model cohort should be assumed similar for people with and without prior exposure to TNFi drugs. In ERG analysis, we assume 59% males, initial age 41 years and weight 73.5 kg, based on means for both arms in the OCTAVE Induction trials. We conduct scenario analysis to assess the impact of age (28 to 58) and body weight (range 70 kg to 80 kg) on the results.

4.3.2.2 Comparators

The model assumes that patients start treatment with tofacitinib or the biologic comparators with an induction phase of treatment. Patients who respond during induction continue to receive maintenance treatment with the same drug (with concomitant use of conventional drugs) until loss of response or an acute exacerbation requiring surgery. Patients who do not respond to induction treatment and those who relapse during maintenance continue to receive conventional treatment alone, until planned or emergency surgery, or death.

Inclusion of comparators in economic analysis

Tables 40 and 41 in the CS (page 130) outline the comparators used in the company's economic analysis:

- **TNFi-naïve subgroup**, all comparators specified in the scope (infliximab, adalimumab, golimumab, vedolizumab, tofacitinib and conventional therapy (CT));
- **TNFi-exposed subgroup**, only vedolizumab, tofacitinib and CT are included. Cost-effectiveness is not reported for infliximab, adalimumab or golimumab.

For patients with prior exposure to TNFi drugs, infliximab and golimumab could not be included in the company's NMA due to a lack of trial evidence (CS section B.2.9.2.1). However, the TNFi-exposed NMA does include adalimumab, so the company could have included adalimumab in the cost-effectiveness analysis for this subgroup. The CS does not give a clear rationale for omitting adalimumab for the TNFi-exposed subgroup.

Clinical experts have advised the ERG that treatment with a TNFi would sometimes be considered for a patient with prior exposure to another TNFi. There is a group of patients who lose response to first TNFi (usually infliximab) for a variety of reasons, such as

pharmacokinetics and anti-drug antibody formation. If they have initially responded and then lost response (secondary loss of response) it would be current practice to switch to a second line TNFi (in-class switch). Those who do not respond to a first line TNFi (primary non-responders), and those who lose response with therapeutic serum trough TNFi levels and without anti-drug antibody formation, are usually switched out-of-class (e.g. to vedolizumab or tofacitinib).

The occurrence of in-class switching is also supported by evidence from the UK IBD Audit: 21% of patients starting adalimumab (17/83) had previously not responded or been intolerant to a TNFi (RCP 2015, page 49).[65]

The ERG considers that adalimumab is a relevant comparator for at least some patients with prior exposure to a TNFi agent. We therefore include adalimumab in ERG analysis for this subgroup. However, we understand that further treatment with a TNFi may not be appropriate for all patients in this subgroup.

Drug use and dosage

SmPC dose regimens and recommendations about when to stop treatment with tofacitinib and biologic comparators are set out in Table 38 (CS page 128). This table also summarises dose assumptions used for costing in the model, see Appendix M (CS M.1.1) for further explanation. Table 38 incorrectly specifies the doses of adalimumab in the model. Based on the licensed dose, patients would receive 160 mg + 80 mg + 2 x 40 mg = 320 mg during the 8-week induction period and 40 mg x 4 = 160 mg per 8 weeks of maintenance. We confirm that the correct doses for adalimumab have been coded in the model.

Dosing and use of conventional drugs are detailed in Table 39 (CS page 129), with further explanation in Appendix M (CS M.1.1). CT is assumed to comprise a combination of aminosalicylates (balsalazide, mesalazine, olsalazine and sulfalazine), corticosteroids (hydrocortisone rectal foam and oral prednisolone) and the immunomodulator azathioprine. Clinical advice to the ERG suggests that the company's assumption of equal usage for the four aminosalicylic acid (5ASA) drugs does not reflect UK practice, as mesalazine is much more commonly prescribed for this patient group. See section 4.3.6.1 (page 159) below for discussion of drug utilisation and costing assumptions.

Stopping rules for drug treatment

- *Discontinuation due to lack of response to induction therapy*
  CS Table 38 summarises SmPC recommendations about when to stop tofacitinib and biologic drug treatment. These recommendations relate to early assessment of response following induction treatment (from 8 to 16 weeks after initiation). In contrast, the clinical trials provide evidence of response at 6 weeks for golimumab and vedolizumab and at 8 weeks for other comparators, and the model assumes a fixed 8-week induction period followed by immediate cessation of treatment for patients whose disease does not show a response in this time. *If in practice clinicians assess response to induction later than 8 weeks, the average cost of induction therapy will be higher than that estimated by the company model. However, effectiveness may also be higher if some patients have a late response to induction. The direction and magnitude of the bias from assuming a fixed 8-week period of induction for all comparators is unclear.*

- *Discontinuation due to loss of response during maintenance*
  Guidance for the TNF-alpha inhibitors (TA329) and vedolizumab (TA342) recommend annual assessment of response, with treatment continuing only if there is clear evidence of ongoing benefit. Clinical advice to the ERG is that the benefit of biologic treatment is usually considered annually, in line with NICE guidance. However, treatment would usually be withdrawn earlier for patients who lose response, as the patient will seek an appointment when symptoms recur or get worse and this will trigger consideration of changing or stopping treatment.

  The company model applies a constant risk of relapse across each 8-week cycle of maintenance, with treatment stopping immediately when patients lose response. Thus, it assumes that maintenance treatment is stopped within 8 weeks of a loss of response. To achieve this, all patients on maintenance treatment must have fast access to clinical assessment on relapse or be seen routinely every 8 weeks. The company model assumes an average of 2 outpatient visits for patients in remission on maintenance treatment and 4.5 visits per year for patients with a response but no remission.

  The ERG considers that the assumption that treatment will be withdrawn following relapse reflects UK practice. However, we have concerns that the costs of monitoring and follow-up in the company's model do not reflect the full cost of ensuring that treatment can be withdrawn within 8 weeks of a relapse. We consider a scenario with

additional costs for outpatient visits to enable treatment cessation within 8 weeks of a relapse - see section 4.3.6.3 below.

- *Trial of withdrawal for patients in stable remission on maintenance treatment*
  TA329 and TA342 also recommend a trial of withdrawal for patients with stable remission after 12 months of treatment, with the option to restart treatment following relapse. The company model does not reflect these recommendations, as maintenance treatment is assumed to continue for as long as patients have a response. We have been advised that in practice, patients in sustained clinical remission are more likely to continue maintenance treatment, as clinicians and patients are reluctant to stop a drug that appears to be working.

- *Other causes of treatment discontinuation*
  The model assumes that all drug treatment, including conventional therapy, stops after emergency or elective surgery.

  The only AEs included in the model were serious infections (SI) (see section 4.3.4.2 below) but the model assumes that treatment continues following SI. The company's NMA of safety outcomes from the induction trials includes discontinuation due to AEs (CS B.2.10.8.2 pages 110 to 112).

  [redacted]

  However, these results were not used in the model, as the company argue that this would lead to double counting because definitions of adverse events in OCTAVE and other trials included worsening of ulcerative colitis, which is already accounted for. The company state that risks of discontinuation due to AE or other causes are low and likely to be outweighed by discontinuation due to lack of efficacy (CS B.3.2.5).

  Clinical advice suggests that tofacitinib or biological treatment would be temporarily withheld following serious infection. If the drug had been clinically effective prior to the infection, withholding the drug until the infection has cleared, and then re-starting the drug again would be an option: e.g. for infections such as tonsillitis, pneumonia and urinary infections. If the infection was opportunistic or severe, such as disseminated herpes virus or meningitis, it is likely that the drug would be permanently stopped. Other SAEs likely to result in treatment cessation include malignancy (e.g. lymphoma), a major cardiovascular event, severe infusion reactions, drug-induced lupus reactions,

hypersensitivity reactions, neurological events (such as demyelination, neuropathy, focal neurology) and joint pains. Some rashes also warrant cessation, especially psoriasis-like eruptions. Leucopenia would always require dose reduction or temporary cessation.

The ERG considers it likely that including discontinuation due to AE from the NMA in the model would cause some degree of double-counting. However, the assumption of no discontinuation due to serious infections or other AEs is also unrealistic and likely to introduce bias.

Surgical treatment

Unlike previous NICE TAs for ulcerative colitis - TA329 and TA342 - surgery is not specified as a comparator in the scope for this current appraisal. This reflects the TA329 and TA342 committee conclusions that patients and clinicians would rather avoid or delay surgery because of adverse effects on wellbeing, potential for complications and the irreversible nature of the intervention that were not captured in the economic evaluations. The company model treats elective surgery as an option for patients with moderately to severely active ulcerative colitis treated with conventional treatment alone. The model also includes a risk of acute exacerbation requiring emergency surgery for patients not in remission (active disease or response without remission).

Drug sequencing

The CS presents results for one line of treatment with tofacitinib or biological comparator, followed by CT or surgical treatment. However, the model includes the facility to compare scenarios with two lines of active tofacitinib/biological treatment before CT/surgical treatment, as in the analysis by Wu et al. (2018) described above (4.2).[63] Our clinical advisors have indicated that after an initial trial of CT alone, patients with moderately to severely active ulcerative colitis would start treatment with a TNFi agent (usually infliximab). Patients without a response in the induction phase and those who lose response on maintenance treatment would then either switch within-class to another TNFi (adalimumab or golimumab) or outside-class (currently vedolizumab).

We conduct scenario analysis to compare the cost-effectiveness of sequenced treatment with biologic/tofacitinib for people without prior TNFi exposure, including in-class switching (e.g. infliximab-adalimumab), step up (e.g. infliximab-vedolizumab, infliximab-tofacitinib) and step-down (e.g. vedolizumab-infliximab, tofacitinib-infliximab) strategies: see 4.4.3.

4.3.3 Model structure

The company describes the structure and key features of their model in CS Section B.3.2.2. The model structure is similar to that in previous ulcerative colitis appraisals TA329 and TA342. It is a Markov cohort model, with a cycle length of 8 weeks and patient lifetime horizon. The half-cycle correction is not incorporated. Costs and QALYs are discounted at an annual rate of 3.5%. The company's illustration of the model is reproduced in Figure 7.

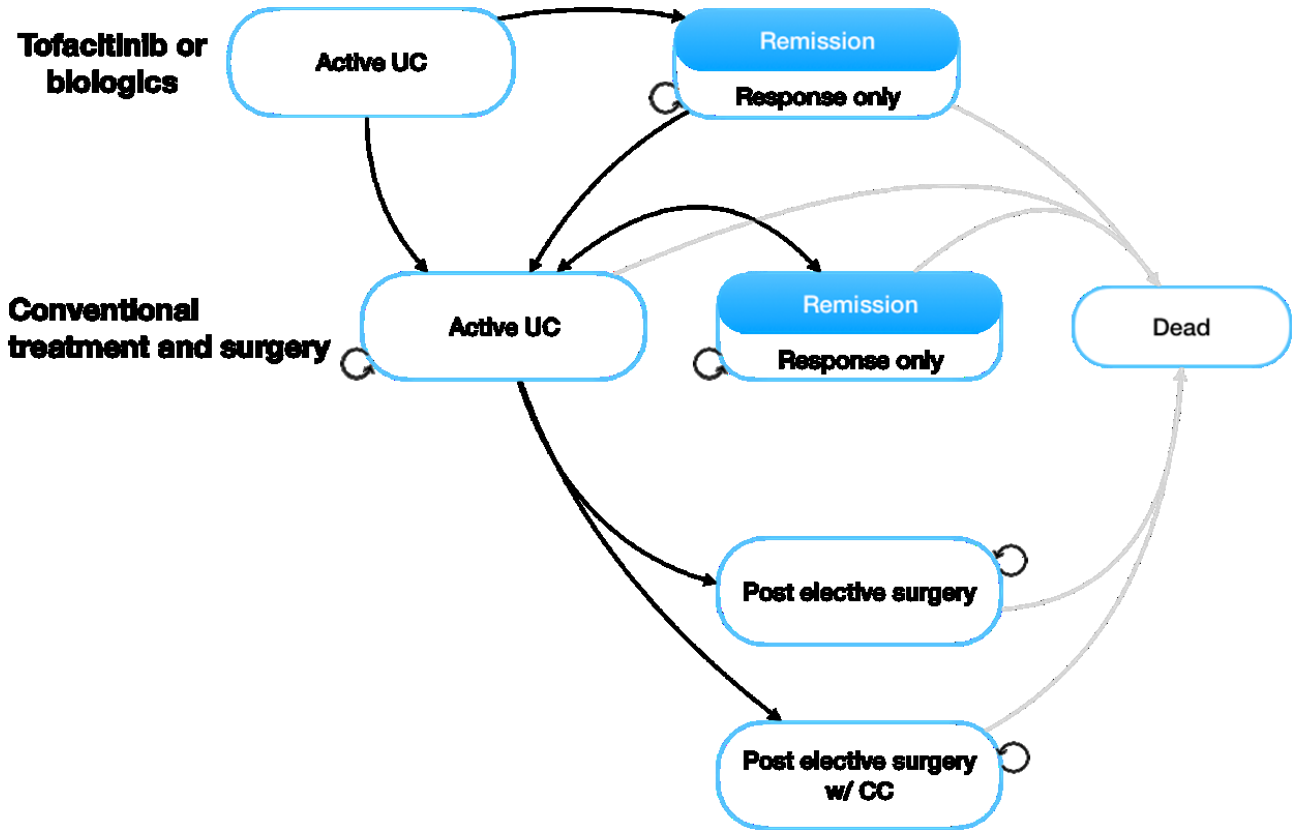


Figure 7. Company's model structure (Figure 31, CS B.3.2.2)

w/ CC = with colectomy complications; UC = Ulcerative Colitis

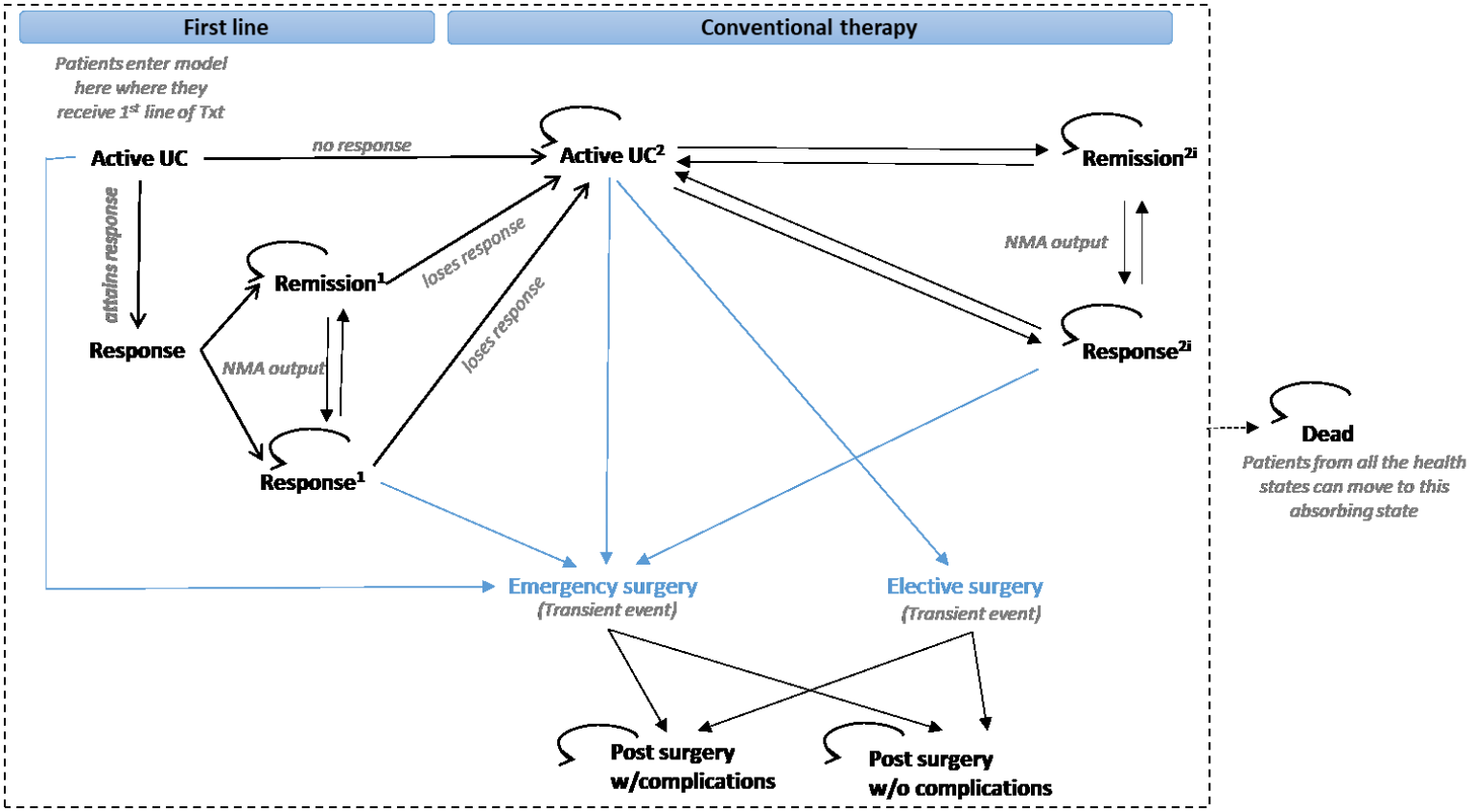
The model consists of nine health states, defined by stage of treatment (first line treatment with tofacitinib or biologic; conventional treatment; or post-surgery) and level of disease control (active UC; clinical response without remission; or remission), which we describe in Table 55. The transitions between the health states are further illustrated in Figure 8.

The company summarise key model assumptions and compare against previous ulcerative colitis appraisals in Tables 37 and 60 of the CS (pages 126 and 153 respectively). We critique of the model features and base case assumptions in section 4.3.7 below.

Table 55: Description of the model health states

	Health states	Description
Tofacitinib / biologic	1. Active UC	Patients enter the model with moderately to severely active ulcerative colitis following intolerance, inadequate response or loss of response to conventional therapy or a TNFi-alpha inhibitor. They commence treatment with an 8-week induction phase of treatment with tofacitinib or a biologic comparator.
	2.Remission	Of those who respond to induction treatment, a proportion attain remission (using clinical definitions of remission and response). Patients continue to receive maintenance treatment so long as they remain in response. For each 8-week maintenance cycle, the proportions of patients with a response and the proportion of responders in remission are estimated from the NMA.
	3.Response only	
Conventional treatment	4.Active UC	Patients transition to the Active UC state on conventional treatment following: • Non-response to tofacitinib/biologic induction • Loss of response in tofacitinib/biologic maintenance For the CT comparator arm, patients start in this state.
	5.Remisson	Patients may attain response with or without remission while on conventional therapy. Transitions between active UC, remission and response only health states continue to occur while patients receive ongoing conventional treatment.
	6.Response only	
Surgery	Emergency surgery *	In each model cycle, a proportion of patients who are not in remission (health states 3, 4 and 6) require emergency surgery due to an acute exacerbation.
	Elective surgery *	A proportion of patients in the Active UC health state are assumed to undergo elective surgery in each cycle.
	7.Post surgery without complications	Surgery is associated with perioperative risks of complications and mortality. Patients who survive surgery transition to one of two health states: with- or without long-term complications.
	8.Post surgery with complications	
	9.Dead	Absorbing state; the model accounts for: • Death from UC only occurs from surgery • Death from other causes (background mortality) occurs from all the health states

*The model treats surgery as a transient event: it is *NOT* a health state



[1] Patients who maintain response here continue to receive the 1st line of treatment in the maintenance phase

[2] Patients who fail to respond receive conventional therapy as 2nd line of treatment

[2i] Patients who maintain response here continue to receive the 2nd line of treatment which is CT in the maintenance phase

Figure 8. ERG illustration of patient transition in the model

The model uses three sets of input parameters:

- **Clinical inputs** that govern rates of response and remission and adverse event rates for comparator treatments, as well as the incidence and complication/mortality rates for surgery;
- **Utilities** for health states and disutilities for adverse events;
- **Resource use and costs** for drug acquisition and administration; monitoring and follow up, treatment of serious infections and surgery.

Values and sources of these parameters are summarised in Table 59 of the CS (page 149). We discuss and critique the parameter sources in sections 4.3.4, 4.3.5 and 4.3.6 below.

4.3.4 Clinical parameters

4.3.4.1 Response and remission

Choice of NMA models for economic analysis

The model uses NMA results to estimate the proportions of patients achieving clinical response and clinical remission in the induction and maintenance phases of treatment. The NMA results used in the company base case are reported in Tables 25 and 26 of the CS (pages 95 and 96). These correspond to the economic model inputs shown in Tables 43 and 45, respectively (CS pages 131 and 134).

See section 3.1.7 for the ERG summary and critique of the NMAs. We highlight key issues related to the company's choice of NMA models to use in their economic analysis.

- **Definitions of response and remission**
  The model uses locally read clinical response and clinical remission outcomes from OCTAVE and other trials (see Table 10 above for outcome definitions). The primary outcome for the OCTAVE trials - remission based on centrally-read endoscopic sub-scores – was not available from other studies in the networks. The company argue that local reading is "closer to real-world data", because clinicians make their own assessment of endoscopy results to inform treatment decisions (CS B.2.3.1.2.4). The NMA sensitivity analysis of centrally-read outcomes are gives similar results to the locally-read analysis. *The ERG agrees that locally-read clinical response/remission results are most relevant for the economic analysis.*

- **Choice of fixed effects versus random effects**

  The company state that their choice of NMA models was based on DIC measures of model fit, but that they preferred the simpler fixed effect approach when DIC statistics were similar (CS B.2.9.2.1.1). Table 56 below summarises the NMA models chosen for the company base case analysis.

Table 56 Selection of response/remission NMA models

	Patient subgroup	Induction	Maintenance
Company base case	TNFi-naive	Random effects	Fixed effects
	TNFi-exposed	Fixed effects	Fixed effects
ERG preference	TNFi-naive	Random effects	Random effects
	TNFi-exposed	Random effects	Fixed effects *

* Random effects model would not run for the maintenance NMA

The ERG has a general preference for the random effect NMA models, as we believe that the fixed effect models may underestimate uncertainty due to heterogeneity between the studies. We test the impact of different NMA models on cost-effectiveness results in section 4.4.3 below.

- **Combination of TNFi-failed and TNFi-exposed subgroups**

  The base case NMAs combine outcomes for subgroups defined as TNFi-failed for vedolizumab with TNFi-exposed subgroups for tofacitinib and adalimumab (CS Table 22). The company conducted a sensitivity analysis for the TNFi-failure subgroup, which reduced the probit score for tofacitinib by -0.13 in the induction phase, bringing it closer to vedolizumab. (CS Table 28). They reported that results were not available for adalimumab and that the analysis could not be run for the maintenance phase. Therefore, the TNFi-failure NMA sensitivity analysis does not provide the required input parameters and was not used in the economic model.

  The ERG considers that combining results for TNFi-failed and TNFi-exposed subgroups is a potential source of bias in favour of tofacitinib. We conduct a scenario analysis using a more like-for-like comparison between tofacitinib and vedolizumab, using data for the TNFi-failed subgroups from the OCTAVE and GEMINI trials (see Table 18 *in section 3.1.7*

Transformation of NMA results to transition probabilities

Results of the clinical response/remission NMAs were transformed from the probit scale to the natural scale and converted to absolute probabilities for use in the model. The probability of response is calculated using the P = 1- ϕ(θ) formula, where ϕ is the inverse of the cumulative normal distribution and θ is the sum of the probit scores for placebo and active treatment. The probability of remission is calculated using P = 1- ϕ(θ) formula, where ϕ is the inverse of the cumulative normal distribution and θ is the sum of the probit scores for placebo, active treatment, and remission.

For the induction phase, the proportions of the cohort with active disease, response but no remission and remission at the end of the first 8-week model cycle are shown in Table 57 below, by treatment and TNFi exposure subgroups.

Table 57 Distribution of cohort by health state at end of induction

	TNFi naïve subgroup			**TNFi exposed subgroup**		
	Active UC	Response only	Remission	Active UC	Response only	Remission
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NA, results not available from network meta-analysis

Some further assumptions are needed to calculate 8-week transition probabilities from the 52-week NMA response/remission rates. The company describes the approaches taken in previous NICE technology appraisals in section B.3.3.1.2 (page 132) of the CS.

- In the TA329 MTA (adalimumab, infliximab and golimumab), the assessment group had access to mid-point response and remission data for the maintenance period.[25] They used these data to estimate transition probabilities for two phases of maintenance - week 8 to 32 and week 32 to 52. The results are generally more favourable for the TNFi drugs in the second period than in the first.
- In the TA342 STA (vedolizumab), the company used a calibration approach to fit transition probabilities to the 52 week NMA results. This involved applying certain constraints, such as that no more than 20% of people with mild disease would enter

remission. This approach was criticised by the TA342 ERG for using arbitrary constraints and assumptions.

In the present appraisal, the company note that they considered both of these approaches, but without success: due to a lack of mid maintenance period results for some comparators; and a failure to accurately predict the target data with calibration.

Instead, the company take a simpler approach by assuming constant risks within and beyond the one-year trial data. The probability of loss of response is calculated from the probability of no response over 52 weeks from the NMA (1 minus the probability of response), adjusted to the 8-week model cycle. Members of the cohort who maintain a response in each cycle are then split between remission and response only health states using a fixed proportion (the ratio of 52-week probabilities of response with and without remission). The resulting estimates of the 8-week probabilities of loss of response and the proportions of patients in response with and without remission are shown in Table 58.

Table 58 Parameters used to model change of health state during maintenance

	TNFi naïve subgroup		TNFi exposed subgroup	
	Probability of losing response (per 8 weeks)	Percentage of responders in remission	Probability of losing response (per 8 weeks)	Percentage of responders in remission
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]
Golimumab 50mg	[redacted]	[redacted]	[redacted]	[redacted]
Golimumab 100mg	[redacted]	[redacted]	[redacted]	[redacted]
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib 5mg	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib 10mg	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab Q8W	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab Q4W	[redacted]	[redacted]	[redacted]	[redacted]
Conventional	[redacted]	[redacted]	[redacted]	[redacted]

Adapted from CS Table 45 page 134.

These calculations are mathematically correct, but we emphasise that they rely on assumptions of a constant risk of loss of response and constant ratio of patients in remission and response throughout maintenance treatment. Clinical advice to the ERG is that these assumptions might not be realistic. Experience with TNFi agents suggests that most serious exacerbations requiring drug change occur in the first year of treatment. Loss of response continues after a year of therapy but tails off in the second and subsequent years. Further,

the proportion of patients with a response and in remission is likely to increase over time, because responders without remission are more likely to stop or switch therapy (or have surgery) whereas those in remission will continue. Thus, the only-responders will tend to drop out faster than those in remission.

Similar concerns were raised by the NICE committee for TA329, which noted a discrepancy between modelled estimates of treatment duration and expert advice that of patients who start a TNFi, one third to one half are expected to continue therapy in the long term (paragraph 4.71).[8]

Results from the OCTAVE Open study are suggestive of similar trends in long-term maintenance of response and remission with tofacitinib (CS B.2.6.3.1 and Appendix L Table 233).

[redacted]

We conclude that the model assumption of constant risk of loss of response for patients on maintenance treatment does not reflect clinical experience. Extrapolation of relapse and discontinuation rates from the maintenance trials is likely to underestimate the average duration of treatment and hence both the costs and QALYs of active treatments. However, it is not possible to estimate the net direction of bias in ICERs between comparators, because trends in long-term risks may vary between TNFi drugs, vedolizumab and tofacitinib.

4.3.4.2 Adverse events: serious infection rates

The company conducted three NMAs on safety, based on data from induction phase RCTs, as described in sections 3.1.7 above (CS B.2.10.8.1). These include discontinuations due to adverse events and incidence of serious adverse events (SAE), but the company model only uses results from the serious infection (SI) NMA (CS B.3.3.3).

Exclusion of other serious adverse events

The company explain that they excluded adverse events other than serious infections because the most common SAEs reported in the trials were GI events, events related to ulcerative colitis, or “worsening of disease”, which may already be accounted for in the model through loss of response and remission, as described above.

Advice from our clinical expert suggests that there are other SAEs, such as malignancy and cardiac events, which though small in number are significant for patients and incur considerable cost to the NHS. This observation is in line with the approach taken in NICE TA 342 which included TB, malignancy (due to lymphoma), acute hypersensitivity reactions and skin site reactions, in addition to SIs.

We agree that there would have been a risk of double-counting the costs and effects of ulcerative colitis exacerbations had all SAEs had been included in the model. The omission of non-infection SAEs does introduce a risk of bias but given the frequency of these events this omission is unlikely to change the cost-effectiveness results.

NMA method for serious infections

The company applied a binomial logit NMA model to estimate the risk of serious infections in the induction trials (CS Table 34 page 111). They chose the random effects model for their base case because the DIC statistic was lower than for the fixed effects model.

The company acknowledges substantial uncertainty in the precision of estimates from the SI NMA, which gave a very high upper limit to the credible interval for all comparators and for tofacitinib in particular because there were no cases of serious infection in the placebo arms of the tofacitinib induction trials. The company note that if the credible interval limits for the SI risks are used in deterministic sensitivity analysis, this parameter has the greatest impact on the ICERs. They argue that this would be misleading and instead apply arbitrary limits around the SI risk for tofacitinib of 0% to 50% increase from placebo.

Whilst the ERG agrees that there is considerable uncertainty associated with the risk of serious infections, we have reservations about the company's approach to estimating this parameter (discussed in detail in section 3.3.10.2). Our verification checks indicated an even higher level of uncertainty around tofacitinib estimates, and we were unable to replicate the company's base case NMA values. We therefore applied a frequentist NMA approach to estimate the risk of serious infection, which we use as a scenario in ERG analysis (details in section 4.4.3).

Transformation of NMA results to SI probabilities

Table 59 shows the probabilities of serious infections used in the company base case, with ranges for sensitivity analysis. The probabilities are estimated from incidence during the induction phase (assumed to be 8 weeks), which is assumed to apply to each subsequent 8-week cycle of maintenance treatment. Except for the tofacitinib arm, the central estimates

from the company's NMA are similar to the ERG frequentist estimates, but the latter approach give more plausible ranges of uncertainty.

Table 59 Probabilities of serious infections used in model (per 8-week cycle)

Treatment	Company (Bayesian RE)			ERG (Frequentist RE)		
	Base case	Lower limit	Upper limit	Base case	Lower limit	Upper limit
Placebo	[redacted]	[redacted]	[redacted]	0.67%		
Adalimumab	[redacted]	[redacted]	[redacted]	0.58%	0.08%	4.15%
Golimumab	[redacted]	[redacted]	[redacted]	0.11%	0.01%	1.25%
Infliximab	[redacted]	[redacted]	[redacted]	0.44%	0.05%	3.90%
Tofacitinib *	[redacted]	[redacted]	[redacted]	1.90%	0.29%	12.57%
Vedolizumab	[redacted]	[redacted]	[redacted]	0.15%	0.01%	1.89%

* By assumption, the company limits range for tofacitinib sensitivity analysis

The company made a number of assumptions in relation to serious infections. First, the risk of serious infection is assumed to be same regardless of patients' prior experience of treatment with TNFi-agents. The duration of serious infections is also assumed to be the same for all comparators: the model applies a disutility for the duration of the 8-week cycle in which the infection occurs. These are simplifying assumptions that appear reasonable.

A rather stronger assumption is that the risk of serious infection is constant throughout treatment (i.e. probability of SI is same in the induction and maintenance phases and regardless of the length of maintenance). The company test this assumption with a scenario in which serious infections are only assumed to occur only in the induction phase.

4.3.4.3 Incidence of emergency and elective surgery

The company conducted a focused search to identify estimates for the probability of colectomy and related complications (see CS section B.3.3.2 and CS Appendix M, section M.3). Misra et al. (2016)[66] is chosen to inform estimates of the cumulative risks of emergency and elective surgery in the base case – see Table 60.The company argues that this study is the most appropriate source as it: was based on a retrospective analysis of UK Hospital Episode Statistics (HES) for ulcerative colitis with a follow-up of 15 years since diagnosis; consisted of a larger and more contemporary cohort; excluded surgery due to colorectal cancer and provided a split for elective and emergency surgery rates.

Table 60 Colectomy rates used in the base case model

	Cumulative risks	Risk per cycle	Value used in the model
Colectomy	6.9%	0.073%	--
Elective colectomy	5.5%	0.058%	0.058%
Emergency colectomy	2.0%	0.021%	0.021%

Source: Misra et al. 2016 (UK HES Data)[66]

The CS reports on 3 other studies: Chhaya et al. 2015; Solberg et al. 2009 (used in TA329 AG model) and Frolkis et al. 2013 (used in TA342).[67-69] The Company use estimates from Frolkis et al. to inform their sensitivity analysis.

The ERG agrees with the company's selection of the Misra et al. study for the base case estimate of surgery risks. For completeness, we test rates from Chhaya et al. in scenario analysis, although we consider it unlikely to influence the results.

4.3.4.4 Colectomy complications and mortality

Perioperative complications

In their base case, based on UK Inflammatory Bowel Disease (IBD) 2014 audit, the company assumed that 32% of patients who underwent elective surgery and 35% of patients who underwent emergency surgery had perioperative complications.[64] Although the rates were doubled in the sensitivity analysis, they did not influence the base case results.

Post-operative complications

The company also included an ongoing risk of pouchitis after elective or emergency surgery. The base case risk was 1.46% per 8-week cycle, based on a Belgian study by Ferrante et al. (2007).[26] The risk was varied in company sensitivity analysis based on a Japanese study by Suzuki et al (2014) [37]. *To explore the sensitivity of results to pouchitis risk, we conducted a scenario analysis similar to that by Tappenden et al. (2016)[54] using a Japanese study by Arai et al. (2010)[70] which reported overall incidence of early and late complications (see section 4.4). It is worth noting that we do not anticipate change in this parameter to have a substantial impact on the base case results.*

Perioperative mortality

The company assumed the same perioperative mortality rate for patients undergoing elective and emergency surgery. In the base case, the mortality risk per operation was estimated to be 2.8% based on the reduction in overall mortality by 19% between round 3

and round 4 of the IBD audit.[64] Our clinical advisor has noted that although the overall surgical mortality may be around 2.8%, emergency surgery will carry a higher risk. The ECCO guidelines quote a mortality rate of 5-8% for emergency surgery and <1% for timely elective surgery in "specialised centres". We view the company's approach of taking an average rate across elective and emergency surgery as a reasonable simplification.

4.3.4.5 All-cause mortality

The model assumes that ulcerative colitis and treatment does not have any influence on mortality, with the exception of perioperative deaths. All-cause mortality risks, adjusted for age and gender-mix, for the general population from the UK Life tables are applied to patients in pre- and post-surgery states. The same approach was used in the assessment group model for TA329, although in TA342 the company applied state-specific relative risks to include an excess risk of death due to ulcerative colitis: 1.9 for moderately to severely active ulcerative colitis and 1.3 for post-surgery ulcerative colitis states .[9] *We consider that the approach in the current appraisal is acceptable. Although there are additional mortality risks not reflected in the model – e.g. for colorectal cancer – the relative risk estimates are likely to include perioperative deaths already accounted for.*

4.3.5 Health related quality of life

The model includes 7 utility parameters:

- A baseline utility for people without ulcerative colitis, adjusted for age and gender;
- 4 multipliers to reflect reduced utility (compared with no ulcerative colitis) for the health states:
    - Active ulcerative colitis;
    - Clinical response without clinical remission
    - Clinical remission;
    - Post-surgery.
- A utility multiplier for the effect of surgical complications;
- And a utility multiplier for the adverse effect of serious infections.

Parameter estimates were obtained from a systematic review of the literature on utility in ulcerative colitis (CS B.3.4.3 and Appendix H) and analysis of EQ-5D utility data from the OCTAVE trials (CS B.3.4.1 and Appendix M).

Utility estimates from published literature

The company conducted a systematic search for utility estimates (CS B.3.4.3 and Appendix H). We consider that the search strategy was satisfactory. As the search was conducted

over six months ago, we updated it, but did not identify any additional relevant studies. The company included 115 studies in their review, 44 of which reported EQ-5D utilities (Table 185, CS Appendix H). In the main submission, the company focus on 11 published studies reporting EQ-5D utility estimates for more than one relevant health state, in addition to economic analyses conducted for NICE TA329 (Archer et al. 2016)[25] and TA342 (Takeda 2014)[48] (see CS Table 50 B. 3.4.3 page 141). Utility parameters from published sources used in the company analysis are shown in Table 61. The company use estimates from Woehl et al. (2008)[71] in their base case and estimates from Swinburn et al. (2012),[72] in order to align with previous NICE technology appraisals for ulcerative colitis (TA329 and TA342).

Table 61 Utility parameters from the literature used in model

Source	Health state	Utility	ERG comments
Ara & Brazier[73]	No disease	Initial values TNFi-naive 0.8968 Prior TNFi 0.8960 Declines over time	Depends on age and gender of cohort. Formula derived from Health Survey for England 2003 and 2006 EQ-5D-3L (n=25,080). Regression coefficients not included in PSA.
Woehl et al.[71]	Active UC	0.4713	Utility multipliers calculated with respect to remission state. Used to adjust 'no disease' in company base case.
	Response	0.8736	
	Remission	1.0000	
	Post-surgery	0.8161	
Swinburn et al.[72]	Active UC	0.6317	Utility multipliers with respect to remission state. Active UC mean of 'severe' and 'moderate' utilities. Used in company scenario analysis.
	Response	0.8944	
	Remission	1.0000	
	Post-surgery	0.6596	
Diamantopoulos[74]	Serious infections	0.9858	Utility multiplier with respect to remission state
Kosmas (2015)[75]	Post-surgery complication	0.7889	Utility multiplier with respect to post-surgery state

Utility estimates from the OCTAVE trials

EQ-5D outcomes from the OCTAVE 1 and 2 induction trials and the OCTAVE sustain maintenance trial are outlined in CS B.2.6.1.2 and B.2.6.2.2, with further information in Table 218 (CS L.1.4) and Figures 54 to 61 (CS M.4). We discuss EQ-5D results from the OCTAVE induction and maintenance trials in section 3.3.7 above. To summarise, patients randomised in OCTAVE 1 and 2 were given an EQ-5D-3L questionnaire at baseline, 2 and 8 weeks, and patients in OCTAVE Sustain were given the questionnaire at baseline, 4, 8, 16, 24, 32, 40 and 52 weeks. Utility scores were calculated using UK preference weights, so are consistent with the NICE Reference Case.[76]

The CS reports analysis of EQ-5D data from the OCTAVE Induction trials to assess change in utility over time based on final health state at week 8. It is stated that the analysis was conducted separately for the TNFi naïve and exposed subgroups, using the full analysis dataset and a 'non-responder imputation method' (CS M.4). The company concluded that this analysis showed 'homogeneity' in mean EQ-5D index by final health state, although no statistical analysis was presented to support these claims. The company then used simple methods to estimate utility parameters from the OCTAVE data, which they used in scenario analysis their original submission (see Table 62 below).

Table 62 Simple estimates of health state utilities from OCTAVE EQ-5D data

Health state	N	Assumed utility	Assumed range (Min-Max)	Comments / assumption
Active UC	[redacted]	[redacted]	[redacted]	Mean of EQ-5D scores at baseline for participants in OCTAVE Induction 1 and 2 trials
Response no remission	[redacted]	[redacted]	[redacted]	Mean area under EQ-5D curves over one year for OCTAVE Sustain participants in remission or response-no-remission states at end of trial (see CS M.4 Table 238)
Remission	[redacted]	[redacted]	[redacted]	

Adapted from CS B.3.4.1 Table 49

In response to clarification questions, the company conducted further analysis of OCTAVE trial data. Linear mixed effect models were applied, grouping patients by health state (clinical remission, clinical response but not remission, active UC) at the trial endpoints (week 8 for OCTAVE 1 and 2, and week 52 for OCTAVE Sustain). Covariates tested included baseline EQ-5D, treatment, prior TNFi exposure, corticosteroid use at baseline, geographic region. We reproduce the results from the company response to clarification question B2 in Table 63 below.

The order of health state mean utilities are logical: for each trial dataset, estimates are highest for patients in remission and lowest for patients without a response. The company note that the mean utility estimates for each health state are higher in the maintenance trial than in the induction trials (although the confidence intervals overlap). This might support the view that primary non-responders (participants in the induction trials who had not had a response by week 8 and were excluded from the maintenance trial) are different to secondary non-responders (participants who started maintenance therapy with a response but lost this over the year of follow up). The company use these results to conduct two scenario analyses around their base case analysis, see Table 64. However, the company emphasise that both these regression-based estimates and their earlier simple estimates of

health state utility values do not sufficiently address the difficulties relating to the re-randomisation design of the OCTAVE Sustain study.

Table 63 Linear mixed model estimates of utility by health state from OCTAVE EQ-5D data (reproduced from Table 11 clarification response question B2)

Efficacy endpoint [a]	OCTAVE Induction 1 and 2 Values at week 8		OCTAVE Sustain Values at week 52	
	N [b]	Adjusted mean (95% CI) [c]	N [d]	Adjusted mean (95% CI) [e]
Non-clinical response	[redacted]	[redacted]	[redacted]	[redacted]
Clinical response (but not clinical remission)	[redacted]	[redacted]	[redacted]	[redacted]
Clinical remission	[redacted]	[redacted]	[redacted]	[redacted]

[a]Efficacy endpoints are based on NRI and Local Read of Endoscopy.
[b] N = number of subjects with non-missing EQ-5D data at week 8
[c] Adjusted mean derived from the linear mixed effects model: Score = Treatment + Prior treatment with TNFi therapy + Corticosteroid use at baseline + Geographic region + Week + Treatment*Week + Baseline EQ-5D with subjects as random effect
[d] N = number of subjects with non-missing EQ-5D data at week 52
[e] Adjusted mean derived from the linear mixed effects model: Score = Treatment + Induction Treatment + Baseline Remission Status + Week + Treatment*Week + Baseline EQ-5D with subjects as random effect.
Abbreviations: CI, confidence interval

Table 64 Additional company scenarios for OCTAVE utility estimates (Clarification Response question B2)

Scenario	Health state	Induction (first cycle)	Maintenance
1	Active UC	Week 8: [redacted]	Week 8: [redacted]
	Response only	Week 8: [redacted]	Week 52: [redacted]
	Remission	Week 8: [redacted]	Week 52: [redacted]
2	Active UC	Week 8: [redacted]	Week 52: [redacted]
	Response only	Week 8: [redacted]	Week 52: [redacted]
	Remission	Week 8: [redacted]	Week 52: [redacted]

The company's simple and regression-based analyses of EQ-5D data from the OCTAVE trials are problematic as sources of utility parameters for the economic model. They are relevant to the decision problem and clinical evidence, but the re-randomisation design and lack of intermediate assessments of clinical response and remission between week 8 and week 52 complicate the interpretation of results. We therefore agree with the company that the utility estimates by Woehl et al. [71] provide a more appropriate source for base case parameters that are consistent with previous NICE appraisals for ulcerative colitis. We use these estimates in ERG preferred analyses, but also test scenarios based on the company's OCTAVE analyses and published sources (Swinburn et al.) [72].

4.3.6 Resource use and costs

4.3.6.1 Drug acquisition

The assumptions underlying drug cost calculations are outlined in section B.3.5.1 (CS pages 143 to 145). Further detail is given in Appendix M (CS M.5).

Tofacitinib and biologic comparators

Table 52 (CS page 144) lists total costs per 8-week cycle for induction and maintenance treatment for tofacitinib and the biologic drugs. However, we note that several of the per-cycle costs in this table do not match the figures in the company's model - see Table 65 below for the drug acquisition costs from the model.

Table 65 Drug acquisition cost for tofacitinib and biologics

Drug	Induction (per 8 weeks)		Maintenance (per 8 weeks)	
	Dose	Cost	Dose	Cost
Tofacitinib [a]	10 mg twice daily	[redacted]	5 mg twice daily [d]	[redacted]
			10 mg twice daily	[redacted]
Adalimumab	160 mg week 0, 80 mg week 2 & 40 mg week 4 & 6	£2,817	40 mg every other week [d]	£1,409
			40 mg every week	£2,817
			27% every week [e]	£1,789
Golimumab [b]	200 mg week 0, 100 mg week 2 & 50 mg week 6	£3,052	50 mg every 4 weeks [d]	£1,526
			100 mg every 4 weeks	£1,526
Infliximab [c] (biosimilar)	5 mg/kg week 0, 2 & 6	£5,269 (£4,742)	5 mg/kg every 8 weeks	£1,756 (£1,581)
Vedolizumab	300 mg week 0, 2 & 6	£6,150	300 mg every 8 weeks [d]	£2,050
			300 mg every 4 weeks	£4,100

a Includes confidential PAS discount for tofacitinib. [redacted]

b Costs for golimumab assume provision of 100 mg dose at same cost as 50 mg dose as agreed in patient Access Scheme (TA329)

c Base case analysis assumes use of infliximab biosimilar (Remsima or Inflectra). Costs allow for wastage (no vial sharing) estimated by simulated distribution of body weight based on means and standard deviations for patients at baseline in the OCTAVE Induction trials.

d Base case analyses in bold. Alternative doses used in scenario analysis.

e Following assumption by ERG in TA329: in maintenance, 73% of patients have 40 mg of adalimumab every other week and 27% of patients have 40 mg of every week

The model includes a confidential Patient Access Scheme (PAS) discount for tofacitinib and the golimumab PAS agreement to supply 100 mg tablets at the same price as 50 mg tablets (TA329). All other drugs are at list price. We note that there is a PAS discount in place for

vedolizumab that is not factored into these costs. We present results including the vedolizumab PAS discount in a confidential addendum to this report.

In addition to the standard dose used in the base case calculations (in bold in the above table), the company presents scenarios for higher maintenance doses for tofacitinib (20 mg per day), adalimumab (27% of patients have 40 mg every week), golimumab (100 mg every 4 weeks) and vedolizumab (300 mg every 4 weeks). In the base case, the company assumed use of a biosimilar for infliximab (Remsima or inflectra). We have been advised that a significant minority of patients on infliximab will be on 6-weekly dosing (around 25-30%, compared with more than 50% on 8-weekly dosing). However, as the cost-effectiveness results are not sensitive to an increase in the cost of infliximab, we do not explore this further.

Cost calculations in the model are correct based on the stated assumptions about dosage and current NHS list prices (MIMS June 2018). Estimates are similar to those in the company model for TA342 (vedolizumab), with the exception of the induction cost for golimumab (in TA342 the company assumed 6, 50 mg doses). We consider the assumption of 3 100mg and 1 50mg dose (as in the current company's submission) to be more reasonable.

Conventional treatment

The costs of conventional drug treatment as a comparator and concomitant with biologic or tofacitinib are summarised in Table 53 of the CS (page 145). These costs match those used in the company base case model, with the exception of azathiopine which is costed in the model allowing for wastage. We summarise the costs used in the company base case analysis in Table 66 below.

Estimated usage is based on reported concomitant medication in the 2016 RCP audit of biological treatment for IBD.[64] The company assumes that for patients on conventional therapy alone, the proportions of patients prescribed the three main classes of drugs (aminosalicylates, corticosteroids and immunomodulators) are similar to reported use at initiation of biological therapy in the audit (50.3%, 47.9% and 46.4% respectively). Concomitant usage rates were based on reported use after three months of biological treatment (46.4%, 20.1% and 37.3% respectively). Azathioprine is excluded from the estimated cost of conventional therapy concomitant with tofacitinib, as this combination is not recommended. Further assumptions were made about usage within the drug classes and dosage – see Appendix M (CS M.5).

Table 66 Drug acquisition cost for conventional treatment

Drug	Per 8 weeks		Usage (% of patients) [c]		
	Dose	Cost	CT alone	With biologic	With tofacitinib [b]
Aminosalicylates					
Balsalazide	1.5 g twice daily	£52.42	12.6%	11.6%	11.6%
Mesalazine	1.2 g daily	£54.90	12.6%	11.6%	11.6%
Olsalazine	500 mg twice daily	£300.53	12.6%	11.6%	11.6%
Sulfasalazine	0.5 g twice daily	£6.87	12.6%	11.6%	11.6%
Corticosteroids					
Prednisolone	20 mg daily	£6.79	44.1%	19.9%	19.9%
Hydrocortisone [d]	Once every other day	£18.66	3.8%	0.6%	0.6%
Immunomodulators					
Azathioprine	2 mg/kg daily [a]	£7.48	46.4%	37.2%	0.0%
		Total cost	**£59.30**	**£52.18**	**£49.40**

a Costs for azathioprine allow for wastage estimated by simulated distribution of body weight based on means and standard deviations for patients at baseline in the OCTAVE Induction trials.
b Azathioprine not recommended for concomitant use with tofacitinib
c Usage estimated from RCP national IBD audit (2016): at initiation of biologic treatment for CT alone; after 3 months of biologic treatment for concomitant treatment.
d Rectal foam

The company has assumed equal usage of the four aminosalicylic acid (5ASA) drugs. However, we have been advised that almost all 5ASA use in the UK is mesalazine. Sulphasalazine is restricted to those with joint disease, and Olsalazine and Balsalazide are very rarely prescribed. Given the high cost of olsalzine, this suggests that the cost of 5ASA drugs is over-estimated. However, doses of 5ASA in patients with active disease, such as those starting tofacitinib or biological therapies, are likely to be maximised; e.g. mesalazine 4.8 g per day.

The above estimates of the cost of conventional therapy are lower than those used in the previous NICE appraisal of vedolizumab (TA342): £204.80 for CT alone and £102.40 concomitant with biologic therapy. However, the TA342 estimates were based on expert opinion, with the assumption that CT costs would be halved it taken with a biologic drug. We consider that the estimates in Table 66 are likely to be more reflective of NHS practice, since they are based on national audit data.

Overall, we consider the drug acquisition costs used in the company model to be realistic. We note that there have been some small changes in NHS prices for included drugs; sulfasalazine (£7.83), prednisolone (£0.47) and azathioprine (£2.20) (MIMS June 2018).

These changes result in a very small reduction in the estimated cost of CT alone (£58.02), with biologic drugs (£51.68) and with tofacitinib (£48.86).

Drug wastage calculations

The dose of infliximab and azathioprine are based on body weight. The company apply assumptions about wastage in their cost calculations, assuming no vial sharing. The wastage calculation methods are described in CS Appendix M (section M.5). In the base case, the company uses the method recommended by Hatswell et al. (2016)[77], with the distribution of body weight simulated from means and standard deviations for men and women in the OCTAVE Induction trials.

The ERG agrees with the company's approach to costing wastage for IV drugs. The model includes an option to use mean body weight from the trials, assuming vial sharing, but we do not consider this further as it is not realistic for NHS practice.

4.3.6.2 Drug administration

Vedolizumab and infliximab are administered by IV infusion and require an outpatient appointment with a healthcare professional. The company assumed 3 appointments for induction and 1 for maintenance per 8-week model cycle. The cost per visit was estimated at £137.37, based on the weighted mean for consultant led and non-consultant led, face-to-face, non-admitted, follow-up gastroenterology clinic appointments (NHS Reference Costs 2016-17) – CS Table 54, page 145. This estimate is similar to that used in the NICE appraisal of vedolizumab (TA342).

Adalimumab and golimumab are administered by subcutaneous injection. The company assume that patients can self-administer these treatments at zero cost to the NHS, due to the available of support from the drug manufacturers.

The company conduct sensitivity analysis around the cost of IV administration for vedolizumab and infliximab (varying the cost per dose from £70.20 to £161.72. We consider this range appropriate. We conduct additional scenario analysis to assess the impact of assuming an initiation of self-administration of subcutaneous injections: adding the cost of a non-consultant led clinic attendance (£107) to the cost of induction for adalimumab and golimumab.

4.3.6.3 Monitoring and follow up

Assumptions about the use and cost of monitoring and follow-up are summarised in section B.3.5.2 Tables 55 and 56 (CS pages 146 to 147) – see Table 67 below.

Table 67 Health state resource use and costs

Resource	Unit cost [a]	Resource use per year by health state [b]				
		Active UC	Response only	Remission	Post-surgery (no compl.)	Post-surgery (complications)
Outpatient visits [c]	£137	6.50	4.50	2.00	1.50	1.75
Blood tests [d]	£3.06	6.50	3.90	3.25	1.50	3.25
Endoscopy [e]	£277	2.00	0.50	0.20	1.25	0.65
Hospital episodes [f]	£2,985	1.50	1.20	0.30	0	3.25
Total per year		**£5,944**	**£4,350**	**£1,236**	**£557**	**£10,131**
Cost of surgery [g]		-	-	-	**£6,091**	**£7,295**

a Unit costs from NHS Reference Costs 2016-17
b Resource use from expert opinion in Tsai et al. (2008)[55], except hospital episodes for response only and remission health states from expert advice to company.
c Weighted average for consultant led and non consultant led (WF01A)
d Directly accessed haematology service (DAPS05)
e Diagnostic colonoscopy, 19 years and over (FE32Z)
f Non-elective inpatient (codes not specified)
g Elective proximal and distal colon procedures, 19 years and over with/without complications (FF32 and 33 (see CS Table 58 page 148).

Resource use assumptions were based on opinion from a panel of UK gastroenterologists, reported by Tsai et al. (2008).[55] The company state that they chose this source because the definition of the health states aligns with those used in the model: with Mayo scores similar to those in the OCTAVE trials. The Tsai et al. estimates of resource use have also been used in other NICE appraisals for ulcerative colitis (TQ329 and TA342).

Tsai et al. reported the same rate of 0.30 hospital admissions per year under standard care, for active ulcerative colitis, response only and remission states. The company changed this to assume more hospital episodes per year for the active UC and response only health states based on clinical expert opinion. Clinical advice to the ERG is that this is unrealistic, and that hospital admission is only undertaken for acute severe colitis (which is already included in the model), moderately severe ulcerative colitis not responding to oral prednisolone (which would not be treated with tofacitinib) and post-surgery with complications (admitted about once a year). Some other usage assumptions are also high in a current NHS context, including outpatient visits for patients in remission and post-surgery without complications, and endoscopy for uncomplicated post-surgery.

Health care usage assumptions from Tsai et al. (2008) are consistent with health state definitions in the model and with previous NICE appraisals for ulcerative colitis (TA329 and TA342). However, we have been advised that some estimates of the number of outpatient visits and endoscopies are high, and that the company's additional assumptions about hospital episodes are unrealistically high, particularly as admission for acute exacerbation requiring emergency surgery is already included in the model. We therefore test an alternative resource use scenario, suggested by our clinical expert (Table 68).

Table 68 ERG scenario for resource use by health state

Resource	Unit cost [a]	Resource use per year by health state [b]				
		Active UC	Response only	Remission	Post-surgery (no compl.)	Post-surgery (complications)
Outpatient visits [c]	£137	6.50	4.00	1.00	1.00	2.00
Blood tests [d]	£3.06	6.50	4.0	4.0	1.00	3.25
Endoscopy [e]	£277	2.00	0.50	0.20	0.2	0.65
Hospital episodes [f]	£2,985	0	0	0	0	1.0

We also question whether the assumption that maintenance treatment will always stop within 8 weeks of a loss of response is consistent with the number of outpatient appointments. We test two scenarios to align the costs of assessing patients on maintenance treatment with the model assumption that treatment will always be discontinued within 8 weeks of a relapse:

- *Add one additional outpatient appointment consultation when patients have a relapse while on maintenance treatment. In this case, all patients are assumed to seek and obtain an additional appointment when they experience symptoms.*
- *Assume 6.5 outpatient visits per year for all patients on maintenance treatment. This would be necessary if patients do not seek or cannot obtain an earlier appointment when they experience symptoms of moderately or severely active ulcerative colitis, so routine appointments would be needed to assess patients every 8 weeks.*

The company model omits ongoing costs of stoma care for the post-colectomy health states. This issue was addressed in the NICE vedolizumab appraisal TA342, and the committee concluded that these costs should be included but that the ERG estimate of £315 for a 6-month period was low. We revisited stoma cost estimates by Buchanan et al. (2011)[78] and uprated them for nurse costs (PSSRU 2017) and HCHS inflation for consumables: estimating an annual cost of £1,065.90 per person with a stoma, or £426.36 per person in the post-surgery health states (assuming 40% have a stoma). We include these costs in ERG preferred analysis.

The unit costs for health resources are also reasonable, although we note that the source for the mean cost per hospital episode is unclear (the CS and model does not specify which NHS Reference Cost codes are included). However, in comparison with estimates in TA329 and TA342, some of the unit costs are low. In particular, the estimated costs of surgery are lower than estimates from previous appraisals, which were based on the analysis by Buchanan et al. (2011): £13,176 for Europe or £11,620 in the UK. The model also omits ongoing stoma are costs for stoma care: estimated at £466 by Buchanan et al.[78]

We conduct additional scenario analysis to test the sensitivity of the results to higher estimates of the cost of surgery and the inclusion of stoma care costs in the post-surgery health states.

4.3.6.4 Treatment of serious infections

Finally, company estimates of the costs of treating serious infections are listed in CS Table 57 (page 147). The cost of £2,539 was estimated as a weighted average of inpatient care for six types of infection, with unit costs and incidence based on NHS Reference Cost data (2016-17). The company explored a wide range around this estimate (£722 to £11,471) in sensitivity analysis, which is appropriate given additional uncertainty due to the omission of other types of adverse events.

4.3.7 ERG critique of model assumptions and inputs

We summarise the key model assumptions alongside ERG's critique in Table 69. Broadly, we agree with company's approach albeit a few concerns, as highlighted in the table.

Table 69 Other model features and base case assumptions

Factor	Company justification		ERG comments
Model framework	Markov model	Allows the modelling of recurrent risks, such as response to treatment after induction and maintenance	We agree with the company general approach (Markov cohort structure) and representation of health states and transitions. The model structure and assumptions are similar to TA329.
Time horizon	Patient lifetime	UC is a chronic condition, so a patient lifetime horizon allows calculation of all relevant costs and quality of life impairment	Agree
Cycle length	8 weeks	Based on maintenance phase assessment intervals in the clinical trials of tofacitinib and other comparators. A fixed cycle length was used to allow the flexibility to adding a continuous sequence of treatments.	Agree.
Half cycle correction	Not applied	Relatively short cycle length	Agree

Factor	Company justification		ERG comments
Source of clinical effectiveness estimates	NMA for clinical response and remission (locally read) for subgroups with/without prior exposure to TNFi drugs	Locally read clinical response/remission reflects real-life practise. Choice of NMA models based on DIC statistics, with preference for fixed effects if no difference	Agree with use of locally-read clinical definitions of response and remission in economic model. We prefer random effects models to better reflect uncertainty related to heterogeneity. Combining TNFi-exposed subgroup for tofacitinib with TNFi-failed subgroup for vedolizumab is likely to have biased results for this comparison. We test alternative NMA model in ERG additional analysis, in section 4.4.3.
Calculation of transition probabilities	Outputs from NMA for response and remission transformed to 8-week probabilities	Simple approach; assumes constant risk through maintenance phase and beyond in extrapolation, as well as a constant ratio of response to remission. Company attempted calibration to fit 8-week transitions but this did not work.	We view it as unrealistic to assume constant risk of loss of response. Clinical experience indicates the risk is greatest in the first 6-12 months; and falls thereafter. The proportion of patients with response and remission is likely to increase over time as per our clinical advice. This is because responders (without remission) are more likely to stop or switch therapy whereas those in remission would continue with treatment. However, in the absence of evidence it is difficult to adapt the model

Factor	Company justification		ERG comments
Treatment waning of effects and discontinuation	Treatment effect was assumed to be maintained with ongoing treatment. Non-responders are given conventional therapy as second-line	Follows the approach taken in the independent economic analysis in NICE TA329	Agree with discontinuation for failure to respond in induction or loss of response in maintenance. We note this assumes that in practice patients who experience exacerbations of symptoms can be assessment and, if appropriate, treatment stopped within 8 weeks. The model does not reflect NICE recommendations for annual assessment of benefit and need for continued treatment in previous appraisals TA329 and TA342. Clinical advice suggests that withdrawal of treatment for patients in remission is unlikely in practice, and the effects of this are difficult to quantify given the model structure and limited evidence over long-term maintenance of remission.
Continuation of conventional therapy	Patients on CT and/or those who previously achieved but lost response to biologics were assumed to continue on CT irrespective of disease state	Simplifying assumption consistent with previous TAs and published literature [25]	Agree
Surgery	A proportion of non-responders and those who discontinue CT undergo elective colectomy. Patients from all health states (except remission) may undergo emergency surgery.	Consistent with clinical practice	Agree

Factor	Company justification		ERG comments
Risk of surgery	Assumed to be time-independent	Consistent with prior TAs; the base case model combined existing evidence with population in the model	Agree. Note surgery is treated as a transient event rather than a health state.
Source of utilities	Background utility ('no disease') based on EQ-5D by age and gender in general population. Health state utilities (EQ-5D) for pre and post-surgical states from Woehl *et al.* 2008. Sensitivity analyses using OCTAVE trial EQ-5D estimates and Swinburn et al.	Woehl *et al.* used as base case in previous TAs, with Swinburn in scenario analysis. Use of age/gender dependent background utility consistent with scenarios in previous TAs. Results consistent in scenarios with simple and regression-based utility estimates from trial EQ-5D data	Agree with the company's approach for the background utility estimates. We also agree with the use of Woehl *et al.* estimates of health state utilities, for consistency with other TA. Improved analysis of trial EQ-5D and scenario analysis in response to clarification questions. But we agree that the re-randomisation design of the maintenance trial complicates interpretation of within-trial utility estiamtes. We conduct additional scenario analysis in section 4.4.3
Source of unit costs	NHS reference costs, eMIT and MIMS for drug costs	Consistent with the NICE reference case	Agree
Biologic treatments	Golimumab formulation	It was assumed that the 100 mg vials of golimumab were used in induction (2x100 mg vial at week 0 and 1x100 mg vial at week 2) and the 50 mg vials were used for the maintenance dose (1x50 mg vial Q4W)	Agree

Factor	Company justification		ERG comments
Conventional therapy	The RCP audit data about use of conventional drugs by drug class at biologic initiation assumed to reflect CT alone for active UC	Assumption in absence of evidence on the CT mix	Agree
	Assumed equal use of 4 drugs in aminosalicylate class (balsalazide, mesalazine olsalazine & sulfasalazine)	Assumption in absence of evidence	Advice to ERG is that most patients receive mesalazine in UK. Doses for active UC higher than specified in company base case. Net effect on costs in base case likely to be neutral.
	Hydrocortisone was considered as a topical treatment (rectal foam); prednisolone was assumed to represent the oral corticosteroid treatment group and beclomethasone is used as add-on treatment to 5-ASA. Azathioprine assumed to represent the immunomodulator group	Simplifying assumptions	Agree
Concomitant medication	Use of conventional drugs concomitant with biologics/ tofacitinib based on 3-months follow-up in RCP audit. Azathioprine was excluded from concomitant use with tofacitinib	The evidence at 3-months follow-up were assumed to be reflective of continuous concomitant use.	Agree

Factor	Company justification		ERG comments
Administration cost for injections	No administration cost for self-administered sub-cutaneous injections assumed	Consistent with clinical practice	Agree. We conducted additional scenario analysis to assess the impact of assuming one outpatient consultation to support initiation of self-administered injections to the cost of induction for adalimumab and golimumab in section 4.4.3.
Health state resource use	Mostly based on Tsai *et al.* 2008, except increased frequency of hospitalisation was assumed for more severe disease	Consistent with structure of economic model and previous Tas. Gradient of hospitalisation with disease severity is realistic	Agree with use of Tsai et al. as base case. But clinical advice to ERG suggests frequency of outpatient visits and endoscopy exceed current UK practice and additional assumptions about hospital episodes are unrealistic. We test alternative resource use scenario in section 4.4.3 We also conduct scenario analysis to assume additional outpatient consultations to achieve 8-weekly assessment of response and cessation of treatment if indicated (see section 4.4.3) The company excludes cost of stoma care and the estimated cost of surgery is low compared with previous appraisals. The test the inclusion of stoma care costs and higher surgery costs in additional analysis, section 4.4.3

Factor	Company justification		ERG comments
Cost of serious infection	The cost of a serious infection was considered to be a weighted average of six types of infection: sepsis, tuberculosis, pneumonia, skin and soft tissue infection, bone and joint infection and urinary tract infection	Simplifying assumption in the absence of other evidence	Agree
Adverse events	Costs and utility loss associated with serious infection risk included	Evidence on the incidence of serious infections was available for all drugs. SIs are often associated with immunosuppressants. In the base case, the range of SIs with tofacitinib was assumed to increase between 0-50% from the base case value	There is a lot of uncertainty associated with SIs due to the rarity of events. Tofacitinib had the highest number of serious infections whilst golimumab had the lowest, We detail our concerns in section 3.1.7 and 4.3.4.2 and conduct additional analysis using an alternative frequentist NMA in section 4.4.3.
Mortality	Death from surgery and other cause mortality (as general population)	Consistent assumption on death from surgery as in TA329. Evidence on death from other cause in UC is sparse.	Agree

Source: CS Table 37 and Table 60

4.3.8 Model validation

The company describes their approach to model validations in CS section B.3.10. They state that they engaged UK clinical experts, statisticians and health economists to validate model inputs and assumptions in a UK advisory board meeting. Further details on the key aspects of validation are summarised in CS Table 78.

The CS stated that clinical experts validated model methods pertaining to the patient population; subgroup analysis by prior TNFi-exposure; time on treatment and discontinuation rates; costs (including monitoring cost for tofacitinib, health state costs and resource use, including rate of hospitalisation); emergency surgery; quality of life and maintenance dose of tofacitinib. The experts are reported to agree with the company's assumptions in most of these aspects, except for:

- **Patient population:** Although the baseline characteristics of the patient population in OCTAVE reflect UK practice, the duration of disease in OCTAVE trials (which was 6-7 years) is longer than that in clinical practice (which is ~2-4 years).
- **Health state unit costs and resource use, including rate of hospitalisation:** Tsai et al. was confirmed to reflect an accurate representation of unit costs and resource use as per clinical practice. However, the experts suggested that the model base-case assumptions relating to annual medical resource use (CS Table 55) underestimated the resource use per patient per year.
- **Tofacitinib maintenance dose:** Experts observed that the company assumption relating to [redacted] of patients benefitting from maintenance dose of 10mg twice daily may not be limited to patients in the TNFi-exposed group only.

The economic model was quality checked by health economists. For face validity, the company compared the proportion of patients in response and remission predicted by the model against the estimated values from the NMA, shown below in Figure 9.

Further, the model results were compared with previous TA329; however, the CS did not report any comparison of the results in TA329 with those in the current appraisal. We discuss this in detail in section 4.4.1. For internal validity, the CS stated that a second modeller reviewed the model; conducted extreme value tests alongside inspecting model code, formulae and references. An independent health economist was reported to have reviewed the model structure, parameter inputs and core model assumptions.

Figure 9 NMA results and model predictions of patient allocation and treatment survival

Source: CS Appendix M.2

4.3.9 Company cost effectiveness results

4.3.9.1 Base case deterministic results

The company present their base case results in CS section B.3.7, page 155. These incorporate the confidential PAS discount for tofacitinib but not the PAS discount for vedolizumab. The base case assume use of biosimilar drugs for infliximab. We report results including all available PAS discounts in a confidential addendum to this report.

People without prior exposure to TNF-alpha inhibitors

Results for the subgroup with no prior TNFi exposure are shown in Table 70.

- Adalimumab, golimumab and infliximab are dominated by tofacitinib – they are estimated to cost more and produce fewer QALYs;
- Tofacitinib gives a mean QALY gain of [redacted] QALYs for a mean additional cost of [redacted] compared with conventional therapy: giving an incremental cost-effectiveness ratio (ICER) of £8,554 per QALY gained;
- Compared with tofacitinib, vedolizumab gives an additional QALY gain of [redacted] QALYs for an additional cost of [redacted]: an ICER of £615,057 per QALY gained.

Table 70 Cost effectiveness: Company base case, no prior TNFi (with tofacitinib PAS)

Strategy	Total		Incremental analysis			Pairwise ICERs tofacitinib vs. comparator (£/QALY)
	QALYs	Costs (£)	QALYs	Costs (£)	ICER (£/QALY)	
Conventional	[redacted]	[redacted]	-	-	-	£8,554
Adalimumab	[redacted]	[redacted]	-	-	Dominated	Dominated
Golimumab	[redacted]	[redacted]	-	-	Dominated	Dominated
Infliximab	[redacted]	[redacted]	-	-	Dominated	Dominated
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£8,554	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£615,057	£615,057

Reproduced from CS B.3.7.1 Table 61 page 155

People with prior TNF-alpha inhibitor exposure

Company base case results for the subgroup of people with prior TNFi exposure are shown in Table 71. The company omits adalimumab as a comparator in this subgroup. Clinical response/remission rates are not available for this subgroup for infliximab or golimumab.

- Compared with conventional therapy, tofacitinib gives a mean gain of [redacted] QALYs for an additional cost of [redacted]: resulting in an ICER of £10,302 per QALY gained;
- Compared with tofacitinib, vedolizumab gives an additional QALY gain of [redacted] QALYs for an additional cost of [redacted]: giving an ICER of over £7.8m per QALY gained.

Table 71 Cost effectiveness: Company base case, prior TNFi exposure (tofacitinib PAS)

Strategy	Total		Incremental analysis			Pairwise ICERs tofacitinib vs. comparator (£/QALY)
	QALYs	Costs (£)	QALYs	Costs (£)	ICER (£/QALY)	
Conventional	[redacted]	[redacted]	-	-	-	£10,302
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£10,302	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£7,838,238	£7,838,238

Reproduced from CS B.3.7.1 Table 62 page 155

Disaggregated model results

The company report QALY and cost results from the model disaggregated by stages of treatment and health state in Appendix J to the CS (pages 382 to 388). We show key results for assessment of the face validity of the model in Table 72 and Table 73 below.

Table 72 shows the break down for patients with no prior exposure to TNF-alpha inhibitors. Predicted survival is very similar for the alternative treatments, at around [redacted] years from model entry (up to age [redacted] years). For all comparators, a large proportion of the estimated lifetime is spent with active ulcerative colitis, under management with conventional drug treatments. After discounting, life expectation is about [redacted] years, with very little difference between the comparators. QALY differences between treatments are slightly larger (from [redacted] to [redacted] discounted QALYs), due to estimated effects on rates of response and remission for the TNF-alpha inhibitors, tofacitinib and vedolizumab. Cost differences between the comparators are largely driven by the cost of the initial drug treatment, which are offset to some degree by savings in the cost of monitoring and managing the condition for the more effective drugs. Other cost differences are small.

Disaggregated results for patients with prior TNFi exposure are shown in Table 73. Modelled health outcomes are less favourable for the TNFi-exposed subgroup than for the TNFi naive subgroup, reflecting the lower response and remission rates from the NMA

Table 72 Disaggregated model results: company base case, no prior TNFi exposure

	CT	Adalimumab	Golimumab	Infliximab	Tofacitinib	Vedolizumab
Years of treatment (undiscounted)						
Initial treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Years by health state (undiscounted)						
Active UC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response nr	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Remission	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post surgery wc	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Life years (discounted)						
Active UC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Remission	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post surgery wc	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs (discounted)						
Active UC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Remission	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post surgery wc	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Costs (discounted)						
Initial treatment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse events	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Health state	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

UC, ulcerative colitis; nr, no remission; wc, with complications

Table 73 Disaggregated model results: company base case, prior TNFi exposure

	Conventional	Tofacitinib	Vedolizumab
Years of treatment (undiscounted)			
Initial treatment	[redacted]	[redacted]	[redacted]
Conventional	[redacted]	[redacted]	[redacted]
Post surgery	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]
Years by health state (undiscounted)			
Active UC	[redacted]	[redacted]	[redacted]
Response nr	[redacted]	[redacted]	[redacted]
Remission	[redacted]	[redacted]	[redacted]
Post surgery	[redacted]	[redacted]	[redacted]
Post surgery wc	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]
Life years (discounted)			
Active UC	[redacted]	[redacted]	[redacted]
Response nr	[redacted]	[redacted]	[redacted]
Remission	[redacted]	[redacted]	[redacted]
Post surgery nc	[redacted]	[redacted]	[redacted]
Post surgery c	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]
QALYs (discounted)			
Active UC	[redacted]	[redacted]	[redacted]
Response nr	[redacted]	[redacted]	[redacted]
Remission	[redacted]	[redacted]	[redacted]
Post surgery nc	[redacted]	[redacted]	[redacted]
Post surgery c	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]
Costs (discounted)			
Initial treatment	[redacted]	[redacted]	[redacted]
Conventional	[redacted]	[redacted]	[redacted]
Adverse events	[redacted]	[redacted]	[redacted]
Surgery	[redacted]	[redacted]	[redacted]
Health state	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

UC, ulcerative colitis; nr, no remission; wc, withouth complications

4.3.9.2 Deterministic sensitivity analyses

The CS presents the parameters and ranges included in their Deterministic Sensitivity Analysis (DSA) in CS Table 59. Parameters for safety, efficacy and utilities were varied using confidence intervals and published literature. For certain parameters such as risk of serious infections, the company conducted exploratory scenarios based on assumptions (see section 4.3.4.2 above). Results of the DSA are tabulated in CS Table 69 and CS Table 70 and presented as tornado plots in CS Figure 37 and CS Figure 38. The tornado plots for both TNFi- naïve and TNFi-exposed subgroups compare tofacitinib against conventional therapy alone. These show that the costs of serious infections, costs of conventional treatment and response estimates for the maintenance phase are key drivers of model results. Other parameters such as risk of colectomy, health state related resource use, response estimates in induction also influence the base case results, but to a lesser extent. The company has not presented tornado plots comparing tofacitinib with other comparators. In particular, the comparison with vedolizumab is important as the effectiveness of the two drugs are comparable. This makes it difficult to draw any robust conclusions from the DSA results. We address this issue in ERG additional analyses in section 4.4.

4.3.9.3 Probabilistic Sensitivity Analysis

The company conducted a probabilistic sensitivity analysis (PSA) on their base-case model to assess parameter uncertainty. Assumptions used to characterise uncertainty are described in CS Section B.3.6.1. Briefly, the company uses CODA samples for safety and efficacy parameters obtained from the NMA. We view this approach as appropriate as this preserves the joint posterior distribution and any correlation of treatment effects in the simulated outputs. Beta distributions are used for colectomy rates, perioperative complications and mortality, post-surgery complications, mortality and utility estimates. Parameters for costs and resource use are assigned gamma distribution. We consider that the parameters are assigned appropriate distributions and the PSA is correctly implemented. The results of the PSA are presented in CS Table 67 and CS Table 68; scatter plots are presented in CS Figure 33 and CS Figure 34; and cost effectiveness acceptability curves (CEACs) are in CS Figure 35 and CS Figure 36. The overall conclusion of the PSA results are similar to the base case results; however, in both the sub groups, total QALYs and costs are higher in the PSA results compared to the base case results. The company attributes this difference in PSA and base case results to the CODA samples used in the PSA. The CS states that at a willingness-to-pay threshold of £20,000 per QALY, tofacitinib had the highest probability of being cost-effective amongst the comparators at 80.5% in the TNFi-naïve group and 56.3% in the TNFi-exposed group, respectively.

4.3.9.4 Scenario Analysis

The company conducted a range of scenario analyses to assess the impact of key variables on the model outcomes. We were unable to replicate the following scenarios as the CS did not provide sufficient explanation: NMA results for the ITT population, maintenance dose mix of tofacitinib and centrally read NMA results. The company provided further information in their response to clarification question B6. They also acknowledged an error in incremental QALYs and incremental costs for the scenario relating to mix maintenance dose of tofacitinib in TNFi-naïve subgroup (CS Table 65) which they corrected in their response. Despite incorporating the changes suggested by the company, we were unable to replicate the company's cost-effectiveness results pertaining to scenario using central read NMA results. We present our results for this scenario in section 4.4.2. A summary of the company's scenarios, alongside their justifications and results obtained are presented in Table 74. The company concluded that the cost effectiveness results in both the sub-groups- TNFi-naïve and TNFi-exposed were predominantly influenced by change in utility estimates.

The ERG considers that the company has been selective in the scenarios that they present to explore the robustness of their base case cost-effectiveness results. In particular, they do not explore the impact of key assumptions such as inclusion of costs associated with stoma care, cost-effectiveness results from alternative NMA models. We extend the range of scenario analyses in ERG additional analyses below.

Table 74 Company scenario analyses

Company scenarios	Brief rationale/assumption	ICERs for Tofacitinib vs CT (£/QALY) TNFi-naïve	TNFi-exposed
Base case		**£8,554**	**£10,302**
Overall ITT population			£7,805
Tofacitinib maintenance dose mix	[redacted] of patients receiving 5mg; [redacted] of patients receiving 10mg	£12,628	£13,947
Fixed baseline utility instead of age-adjusted	Assumption that patient quality of life stays constant over time.	£8,760	£10,589
OCTAVE trial utilities	EQ-5D data were collected in Tofacitinib Phase III clinical trials	£15,508	£18,276
Swinburn utilities	To compare with previous analyses	£11,932	£14,487
Emergency surgery from any state	Due to the uncertainty on the likely protection from acute events based on the level of response/remission, patients are assumed to undergo emergency surgery regardless of state membership	£8,194	£9,962
Emergency surgery only from active UC	As above but assuming response to treatment offers the same level of protection from acute events, as remission	£8,652	£10,475
No emergency surgery	As above, but assuming no emergency surgery in the model	£8,710	£10,593
Central read NMA results	Central read was the primary endpoint in OCTAVE trials.	£9,469	£10,793
Discounting every cycle	It tested the sensitivity of the model when the discounting of outcomes is applied every 8 weeks.	£8,606	£10,398
Adalimumab maintenance 73% 40 mg Q2W and 27% 40 mg QW	Dose escalation of adalimumab was assumed in Archer *et al*.	£8,554	--
Golimumab 100 mg every 4 weeks in maintenance	A 100 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients, such as those who have experienced a decrease in their response	£8,554	--
Vedolizumab 300 mg every 4 weeks in maintenance	A 300 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients who have a body weight ≥ 80 kg	£8,554	Dominated

Source: CS Table 63 to 66; 71 to 77

4.4 Additional work undertaken by the ERG

4.4.1 ERG model validation

4.4.1.1 Model verification procedures

We checked the economic model for transparency and validity. The visual basic code used within the model was accessible. The NMA code in WinBUGs was provided in Appendix D.1.3.4.

We conducted a range of 'white box' tests to verify model inputs, calculations and outputs:

- Cross-checking of all parameter inputs against values in the CS and cited sources;
- Checking the individual equations within the model;
- A range of extreme value and logic tests to check the plausibility of changes in results when parameters are changed
- Checking the VBA code for treatment sequencing
- Checking all model outputs against results cited in the CS, including the base case, PSA, DSA and manually ran all the scenarios
- Running the NMA code in WinBUGs to replicate selected results (see section 3.1.7).

In addition, we checked the model calculations of patient transitions through the health states, costs and QALYs by re-coding the model independently based on the inputs from the company's submitted model.

Overall, we found the economic model to be of a good quality, with very few errors in input parameters, logic or coding. We identified a few small errors that we correct in section 4.4.2 below, which did not make any substantive difference to the results. We were also successful in replicating outputs from most of the company's NMA models, with the exception of the serious infection NMA (section 3.1.7).

4.4.1.2 External validity

We have tabulated the model predictions against the observed clinical data for the maintenance phase, in Table 75 below. While the model results appear comparable with the clinical data for the tofacitinib arm in the TNFi-naïve group, there are large differences in the estimates for TNFi- exposed subgroup for this arm, along with the placebo arms for both induction and maintenance phases.

Table 75 Comparison of the predicted model results of Tofacitinib and Placebo (CT) against the observed clinical data – INDUCTION Phase

Study	Treatment	TNFi-naive		TNFi-exposed	
		Clinical response	Clinical remission	Clinical response	Clinical remission
OCTAVE Induction 1	Placebo	[redacted]	[redacted]	[redacted]	[redacted]
	Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]
OCTAVE Induction 2	Placebo	[redacted]	[redacted]	[redacted]	[redacted]
	Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]
Model	Placebo	[redacted]	[redacted]	[redacted]	[redacted]
	Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix J.1.2. Table 199

4.4.1.3 Cross validation

In section 4.2 above (page 134), we state that the CS reported previous economic models, including published literature and analyses conducted by ERGs for previous NICE TAs, for patients in ulcerative colitis. Whilst we acknowledge that there are methodological differences between the economic models across these studies, nonetheless we view that they provide sources for cross-validation of results from the company base-case analysis. Of the reported studies, we cross-validate the modelled findings of the current appraisal with 2 previous NICE TAs (TA342 and TA329) and 1 published study as summarised in Table 76. The most relevant analysis for the current appraisal is the final version from the NICE TA of vedolizumab (TA342). This appraisal relates to same patient population as the current appraisal and comparators overlap, except Tofacitinib and surgery.

Table 76 Comparison of modelled outcomes

Study name (time horizon)	QALYs		Life years	
Current appraisal (lifetime)	***TNFi- naive***	***TNFi-exposed***	***TNFi- naive***	***TNFi-exposed***
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]
TA342 (10 years)	Ada: 5.76	Ved: 5.46	Not reported	Not reported
	Gol: 5.79	CT: 5.37		
	Inf: 5.82	Surgery: 4.28		
	Surgery: 4.28			
	Ved: 5.90			
	CT: 4.28			
TA329 (Lifetime, AG model)	***Moderate to severe UC who failed at least 1 prior therapy***			
	Ada: 10.82		Not reported	
	Inf:10.81			
	Gol: 10.63			
	CT: 10.47			
Wu et al. (lifetime)	***Moderate to severe UC***			
	CT:10.49		Not reported	
	Ved→CT: 11.48			
	Tof→CT: 11.51			
	Inf→CT: 10.87			
	Gol→CT:10.89			
	Ada→CT: 10.71			
	Ved→Tof→CT: 12.37			
	Inf→Tof→CT:11.81			
	Gol→Tof→CT:11.83			
	Ada→Tof→CT:11.67			
	Tof→Ved→CT:12.37			
	Tof→Inf→CT:11.84			
	Tof→Gol→CT:11.86			
	Tof→Ada→CT:11.70			

4.4.2 ERG corrections to company model

We identified a few errors in the company's model, as shown in Table 77 below. The company corrected issue 2(ii) and provided further information to address issue 2(i) as response to the clarification questions. However, the ERG was unable to replicate the company's results for scenario in issue 2(i), although the differences in ICERs, obtained by the company and ERG, were minimal. The ERG implemented the corrections in Issues 1 and 3. These are discussed in the following sub-sections.

Table 77 ERG corrections to company model

<table>
<tr><th>Aspect of model</th><th>Problem</th><th>ERG Correction</th></tr>
<tr><td rowspan="2">1.Cost calculations</td><td>i. Cost of elective surgery with complications: the company used the cost of surgery without complications</td><td>Recoded column FA in 'Engine2L' sheet</td></tr>
<tr><td>ii. Cost of CT : We noted a few small changes in prices for sulfasalazine (£6.87), prednisolone (£0.91) and azathioprine (£2.17)</td><td>Values used by the ERG:
Sulfasalazine: £7.83;
Prednisolone: £0.47;
Azathioprine: £2.20

(MIMS June 2018)</td></tr>
<tr><td rowspan="2">2. Scenario analysis</td><td>i. Centrally read NMA results: ERG was unable to replicate the cost-effectiveness results presented by the company in CS Table 72 (scenario 7) and CS Table 76 (scenario 7)</td><td>We were unable to replicate the ICERs for tofacitinib vs CT (£/QALYs) reported by the company for this scenario (shown below)
<table>
<tr><td></td><td>Company</td><td>ERG</td></tr>
<tr><td>TNFi-naïve</td><td>£9,469</td><td>£9,524</td></tr>
<tr><td>TNFI-exp</td><td>£10,793</td><td>£10,789</td></tr>
</table></td></tr>
<tr><td>ii. CS Table 65: Error in incremental costs and incremental QALYs</td><td>Company corrected this as response to clarification question B6 (b). The corrections did not change the ICER.</td></tr>
<tr><td>3.Weight - wastage</td><td>i. Error in estimation of weight – wastage in cell N17:N18 and cell Q17:Q18 in sheet!Cost_Drug</td><td>Recoded the cells in sheet!Cost_Drug. The corrections do not have any impact on the base case CE results as these use 'fitting distribution' approach for wastage calculation.</td></tr>
</table>

4.4.2.1 Results for TNFi-naive subgroup

Making the corrections in Table 77 to the company's base case model resulted in a small increase in the ICERs for people without prior exposure to TNFi (Table 78**)**. The results were robust to deterministic and probabilistic sensitivity analysis and scenario analyses.

Table 78 Deterministic company base case (ERG corrected) -TNFi-naive (tofacitinib PAS)

Strategy	Total		Incremental analysis			Pairwise ICERs TOFA vs. comparator (£/QALY)
	QALYs	Costs (£)	QALYs	Costs (£)	ICER (£/QALY)	
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£8,564
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Tofa. dominant
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Tofa. dominant
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Tofa. dominant
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£8,564	N/A
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£615,077	£615,077

Table 79 DSA results company base case (ERG corrected) - TNFi-naïve (tofacitinib PAS)

	ICER TOFA vs. CT (£/QALY)	
Base case	**£8,564**	
Parameter	**Low limit**	**High limit**
Serious infection costs	£7,622	£13,191
Conventional treatment costs (min-max)	£9,559	£4,137
Response/remission treatment effect - maintenance	£6,292	£11,920
Colectomy risk (No risk - Frolkis 10y)	£7,388	£11,109
Health-state related resource use per patient per year	£8,334	£10,994
Response/remission treatment effect - induction	£7,609	£10,180
Serious infection risk	£7,259	£9,382
Hospitalisation cost	£9,850	£7,604
Pre-surgery health state utilities	£8,105	£9,493
OP visit + blood test costs	£9,140	£8,353
Endoscopy cost	£9,067	£8,082
Remission (z) - maintenance	£8,315	£8,838
Post-surgery health state utilities	£8,511	£8,617
Periorative mortality risk (0 - 3%)	£8,587	£8,559
Remission (z) - induction	£8,545	£8,581
Post-operative pouchitis (0.7 - 2%)	£8,576	£8,552
Colectomy cost	£8,573	£8,553
Serious infection utility reduction (0% - 3%)	£8,555	£8,572
Periorative complications (No risk - double the risk)	£8,566	£8,561
Post-surgery complication utility weight reduction (0% - 40%)	£8,566	£8,561
OP administration cost (£70 - £161)	£8,564	£8,564

Table 80 Probability of being cost-effective - TNFi-naïve subgroup

Treatments	£20k per QALY WTP	£30k per QALY WTP
Conventional	[redacted]	[redacted]
Adalimumab	[redacted]	[redacted]
Golimumab	[redacted]	[redacted]
Infliximumab	[redacted]	[redacted]
Tofacitinib	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]

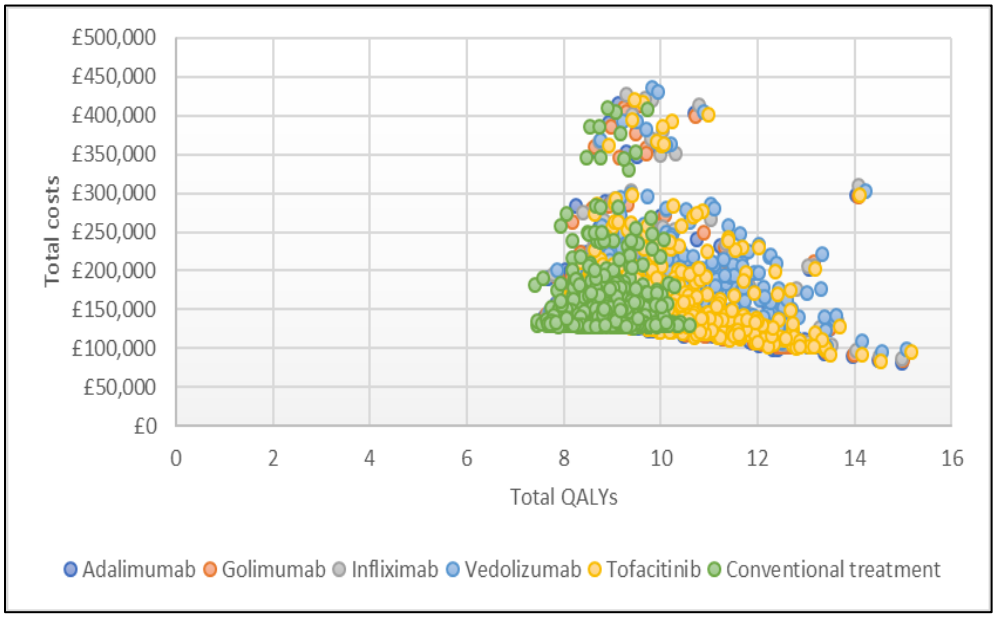


Figure 10 PSA scatter plot - TNFi-naïve subgroup

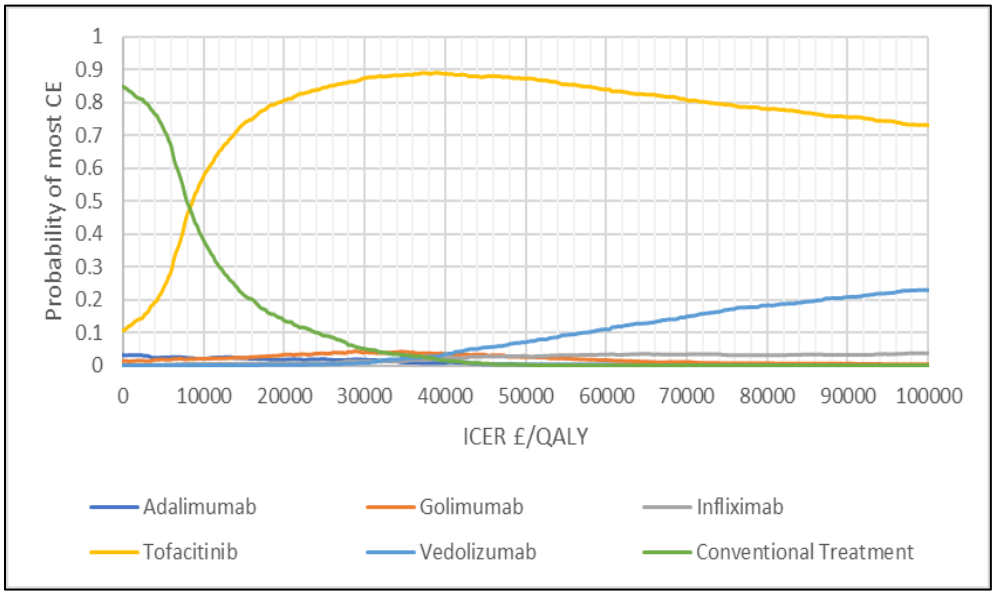


Figure 11 Cost effectiveness acceptability plane - TNFi-naïve subgroup

Table 81 Scenario analyses, company base case (ERG corrected) – TNFi-naive subgroup

Scenarios	Assumption	ICER for tofacitinib vs.	
		CT	Vedolizumab
Base case		**£8,564**	**£615,077**
Tofacitinib maintenance dose mix	[redacted] of patients receiving 5mg; [redacted] of patients receiving 10mg	£12,637	Tofacitinib dominant
Fixed baseline utility instead of age-adjusted	Assumption that patient quality of life stays constant over time.	£8,770	£634,346
OCTAVE trials utilities	EQ-5D data were collected in Tofacitinib Phase III clinical trials	£15,525	£1,079,814
Swinburn utilities	To compare with previous analyses	£11,945	£853,228
Emergency surgery from any state	Due to the uncertainty on the likely protection from acute events based on the level of response/remission, patients are assumed to undergo emergency surgery regardless of state membership	£8,204	£606,872
Emergency surgery from active UC only	As above but assuming response to treatment offers the same level of protection from acute events, as remission	£8,661	£618,151
No emergency surgery	As above, but assuming no emergency surgery in the model	£8,719	£618,068
Central read NMA	Central read was the primary endpoint in OCTAVE trials.	£9,534	£187,809
Discounting every cycle	It tested the sensitivity of the model when the discounting of outcomes is applied every 8 weeks.	£8,616	£617,451
Vedolizumab dose 300 mg Q4W	A 300 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients who have a body weight ≥ 80 kg	£8,564	Tofacitinib dominant

4.4.2.2 Results for TNFi-exposed subgroup

Table 82 Deterministic company base case (ERG corrected), TNFi-exposed (TOF PAS)

Strategy	Total		Incremental analysis			Pairwise ICERs TOF vs. comparator (£/QALY)
	QALYs	Costs (£)	QALYs	Costs (£)	ICER (£/QALY)	
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£10,311
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Tofa. dominant
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£10,311	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£7,838,381	£7,838,381

Table 83: DSA results for TNFi-exposed subgroup (compared to CT)

	ICER (£/QALY)	
Base case	**£10,311**	
Parameter	**Low limit**	**High limit**
Serious infection costs	£9,376	£14,909
Conventional treatment costs (min-max)	£11,302	£5,903
Response/remission treatment effect - maintenance	£7,825	£13,342
Health-state related resource use per patient per year	£9,531	£12,383
Colectomy risk (No risk - Frolkis 10y)	£9,108	£11,909
Serious infection risk	£9,013	£11,126
Response/remission treatment effect - induction	£9,461	£11,501
Hospitalisation cost	£11,481	£9,439
Pre-surgery health state utilities	£9,751	£11,374
Remission (z) - maintenance	£9,758	£10,946
OP visit + blood test costs	£10,857	£10,112
Endoscopy cost	£10,818	£9,827
Post-surgery health state utilities	£10,250	£10,373
Remission (z) - induction	£10,250	£10,371
Periorative mortality risk (0 - 3%)	£10,339	£10,305
Post-operative pouchitis (0.7 - 2%)	£10,323	£10,299
Colectomy cost	£10,321	£10,301
Serious infection utility reduction (0% - 3%)	£10,301	£10,321
Post-surgery complication utility weight reduction (0% - 40%)	£10,314	£10,308
Perioperative complications (No risk - double the risk)	£10,314	£10,309
OP administration cost (£70 - £161)	£10,311	£10,311

Table 84: Probability of being cost-effective - TNFi-exposed subgroup

Treatments	£20k per QALY WTP	£30k per QALY WTP
Tofacitinib	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]
CT	[redacted]	[redacted]

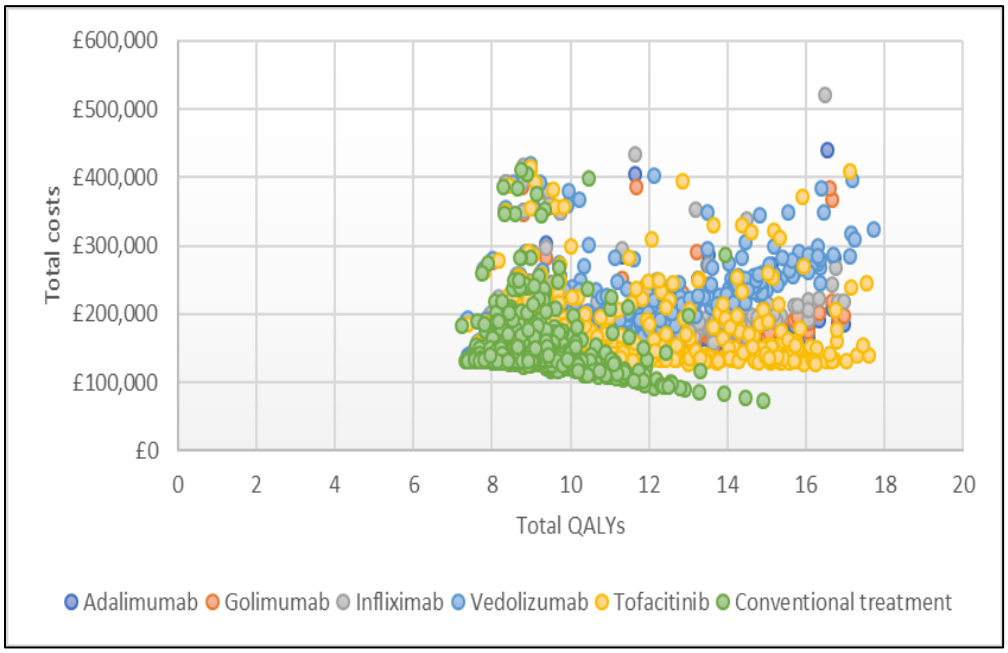


Figure 12 PSA scatter plot for TNFi-exposed subgroup

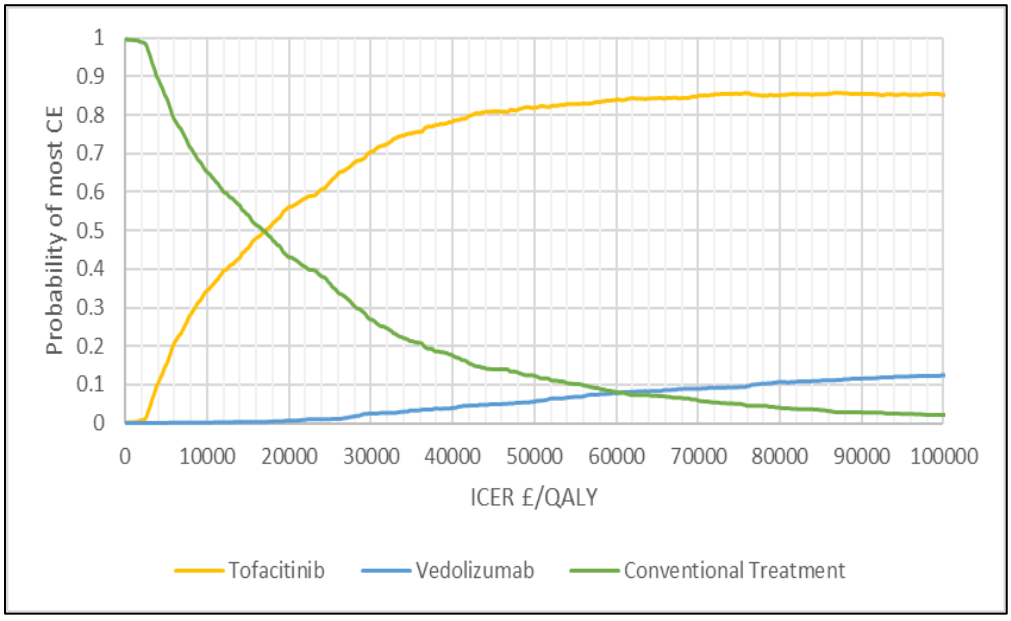


Figure 13 Cost effectiveness acceptability plane for TNFi-exposed subgroup

Table 85 Scenario analyses, company base case (ERG corrected) – TNFi-exposed

Scenarios	Assumption	ICER for Tofacitinib vs.	
		CT	**Vedolizumab**
Base case		**£10,311**	**£7,838,381**
Tofacitinib maintenance dose mix	[redacted] of patients receiving 5mg; [redacted] of patients receiving 10mg	£13,956	Tofacitinib dominant
Fixed baseline utility instead of age-adjusted	Assumption that patient quality of life stays constant over time.	£10,599	£6,502,288
OCTAVE trials utilities	EQ-5D data were collected in Tofacitinib Phase III clinical trials	£18,292	Tofacitinib dominant
Swinburn utilities	To compare with previous analyses	£14,501	£7,087,359
Emergency surgery from any state	Due to the uncertainty on the likely protection from acute events based on the level of response/remission, patients are assumed to undergo emergency surgery regardless of state membership	£9,971	£7,612,076
Emergency surgery from active UC only	As above but assuming response to treatment offers the same level of protection from acute events, as remission	£10,485	£6,780,235
No emergency surgery	As above, but assuming no emergency surgery in the model	£10,603	£6,781,118
Central read NMA	Central read was the primary endpoint in OCTAVE trials.	£10,798	Tofacitinib dominant
Discounting every cycle	It tested the sensitivity of the model when the discounting of outcomes is applied every 8 weeks.	£10,408	£8,260,662
Vedolizumab dose 300 mg Q4W	A 300 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients who have a body weight ≥ 80 kg	£10,311	Tofacitinib dominant

4.4.3 ERG additional analysis

Table 86 below summarises ERG assumptions and scenario analyses that we ran to further explore uncertainties over the model results.

Table 87 shows the cumulative impact of applying ERG preferred assumptions to the company base case model. The change that has the biggest impact is the use of alternative NMA models to populate the input parameters for serious infections. This is consistent with the company's observation based on their sensitivity analyses. Varying the age of the patients; using different NMA models for clinical response and remission and adding the costs of stoma care have little impact on the results. Collectively our preferred assumptions give very similar results to the company's model. TNF-inhibitors remain dominated (with higher costs and fewer QALYs) than tofacitinib in both the sub-groups across the range of assumptions tested. The pairwise ICERs for tofacitinib compared with vedolizumab mostly fall in the south-west quadrant (meaning tofacitinib is less effective but also less costly than vedolizumab), under our preferred set of assumptions vedolizumab is dominated by tofacitinib. However, we note again that these results do not take account of the PAS discount for vedolizumab. Final results including all PAS discounts are provided in the confidential addendum to this report.

We performed a range of additional scenario analyses on the ERG preferred base case, as specified in Table 86. Results are summarised in Table 88 and Table 89 below, with full cost-effectiveness results in Appendix 9.3. In the TNFi-naïve subgroup, Tofacitinib dominated the TNFi-agents (adalimumab, infliximumab and golimumab) across all the scenarios. The ICERs for Tofacitinib vs CT were most sensitive to sources for utilities and assumptions about health service use but remained below £20,000 per QALY for all scenarios. For Tofacitinib vs Vedolizumab, the ICERs moved between the south–east (indicating, tofacitinib was cheaper and more effective than vedolizumab) and south west quadrants (indicating, tofacitinib was cheaper and less effective compared with vedolizumab).

Similarly the TNFi-experienced subgroup, tofacitinib dominated TNFi- agents in all scenarios. The ICER for tofacitinib vs CT remained low, reaching a maximum of £21,376 per QALY with OCTAVE EQ-5D utility estimtes Tofacitinib dominated vedolizumab across all the scenarios, except in the the company's preferred NMA models for response and remission (favouring fixed effect models).

Table 86 ERG preferred assumptions and scenarios

Aspect of the model		Company base case	ERG preferred	ERG scenarios		Reason for analysis
Patients	Age (yrs)	TNFi-naïve: 41.5 TNFi-exposed: 40.9	Average of all patients in OCTAVE 1 and 2: 41	Range: 28-52		To explore the impact of patient characteristics on the cost-effectiveness results
	Weight (kgs)	TNFi-naïve: 74.6 TNFi-exposed: 72.5	Average for all patients in OCTAVE 1 and 2: 73.5	Range: 70-80 kg		
Comparator	TNFi-exposed	Excludes adalimumab	Include adalimumab			NMA results available for adaliumuab in TNFi-exposed group. Clinical advice that some patients would switch to another TNFi
	Treatment sequencing	The base case includes only 1st line and 2nd line treatments	No change	INF-ADA-CT INF-VED-CT INF-TOF-CT VED-ADA-CT TOF-ADA-CT	GOL-ADA-CT GOL-VED-CT GOL-TOF-CT ADA-VED-CT ADA-TOF-CT	To test effect of switching within or between classes and compare 'step-up' and 'step-down' strategies
NMA models	Remission and response rates	Use FE models except for TNFi-naive induction (FE better fit)	Use RE except for TNFi-experienced maintenance (RE would not run)	FE for both subgroups, induction and maintenance		ERG prefers RE models, given study heterogeneity
		Combined TNFi-failed for vedolizumab with TNFi-exposed for tofacitinib and adalimumab	No change	Use TNFi-failed for both vedolizumab and tofacitinib with TNFi-experienced for adalimumab		To provide a more like-for-like comparison between tofacitinib and vedolizumab - main competitors.
	Serious infections	Bayesian random effect model	Frequentist random effects NMA model	Bayesian random effect model		Due to rarity and null events credible intervals for Bayesian NMA are implausibly wide.
Utilities	Sources for pre and post-surgery health states	Background age/gender specific general population EQ-5D for remission. Utility multipliers for other heallth states from Woehl et al. 2008	Same as company	• Swinburn et al. • OCTAVE 8 weeks • OCTAVE 52 weeks		Woehl et al. used in previous TAs. For scenario analysis, we use results analysis of EQ-5D data from OCTAVE provided in company clarification response

Resource use and costs	Drug stopping rule	8 weekly loss of response or surgery	Same as company	Additional OP visits to assess response within 8 weeks	Include costs required to allow rapid assessment and change of therapy following exacerbations
	Conventional drug usage	Estimated from RCP IBD audit 2016	Same as company	Patient use of mesalazine: 50.3% (CT), 46.2% (concurrent). No other aminoslicylates	Clinical expert advice
	Health state resource use	Based on Tsai et al. plus additional admissions	Same as company	Reduced admissions, outpatient follow up and endoscopy	To reflect advice on current NHS clinical practice
	Drug administration costs	OP visit for IV infusion (infliximab, vedolizumab) No administration cost for self-administered subcutaneous injections (golimumab, adalimumab)	Same as company	Assume 1 OP visit at start of treatment for training on subcutaneous injections	Company states that support for self-administration of injections is provided by manufacturers. But this may not always be available in NHS.
	Hospitalisation and surgery costs	NHS Reference costs 2016-17 for colectomy procedures	NHS Reference costs + cost of stoma care post-surgery (Buchanan et al. uprated for inflation)	Buchannan et al. estimate of surgery cost (uprated to 2016/17 prices) – includes repeat procedures	Stoma costs To align with previous TA 342
Surgery	Incidence rate	Misra et al. (UK HES Data)	Same as company	Chhaya et al.	Exploratory analyses
	Complications	IBD audit	Same as company	Tappenden et al.: Probabiity of perioperative complications (elective 0.2386; emergency 0.2614), probability of post-surgery complications (0.173)	To align with previous TA 342

Table 87 Cumulative effect of ERG preferred assumptions

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
Company base case (ERG corrected)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£8,564
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	£615,077 (SW)
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£10,311
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	£7,838,381 (SW)
Average age: 41 years	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£8,562
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	-
	Vedolizumab	[redacted]	[redacted]	£614,916 (SW)
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£10,304
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	-
	Vedolizumab	[redacted]	[redacted]	£7,798,892 (SW)
+ ERG preferred NMAs for remission and response	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£8,584
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	£590,046 (SW)
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£10,148
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
+ Frequentist NMA for serious infections	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,886
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£607,642 (SW)
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,458
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant
+ Cost of stoma-care = ERG preferred analysis	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,815
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	£607,571 (SW)
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,389
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

Table 88 Scenario analyses, ERG base case (Tofacitinib PAS) – TNFi-naive subgroup

Scenarios	ICER for tofacitinib versus	
	CT	Vedolizumab
ERG preferred base case	**£7.815**	**£607,571**
Age: 28 years	£7,644	£589,024
Age: 52 years	£8,019	£628,794
Weight: 70 kg	£7,827	£607,395
Weight: 80 kg	£7,819	£607,504
NMA: FE for response and remission	£7,793	£633,458
NMA: TNFi-failed (Ved) + TNFi-exp (tof and ada)	Not relevant	Not relevant
NMA: FE for Serious Infections	£8,513	£589,976
Utility: Swinburn et al.	£10,898	£845,865
Utility: OCTAVE 8 weeks	£17,764	£1,360,239
Utility: OCTAVE 52 weeks	£18,256	£1,373,067
Drug stopping: 6.5 OP visits for all patients in maintenance	£9,090	£608,793
Reduced health state resource use (clinical scenario)	£13,938	£613,289
CT drug usage	£7,827	£607,576
Drug admin cost for subcutaneous injection	£7,815	£607,571
Stoma care costs (£81.66 per cycle based on TA342)	£7,804	£607,561
Surgery costs (based on Buchannan et al.)	£7,764	£607,522
Surgery incidence rate (based on Chhaya et al.)	£7,980	£611,440
Surgery complications (based on Tappenden et al.)	£7,556	£605,226
Treatment sequencing	£13,951	£614,361

Table 89 Scenario analyses, ERG base case (Tofacitinib PAS) – prior TNFi experience

Scenarios	ICER for tofacitinib versus	
	CT	Vedolizumab
ERG preferred base case	£9,389	Tofacitinib dominant
Age: 28 years	£9,170	Tofacitinib dominant
Age: 52 years	£9,648	Tofacitinib dominant
Weight: 70 kg	£9,401	Tofacitinib dominant
Weight: 80 kg	£9,394	Tofacitinib dominant
NMA: FE for response and remission	£9,541	£8,801,245
NMA: TNFi-failed (Ved) + TNFi-exp (tof and ada)	£9,669	£2,521,513
NMA: FE for Serious Infections	£10,080	Tofacitinib dominant
Utility: Swinburn et al.	£13,198	Tofacitinib dominant
Utility: OCTAVE 8 weeks	£21,376	Tofacitinib dominant
Utility: OCTAVE 52 weeks	£21,283	Tofacitinib dominant
Drug stopping: 6.5 OP visits for all patients in maintenance	£10,597	Tofacitinib dominant
Reduced health state resource use (clinical scenario)	£14,950	Tofacitinib dominant
CT drug usage	£9,402	Tofacitinib dominant
Drug admin cost for subcutaneous injection	£9,389	Tofacitinib dominant
Stoma care costs (£81.66 per cycle based on TA342)	£9,379	Tofacitinib dominant
Surgery costs (based on Buchannan et al.)	£9,341	Tofacitinib dominant
Surgery incidence rate (based on Chhaya et al.)	£9,558	Tofacitinib dominant
Surgery complications (based on Tappenden et al.)	£9,134	Tofacitinib dominant
Treatment sequencing	£9,389	Tof-Ada-CT dominant

5 End of life

The NICE end of life treatment criteria are not applicable and are not included in the CS.

6 Innovation

The CS highlights six aspects of tofacitinib therapy for moderately to severely active ulcerative colitis in making the case for innovation (CS B.2.12). These six aspects are:

- Tofacitinib is the first in a new class of treatments and has a novel mechanism of action (inhibitor of JAKs).
- Tofacitinib is an oral therapy in contrast to the available biologic therapies for people with moderately to severely active ulcerative colitis which are administered either as infusion or by subcutaneous injection.
- Tofacitinib is a small synthetic molecule which means the formation of anti-drug antibodies (which reduce the efficacy of large protein biologics such as the TNF-alpha inhibitors) is not likely to occur, the risk of immunogenicity is reduced, and therapeutic drug monitoring is not required.
- Tofacitinib is a monotherapy, which would be expected to have a more favourable safety profile than combination therapies of a biologic therapy plus immunomodulatory agent. (NB the ERG notes that in the company's safety NMA, tofacitinib had the second-highest probability of serious adverse events after placebo (section 3.3.10.2).
- Tofacitinib treatment may be interrupted without the expectation of a reduced response
- Tofacitinib has a rapid onset of action.

7 DISCUSSION

7.1 Summary of clinical effectiveness issues

Choice of model fit for clinical response and clinical remission NMAs

Heterogeneity was present among the studies included in the NMA. Heterogeneity was due to differences in the designs of the included studies and differences between the baseline characteristics of the patients recruited to the trials included in the NMAs. In some cases the fit of the fixed-effect and random-effects models were comparable in terms of fit and the company chose the fixed-effects model in these circumstances. The ERG would have chosen the

random effects model to account for between study heterogeneity and provide a more conservative analysis.

Choice of model fit for serious infection NMA

Serious infections NMA under both random-effects (company choice) and fixed-effect (ERG alternative) resulted in very wide credible intervals. There are two issues with the available tofacitinib data on serious infections. Firstly the number of serious infections that occurred in the Phase II tofacitinib trial is higher than for the other OCTAVE trials. Secondly, in the Phase II trial and both of the OCTAVE Induction trials there were no serious infection events in the placebo arms. The ERG therefore ran an alternative NMA using a frequentist framework that allows for a value of 0.5 to be added to zero cells. Whilst adding a value to a zero cell is controversial this analysis does not adversely impact the confidence intervals for tofacitinib on account of the absence of serious infections among any of the placebo arms in the OCTAVE trials programme.

Absence of maintenance phase safety NMAs

No NMA for safety outcomes was conducted for the maintenance phase. The ERG believe this could have been achieved by using the mFAS population of OCTAVE Sustain. Whilst the use of mFAS would have aligned the re-randomised studies these would still have to be combined with data from the studies with a treat-through design and hence would only have been a partial solution.

No exploration of correction for different durations of induction and maintenance phases or differences between studies with a re-randomisation design.

Not all studies included in the NMAs had the same induction and maintenance phase durations as the OCTAVE tofacitinib studies. In particular the studies of golimumab and vedolizumab had a shorter induction phase (6 weeks versus OCTAVE studies 8 weeks) and the maintenance phases of the adalimumab (44 weeks), infliximab (46 weeks) and vedolizumab (46 weeks) studies were shorter than those of the tofacitinib (52 weeks) and golimumab studies (54 weeks). It is possible that there could be a bias against studies with a shorter induction period (if a higher response could be possible if measured at 8 weeks instead of 6 weeks). If this were the case this would bias against golimumab and vedolizumab in the induction phase. Similarly it is possible that there could be a bias in favour of studies with a shorter maintenance period (if

fewer responders lose response in the shorter time frame). If this were the case the bias would work against tofacitinib which has one of the longer maintenance phases.

7.2 Summary of cost effectiveness issues

Baseline characteristics of patient population included in the economic model

For their base case analyses, patient characteristics (including initial age, weight and gender mix) for the two sub-groups of TNFi-naïve and TNFI-exposed are based on means from the tofacitinib arms in the OCTAVE Induction trials. We view that these baseline characteristics should be assumed similar for people with and without prior exposure to TNFi drugs. We explore this in our additional analyses.

Analysis for the whole population: ITT NMA

The company has conducted an ITT NMA for the whole population and performed a cost-effectiveness analysis with the ITT population. We do not consider this scenario to be reliable because of the high level of uncertainty in the underlying NMA. The scenario also omits relevant comparators (the TNFi drugs), so does not address the specified decision problem. The ERG, therefore, focuses on separate analyses for the two TNFi exposure subgroups which is consistent with committee considerations in the NICE appraisal of vedolizumab (TA342).

Comparator

- *Exclusion of adalimumab in TNFi-exposed sub group*

The company excludes adalimumab, infliximab and golimumab as comparators for patients with prior exposure to a TNFi. Whilst clinical response and remission rates are not available for infliximab or golimumab in this sub group, but they are available for adalimumab. Further, the occurrence of in-class switching is also supported by evidence from the UK IBD Audit: 21% of patients starting adalimumab (17/83) had previously not responded or been intolerant to a TNFi. So, the ERG considers adalimumab as a relevant comparator for at least some patients with prior exposure to a TNFi agent. We therefore include adalimumab in ERG analysis for this subgroup. However, we understand that further treatment with a TNFi may not be appropriate for all patients in this subgroup.

- *Conventional therapy*

The company assumes equal use of 4 drugs in aminosalicylate class (balsalazide, mesalazine

olsalazine & sulfasalazine). However, clinical advice to ERG is that most patients receive mesalazine in UK and the doses for active ulcerative colitis are potentially higher than specified in company base case. We view that the net effect on costs from incorporating the changes in base case is likely to be neutral.

Treatment waning of effects and discontinuation

The company assumes treatment effect to be maintained with ongoing treatment and non-responders are given conventional therapy as second-line. The ERG agrees with company's approach to allow discontinuation for failure to respond in induction or loss of response in maintenance, based on the independent economic analysis in NICE TA329. We note this assumes that in practice, patients who experience exacerbations of symptoms can be assessed and, if appropriate, treatment stopped within 8 weeks. However, the model does not reflect NICE recommendations for annual assessment of benefit and need for continued treatment in previous appraisals TA329 and TA342. Clinical advice suggests that withdrawal of treatment for patients in remission is unlikely in practice, and the effects of this are difficult to quantify given the model structure and limited evidence over long-term maintenance of remission.

The company model applies a constant risk of relapse across each 8-week cycle of maintenance, with treatment stopping immediately when patients lose response. Thus, it assumes that maintenance treatment is stopped within 8 weeks of a loss of response. We consider this assumption to reflect UK practice. However, we have concerns that the costs of monitoring and follow-up in the company's model do not reflect the full cost of ensuring that treatment can be withdrawn within 8 weeks of a relapse. We address this by considering additional costs for outpatient visits to enable treatment cessation within 8 weeks of a relapse in our additional analyses.

Source of clinical effectiveness estimates

- *Choice of NMA models for economic analysis*

In general, we agree with company's approach to use locally-read clinical definitions of response and remission in economic model. Whilst, they state that their choice of NMA models was based on DIC measures of model fit, but they preferred the simpler fixed effect approach when DIC statistics were similar. The ERG has a general preference for the random effect NMA models, as we believe that the fixed effect models may underestimate uncertainty due to

heterogeneity between the studies. We test the impact of different NMA models on cost-effectiveness results in our additional analyses.

- *Combination of TNFi-failed and TNFi-exposed subgroups*

The base case NMAs combine outcomes for subgroups defined as TNFi-failed for vedolizumab with TNFi-exposed subgroups for tofacitinib and adalimumab. We consider that combining results for TNFi-failed and TNFi-exposed subgroups is a potential source of bias in favour of tofacitinib. We conduct a scenario analysis using a more like-for-like comparison between tofacitinib and vedolizumab, using data for the TNFi-failed subgroups from the OCTAVE and GEMINI trials.

- *Transformation of NMA results to transition probabilities*

The company transformed the results of the clinical response/remission NMAs from the probit scale to the natural scale and converted to absolute probabilities for use in the model. They take a simpler approach by assuming constant ratio of patients in remission and response throughout maintenance phase and beyond in extrapolation. Clinical advice to the ERG is that these assumptions might not be realistic as clinical -experience indicates the risk is greatest in the first 6-12 months; and falls thereafter. The proportion of patients with response and remission is likely to increase over time as per our clinical advice. This is because responders (without remission) are more likely to stop or switch therapy whereas those in remission would continue with treatment. However, in the absence of evidence it is difficult to adapt the model. Therefore, we conclude that the model assumption of constant risk of loss of response for patients on maintenance treatment does not reflect clinical experience. Extrapolation of relapse and discontinuation rates from the maintenance trials is likely to underestimate the average duration of treatment and hence both the costs and QALYs of active treatments. However, it is not possible to estimate the net direction of bias in ICERs between comparators, because trends in long-term risks may vary between TNFi drugs, vedolizumab and tofacitinib.

- *Exclusion of other serious adverse events*

The company excluded adverse events other than serious infections We agree that there would have been a risk of double-counting the costs and effects of ulcerative colitis exacerbations had all SAEs had been included in the model. Although, the omission of non-infection SAEs does introduce a risk of bias but given the frequency of these events this is unlikely to change the cost-effectiveness results.

- *NMA method for serious infections*

The company applied a binomial logit NMA model to estimate the risk of serious infections in the induction trials and chose the random effects model for their base case. Whilst the ERG agrees that there is considerable uncertainty associated with the risk of serious infections, we have reservations about the company's approach to estimating this parameter. Our verification checks indicated an even higher level of uncertainty around tofacitinib estimates, and we were unable to replicate the company's base case NMA values. We therefore applied a frequentist NMA approach to estimate the risk of serious infection, which we use as a scenario in ERG analysis

- All-cause mortality

The model adjusted mortality risks for age and gender mix for the general population and applied these to patients in pre-and post-surgery states. They assumed that, except for perioperative deaths, ulcerative colitis and treatment do not influence mortality. In general, we view this approach as reasonable, although there are additional mortality risks not reflected in the model – e.g. for colorectal cancer –although the relative risk estimates are likely to include perioperative deaths already accounted for.

Health Related Quality of life

The company's simple and regression-based analyses of EQ-5D data from the OCTAVE trials are problematic as sources of utility parameters for the economic model. They are relevant to the decision problem and clinical evidence, but the re-randomisation design and lack of intermediate assessments of clinical response and remission between week 8 and week 52 complicate the interpretation of results. We therefore agree with the company that the utility estimates by Woehl et al.[71] provide a more appropriate source for base case parameters that are consistent with previous NICE appraisals for ulcerative colitis. We use these estimates in ERG preferred analyses, but also test scenarios based on the company's OCTAVE analyses and published sources (Swinburn et al.).[72]

Resource use and costs

- *Drug acquisition*

We consider the drug acquisition costs used in the company model to be realistic, although there have been some small changes in NHS prices for included drugs; sulfasalazine (£7.83),

prednisolone (£0.47) and azathioprine (£2.20) (MIMS June 2018). These changes result in a very small reduction in the estimated cost of CT alone (£58.02), with biologic drugs (£51.68) and with tofacitinib (£48.86).

- *Drug administration*

Adalimumab and golimumab are administered by subcutaneous injection. The company assumed at zero cost to the NHS for self-administering these drugs. So, we conduct additional scenario analysis to assess the impact of assuming an initiation of self-administration of subcutaneous injections by adding the cost of a non-consultant led clinic attendance (£107) to the cost of induction for adalimumab and golimumab in our additional analyses.

- *Monitoring and follow up*

The company made health care usage assumptions from Tsai et al. (2008) which are consistent with health state definitions in the model and with previous NICE appraisals for ulcerative colitis (TA329 and TA342). We agree with the use of Tsai et al. as base case. But clinical advice to ERG suggests frequency of outpatient visits and endoscopy exceed current UK practice and additional assumptions about hospital episodes are unrealistic. We test alternative resource use scenario in our additional analyses.

We also question whether the assumption that maintenance treatment will always stop within 8 weeks of a loss of response is consistent with the number of outpatient appointments. To explore this, we conduct two scenario analyses to align the costs of assessing patients on maintenance treatment with the model assumption that treatment will always be discontinued within 8 weeks of a relapse.

The company excludes cost of stoma care and the estimated cost of surgery is low compared with previous appraisals. We test the inclusion of stoma care costs and higher surgery costs in our additional analysis.

8 REFERENCES

1. Dignass A, Eliakim R, Magro F, et al. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis. Part 1: Definitions and diagnosis. *J Crohns Colitis* 2012;6:965-90.
2. Ungaro R, Mehandru S, Allen PB, et al. Ulcerative colitis. *Lancet* 2017;389:1756-70.
3. Mowat C, Cole A, Windsor A, et al. Guidelines for the management of inflammatory bowel disease in adults. *Gut* 2011;60(5):571-607.
4. Magro F, Gionchetti P, Eliakim R, et al. Third European evidence-based consensus on diagnosis and management of ulcerative colitis. Part 1: Definitions, diagnosis, extra-intestinal manifestations, pregnancy, cancer surveillance, surgery, and ileo-anal pouch disorders. *J Crohns Colitis* 2017;11(6):649-70.
5. Harbord M, Eliakim R, Bettenworth D, et al. Third European evidence-based consensus on diagnosis and management of ulcerative colitis. Part 2: Current management. *J Crohns Colitis* 2017;11:769-84.
6. NICE (National Institute for Health and Care Excellence). Ulcerative colitis: management. Clinical guideline cg166: NICE, 2013:1-36.
7. Burisch J, Jess T, Martinato M, et al. The burden of inflammatory bowel disease in Europe. *J Crohns Colitis* 2013;7(4):322-27.
8. NICE (National Institute for Health and Care Excellence). Infliximab, adalimumab and golimumanb for treating moderately to severely active ulcerative colitis after the failure of conventional therapy. NICE Guidance TA329. London: NICE, 2015.
9. NICE (National Institute for Health and Care Excellence). Vedolizumab for treating moderately to severely active ulcerative colitis. NICE Guidance TA342. London: NICE, 2015.
10. EMA (European Medicines Agency). Summary of Product Characteristics. Xeljanz (tofacitinib) 2018:1-61.
11. Sandborn WJ, Ghosh S, Panes J, et al. Tofacitinib, an oral Janus Kinase inhibitor in active ulcertaive colitis. *NEJM* 2012;367:616-24.
12. EMA (European Medicines Agency). Guideline on the development of new medicinal products for the treatment of ulcerative colitis [draft version]: Committee for Medicinal Products for Human Use (CHMP), 2016:1-16.
13. Mukherjee A, Hazra A, Smith MK, et al. Exposure-response characterization of tofacitinib efficacy in moderate to severe ulcerative colitis: Results from a dose-ranging phase 2 trial. *Br J Clin Pharmacol* 2018;84(6):1136-45.
14. Bonovas S, Lytras T, Nikolopoulos G, et al. Systematic review with network meta-analysis: comparative assessment of tofacitinib and biological therapies for moderate-to-severe ulcerative colitis. *Aliment Pharmacol Ther* 2018;47(4):454-65.
15. Panés J, Vermeire S, Lindsay JO, et al. Tofacitinib in patients with ulcerative colitis: Health-related quality of life in phase 3 randomised controlled induction and maintenance studies. *J Crohns Colitis* 2018;12(2):145-56.
16. Motoya S, Watanabe M, Kim HJ, et al. Tofacitinib induction and maintenance therapy in East Asian patients with active ulcerative colitis: Subgroup analyses from three phase 3 multinational studies. *Intest Res* 2018;16(2):233-45.
17. Sands BE, Moss AC, Armuzzi A, et al. Sa1761 - Efficacy and safety of dose escalation to tofacitinib 10 mg bid for patients with ulcerative colitis following loss of response on tofacitinib 5 mg bid maintenance therapy: Results from OCTAVE Open. *Gastroenterology* 2018;154 (6) Suppl 1:S-385.

18. Sands BE, Taub PR, Feagan BG, et al. Sa1750 - The effect of tofacitinib on serum lipids and cardiovascular safety in patients with ulcerative colitis - results from the tofacitinib ulcerative colitis clinical program. *Gastroenterology* 2018;154 (6) Suppl 1:S-380.
19. Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as induction and maintenance therapy for ulcerative colitis. *NEJM* 2017;376:1723-36.
20. Pfizer Inc. Interim Clinical Study Report: A Multi-Center, Open-Label Study of CP-690,550 in Subjects With Moderate to Severe Ulcerative Colitis: OCTAVE Open 2016
21. Danese S, Fiocchi C. Ulcerative colitis. *NEJM* 2011;365:1713-25.
22. FDA (Food and Drug Administration). Ulcerative colitis: Clinical trial endpoints. Guidance for industry [draft version]: US Food and Drug Administration, Center for Drug Evaluation and Reserach (CDER), 2016:1-19.
23. Mazzuoli S, Guglielmi FW, Antonelli E, et al. Definition and evaluation of mucosal healing in clinical practice. *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver* 2013;45(12):969-77. doi: 10.1016/j.dld.2013.06.010 [published Online First: 2013/08/13]
24. Khan HM, Mehmood F, Khan N. Optimal management of steroid-dependent ulcerative colitis. *Clinical and experimental gastroenterology* 2015;8:293-302. doi: 10.2147/ceg.s57248 [published Online First: 2015/12/10]
25. Archer R, Tappenden P, Ren S, et al. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262): clinical effectiveness systematic review and economic model. *Health Technology Assessment* 2016;20(39)
26. Ferrante M, Declerck S, De Hertogh G, et al. Outcome after proctocolectomy with ileal pouch-anal anastomosis for ulcerative colitis. *Inflamm Bowel Dis* 2007;14(1):20-28.
27. Feagan BG, Patel H, Colombel JF, et al. Effects of vedolizumab on health-related quality of life in patients with ulcerative colitis: results from the randomised GEMINI 1 trial. *Alimentary Pharmacology & Therapeutics* 2017;45(2):264-75. doi: https://dx.doi.org/10.1111/apt.13852
28. LeBlanc K, Mosli MH, Parker CE, et al. The impact of biological interventions for ulcerative colitis on health-related quality of life. *The Cochrane database of systematic reviews* 2015(9):Cd008655. doi: 10.1002/14651858.CD008655.pub3 [published Online First: 2015/09/24]
29. Stark RG, Reitmeir P, Leidl R, et al. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. *Inflamm Bowel Dis* 2010;16(1):42-51. doi: 10.1002/ibd.20989 [published Online First: 2009/05/29]
30. Yarlas A, Bayliss M, Cappelleri JC, et al. Psychometric validation of the SF-36((R)) Health Survey in ulcerative colitis: results from a systematic literature review. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2018;27(2):273-90. doi: 10.1007/s11136-017-1690-6 [published Online First: 2017/08/30]
31. Ware J E .Jr, Kosinski M, Bjorner JB, et al. User's manual for the SF-36v2 Health Survey. 2nd ed. Lincoln, RI: QualityMetric Incorporated 2007.
32. Reilly M C, Zbrozek A S, Dukes E M. The validity and reproducibility of a work productivity and activity impairment instrument. *PharmacoEconomics* 1993;4(5):353.
33. Travis SP, Feagan BG, Peyrin-Biroulet L, et al. Effect of adalimumab dose escalation on clinical, healthrelated quality of life, treatment satisfaction and work productivity outcomes among patients with ulcerative colitis in a clinical practice setting: Results from inspirada. *Gastroenterology* 2017;152 (5 Supplement 1):S742-S43.
34. Binion DG, Louis E, Oldenburg B, et al. Effect of adalimumab on work productivity and indirect costs in moderate to severe Crohn's disease: a meta-analysis. *Canadian journal*

of gastroenterology = Journal canadien de gastroenterologie 2011;25(9):492-6. [published Online First: 2011/09/14]
35. Mshimesh BR. Efficacy and safety of adalimumab versus infliximab in patients suffered from moderate to severe active ulcerative colitis. *Asian Journal of Pharmaceutical and Clinical Research* 2017;10(3):300-07.
36. Panaccione R, Ghosh S, Middleton S, et al. Combination therapy with infliximab and azathioprine is superior to monotherapy with either agent in ulcerative colitis. *Gastroenterology* 2014;146(2):392-400.e3. doi: 10.1053/j.gastro.2013.10.052 [published Online First: 2014/02/12]
37. Suzuki Y, Motoya S, Hanai H, et al. Efficacy and safety of adalimumab in Japanese patients with moderately to severely active ulcerative colitis. *Journal of gastroenterology* 2014;49(2):283-94. doi: 10.1007/s00535-013-0922-y [published Online First: 2013/12/24]
38. Rutgeerts P, Sandborn WJ, Feagan BG, et al. Infliximab for induction and maintenance therapy for ulcerative colitis. *The New England journal of medicine* 2005;353(23):2462-76. doi: 10.1056/NEJMoa050516 [published Online First: 2005/12/13]
39. Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab maintains clinical response in patients with moderate-to-severe ulcerative colitis. *Gastroenterology* 2014;146(1):96-109.e1. doi: 10.1053/j.gastro.2013.06.010 [published Online First: 2013/06/19]
40. Hibi T, Imai Y, Senoo A, et al. Efficacy and safety of golimumab 52-week maintenance therapy in Japanese patients with moderate to severely active ulcerative colitis: a phase 3, double-blind, randomized, placebo-controlled study-(PURSUIT-J study). *Journal of gastroenterology* 2017;52(10):1101-11. doi: 10.1007/s00535-017-1326-1 [published Online First: 2017/03/23]
41. Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *The New England journal of medicine* 2013;369(8):699-710. doi: 10.1056/NEJMoa1215734 [published Online First: 2013/08/24]
42. Jiang XL, Cui HF, Gao J, et al. Low-dose Infliximab for Induction and Maintenance Treatment in Chinese Patients With Moderate to Severe Active Ulcerative Colitis. *Journal of clinical gastroenterology* 2015;49(7):582-8. doi: 10.1097/mcg.0000000000000319 [published Online First: 2015/04/07]
43. Kobayashi T, Suzuki Y, Motoya S, et al. First trough level of infliximab at week 2 predicts future outcomes of induction therapy in ulcerative colitis-results from a multicenter prospective randomized controlled trial and its post hoc analysis. *Journal of gastroenterology* 2016;51(3):241-51. doi: 10.1007/s00535-015-1102-z [published Online First: 2015/07/15]
44. Rodgers M, Epstein D, Bojke L, et al. Etanercept, infliximab and adalimumab for the treatment of psoriatic arthritis: a systematic review and economic evaluation. *Health technology assessment (Winchester, England)* 2011;15(10):i-xxi, 1-329. doi: 10.3310/hta15100 [published Online First: 2011/02/22]
45. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in medicine* 2004;23(9):1351-75. doi: 10.1002/sim.1761 [published Online First: 2004/04/30]
46. Dias S, Welton NJ, Sutton AJ, et al. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework For Pairwise And Network Meta-Analysis Of Randomised Controlled Trials. Report by the Decision Support Unit: University of Sheffield, 2011 (last updated 2016):1-98.
47. Thorlund K, Druyts E, Toor K, et al. Incorporating alternative design clinical trials in network meta-analyses. *Clinical epidemiology* 2015;7:29-35. doi: 10.2147/clep.s70853 [published Online First: 2015/01/08]

48. Takeda. Vedolizumab (Entyvio ®) for the treatment of adults with moderate to severe active ulcerative colitis. Manufacturer submission to NICE TA342., 2014.
49. Thorlund K, Druyts E, Toor K, et al. Comparative efficacy of golimumab, infliximab, and adalimumab for moderately to severely active ulcerative colitis: a network meta-analysis accounting for differences in trial designs. *Expert review of gastroenterology & hepatology* 2015;9(5):693-700. doi: 10.1586/17474124.2015.1024657 [published Online First: 2015/03/13]
50. Sandborn WJ, van Assche G, Reinisch W, et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. *Gastroenterology* 2012;142(2):257-65.e1-3. doi: 10.1053/j.gastro.2011.10.032 [published Online First: 2011/11/09]
51. NICE (National Institute for Health and Care Excellence). Guide to the methods of technology appraisal London, 2013.
52. Buckland A, Bodger K. The cost-utility of high dose oral mesalazine for moderately active ulcerative colitis. *Alimentary pharmacology & therapeutics* 2008;28(11-12):1287-96.
53. Essat M, Tappenden P, Ren S, et al. Vedolizumab for the treatment of adults with moderate-to-severe active ulcerative colitis: an evidence review group perspective of a NICE single technology appraisal. *Pharmacoeconomics* 2016;34(3):245-57.
54. Tappenden P, Ren S, Archer R, et al. A model-based economic evaluation of biologic and non-biologic options for the treatment of adults with moderately-to-severely active ulcerative colitis after the failure of conventional therapy. *Pharmacoeconomics* 2016;34(10):1023-38.
55. Tsai H, Punekar Y, Morris J, et al. A model of the long-term cost effectiveness of scheduled maintenance treatment with infliximab for moderate-to-severe ulcerative colitis. *Alimentary pharmacology & therapeutics* 2008;28(10):1230-39.
56. Wilson MR, Bergman A, Chevrou-Severac H, et al. Cost-effectiveness of vedolizumab compared with infliximab, adalimumab, and golimumab in patients with ulcerative colitis in the United Kingdom. *Eur J Health Econ* 2018;19(2):229-40. doi: https://dx.doi.org/10.1007/s10198-017-0879-5
57. Ali T, Skup M, Yang M, et al. Cost-Effectiveness of adalimumab in moderately to severely active ulcerative colitis. *American Journal of Gastroenterology* 2012;107:S641.
58. Mukhekat V, Ahmad I, Merali Y, et al. PTU-078 The Application Of A Markov Model For Evauating Asa Therapy For Ulcerative Colitis. *Gut* 2014;63(Suppl 1):A73-A73.
59. Wilson M, Kerrigan M, Smyth M, et al. Cost-effectiveness of vedolizumab compared with conventional therapy for treatment of moderately-to-severely active ulcerative colitis in the United Kingdom. *Gut* 2016;65:A260-A60.
60. Wilson M, Mody R, Ursan I, et al. Sa1019 Cost-Effectiveness of Vedolizumab Compared With Conventional Therapy and Biologics for Treatment of Moderately to Severely Active Ulcerative Colitis in the United States. *Gastroenterology* 2015;148(4):S-198-S-99.
61. Yang H, Curry A, Wang S, et al. Cost-effectiveness of adalimumab in moderately to severely active ulcerative colitis. *United European Gastroenterology Journal* 2014;2:A374.
62. Wilson MR, Azzabi Zouraq I, Chevrou-Severac H, et al. Cost-effectiveness of vedolizumab compared with conventional therapy for ulcerative colitis patients in the UK. *ClinicoEcon* 2017;9:641-52. doi: https://dx.doi.org/10.2147/CEOR.S135609
63. Wu B, Wang Z, Zhang Q. Cost-Effectiveness of Different Strategies for the Treatment of Moderate-to-Severe Ulcerative Colitis. *Inflamm Bowel Dis* 2018;27:27. doi: https://dx.doi.org/10.1093/ibd/izy114
64. Royal College of Physicians (RCP). National clinical audit of biological therapies: UK inflammatory bowel disease (IBD) audit. Annual Report September 2016, 2016.
65. Royal College of Physicians (RCP). National clinical audit of biological therapies: UK inflammatory bowel disease (IBD) audit. Adult report. London, 2015.

66. Misra R, Askari A, Faiz O, et al. Colectomy rates for ulcerative colitis differ between ethnic groups: results from a 15-year nationwide cohort study. *Canadian Journal of Gastroenterology and Hepatology* 2016;2016
67. Chhaya V, Saxena S, Cecil E, et al. The impact of timing and duration of thiopurine treatment on colectomy in ulcerative colitis: a national population-based study of incident cases between 1989–2009. *Alimentary pharmacology & therapeutics* 2015;41(1):87-98.
68. Frolkis AD, Dykeman J, Negrón ME, et al. Risk of surgery for inflammatory bowel diseases has decreased over time: a systematic review and meta-analysis of population-based studies. *Gastroenterology* 2013;145(5):996-1006.
69. Solberg IC, Lygren I, Jahnsen J, et al. Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (IBSEN Study). *Scandinavian journal of gastroenterology* 2009;44(4):431-40.
70. Arai K, Koganei K, Kimura H, et al. Incidence and outcome of complications following restorative proctocolectomy. *The American journal of surgery* 2005;190(1):39-42.
71. Woehl A, Hawthorne A, McEwan P. The relation between disease activity, quality of life and health utility in patients with ulcerative colitis. *Gut* 2008;57:A153-A53.
72. Swinburn P, Elwick H, Bean K, et al. PTU-127 The impact of surgery on health related quality of life in ulcerative colitis. *Gut* 2012;61(Suppl 2):A237-A37.
73. Ara R, Brazier J. Populating an economic model with health state utility values: moving toward better practice. *Value in Health* 2010;13(5):509-18.
74. Diamantopoulos A, Finckh A, Huizinga T, et al. Tocilizumab in the treatment of rheumatoid arthritis: a cost-effectiveness analysis in the UK. *Pharmacoeconomics* 2014;32(8):775-87.
75. Kosmas C, Kerr C, Marsh S, et al. Health Related Quality of Life (Hrqol) In Patients Who Have Undergone Colectomy for Ulcerative Colitis: Impacts of Complications Post-Surgery. *Value in Health* 2015;18(7):A630.
76. NICE (National Institute for Health and Care Excellence). Guide to the methods of technology appraisal 2013. *NICE, London* 2013
77. Hatswell AJ, Porter J, Lee D, et al. The cost of costing treatments incorrectly: errors in the application of drug prices in economic evaluation due to failing to account for the distribution of patient weight. *Value in Health* 2016;19(8):1055-58.
78. Buchanan J, Wordsworth S, Ahmad T, et al. Managing the long term care of inflammatory bowel disease patients: The cost to European health care providers. *Journal of Crohn's and Colitis* 2011;5(4):301-16.

9 APPENDICES

Appendix 1 Additional results tables for subgroup analyses by TNFi-exposure status

Results according to subgroups by TNFi-exposure status for the outcomes of remission (primary outcome), mucosal healing, and Sustained corticosteroid-free remission among patients in remission at baseline are presented below in Table 90 to Table 94.

Remission

Table 90 Proportion of patients in remission in OCTAVE Induction 1 and 2 according to prior TNFi treatment (FAS, NRI, central and local reads)

Subgroup: Prior-TNFi treatment	Tofacitinib 10 mg	Placebo	Difference (95% CI); p-value	p-value for heterogeneity
	n/N (%)	n/N (%)		
OCTAVE 1, week 8				
TNFi-naïve Central read	56/222 (25.2)	9/57 (15.8)	9.4 (−1.6, 20.5); p=0.1328	0.1034
TNFi-exposed Central read	32/254 (12.6)	1/65 (1.5)	11.1 (6.0, 16.1); p=0.0090	
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed Local read	[redacted]	[redacted]	[redacted]	
OCTAVE 2, week 8				
TNFi-naïve Central read	43/195 (22.1)	4/47 (8.5)	13.5 (3.7, 23.4); p=0.0352	0.0956
TNFi-exposed Central read	28/234 (12.0)	0/65 (0.0)	12.0 (7.8, 16.1); p=0.0034	
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed Local read	[redacted]	[redacted]	[redacted]	
OCTAVE 1 & 2 pooled data, week 8				
TNFi-naïve Central read	[redacted]	[redacted]	[redacted]	NR
TNFi-exposed Central read	[redacted]	[redacted]	[redacted]	NR
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]	NR
TNFi-exposed	[redacted]	[redacted]	[redacted]	NR

Local read				

Source: CS Appendix E Table 121

Table 91 Proportion of patients in remission in OCTAVE Sustain at week 52 according to prior TNFi treatment (FAS, NRI, central and local reads)

Subgroup Prior-TNFi treatment	TOF 5 mg n/N (%))	PBO n/N (%)	Difference vs placebo (95% CI); p-value	TOF 10 mg n/N (%)	Difference vs placebo (95% CI); p-value
TNFi-naïve Central read	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed Central read	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed Local read	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix E Table 125

Mucosal healing

Table 92 Proportion of patients with mucosal healing in OCTAVE Induction 1 and 2 according to prior TNFi treatment (FAS, NRI, central and local reads)

Subgroup: Prior-TNFi treatment	Tofacitinib 10 mg n/N (%)	Placebo n/N (%)	Difference (95% CI); p-value	p-value for heterogeneity
OCTAVE 1, week 8				
TNFi-naïve Central read	88/222 (39.6)	15/57 (26.3)	13.3 (0.2, 26.4); p=0.0630	0.1169
TNFi-exposed Central read	61/254 (24.0)	4/65 (6.2)	17.9 (10.0, 25.7); p=0.0014	
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed	[redacted]	[redacted]	[redacted]	

Local read				
OCTAVE 2, week 8				
TNFi-naïve Central read	71/195 (36.4)	9/47 (19.1)	17.3 (4.1, 30.4); p=0.0239	0.3958
TNFi-exposed Central read	51/234 (21.8)	4/65 (6.2)	15.6 (7.8, 23.5); p=0.0040	
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed Local read	[redacted]	[redacted]	[redacted]	
OCTAVE 1 & 2 pooled data, week 8				
TNFi-naïve Central read	[redacted]	[redacted]	[redacted]	NR
TNFi-exposed Central read	[redacted]	[redacted]	[redacted]	NR
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]	NR
TNFi-exposed Local read	[redacted]	[redacted]	[redacted]	NR

Source: CS Appendix E Table 122

Table 93 Proportion of patients with mucosal healing in OCTAVE Sustain at week 52 according to prior TNFi treatment (FAS, NRI, central and local reads)

Subgroup Prior-TNFi treatment	TOF 5 mg n/N (%))	PBO n/N (%)	Difference vs placebo (95% CI); p-value	TOF 10 mg n/N (%)	Difference vs placebo (95% CI); p-value
TNFi-naïve Central read	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed Central read	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-naïve Local read	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TNFi-exposed Local read	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix E Table 126

Sustained corticosteroid-free remission among patients in remission at baseline

Table 94 Proportion of patients in remission at baseline who had sustained corticosteroid-free remission in OCTAVE Sustain at week 52 according to prior TNFi treatment (FAS, NRI, central and local reads)

Subgroup Prior-TNFi treatment (Yes/No)	TOF 5 mg n/N (%))	PBO n/N (%)	Difference vs placebo (95% CI); p-value	TOF 10 mg n/N (%)	Difference vs placebo (95% CI); p-value
TNFi-naïve Central read					
TNFi-exposed Central read					
TNFi-naïve Local read					
TNFi-exposed Local read					

Source: CS Appendix E Table 129

Appendix 2 NMA sensitivity analyses, additional tables

The tables below are condensed versions of tables that are reported in full in CS Appendix D.1.3.5. Sensitivity analyses were conducted as follows:

- Using centrally read endoscopic subscores for the clinical response, clinical remission and mucosal healing outcomes instead of locally read endoscopic subscores
- Excluding studies in which the majority of participants were Asian. These studies were Suzuki 2014, Mshimesh 2017, Jiang 2015, Kobayashi 2015 and Pursuit J. The CS states that this "*sensitivity analysis is aligned with the base-case assumptions made in the NMA supporting TA329*".
- Limiting the data from the OCTAVE trials and the ULTRA 2 study to patients with prior TNFi failure (i.e. a subset of the base case data which included all patients with prior TNFi-exposure)
- Conducting an overall ITT analysis in which data were not divided into two subgroups by TNFi-exposure status (i.e. combined analysis regardless of prior TNFi-exposure status).

Results for are presented below in Table 95 to Table 102.

Table 95 Clinical response and clinical remission NMA sensitivity analyses – Induction phase, Comparator vs PBO

Comparator	Sensitivity Analyses								
	Centrally read endoscopic subscores			Exclusion of Asian studies			TNFi-exposed using TNFi-failures		
	OR, median (95%CrI)		SUCRA	OR, median (95%CrI)		SUCRA	OR, median (95%CrI)		SUCRA
	Clinical response	Clinical remission		Clinical response	Clinical remission		Clinical response	Clinical remission	
TNFi-naïve subgroup									
PBO			[redacted]			[redacted]			
TOF	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
INF	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
GOL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
VED	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
TNFi-exposed subgroup									
PBO			[redacted]						[redacted]
TOF	[redacted]	[redacted]	[redacted]				[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]						
VED	[redacted]	[redacted]	[redacted]				[redacted]	[redacted]	[redacted]

Source: CS Appendix D Table 101; Table 103, Table 104

Table 96 Clinical response and clinical remission NMA sensitivity analyses – Maintenance phase, Comparator vs PBO

Comparator	Sensitivity Analyses								
	Centrally read endoscopic subscores			Exclusion of Asian studies			TNFi-exposed using TNFi-failures		
	OR, median (95%CrI)		SUCRA	OR, median (95%CrI)		SUCRA	OR, median (95%CrI)		SUCRA
	Clinical response	Clinical remission		Clinical response	Clinical remission		Clinical response	Clinical remission	
TNFi-naïve subgroup									
PBO									
TOF 5 mg									
TOF 10 mg									
INF									
ADA									
GOL 50 mg									
GOL 100 mg									
VEDQ8W									
VED Q4W									
TNFi-exposed subgroup									
PBO									
TOF 5 mg									
TOF 10 mg									
ADA									
VED Q8W									
VED Q4W									

Source: CS Appendix D Table 102, 103, 104

A sensitivity analysis for the overall ITT population (i.e. combining TNFi-naïve and TNFi-exposed participants) was also reported.

Table 97 Overall ITT scenario analysis NMA results – comparative effects and probabilities of achieving clinical response and clinical remission

Comparator	Comparator vs PBO		TOF vs comparator		Absolute probability		SUCRA
	Odds ratio, median (95%CrI)		Odds ratio, median (95%CrI)				
	Clinical response	Clinical remission	Clinical response	Clinical remission	Clinical response	Clinical remission	
Induction phase							
PBO							

TOF	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
INF	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Maintenance phase (re-randomised responder trials only)							
PBO			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF 5 mg	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 50 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL 100 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED Q8W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED Q4W	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix D Table 106

Table 98 Mucosal healing NMA sensitivity analyses – Induction phase, Comparator vs placebo

	Sensitivity Analyses					
	Centrally read endoscopic subscores		**Exclusion of Asian studies**		**TNFi-exposed using TNFi-failures**	
Comparator	**OR, median (95%CrI)**	**SUCRA**	**OR, median (95%CrI)**	**SUCRA**	**OR, median (95%CrI)**	**SUCRA**
TNFi-naïve subgroup						
PBO		[redacted]		[redacted]		
TOF	[redacted]	[redacted]	[redacted]	[redacted]		
INF	[redacted]	[redacted]	[redacted]	[redacted]		
ADA	[redacted]	[redacted]	[redacted]	[redacted]		
GOL	[redacted]	[redacted]	[redacted]	[redacted]		
VED	[redacted]	[redacted]	[redacted]	[redacted]		
TNFi-exposed subgroup						
PBO		[redacted]				[redacted]

TOF	[redacted]	[redacted]			[redacted]	[redacted]
ADA	[redacted]	[redacted]				
VED	[redacted]	[redacted]			[redacted]	[redacted]

Source: CS Appendix D Table 112, 114, 115

Table 99 Mucosal healing NMA sensitivity analyses – Maintenance phase, Comparator vs placebo

Comparator	Sensitivity Analyses					
	Centrally read endoscopic subscores		Exclusion of Asian studies		TNFi-exposed using TNFi-failures	
	OR, median (95%CrI)	SUCRA	OR, median (95%CrI)	SUCRA	OR, median (95%CrI)	SUCRA
TNFi-naïve subgroup						
PBO		[redacted]		[redacted]		
TOF 5 mg	[redacted]	[redacted]	[redacted]	[redacted]		
TOF 10 mg	[redacted]	[redacted]	[redacted]	[redacted]		
INF	[redacted]	[redacted]	[redacted]	[redacted]		
ADA	[redacted]	[redacted]	[redacted]	[redacted]		
GOL 50 mg	[redacted]	[redacted]	[redacted]	[redacted]		
GOL 100 mg	[redacted]	[redacted]	[redacted]	[redacted]		
VEDQ8W	[redacted]	[redacted]	[redacted]	[redacted]		
VED Q4W	[redacted]	[redacted]	[redacted]	[redacted]		
TNFi-exposed subgroup						
PBO		[redacted]				[redacted]
TOF 5 mg	[redacted]	[redacted]			[redacted]	[redacted]
TOF 10 mg	[redacted]	[redacted]			[redacted]	[redacted]
ADA	[redacted]	[redacted]			[redacted]	[redacted]
VED Q8W	[redacted]	[redacted]			[redacted]	[redacted]
VED Q4W	[redacted]	[redacted]			[redacted]	[redacted]

Source: CS Appendix D Table 113, 114, 115

Adverse events

Sensitivity analysis was conducted using a network from which the Asian studies (Suzuki 2014, Kobayashi 2015 and Mshimesh 2017) were excluded and also the tofacitinib phase II study (Sandborn 2012). The exclusion of the Asian studies also caused the loss of the UC-SUCCESS trial (azathioprine versus infliximab) from the network as it could no longer be connected to the network of evidence (Figure 14).

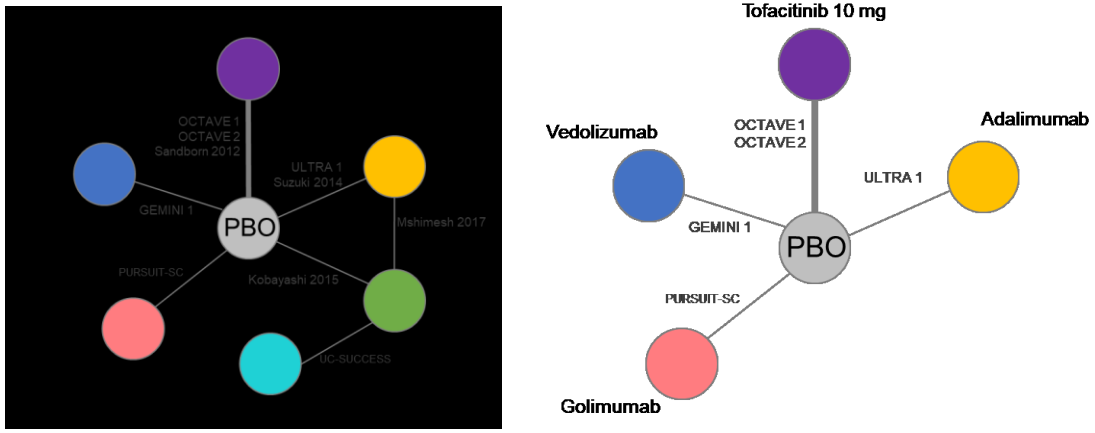


Figure 14 Base-case safety evidence network (left) and sensitivity analysis network (right)

Results are shown in the tables below.

Table 100 Induction phase sensitivity analysis NMA results – comparative effects and probabilities of discontinuing due to AEs

Comparator	Comparator vs PBO		TOF vs comparator	Absolute probability, median (95% CrI)	SUCRA
	Treatment effect (logit scale), median (95% CrI)	Odds ratio, median (95% CrI)	Odds ratio, median (95% CrI)		
PBO			[redacted]	[redacted]	[redacted]
TOF	[redacted]	[redacted]		[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix D Table 116

Table 101 Induction phase sensitivity analysis NMA results – comparative effects and probabilities of serious AEs

Comparator	Comparator vs PBO		TOF vs comparator	Absolute probability, median (95% CrI)	SUCRA
	Treatment effect (logit scale), median (95% CrI)	Odds ratio, median (95% CrI)	Odds ratio, median (95% CrI)		
PBO			[redacted]	[redacted]	[redacted]
TOF	[redacted]	[redacted]		[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix D Table 117

Table 102 Induction phase sensitivity analysis NMA results – comparative effects and probabilities of serious infections

Comparator	Comparator vs PBO		TOF vs comparator	Absolute probability, median (95% CrI)	SUCRA
	Treatment effect (logit scale), median (95% CrI)	Odds ratio, median (95% CrI)	Odds ratio, median (95% CrI)		
PBO			[redacted]	[redacted]	[redacted]
TOF	[redacted]	[redacted]		[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VED	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix D Table 118

Appendix 3 Health economics: results of ERG scenario analyses

Table 103 ERG base case: scenarios on patient age (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
Patient age: 28 years	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,644
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£589,024
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,170
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant
Patient age: 52 years	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£8,019
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£628,794
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,648
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

Table 104 ERG base case: scenarios on weight * (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
Patient weight: 70 kgs	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,827
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£607,395
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,401
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant
Patient weight: 80 kgs	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,819
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£607,504
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,394
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

* Scenario based on “Use average of OCTAVE” option for wastage calculations

Table 105 ERG base case: scenarios on NMA models (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
Network meta-analysis (company preferred response/remission)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,793
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£633,458
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,541
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	£8,801,245
TNFi-failed (ved) + TNFi-exposed for tof and ada	***TNFi-exposed***			
	Conventional	[redacted]	[redacted]	£9,669
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£2,521,513
Network meta-analysis (FE for serious infections)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£8,513
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£589,976
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£10,080
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

Table 106 ERG base case: scenarios on utility (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
Utility source: Swinburn	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£10,898
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£845,865
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£13,198
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	--
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant
Utility source: OCTAVE 8 weeks	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£17,764
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£1,360,239
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£21,376
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant
Utility source: OCTAVE 52 weeks	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£18,256
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£1,373,067
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£21,283
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

Table 107 ERG base case: scenarios on resource use (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
Drug stopping rule (6.5 outpatient visit for all patients in maintenance)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£9,090
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	-
	Vedolizumab	[redacted]	[redacted]	£608,793
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£10,597
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	-
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant
Reduced health state resource use (clinical practice scenario)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£13,938
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	-
	Vedolizumab	[redacted]	[redacted]	£613,289
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£14,950
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	-
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

Table 108 ERG base case: scenarios on drug costs (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
CT drug usage (mesalazine only)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,827
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	-
	Vedolizumab	[redacted]	[redacted]	£607,576
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,402
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	-
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant
Drug admin cost (for subcutaneous injection for ada and gol.)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,815
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£607,571
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,389
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

Table 109 ERG base case: scenarios on surgery cost (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
Stoma care costs (£81.66 per 8 week cycle TA342)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,804
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£607,561
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,379
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant
Surgery costs (£13,156 Buchannan et al.)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,764
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£607,522
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,341
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

Table 110 ERG base case: scenarios on surgery risks (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
Surgery incidence rate (Chhaya et al.)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,980
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£611,440
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,558
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant
Surgery complications (Tappenden et al.)	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	£7,556
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Golimumab	[redacted]	[redacted]	Tofacitinib dominant
	Infliximab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	£605,226
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	£9,134
	Adalimumab	[redacted]	[redacted]	Tofacitinib dominant
	Tofacitinib	[redacted]	[redacted]	
	Vedolizumab	[redacted]	[redacted]	Tofacitinib dominant

Table 111 ERG base case: drug sequencing scenarios (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Pairwise ICER (£/QALY) (Tof vs comparator)
Treatment sequencing	***TNFi- naive***			
	Conventional	[redacted]	[redacted]	[redacted]
	Inf-Ada-CT	[redacted]	[redacted]	[redacted]
	Inf-Ved-CT	[redacted]	[redacted]	[redacted]
	Inf-Tof-CT	[redacted]	[redacted]	[redacted]
	Tof-Ada-CT	[redacted]	[redacted]	[redacted]
	Ved-Ada-CT	[redacted]	[redacted]	[redacted]
	TNFi- naive			
	Conventional	[redacted]	[redacted]	[redacted]
	Gol-Ada-CT	[redacted]	[redacted]	[redacted]
	Gol-Ved-CT	[redacted]	[redacted]	[redacted]
	Gol-Tof-CT	[redacted]	[redacted]	[redacted]
	Tof-Ada-CT	[redacted]	[redacted]	[redacted]
	Ved-Ada-CT	[redacted]	[redacted]	[redacted]
	TNFi-exposed			
	Conventional	[redacted]	[redacted]	[redacted]
	Tof-Ada-CT	[redacted]	[redacted]	[redacted]
	Ved-Ada-CT	[redacted]	[redacted]	[redacted]

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Tofacitinib for previously treated active ulcerative colitis [ID1218]

You are asked to check the ERG report from SHTAC to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on 6 August 2018** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 Typographical errors

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
ERG report page 135, section 4.2 states; "*The analysis by Wu et al. indicated that one of the treatment sequences shown in Table 52 would be optimal in the UK context, depending on the incremental cost-effectiveness ration (ICER) threshold."*	Suggest changing to: "*The analysis by Wu et al. indicated that one of the treatment sequences shown in Table 52 would be optimal in the UK context, depending on the incremental cost-effectiveness ratio~~n~~ (ICER) threshold."*	Typographical error No impact on clinical and cost-effectiveness conclusions presented.	This typographical error has been corrected.
ERG report page 139, section 4.3.2 states; *"The gender distribution in the audit was 59% (529/903), similar to that in the OCTAVE trials."*	Suggest changing to: *"The gender distribution in the audit was 59% males (529/903), similar to that in the OCTAVE trials."*	Typographical error No impact on clinical and cost-effectiveness conclusions presented.	Omitted word has been inserted.
ERG report page 157, section 4.3.5 states; *"EQ-5D outcomes from the OCTAVE 1 and 2 induction trials and the OCTAVE sustain maintenance trial are outlined in CS B.2.6.1.2 and B.2.6.2.2, with further information in Table 218 (CS L.1.4) and Figures 54 to 61 (CS M.4)"*	Suggest changing to: *"EQ-5D outcomes from the OCTAVE 1 and 2 induction trials and the OCTAVE Sustain maintenance trial are outlined in CS B.2.6.1.2 and B.2.6.2.2, with further information in Table 218 (CS L.1.4) and Figures 54 to 61 (CS M.4)"*	Typographical error No impact on clinical and cost-effectiveness conclusions presented.	This typographical error has been corrected.
ERG report page 162, section 4.3.6 states; *"Given the high cost of olsalzine, this suggests that the cost of 5ASA drugs is over-estimated."*	Suggest changing to: *"Given the high cost of olsalazine, this suggests that the cost of 5ASA drugs is over-estimated."*	Typographical error No impact on clinical and cost-effectiveness conclusions presented.	This typographical error has been corrected.

ERG report page 164, section 4.3.6 states; *"The Tsai et al. estimates of resource use have also been used in other NICE appraisals for ulcerative colitis (TQ329 and TA342)"*	Suggest changing to: *"The Tsai et al. estimates of resource use have also been used in other NICE appraisals for ulcerative colitis (TA329 and TA342)"*	Typographical error No impact on clinical and cost-effectiveness conclusions presented.	This typographical error has been corrected.
ERG report page 193, section 4.4.3 states: *"The ICER for tofacitinib vs CT remained low, reaching a maximum of £21,376 per QALY with OCTAVE EQ-5D utility estimtes Tofacitinib dominated vedolizumab across all the scenarios, except in the the company's preferred NMA models for response and remission (favouring fixed effect models)."*	Suggest changing to: *"The ICER for tofacitinib vs CT remained low, reaching a maximum of £21,376 per QALY with OCTAVE EQ-5D utility estimtes Tofacitinib dominated vedolizumab across all the scenarios, except in the ~~the~~ company's preferred NMA models for response and remission (favouring fixed effect models)."*	Typographical error No impact on clinical and cost-effectiveness conclusions presented.	This typographical error has been corrected. A second typographical error (incorrect spelling of estimates) has also been corrected.

Issue 2 Missing Commercial or Academic in Confidence marking (CiC/AiC)

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
ERG report page 23 states; *"However, we identified errors in the scenarios relating the use of central read NMA results and tofacitinib maintenance using [redacted] split"*	Please change to: *"However, we identified errors in the scenarios relating the use of central read NMA results and tofacitinib maintenance using [redacted] split"*	Academic sensitive information requiring marking. No impact on clinical and cost-effectiveness conclusions presented	The ERG contacted NICE for clarification on the appropriate marking because (i) This information is not marked AIC in the shaded grey summary box

			at the start of CS B.3 in the version of the CS sent to the ERG; (ii) the ERG found this information marked CIC in the response to clarification question B6(b); (iii) following their factual error check the company have suggested AIC marking. On the advice of NICE this information has now been marked CIC in the ERG report.
Page 40: ERG report Table 6: Summary characteristics of tofacitinib RCTs. The first cell under heading "OCTAVE Open" gives the following data:	Please change to: *"Tofacitinib[b]* *10 mg BID ([redacted])* *5 mg BID ([redacted])"*	Academic sensitive information requiring marking. No impact on clinical and cost-effectiveness	This information, which is not marked AIC in Table 8 of

<table>
<tr>
<td>“Tofacitinib[b]
10 mg BID ([redacted])

5 mg BID ([redacted])”

OCTAVE Open is an ongoing clinical trial and Pfizer plans to publish the results when available and therefore the the n number for the 5mg dose should be marked academic in confidence.</td>
<td></td>
<td>conclusions presented</td>
<td>the CS sent to the ERG, has now been marked AIC in the ERG report.</td>
</tr>
<tr>
<td>Page 42, footnote b: ERG report states: “Note that there appears to be a typographical error in CS Table 8 where the number of patients receiving tofacitinib 10 mg is given as [redacted].”

Pfizer confirms the typographical error in the CS, and requests that n number should be marked as academic in confidence.</td>
<td>Please change to:

“There is a typographical error in CS Table 8 where the number of patients receiving tofacitinib 10 mg is given as n=[redacted].”</td>
<td>Clarification and data is academic in confidence; although incorrect it indicates the magnitude of the population.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>This information, which as noted in the row above was not marked AIC in Table 8 of the CS sent to the ERG, has now been marked AIC in the ERG report.</td>
</tr>
<tr>
<td>Table 14 of the ERG report, page 74.

The values within the ERG network meta-analyses are, as described by the ERG, comparable to the Pfizer analysis and therefore compromise the academic in confidence marked information, which Pfizer intends to publish.</td>
<td>Please change to:

<table>
<tr><th>Comparator</th><th colspan="3">Treatment effect vs placebo, median (95% CrI), probit scale[a]</th></tr>
<tr><th></th><th>Company base-case (fixed effects)</th><th>ERG replication of base-case (fixed effects)</th><th>ERG alternative model selection (random effects)</th></tr>
<tr><td colspan="4">Maintenance phase</td></tr>
<tr><td>Tofacitinib 5 mg</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Tofacitinib 10 mg</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Infliximab 5 mg/kg</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
<td>Academic sensitive information requiring marking.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>The ERG had already requested advice from NICE on the likely need to mark items AIC in ERG tables 14 to</td>
</tr>
</table>

<table>
<tr>
<td>To avoid any publication of NMA results by NICE in advance of the Pfizer publication we would request the entire table 14 to be marked academic in confidence.</td>
<td>
<table>
<tr><td>Adalumimab 40 mg Q2W</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Golimumab 50 mg</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
<td></td>
<td>16, and tables 18 and 19. This has now been done.</td>
</tr>
<tr>
<td>Table 15 of the ERG report, page 74.

The values within the ERG network meta-analyses are as described by the ERG, comparable to the Pfizer analysis and therefore compromise the academic in confidence marked information, which Pfizer intends to publish.

To avoid any publication of NMA results by NICE in advance of the Pfizer publication we would request the entire table 15 to be marked academic in confidence.</td>
<td>Please change to:

<table>
<tr><th rowspan="2">Comparator</th><th colspan="3">Treatment effect vs placebo, median (95% CrI), probit scale[a]</th></tr>
<tr><th>Company base-case (fixed effects)</th><th>ERG replication of base-case (fixed effects)</th><th>ERG alternative model selection (random effects)</th></tr>
<tr><td colspan="4">Induction phase</td></tr>
<tr><td>Tofacitinib 10 mg</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Adalumimab 160/80/40 mg</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Vedolizumab 300 mg</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
<td>Academic sensitive information requiring marking.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>As noted above, this is now marked AIC.</td>
</tr>
<tr>
<td>Table 18 of the ERG report, page 78.

The values within the ERG network meta-analyses are as described by the ERG, comparable to the Pfizer analysis and therefore compromise the academic in confidence marked information, which Pfizer intends to publish.

To avoid any publication of NMA results by NICE in advance of the Pfizer publication we would request the entire table 18 to be marked academic in confidence.</td>
<td>Please change to:

<table>
<tr><th rowspan="2">Comparator</th><th colspan="3">Treatment effect vs placebo, median (95% CrI), probit scale</th></tr>
<tr><th>Company base-case (fixed effects)</th><th>ERG replication of base-case (fixed effects)</th><th>ERG exploratory scenario analysis (fixed effects)</th></tr>
<tr><td colspan="4">Maintenance phase</td></tr>
<tr><td>Tofacitinib 5 mg</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Tofacitinib 10 mg</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Adalumimab 40 mg Q2W</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Vedolizumab 300 mg Q8W</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Vedolizumab 300 mg Q4W</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
<td>Academic sensitive information requiring marking.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>As noted above, this is now marked AIC</td>
</tr>
</table>

<table>
<tr>
<td>Table 19 of the ERG report, page 79.

The values within the ERG network meta-analyses are as described by the ERG, comparable to the Pfizer analysis and therefore compromise the academic in confidence marked information, which Pfizer intends to publish.

To avoid any publication of NMA result by NICE in advance of the Pfizer publication we would request the entire table 19 to be marked academic in confidence.</td>
<td>Please change to:

<table>
<tr><th rowspan="2">Comparator</th><th colspan="2">Treatment effect vs placebo, median (95% CrI)</th></tr>
<tr><th>Company baseline</th><th>ERG replication of company baseline</th></tr>
<tr><td colspan="3">Induction TNFi-exposed, probit scale</td></tr>
<tr><td>Response/remission</td><td>[redacted]</td><td>[redacted])</td></tr>
<tr><td colspan="3">Maintenance TNFi-naïve, probit scale</td></tr>
<tr><td>Response/remission</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td colspan="3">Induction, logit scale</td></tr>
<tr><td>Serious Infections</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Serious adverse events</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
<td>Academic sensitive information requiring marking.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>As noted above, this is now marked AIC</td>
</tr>
<tr>
<td>ERG report page 174, section 4.3.8 states

“Experts observed that the company assumption relating to [redacted] of patients benefitting from maintenance dose of 10mg twice daily may not be limited to patients in the TNFi-exposed group only.”</td>
<td>Please change to:

“Experts observed that the company assumption relating to [redacted] of patients benefitting from maintenance dose of 10mg twice daily may not be limited to patients in the TNFi-exposed group only.”</td>
<td>Academic sensitive information requiring marking.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>As indicated in the first row of this section, this information has now been marked CIC on the advice of NICE.</td>
</tr>
</table>

<table>
<tr>
<td>On page 182, the ERG report states:

Table 74 presents academic in confidence data, which Pfizer would like to have marked accordingly.</td>
<td>Please change to

<table>
<tr><th rowspan="2">Company scenarios</th><th rowspan="2">Brief rationale/assumption</th><th colspan="2">ICERs for Tofacitinib vs CT (£/QALY)</th></tr>
<tr><th>TNFi-naïve</th><th>TNFi-exposed</th></tr>
<tr><td>Base case</td><td></td><td>£8,554</td><td>£10,302</td></tr>
<tr><td>Overall ITT population</td><td></td><td colspan="2">£7,805</td></tr>
<tr><td>Tofacitinib maintenance dose mix</td><td>[redacted] of patients receiving 5mg; [redacted] of patients receiving 10mg</td><td>£12,628</td><td>£13,947</td></tr>
</table></td>
<td>Academic sensitive information requiring marking.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>As indicated in the first row of this section, this information has now been marked CIC on the advice of NICE.</td>
</tr>
<tr>
<td>On page 184, the ERG table 75 presents CiC and AiC marked information.

With the updated Document B and appendices dated 02/07/2018 as required by NICE, Pfizer have revised the CiC marking into AiC marking.</td>
<td>Please change to:

<table>
<tr><th rowspan="2">Study</th><th rowspan="2">Treatment</th><th colspan="2">TNFi-naive</th><th colspan="2">TNFi-exposed</th></tr>
<tr><th>Clinical response</th><th>Clinical remission</th><th>Clinical response</th><th>Clinical remission</th></tr>
<tr><td rowspan="2">OCTAVE Induction 1</td><td>Placebo</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Tofacitinib</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>****</td></tr>
<tr><td rowspan="2">OCTAVE Induction 2</td><td>Placebo</td><td>****</td><td>[redacted]</td><td>****</td><td>[redacted]</td></tr>
<tr><td>Tofacitinib</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td rowspan="2">Model</td><td>Placebo</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Tofacitinib</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table></td>
<td>Academic sensitive information requiring marking.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>CIC marking now changed to AIC marking.</td>
</tr>
<tr>
<td>On page 189, of the ERG report Table 81 presents academic in confidence data, which Pfizer would like to have marked accordingly.</td>
<td>Please change to:

<table>
<tr><th rowspan="2">Scenarios</th><th rowspan="2">Assumption</th><th colspan="2">ICER for tofacitinib vs.</th></tr>
<tr><th>CT</th><th>Vedolizumab</th></tr>
<tr><td>Base case</td><td></td><td>£8,564</td><td>£615,077</td></tr>
<tr><td>Tofacitinib</td><td>[redacted] of patients receiving</td><td>£12,63</td><td>Tofacitini</td></tr>
</table></td>
<td>Academic sensitive information requiring marking.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>As indicated in the first row of this section, this information has now been marked CIC on the advice of</td>
</tr>
</table>

<table>
<tr><td></td><td><table>
<tr><td>maintenance dose mix</td><td>5mg; [redacted] of patients receiving 10mg</td><td>7</td><td>b dominant</td></tr>
</table></td><td></td><td>NICE.</td></tr>
<tr><td>On page 192, the ERG reportTable 85 presents academic in confidence data, which Pfizer would like to have marked accordingly.</td><td><table>
<tr><th>Scenarios</th><th>Assumption</th><th colspan="2">ICER for Tofacitinib vs.</th></tr>
<tr><th></th><th></th><th>CT</th><th>Vedolizumab</th></tr>
<tr><td>Base case</td><td></td><td>£10,311</td><td>£7,838,381</td></tr>
<tr><td>Tofacitinib maintenance dose mix</td><td>[redacted] of patients receiving 5mg; [redacted] of patients receiving 10mg</td><td>£13,956</td><td>Tofacitinib dominant</td></tr>
</table></td><td>Academic sensitive information requiring marking.
No impact on clinical and cost-effectiveness conclusions presented</td><td>As indicated in the first row of this section, this information has now been marked CIC on the advice of NICE.</td></tr>
</table>

Issue 3 Minor factual inaccuracies and clarifications

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 15 of the ERG report states: "*At week 52 in the OCTAVE Sustain maintenance trial the results for clinical remission also favoured tofacitinib (difference versus placebo 35.1%, 95% CI 26.7 to 43.5, p<0.0001 using locally read data).*" Pfizer believes the data here to be incomplete as it only contains results for the 10mg BID dose and omits the results for the 5mg BID dose.	Suggest changing to: *"At week 52 in the OCTAVE Sustain maintenance trial the results for clinical remission also favoured tofacitinib (difference versus placebo 35.1%, 95% CI 26.7 to 43.5, p<0.0001 (10mg BID); 26.8%, 95% CI 18.5 to 35.1, p<0.0001(5mg BID), using locally read data).*"	To provide the reader with the factually complete results for both doses. No impact on clinical and cost-effectiveness conclusions presented.	For completeness the ERG has provided the results for the 5mg tofacitinib dose as suggested.
Pages 15 to 16 of the ERG report states: *"Safety data for tofacitinib in patients with moderate to severely active ulcerative colitis*	Suggest changing to: *"Safety data for tofacitinib in patients with moderate to severely active ulcerative colitis*	Factual accuracy and clarity of the document. No impact on clinical and cost-	For clarity the ERG has stated that the comparisons between tofacitinib and placebo

comes from the Phase II trial, the three Phase III OCTAVE trials and the ongoing OCTAVE Open extension study. Rates of adverse events of any type were broadly similar for the tofacitinib and placebo arms within each trial with serious adverse events affecting fewer than 10% of patients." Pfizer believes this requires clarification as to the trials to which the data refers, as OCTAVE Open was not placebo controlled.	*comes from the Phase II trial, the three Phase III OCTAVE trials and the ongoing OCTAVE Open extension study. Rates of adverse events of any type were broadly similar for the tofacitinib and placebo arms within OCTAVE Induction 1 and 2 and OCTAVE Sustain, with serious adverse events affecting fewer than 10% of patients in these trials."*	effectiveness conclusions presented	arms only apply to the OCTAVE Induction 1 and 2 and the OCTAVE Sustain trials (i.e. not OCTAVE Open which was not placebo controlled).
Page 33 of the : ERG report states: *"The primary outcome in the phase 3 OCTAVE trials was clinical remission whilst the primary outcome in the phase 2 trial was clinical response. HRQoL was a secondary outcome in all the tofacitinib trials, and mucosal healing was a secondary outcome in the phase 3 trials."* Pfizer believes this requires correction and clarification, as the primary endpoint is inaccurate, being the more stringent endpoint of remission, and the description of the secondary outcomes is incomplete.	Suggest changing to: "*The primary outcome in the phase 3 OCTAVE trials was remission whilst the primary outcome in the phase 2 trial was clinical response. HRQoL was a secondary outcome in all the tofacitinib trials, and mucosal healing was a key secondary outcome in the phase 3 trials."*	Factual accuracy of the document. No impact on clinical and cost-effectiveness conclusions presented	The ERG has corrected this typographical error.
Page 38 of the ERG report states: *"Prohibited therapies included TNFi therapies within 8 weeks of baseline; azathioprine, methotrexate, and 6-mercaptopurine within 2 weeks; and ciclosporin and intravenous corticosteroids (CS Tables 9 and 10)."* Pfizer believes that this requires clarification to reflect the scope of the decision problem	Suggest changing to: *"Prohibited therapies included TNFi therapies within 8 weeks of baseline; azathioprine, methotrexate, and 6-mercaptopurine within 2 weeks; anti-adhesion molecule therapy taken within 1 year; and ciclosporin and intravenous corticosteroids (CS Tables 9 and 10)."*	For clarity that protocol reflects scope of the decision problem. No impact on clinical and cost-effectiveness conclusions presented	The list reported by the ERG is not exhaustive therefore this is not a factual inaccuracy. However, in the interests of clarity the ERG has amended the text as suggested by the

which includes further biologics.			company.
Page 38 of the ERG report states: *"The OCTAVE induction and maintenance trials conform to a re-randomisation design."* Further in the paragraph the ERG outlines the alternative trial design (treat-through), and as the paragraph stands it can be interpreted that the OCTAVE re-randomisation design is by choice, whereas it was advocated by the regulatory authorities to exposure time of inadequate treatments, such as placebo.	Suggest changing to: *"The OCTAVE induction and maintenance trials conform to a re-randomisation design, as requested by regulatory authorities."*	For clarity that protocol reflects scope of the decision problem. No impact on clinical and cost-effectiveness conclusions presented	Not a factual inaccuracy. This information is not provided in the CS.
Page 53 of the ERG report states: "The company also defined *key secondary outcomes: mucosal healing (OCTAVE 1 and 2: week 8; OCTAVE Sustain: week 52), and for OCTAVE Sustain only, sustained corticosteroid-free remission among patients in remission at baseline (week 52)."* The timings for assessment of sustained corticosteroid-free remission are incomplete.	Suggest changing to: *"The company also defined key secondary outcomes: mucosal healing (OCTAVE 1 and 2: week 8; OCTAVE Sustain: week 52), and for OCTAVE Sustain only, sustained corticosteroid-free remission among patients in remission at baseline (measured at weeks 24 and 52)."*	Factual accuracy of the document. No impact on clinical and cost-effectiveness conclusions presented	The ERG agrees the text is incomplete and has updated this as suggested by the company.
Page 55 of the ERG report Table 10. Clinical effectiveness outcomes and outcome definitions of the OCTAVE RCTs. The cell for a key secondary outcome: sustained corticosteroid-free remission and when assessed for OCTAVE Sustain gives the following data: *"52."* This is inaccurate as this was measured at two	Suggest changing to: *"24, 52."*	Factual accuracy of the document. No impact on clinical and cost-effectiveness conclusions presented	The ERG agrees the text is incomplete and has updated this as suggested by the company.

time points in the study.			
ERG report page 66, section 3.1.7 states; *"Where there was a difference in the deviance information criterion (DIC) of less than 3, the company favoured the fixed effects model."* The statement made by the ERG does not fully convey the full extent of the company's decision making for the model choice (page 88-89 of CS), which was more considered than using the DIC measures as a proxy.	Suggest changing to: *"Where models were comparable in terms of results and goodness of fit, the company favoured the fixed effects model."*	Factual accuracy of the document. Potential impact on clinical conclusions and minimal impact on cost-effectiveness conclusions presented	No change made. The ERG cannot find evidence for the company's decision making for model choice on pages 88-89 (text both before and after CS Table 22) of the CS. Furthermore, the detailed description of the methods of network meta-analysis in Appendix D.1.3.3 states "Where the difference in DIC suggested indifference, the simpler fixed effect model was preferred". The hints in the text that the company may have intended for NMA model results to be similar in terms of results (B.2.9.2.1.1 and D.1.3.3) are not supported by any evidence (the CS only presents results from the favoured model).
ERG report page 73, section 3.1.7 states; *"Similarly, the company preferred the fixed effects model in the maintenance phase TNFi-naïve population for clinical response/remission."*	Suggest changing to: *"Similarly, the company preferred the fixed effects model in the maintenance phase TNFi-naïve population for clinical response/remission as it deemed the random effect results implausibly imprecise because*	Factual accuracy of the document. Potential impact on clinical conclusions and minimal impact on cost-effectiveness	The ERG agrees that the company have stated the results were implausibly imprecise and no treatment was predicted to be significantly better

The statement made by the ERG does not fully convey the full extent of the company's decision making for the model choice as detailed in the CS (page 89 of CS).	*no treatment was predicted to be significantly better than placebo."*	conclusions presented	than placebo. The text has therefore been amended albeit the ERG is not convinced that this is an argument in favour of choosing the fixed-effect model.
Pages 129 to 130 of the ERG report states: *"Rates of adverse events of any type were broadly similar for the tofacitinib and placebo arms within each trial. Serious adverse events affected fewer than 10% of patients in the tofacitinib trials."* Pfizer believes that this requires clarification as OCTAVE Open was not placebo controlled	Suggest changing to: *"Rates of adverse events of any type were broadly similar for the tofacitinib and placebo arms within OCTAVE Induction 1 and 2 and OCTAVE Sustain. Serious adverse events affected fewer than 10% of patients in these trials."*	Factual accuracy of the document. No impact on clinical and cost-effectiveness conclusions presented	For clarity of the document the ERG agrees with the company's suggested wording.
ERG report page 141, section 4.3.2 states; *"These recommendations relate to early assessment of response following induction treatment (from 8 to 16 weeks after initiation)."* Table 39 on page 122 of the CS presents the licenced doses and stopping rules according to SmPC. The SmPC of adalimumab suggests to cease treatment if response has not been achieved within 2-8 weeks of initiating treatment. Therefore the range quoted in the ERG report should state 2-16 weeks to encompass all stopping times across treatments as per the SmPCs.	Suggest changing to: *"These recommendations relate to early assessment of response following induction treatment (from 2 to 16 weeks after initiation)."*	Factual accuracy of the document. No impact on clinical and cost-effectiveness conclusions presented	The ERG agrees and the text has been corrected as suggested.

ERG report page 149, section 4.3.4 states; *"The company state that their choice of NMA models was based on DIC measures of model fit, but that they preferred the simpler fixed effect approach when DIC statistics were similar (CS B.2.9.2.1.1)"* The statement made by the ERG does not fully convey the full extent of the company's decision making for the model choice (page 88-89 of CS), which was more considered than simply using the DIC measures as a proxy.	Suggest changing to: *"The company state that their choice of NMA models was based on model fit statistics and assessment of result outputs. For the induction phase the company chose fixed effects over random effects as models were comparable in terms of results and goodness of fit. For the maintenance phase the company chose the fixed effect model as it deemed the random effect results implausibly imprecise because no treatment was predicted to be significantly better than placebo. (CS B.2.9.2.1.1)"*	Factual accuracy of the document. Potential impact on clinical conclusions and minimal impact on cost-effectiveness conclusions presented	As the ERG has stated above (in relation to ERG report p.66) Appendix D.1.3.3 states "Where the difference in DIC suggested indifference, the simpler fixed effect model was preferred" but the ERG agrees that the company's stated reason for choosing the fixed-effect model for the maintenance phase was the "implausibly imprecise" results obtained from the random effects model. The ERG has therefore amended the text as follows: "The company state that their choice of NMA models was based on ~~DIC measures of~~ model fit statistics. For the induction phase the results and model fit for the fixed and random effects models were comparable for both patients subgroups. In the TNFi-naïve subgroup the model fit diagnostics were slightly better for the random effects model so this was preferred. For

<table>
<tr>
<td></td>
<td></td>
<td></td>
<td>the TNFi-exposed subgroup ~~, but that~~ they preferred the simpler fixed effect approach ~~when~~ because the DIC statistics were similar (CS B.2.9.2.1.1). In the maintenance phase the fixed effect models were preferred because the company deemed the random effects results implausibly imprecise with no treatment predicted to be significantly better than placebo.”</td>
</tr>
<tr>
<td>Table 75 of page 184 in the ERG report presents incorrect information sourced from the CS table 199 (Appendix J.1.2), which also presents the incorrect values.

As a result the ERG conclusions on the comparability of trial results versus model predictions are no longer applicable.

Table 90 on page 212 of the ERG report presents the correct values.</td>
<td>Please consider deleting “While the model results appear comparable with the clinical data for the tofacitinib arm in the TNFi-naïve group, there are large differences in the estimates for TNFi- exposed subgroup for this arm, along with the placebo arms for both induction and maintenance phases.”

We also request that table 75 is changed to:
<table>
<tr><th rowspan="2">Study</th><th rowspan="2">Treatment</th><th colspan="2">TNFi-naive</th><th colspan="2">TNFi-exposed</th></tr>
<tr><th>Clinical response</th><th>Clinical remission</th><th>Clinical response</th><th>Clinical remission</th></tr>
<tr><td rowspan="2">OCTAVE Induction</td><td>Placebo</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Tofacitinib</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>****</td></tr>
</table>
</td>
<td>Factual accuracy of the document.

No impact on clinical and cost-effectiveness conclusions presented</td>
<td>Estimates for OCTAVE 1 & 2 in Table 75 of the ERG report were sourced from CS Table 199 in Appendix J.1.2 of the original company submission (which the company has now acknowledged to contain incorrect values).

We have updated our report to reflect the changes suggested by the company.</td>
</tr>
</table>

	1							
	OCTAVE Induction 2	Placebo	****	[redacted]	****	[redacted]		
		Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]		
	Model	Placebo	[redacted]	[redacted]	[redacted]	[redacted]		
		Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]		
Page 203 of the ERG report states: *"Whilst, they state that their choice of NMA models was based on DIC measures of model fit, but they preferred the simpler fixed effect approach when DIC statistics were similar."* The statement made by the ERG does not fully convey the full extent of the company's decision making for the model choice (page 88-89 of CS), which was more considered than simply by using the DIC measures as a proxy							Suggest deleting or changing to: *"Whilst they state that their choice of NMA models was based on model fit statistics and assessment of result outputs; for the induction phase the company chose fixed effects over random effects as models were comparable in terms of results and goodness of fit. For the maintenance phase the company chose the fixed effect model as it deemed the random effect results implausibly imprecise because no treatment was predicted to be significantly better than placebo."*	Factual accuracy of the document. Potential impact on clinical conclusions and minimal impact on cost-effectiveness conclusions presented	In line with earlier responses, the ERG has made an amendment to the text to capture the company's view that the maintenance phase NMA for the TNFi-naïve subgroup under the random effects model gave implausibly imprecise results.

Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE

Tofacitinib for moderately to severely active ulcerative colitis

ERRATUM

Replacement pages for factual inaccuracies in Evidence Review Group report

14 August 2018

Produced by Southampton Health Technology Assessments Centre (SHTAC)

tofacitinib groups with statistically significant differences between the 5mg and the 10mg tofacitinib arms versus placebo.

Remission, mucosal healing and sustained corticosteroid-free remission did not contribute data to the economic model.

Clinical remission is an outcome with an almost identical definition to the primary outcome of remission. The difference being that the rectal bleeding sub-score of the Mayo score does not have to be zero to achieve clinical remission. The outcomes of clinical remission and clinical response contribute data to the economic model.

Using locally read data (which were used in the base case economic evaluation) in OCTAVE 1, the mean difference between the tofacitinib group and the placebo group was 13.3 percentage points (95% CI 6.5 to 20.2, p=0.0017). The corresponding data for OCTAVE 2 were a mean difference from placebo of 15.6 percentage points (95% CI 9.9 to 21.3, p=0.0002). At week 52 in the OCTAVE Sustain maintenance trial the results for clinical remission also favoured tofacitinib (difference versus placebo 35.1%, 95% CI 26.7 to 43.5, p<0.0001 (10mg BID); 26.8%, 95% CI 18.5 to 35.1, p<0.0001 (5mg BID), both using locally read data).

Clinical response at both week 8 (OCTAVE Induction trials) and week 52 (OCTAVE Sustain trial) was also statistically significantly higher among participants who received tofacitinib.

Subgroup analyses according to prior TNFi-exposure status were conducted for the main clinical effectiveness outcomes. The results were consistent regardless of prior TNFi-exposure status.

HRQoL was reported using generic (EQ-5D and SF-36) and disease specific (IBDQ and WPAI-UC) instruments. HRQoL was typically improved by tofacitinib treatment however for some HRQoL measures the ERG was uncertain about the impact of missing data. Data from the EQ-5D-3L did not inform the base-case economic model but were included in a scenario analysis.

Safety data for tofacitinib in patients with moderate to severely active ulcerative colitis comes from the Phase II trial, the three Phase III OCTAVE trials and the ongoing OCTAVE Open extension study. Rates of adverse events of any type were broadly similar for the tofacitinib and

placebo arms within OCTAVE Induction 1 and 2 and OCTAVE Sustain with serious adverse events affecting fewer than 10% of patients. Ulcerative colitis was the most frequent serious adverse event and most other serious adverse events were related to ulcerative colitis. Serious infections were uncommon (data on serious infections were included in the economic model). Overall, and in comparison with evidence from the use of tofacitinib in patients with rheumatoid arthritis, no new safety signals were identified.

There are no head-to-head RCTs of tofacitinib versus the comparators defined in the company's decision problem. Therefore the company used NMA to estimate the relative effectiveness and safety of tofacitinib in both the induction and maintenance phases of treatment in comparison to TNF-alpha inhibitors (infliximab, adalimumab and golimumab), vedolizumab and conventional therapies. The company's systematic review identified 21 RCTs that were considered for inclusion in the NMA. Four of these were the tofacitinib RCTs listed above, a further 14 were included in one or more NMA networks and three studies could not be included in any of the NMA networks.

Table 1 NMAs conducted by the company

	TNFi-naïve population subgroup	**TNFi-exposed population subgroup**
Induction phase	Clinical response and clinical remission Mucosal healing	Clinical response and clinical remission Mucosal healing
	Safety outcomes (discontinuation due to AEs, SAEs, serious infections)	
Maintenance phase	Clinical response and clinical remission Mucosal healing	Clinical response and clinical remission Mucosal healing

The ERG judged the NMAs to be generally well conducted but identified nine issues:

- Use of the probit scale to model clinical response/clinical remission is an improvement on a previous approach in NICE guidance TA342 but a multinomial logit model could have been considered.
- Potential inconsistency in a closed loop of the maintenance TNFi-naïve network was not examined

Table 3 Cost effectiveness: Company base case, with prior TNFi (with tofacitinib PAS)

Strategy	Total		Incremental analysis			Pairwise ICERs tofacitinib vs. comparator (£/QALY)
	QALYs	Costs (£)	QALYs	Costs (£)	ICER (£/QALY)	
Conventional	[redacted]	[redacted]	[redacted]	[redacted]	-	£10,302
Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]	£10,302	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£7,838,238	£7,838,238

A range of uncertainty analyses were conducted by the company, but they have been selective in the scenarios they present

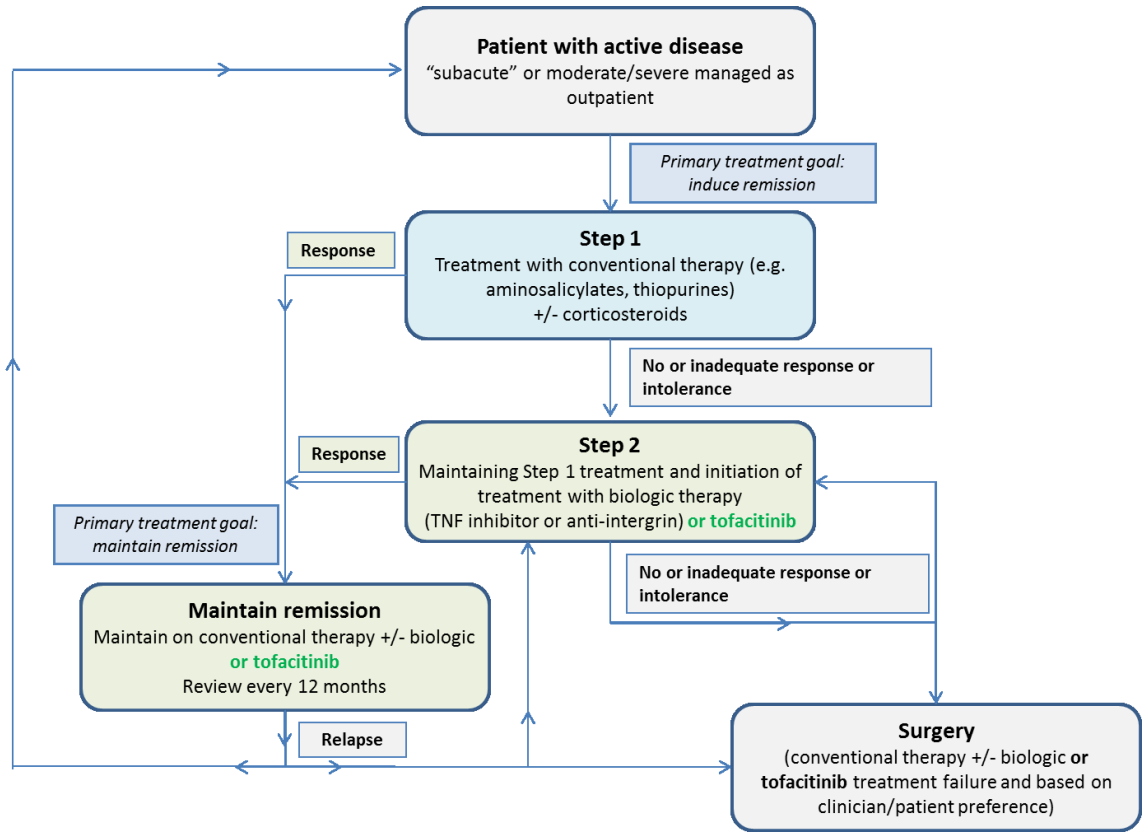
The company performed a range of deterministic-, probabilitistic- and scenario analyses to assess the methodological as well as parameter uncertainty of their base case analyses. The ERG agrees with their assumptions for DSA and PSA and their results, in general. However, we identified errors in the scenarios relating the use of central read NMA results and tofacitinib maintenance using [redacted] split. The company corrected the error in the latter scenario in their response to clarification question. For the scenario analyses, we view that the company has been selective in the scenarios they present.

Commentary on the robustness of submitted evidence

Strengths

- The model structure is consistent and follows the conventional design for ulcerative colitis appraisals.
- The model generally adheres to the NICE scope for this appraisal.
- The perspective of the analysis aligns with the NICE guide to the methods of Technology Appraisal.
- The model uses a lifetime time horizon to allow estimation of all relevant costs and quantity of life impairment.
- The model uses appropriate sources for costs and resource use and in line with other technology appraisals
- The model allows the flexibility to incorporate treatment sequencing which provides a closer reflection of clinical practice.
- The ERG agrees with the company's approach to modelling surgery and its related risks, source of costs and utilities for the base case and mortality.

history and clinical decision making on the appropriateness of therapies, and therefore may not adequately capture the nuances of clinical practice when comparing to the NICE scope" (clarification response A3). In their clarification response the company provided a simplified version of CS Figure 1 in order to better represent the position of tofacitinib in the treatment pathway in relation to the NICE scope (reproduced in Figure 1).



Source: company's clarification response A3

Figure 1 Proposed position of tofacitinib within the treatment pathway

Outcomes

The outcomes included in the CS are clinically meaningful and are consistent with the NICE scope and EMA guidance on methods for clinical trials in ulcerative colitis.[12] The primary outcome in the phase 3 OCTAVE trials was remission whilst the primary outcome in the phase 2 trial was clinical response. HRQoL was a secondary outcome in all the tofacitinib trials, and mucosal healing was a secondary outcome in the phase 3 trials. Details of the outcome selection are discussed further below in section 3.1.4. In summary, the key issues noted by the ERG are:

In all four RCTs the comparator was placebo. All these RCTs were used in support of the company's application for a marketing authorisation and were sponsored by Pfizer, the manufacturer of tofacitinib.

The Phase II trial is not described in detail in the CS but it is included in the company's NMA (CS section B.2.9) and data from this trial are also included in the adverse events section (CS Appendix F Table 166). As the Phase II trial was a small dose-finding study with 194 patients, of whom only 33 received the licensed 10 mg BID dose (company clarification response A16), the CS focuses on the Phase III trials. The ERG agrees that this is reasonable and accordingly the current ERG report also focuses primarily on the Phase III trials.

It was unclear to the ERG from the description of the Phase II trial population reported both in the CS and in the trial publication whether this matched the NICE scope. The company confirmed that it does match the scope, as "*patients were only included if they continued to have moderate to severe disease despite previous treatment*" (clarification response A2). In addition, the company provided a table detailing the failed drug treatments at baseline (clarification response Table 1) and full details of the inclusion and exclusion criteria (clarification response Appendix A).

The number of centres in the studies ranged from 51 (Phase II trial) to 297 (OCTAVE Sustain), but it should be noted that a number of centres in the Phase III trials randomised just one patient (16 centres in OCTAVE 1; 25 centres in OCTAVE 2; and 66 centres in OCTAVE Sustain[19]). While each study included some patients from the UK, this number was low [redacted].

OCTAVE 1 and 2 were double-blind, randomised placebo-controlled tofacitinib induction trials with an 8 week treatment phase, and used identical methods (see Table 6).

In addition to the criteria listed above, patients had to have moderately to severely active disease (6 to 12 on the Mayo score, with a rectal bleeding sub-score of 1 to 3 and an endoscopic sub-score of 2 or 3). Prohibited therapies included TNFi therapies within 8 weeks of baseline; azathioprine, methotrexate, and 6-mercaptopurine within 2 weeks; anti-adhesion-molecule therapy taken within 1 year; and ciclosporin and intravenous corticosteroids (CS Tables 9 and 10). Permitted concomitant medications for ulcerative colitis included oral

Table 6 Summary characteristics of tofacitinib RCTs

Phase II trial[11] (efficacy/dose RCT)		**OCTAVE 1[19] (induction RCT)**		**OCTAVE 2[19] (induction RCT)**		**OCTAVE Sustain[19] (maintenance RCT)**		**OCTAVE Open[20] (extension study)**
Tofacitinib 0.5 mg (n=31) 3 mg BID (n=33) 10 mg BID (n=33) 15 mg BID (n=49)	Placebo (n=48)	Tofacitinib 10 mg BID (n=476)[a]	Placebo (n=122)	Tofacitinib 10 mg BID (n=429)[a]	Placebo (n=112)	Tofacitinib 10 mg BID (n=197) 5 mg BID (n=198)	Placebo (n=198)	Tofacitinib[b] 10 mg BID ([redacted]) 5 mg BID ([redacted])
Design: randomised, double-blind, placebo-controlled trial (2:2:2:3:3 ratio tofacitinib 0.5 mg: 3mg: 10 mg: 15 mg: placebo)		*Design:* identical randomised, double-blind, placebo-controlled trials (4:1 ratio tofacitinib: placebo, stratified according to previous treatment with TNFi therapies, glucocorticoid use at baseline, and geographic region)				*Design*: randomised, double-blind, placebo-controlled trial (1:1:1 ratio tofacitinib 5 mg: tofacitinib 10 mg; placebo)		*Design:* open-label extension
Location: 51 sites worldwide (UK = 2, [redacted][d])		*Location:* 144 sites worldwide (UK = 2, [redacted])		*Location:* 169 sites worldwide (UK = 3, [redacted])		*Location:* 297 sites worldwide (UK = 5, [redacted])		*Location:* 215 sites worldwide (UK = 5)
Inclusion: • age ≥18 years • confirmed diagnosis of UC for ≥3 months		*Inclusion:* • age ≥18 years • confirmed diagnosis of UC for ≥4 months				*Inclusion:* • entry criteria for the Induction trials • completed 8 weeks induction therapy		*Inclusion:* • completed or demonstrated treatment failure in

		2.5 mg per week until the dose was 0 mg).	
Length of follow-up: 8 weeks of treatment and 4 weeks follow-up	*Length of follow-up:* 9 weeks (primary efficacy endpoint at 8 weeks)	*Length of follow-up:* 53 weeks (primary efficacy endpoint at 52 weeks)	*Length of follow-up:* up to 6 years (12-month interim results reported)

Sources Sandborn et al.[11], CS Table 7, 8, 9, 10, 13 and 14, and B.2.6.3.1

BID, twice daily; NR, not reported; UC, Ulcerative colitis.

[a] 15 mg BID tofacitinib treatment was discontinued based on feedback from regulatory authorities (OCTAVE 1: n=38, OCTAVE 2: n=18)

[b] Three subpopulations received tofacitinib 10 mg ([redacted]) in the open label extension study: Induction non-responders tofacitinib 10 mg [redacted]; maintenance completers tofacitinib 10 mg [redacted], maintenance treatment failures 10 mg ([redacted] comprising participants from OCTAVE Induction 1 and 2 who withdrew from OCTAVE Sustain due to treatment failure on tofacitinib (5 mg, [redacted]; 10 mg,[redacted];) or placebo ([redacted]). One subpopulation received tofacitinib 5 mg in the open label extension study: Maintenance: remission tofacitinib 5 mg [redacted]. Note that there appears to be a typographical error in CS Table 8 where the number of patients receiving tofacitinib 10 mg is given as n=[redacted].

[d] [redacted]

3.1.5 Description and critique of company's outcome selection

The outcomes included in the CS match those in the NICE final scope and appear appropriate. However, time to surgical intervention, although specified in the NICE final scope, was not included, as this was not assessed in the OCTAVE trials.

In clinical trials of therapies for ulcerative colitis the Mayo Score is widely used and was used within the OCTAVE trials (CS Section B1.3.1 and CS Table 3). There are four components to the Mayo score, one of which is 'Endoscopic findings'. In the OCTAVE trials the Mayo endoscopic sub-score was assessed both locally (by the study site investigator) and centrally (from a video recording). Consequently the outcomes in the CS that utilise the endoscopic sub-score were reported separately using the local or the central read of the endoscopic data. The ERG notes that the FDA[22] state that central reading is the preferred approach and the OCTAVE clinical trial programme is the first in ulcerative colitis to use central reads (CS Section B.2.3.1.2.4).

The primary outcome in OCTAVE 1 and 2 was remission at week 8 based on centrally read endoscopic Mayo sub-scores, and at week 52 in OCTAVE Sustain (for definition of remission see Table 10). Higher Mayo scores indicate more severe disease. The company also defined key secondary outcomes: mucosal healing (OCTAVE 1 and 2: week 8; OCTAVE Sustain: week 52), and for OCTAVE Sustain only, sustained corticosteroid-free remission among patients in remission at baseline (measured at weeks 24 and 52). Mucosal healing is associated with lower rates of hospitalisation and surgery,[23] while the use of corticosteroids long-term is not suitable due to side effects so a corticosteroid-free remission is important.[24]

Clinical response and clinical remission based on Mayo sores (for definitions see Table 10) were reported for all three trials (OCTAVE 1 and 2: week 8; OCTAVE Sustain: week 52). As can be observed from Table 10 the difference between the primary outcome of remission and the secondary outcome of clinical remission is that for the former the rectal bleeding sub-score must be zero whereas this is not necessary for the outcome of clinical remission. Clinical response and clinical remission were the only clinical effectiveness outcomes included in the economic model (the primary outcome did not contribute to the economic model), as they were thought to ensure comparability with trials of biological therapies for ulcerative colitis.

which has a recall period of 1 week (in OCTAVE 1 and 2 assessed at baseline and week 8; in OCTAVE Sustain assessed at baseline and weeks 24 and 52). Higher scores indicate better HRQoL. A systematic review[30] of the SF-36 in patients with ulcerative colitis suggests that a group-level clinically important difference threshold of 3 points for both summary scores and responder-level thresholds of 3.1 for PCS and 3.8 for MCS based on the SF-36v2 manual.[31]

- The WPAI-UC score, based on a 6-item questionnaire (version 2) assessing work productivity, is also reported by all three OCTAVE RCTs (OCTAVE 1 and 2 at baseline and week 8; OCTAVE Sustain at baseline and week 52). The questionnaire yields four scores expressed as impairment percentages: absenteeism; presenteeism; work productivity loss; non-work activity impairment. A higher score indicates greater impairment.[32] As part of the response to NICE and the ERG's clarification question A12, the company states that it is not aware of any validated MCID for this outcome in patients with ulcerative colitis. However the company also state that extrapolating from Crohn's Disease suggests a 7% decrease is the MCID for the WPAI.[33,34]

Table 10 Clinical effectiveness outcomes and outcome definitions of the OCTAVE RCTs

Outcome	Definition	When assessed, week		Used in Model
		OCTAVE 1 & 2	OCTAVE Sustain	
Primary: Remission based on centrally-read endoscopic sub-scores	Mayo score ≤2, no individual sub-score >1, rectal bleeding sub-score = 0	8	52	No
Key secondary: Mucosal healing	Mayo endoscopic sub-score ≤1	8	52	No
Key secondary: Sustained corticosteroid-free remission among patients in remission at baseline	Remission (as defined above for the primary outcome) plus no treatment with steroids for ≥4 weeks before the 24-week and 52-week visits	Not assessed	24, 52	No
Clinical response	Mayo score decrease from baseline ≥ 3, and ≥ 30%, with a decrease in rectal	Week 8	52	Yes

Table 13 Company choice of base-case and ERG preference

	Company base-case model	ERG favoured model
Clinical response/clinical remission, Induction TNFi-naive	Random effects	Random effects
Clinical response/clinical remission, Induction TNFi-exposed	Fixed effects	Random effects
Clinical response/clinical remission, Maintenance TNFi-naive	Fixed effects	Random effects
Clinical response/clinical remission, Maintenance TNFi-exposed	Fixed effects	Fixed effects
Serious infections, Induction	Random effects	Fixed effects

In the induction phase TNFi-exposed subgroup, the fixed effects model was preferred despite similar DIC and similar total residual deviance. The ERG would have selected the random effects model as the more conservative analysis. Whilst the base case models are presented in the main NMA results (CS Table 25) the alternative model is not reported. We would prefer to have seen this explored as a sensitivity analysis.

Similarly, the company preferred the fixed effects model in the maintenance phase TNFi-naïve population for clinical response/remission as it deemed the random effect results implausibly imprecise because no treatment was predicted to be significantly better than placebo. The ERG would have chosen the random effects model for both the lower DIC and total residual deviance. The ERG would prefer to have seen this explored as a sensitivity analysis.

Finally, the company chose the random effects model for serious infections. In response to a clarification request the company provided the random effect standard deviation (1.82, 95%CrI 0.15, 4.59) (clarification question A22). This wide CrI indicates weak support for the random effects model which has a similar DIC, thus we might have favoured the fixed effects model. The ERG would prefer to have seen the fixed effects model included in a sensitivity analysis.

Table 14 and Table 15 show the results of the ERG validation and exploratory analysis for the response and remission analyses. The ERG ran the same number of chains, burn-in and simulations reported by the company (section D.1.3.3). Models converged and our results concur to two decimal places.

The alternative choice random effects models show wider credible intervals and some variation in the median estimates for adalimumab and golimumab in the maintenance analysis for the TNFi-naïve population as smaller studies are given more weight under the random effects than the fixed effects model.

Table 14 ERG replication and additional analysis on model choice - clinical response and clinical remission for TNFi-naïve subgroup

Comparator	Treatment effect vs placebo, median (95% CrI), probit scale[a]		
	Company base-case (fixed effects)	ERG replication of base-case (fixed effects)	ERG alternative model selection (random effects)
Maintenance phase			
Tofacitinib 5 mg	[redacted]	[redacted]	[redacted]
Tofacitinib 10 mg	[redacted]	[redacted]	[redacted]
Infliximab 5 mg/kg	[redacted]	[redacted]	[redacted]
Adalumimab 40 mg Q2W	[redacted]	[redacted]	[redacted]
Golimumab 50 mg	[redacted]	[redacted]	[redacted]

Source of company base-case (fixed effects) is CS Table 26
[a] On the probit scale, negative coefficients indicate improvement over placebo. Where the upper and lower CrI are both negative, treatments show strong evidence of benefit versus placebo.

Table 15 ERG replication and additional analysis on model choice - clinical response and clinical remission for TNFi-exposed subgroup

Comparator	Treatment effect vs placebo, median (95% CrI), probit scale[a]		
	Company base-case (fixed effects)	ERG replication of base-case (fixed effects)	ERG alternative model selection (random effects)
Induction phase			
Tofacitinib 10 mg	[redacted]	[redacted]	[redacted]
Adalumimab 160/80/40 mg	[redacted]	[redacted]	[redacted]

Vedolizumab 300 mg	[redacted]	[redacted]	[redacted]

Source of company base-case (fixed effects) is CS Table 25
[a] On the probit scale, negative coefficients indicate improvement over placebo. Where the upper and lower CrI are both negative, treatments show strong evidence of benefit versus placebo.

However, when we attempted to replicate the serious infections results there was a higher level of uncertainty around the coefficients particularly for tofacitinib (Table 16). The wider credible intervals persisted under the fixed effects model conducted by the ERG.

Table 16 ERG replication and additional analysis on model choice - serious infections

Comparator	Treatment effect vs placebo, median (95% CrI), logit scale		
	Company base-case (random effects)	ERG replication of base-case (random effects)	ERG alternative model selection (fixed effects)
Tofacitinib 10 mg	[redacted]	[redacted]	[redacted]
Infliximab 10 mg/kg	[redacted]	[redacted]	[redacted]
Adalumimab 160/80/40 mg	[redacted]	[redacted]	[redacted]
Golimumab 200/100 mg	[redacted]	[redacted]	[redacted]
Vedlizumab 300 mg	[redacted]	[redacted]	[redacted]
Azathioprine	[redacted]	[redacted]	[redacted]

Source of company base-case (fixed effects) is CS Table 34

The very wide credible intervals for tofacitinib are caused by the lack of any serious infections across placebo arms in the three tofacitinib studies, hence the difficulty to estimate a relative treatment effect compared to placebo (Table 17). There was also considerable autocorrelation in the tofacitinib coefficient despite thinning and running an extended number of simulations.

The reasons for the difference in our results are unclear, particularly how the company arrived at their estimate for tofacitinib.

case which combined TNFi-exposed data for tofacitinib and adalimumab with TNFi-failure data for vedolizumab. Our scenario analysis at least included comparable data for tofacitinib and vedolizumab.

In the event, as Table 18 shows, use of TNFi-failure data makes little difference to the response/remission results for tofacitinib.

Table 18 ERG scenario analysis using TNFi-failure data from both OCTAVE Sustain and GEMINI 1

Comparator	Treatment effect vs placebo, median (95% CrI), probit scale[a]		
	Company base-case (fixed effects)	ERG replication of base-case (fixed effects)	ERG exploratory scenario analysis (fixed effects)
Maintenance phase			
Tofacitinib 5 mg			
Tofacitinib 10 mg			
Adalumimab 40 mg Q2W			
Vedolizumab 300 mg Q8W			
Vedolizumab 300 mg Q4W			

Source of company base-case (fixed effects) is CS Table 28
[a] on the probit scale, negative coefficients indicate improvement over placebo. Where the upper and lower CrI are both negative, treatments show strong evidence of benefit versus placebo.

3.1.7.4 Baseline response models – uncertainty around absolute probabilities

To estimate absolute probabilities of each event, treatment effects from the NMA were combined with an estimate of the placebo (baseline) response from the placebo arms of included studies. In response to clarification request A17 the company provided the data, priors and output (meanA, precA) in WinBUGs code format for the probit baseline models. We were able to replicate selected median estimates for the baseline calculations. However, despite running the CS code [validated against NICE DSU Technical Support Document (TSD) 2[46]] and data we were unable to replicate the baseline credible intervals used in the

probit or logit models. The company models tended to lead to wider credible intervals compared to our calculations, thus would lead to conservative results. A summary of the differences in our findings is provided in Table 19 below.

Table 19 ERG replication of baseline (placebo) response results

Comparator	Treatment effect vs placebo, median (95% CrI)	
	Company baseline	ERG replication of company baseline
Induction TNFi-exposed, probit scale		
Response/remission	[redacted]	[redacted]
Maintenance TNFi-naïve, probit scale		
Response/remission	[redacted]	[redacted]
Induction, logit scale		
Serious Infections	[redacted]	[redacted]
Serious adverse events	[redacted]	[redacted]

3.1.7.5 Inclusion of the tofacitinib phase II trial

The Sandborn 2012 Phase II (induction) tofacitinib trial[11] is less well described in the CS despite being included in the NMAs. Furthermore, the company state:
All studies, except for one [Sandborn 2012], were conducted in patients with moderately to severely active ulcerative colitis who had an inadequate response to or had failed to tolerate one or more of the following conventional therapies: oral or intravenous corticosteroids, azathioprine, and/or 6-mercaptopurine (CS section B.2.9.1.1).

The ERG thus questioned the eligibility of this trial. The company confirmed that the Phase II trial met the inclusion criteria for the NMA and they also provided selected NMA results obtained with the Phase II trial excluded from the NMA (Table 7 in clarification response A16). These results for response and remission for the TNFi-naïve and TNFi-exposed populations in the induction period were similar to the base case (CS Table 25).

Base case results without the Phase II trial were not provided for the safety outcomes. However, given the relatively high serious infection rate in the tofacitinib arms of the Phase II trial compared to the OCTAVE trials (6% [2/33] patients had an event compared to 1% [6/476] in OCTAVE Induction 1 and none in OCTAVE Induction 2), the Phase II trial may

healing at week 52 and sustained mucosal healing at weeks 24 and 52 were reported for the 5 mg and 10 mg tofacitinib maintenance doses in comparison to the placebo arm of the trial.

Sustained corticosteroid-free remission among those in remission at baseline (a further key secondary outcome) in the OCTAVE Sustain trial, was statistically significantly greater in the tofacitinib 5 mg and 10 mg arms than in the placebo arm.

Clinical remission, which has a very similar definition to the primary outcome of remission, contributed data to the economic model via the NMA. Due to the similarity of outcome definition the results from the OCTAVE trials were almost identical to those reported above for remission, favouring tofacitinib.

The outcome of clinical response also contributes data to the economic analysis via NMA. The percentage difference between the tofacitinib group and the placebo group in favour of tofacitinib was statistically significant in both OCTAVE induction trials and the OCTAVE Sustain maintenance trial and for both the central and locally read data.

HRQoL was reported using both generic (EQ-5D and SF-36) and disease specific (IBDQ and WPAI-UC) instruments. Results showed HRQoL was typically improved by tofacitinib treatment; however, for some HRQoL measures we are uncertain about the impact of the missing data. Data from the EQ-5D-3L do not inform the base-case economic model but were included in a scenario analysis.

Subgroup analyses focused on results according to prior TNFi-exposure. Note that this is a more restricted subgroup than that of prior biologic therapy (which would also include other biological therapies such as vedolizumab) which is listed in the NICE scope. The OCTAVE trials were not powered to test the statistical significance of subgroup analyses so the results should be interpreted cautiously. Overall, the results were consistent regardless of prior TNFi-exposure status.

Safety data for tofacitinib in patients with moderate to severely active ulcerative colitis comes from the Phase II tofacitinib trial, the two OCTAVE Induction trials, the OCTAVE Sustain trial and the ongoing OCTAVE Open extension study. In total tofacitinib has been evaluated in 1157 patients with ulcerative colitis with a maximum exposure to tofacitinib of 4.4 years.

Rates of adverse events of any type were broadly similar for the tofacitinib and placebo arms within OCTAVE Induction 1 and 2 and OCTAVE Sustain. Serious adverse events

For comparison, the median age at diagnosis of ulcerative colitis in the 2016 RCP audit was 32 years (interquartile range (IQR) 24 to 45) and the median age at initiation of biologic treatment was 39 years (IQR 28 to 52).[64] The gender distribution in the audit was 59% males (529/903), similar to that in the OCTAVE trials.

We consider that the gender mix, initial age and weight of the model cohort should be assumed similar for people with and without prior exposure to TNFi drugs. In ERG analysis, we assume 59% males, initial age 41 years and weight 73.5 kg, based on means for both arms in the OCTAVE Induction trials. We conduct scenario analysis to assess the impact of age (28 to 58) and body weight (range 70 kg to 80 kg) on the results.

4.3.2.2 Comparators

The model assumes that patients start treatment with tofacitinib or the biologic comparators with an induction phase of treatment. Patients who respond during induction continue to receive maintenance treatment with the same drug (with concomitant use of conventional drugs) until loss of response or an acute exacerbation requiring surgery. Patients who do not respond to induction treatment and those who relapse during maintenance continue to receive conventional treatment alone, until planned or emergency surgery, or death.

Inclusion of comparators in economic analysis

Tables 40 and 41 in the CS (page 130) outline the comparators used in the company's economic analysis:

- **TNFi-naïve subgroup**, all comparators specified in the scope (infliximab, adalimumab, golimumab, vedolizumab, tofacitinib and conventional therapy (CT));
- **TNFi-exposed subgroup**, only vedolizumab, tofacitinib and CT are included. Cost-effectiveness is not reported for infliximab, adalimumab or golimumab.

For patients with prior exposure to TNFi drugs, infliximab and golimumab could not be included in the company's NMA due to a lack of trial evidence (CS section B.2.9.2.1). However, the TNFi-exposed NMA does include adalimumab, so the company could have included adalimumab in the cost-effectiveness analysis for this subgroup. The CS does not give a clear rationale for omitting adalimumab for the TNFi-exposed subgroup.

Clinical experts have advised the ERG that treatment with a TNFi would sometimes be considered for a patient with prior exposure to another TNFi. There is a group of patients

Stopping rules for drug treatment

- *Discontinuation due to lack of response to induction therapy*
  CS Table 38 summarises SmPC recommendations about when to stop tofacitinib and biologic drug treatment. These recommendations relate to early assessment of response following induction treatment (from 2 to 16 weeks after initiation). In contrast, the clinical trials provide evidence of response at 6 weeks for golimumab and vedolizumab and at 8 weeks for other comparators, and the model assumes a fixed 8-week induction period followed by immediate cessation of treatment for patients whose disease does not show a response in this time. *If in practice clinicians assess response to induction later than 8 weeks, the average cost of induction therapy will be higher than that estimated by the company model. However, effectiveness may also be higher if some patients have a late response to induction. The direction and magnitude of the bias from assuming a fixed 8-week period of induction for all comparators is unclear.*

- *Discontinuation due to loss of response during maintenance*
  Guidance for the TNF-alpha inhibitors (TA329) and vedolizumab (TA342) recommend annual assessment of response, with treatment continuing only if there is clear evidence of ongoing benefit. Clinical advice to the ERG is that the benefit of biologic treatment is usually considered annually, in line with NICE guidance. However, treatment would usually be withdrawn earlier for patients who lose response, as the patient will seek an appointment when symptoms recur or get worse and this will trigger consideration of changing or stopping treatment.

  The company model applies a constant risk of relapse across each 8-week cycle of maintenance, with treatment stopping immediately when patients lose response. Thus, it assumes that maintenance treatment is stopped within 8 weeks of a loss of response. To achieve this, all patients on maintenance treatment must have fast access to clinical assessment on relapse or be seen routinely every 8 weeks. The company model assumes an average of 2 outpatient visits for patients in remission on maintenance treatment and 4.5 visits per year for patients with a response but no remission.

  The ERG considers that the assumption that treatment will be withdrawn following relapse reflects UK practice. However, we have concerns that the costs of monitoring and follow-up in the company's model do not reflect the full cost of ensuring that treatment can be withdrawn within 8 weeks of a relapse. We consider a scenario with

- **Choice of fixed effects versus random effects**

  The company state that their choice of NMA models was based on model fit statistics. For the induction phase the results and model fit for the fixed and random effects models were comparable for both patients subgroups. In the TNFi-naïve subgroup the model fit diagnostics were slightly better for the random effects model so this was preferred. For the TNFi-exposed subgroup they preferred the simpler fixed effect approach because the DIC statistics were similar (CS B.2.9.2.1.1). In the maintenance phase the fixed effect models were preferred because the company deemed the random effects results implausibly imprecise with no treatment predicted to be significantly better than placebo. Table 56 below summarises the NMA models chosen for the company base case analysis.

Table 56 Selection of response/remission NMA models

	Patient subgroup	Induction	Maintenance
Company base case	TNFi-naive	Random effects	Fixed effects
	TNFi-exposed	Fixed effects	Fixed effects
ERG preference	TNFi-naive	Random effects	Random effects
	TNFi-exposed	Random effects	Fixed effects *

* Random effects model would not run for the maintenance NMA

The ERG has a general preference for the random effect NMA models, as we believe that the fixed effect models may underestimate uncertainty due to heterogeneity between the studies. We test the impact of different NMA models on cost-effectiveness results in section 4.4.3 below.

- **Combination of TNFi-failed and TNFi-exposed subgroups**

  The base case NMAs combine outcomes for subgroups defined as TNFi-failed for vedolizumab with TNFi-exposed subgroups for tofacitinib and adalimumab (CS Table 22). The company conducted a sensitivity analysis for the TNFi-failure subgroup, which reduced the probit score for tofacitinib by -0.13 in the induction phase, bringing it closer to vedolizumab. (CS Table 28). They reported that results were not available for adalimumab and that the analysis could not be run for the maintenance phase. Therefore, the TNFi-failure NMA sensitivity analysis does not provide the required input parameters and was not used in the economic model.

4.3.8 Model validation

The company describes their approach to model validations in CS section B.3.10. They state that they engaged UK clinical experts, statisticians and health economists to validate model inputs and assumptions in a UK advisory board meeting. Further details on the key aspects of validation are summarised in CS Table 78.

The CS stated that clinical experts validated model methods pertaining to the patient population; subgroup analysis by prior TNFi-exposure; time on treatment and discontinuation rates; costs (including monitoring cost for tofacitinib, health state costs and resource use, including rate of hospitalisation); emergency surgery; quality of life and maintenance dose of tofacitinib. The experts are reported to agree with the company's assumptions in most of these aspects, except for:

- **Patient population:** Although the baseline characteristics of the patient population in OCTAVE reflect UK practice, the duration of disease in OCTAVE trials (which was 6-7 years) is longer than that in clinical practice (which is ~2-4 years).
- **Health state unit costs and resource use, including rate of hospitalisation:** Tsai et al. was confirmed to reflect an accurate representation of unit costs and resource use as per clinical practice. However, the experts suggested that the model base-case assumptions relating to annual medical resource use (CS Table 55) underestimated the resource use per patient per year.
- **Tofacitinib maintenance dose:** Experts observed that the company assumption relating to [redacted] of patients benefitting from maintenance dose of 10mg twice daily may not be limited to patients in the TNFi-exposed group only.

The economic model was quality checked by health economists. For face validity, the company compared the proportion of patients in response and remission predicted by the model against the estimated values from the NMA, shown below in Figure 9.

Further, the model results were compared with previous TA329; however, the CS did not report any comparison of the results in TA329 with those in the current appraisal. We discuss this in detail in section 4.4.1. For internal validity, the CS stated that a second modeller reviewed the model; conducted extreme value tests alongside inspecting model code, formulae and references. An independent health economist was reported to have reviewed the model structure, parameter inputs and core model assumptions.

Table 74 Company scenario analyses

Company scenarios	Brief rationale/assumption	ICERs for Tofacitinib vs CT (£/QALY)	
		TNFi-naïve	TNFi-exposed
Base case		**£8,554**	**£10,302**
Overall ITT population			£7,805
Tofacitinib maintenance dose mix	[redacted] of patients receiving 5mg; [redacted] of patients receiving 10mg	£12,628	£13,947
Fixed baseline utility instead of age-adjusted	Assumption that patient quality of life stays constant over time.	£8,760	£10,589
OCTAVE trial utilities	EQ-5D data were collected in Tofacitinib Phase III clinical trials	£15,508	£18,276
Swinburn utilities	To compare with previous analyses	£11,932	£14,487
Emergency surgery from any state	Due to the uncertainty on the likely protection from acute events based on the level of response/remission, patients are assumed to undergo emergency surgery regardless of state membership	£8,194	£9,962
Emergency surgery only from active UC	As above but assuming response to treatment offers the same level of protection from acute events, as remission	£8,652	£10,475
No emergency surgery	As above, but assuming no emergency surgery in the model	£8,710	£10,593
Central read NMA results	Central read was the primary endpoint in OCTAVE trials.	£9,469	£10,793
Discounting every cycle	It tested the sensitivity of the model when the discounting of outcomes is applied every 8 weeks.	£8,606	£10,398
Adalimumab maintenance 73% 40 mg Q2W and 27% 40 mg QW	Dose escalation of adalimumab was assumed in Archer *et al*.	£8,554	--
Golimumab 100 mg every 4 weeks in maintenance	A 100 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients, such as those who have experienced a decrease in their response	£8,554	--
Vedolizumab 300 mg every 4 weeks in maintenance	A 300 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients who have a body weight ≥ 80 kg	£8,554	Dominated

Source: CS Table 63 to 66; 71 to 77

4.4 Additional work undertaken by the ERG

4.4.1 ERG model validation

4.4.1.1 Model verification procedures

We checked the economic model for transparency and validity. The visual basic code used within the model was accessible. The NMA code in WinBUGs was provided in Appendix D.1.3.4.

We conducted a range of 'white box' tests to verify model inputs, calculations and outputs:

- Cross-checking of all parameter inputs against values in the CS and cited sources;
- Checking the individual equations within the model;
- A range of extreme value and logic tests to check the plausibility of changes in results when parameters are changed
- Checking the VBA code for treatment sequencing
- Checking all model outputs against results cited in the CS, including the base case, PSA, DSA and manually ran all the scenarios
- Running the NMA code in WinBUGs to replicate selected results (see section 3.1.7).

In addition, we checked the model calculations of patient transitions through the health states, costs and QALYs by re-coding the model independently based on the inputs from the company's submitted model.

Overall, we found the economic model to be of a good quality, with very few errors in input parameters, logic or coding. We identified a few small errors that we correct in section 4.4.2 below, which did not make any substantive difference to the results. We were also successful in replicating outputs from most of the company's NMA models, with the exception of the serious infection NMA (section 3.1.7).

1.1.1.1 External validity

We have tabulated the model predictions against the observed clinical data for the maintenance phase, in Table 75 below.

Table 75 Comparison of the predicted model results of Tofacitinib and Placebo (CT) against the observed clinical data – INDUCTION Phase

Study	Treatment	TNFi-naive		TNFi-exposed	
		Clinical response	Clinical remission	Clinical response	Clinical remission
OCTAVE Induction 1	Placebo	[redacted]	[redacted]	[redacted]	[redacted]
	Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]
OCTAVE Induction 2	Placebo	[redacted]	[redacted]	[redacted]	[redacted]
	Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]
Model	Placebo	[redacted]	[redacted]	[redacted]	[redacted]
	Tofacitinib	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix J.1.2. Table 199

4.4.1.3 Cross validation

In section 4.2 above (page 134), we state that the CS reported previous economic models, including published literature and analyses conducted by ERGs for previous NICE TAs, for patients in ulcerative colitis. Whilst we acknowledge that there are methodological differences between the economic models across these studies, nonetheless we view that they provide sources for cross-validation of results from the company base-case analysis. Of the reported studies, we cross-validate the modelled findings of the current appraisal with 2 previous NICE TAs (TA342 and TA329) and 1 published study as summarised in Table 76. The most relevant analysis for the current appraisal is the final version from the NICE TA of vedolizumab (TA342). This appraisal relates to same patient population as the current appraisal and comparators overlap, except Tofacitinib and surgery.

Table 81 Scenario analyses, company base case (ERG corrected) – TNFi-naive subgroup

Scenarios	Assumption	ICER for tofacitinib vs.	
		CT	Vedolizumab
Base case		**£8,564**	**£615,077**
Tofacitinib maintenance dose mix	[redacted] of patients receiving 5mg; [redacted] of patients receiving 10mg	£12,637	Tofacitinib dominant
Fixed baseline utility instead of age-adjusted	Assumption that patient quality of life stays constant over time.	£8,770	£634,346
OCTAVE trials utilities	EQ-5D data were collected in Tofacitinib Phase III clinical trials	£15,525	£1,079,814
Swinburn utilities	To compare with previous analyses	£11,945	£853,228
Emergency surgery from any state	Due to the uncertainty on the likely protection from acute events based on the level of response/remission, patients are assumed to undergo emergency surgery regardless of state membership	£8,204	£606,872
Emergency surgery from active UC only	As above but assuming response to treatment offers the same level of protection from acute events, as remission	£8,661	£618,151
No emergency surgery	As above, but assuming no emergency surgery in the model	£8,719	£618,068
Central read NMA	Central read was the primary endpoint in OCTAVE trials.	£9,534	£187,809
Discounting every cycle	It tested the sensitivity of the model when the discounting of outcomes is applied every 8 weeks.	£8,616	£617,451
Vedolizumab dose 300 mg Q4W	A 300 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients who have a body weight ≥ 80 kg	£8,564	Tofacitinib dominant

Table 85 Scenario analyses, company base case (ERG corrected) – TNFi-exposed

Scenarios	Assumption	ICER for Tofacitinib vs.	
		CT	Vedolizumab
Base case		**£10,311**	**£7,838,381**
Tofacitinib maintenance dose mix	[redacted] of patients receiving 5mg; [redacted] of patients receiving 10mg	£13,956	Tofacitinib dominant
Fixed baseline utility instead of age-adjusted	Assumption that patient quality of life stays constant over time.	£10,599	£6,502,288
OCTAVE trials utilities	EQ-5D data were collected in Tofacitinib Phase III clinical trials	£18,292	Tofacitinib dominant
Swinburn utilities	To compare with previous analyses	£14,501	£7,087,359
Emergency surgery from any state	Due to the uncertainty on the likely protection from acute events based on the level of response/remission, patients are assumed to undergo emergency surgery regardless of state membership	£9,971	£7,612,076
Emergency surgery from active UC only	As above but assuming response to treatment offers the same level of protection from acute events, as remission	£10,485	£6,780,235
No emergency surgery	As above, but assuming no emergency surgery in the model	£10,603	£6,781,118
Central read NMA	Central read was the primary endpoint in OCTAVE trials.	£10,798	Tofacitinib dominant
Discounting every cycle	It tested the sensitivity of the model when the discounting of outcomes is applied every 8 weeks.	£10,408	£8,260,662
Vedolizumab dose 300 mg Q4W	A 300 mg Q4W maintenance dose was assessed as part of the clinical trials and is recommended for consideration in some patients who have a body weight ≥ 80 kg	£10,311	Tofacitinib dominant

olsalazine & sulfasalazine). However, clinical advice to ERG is that most patients receive mesalazine in UK and the doses for active ulcerative colitis are potentially higher than specified in company base case. We view that the net effect on costs from incorporating the changes in base case is likely to be neutral.

Treatment waning of effects and discontinuation

The company assumes treatment effect to be maintained with ongoing treatment and non-responders are given conventional therapy as second-line. The ERG agrees with company's approach to allow discontinuation for failure to respond in induction or loss of response in maintenance, based on the independent economic analysis in NICE TA329. We note this assumes that in practice, patients who experience exacerbations of symptoms can be assessed and, if appropriate, treatment stopped within 8 weeks. However, the model does not reflect NICE recommendations for annual assessment of benefit and need for continued treatment in previous appraisals TA329 and TA342. Clinical advice suggests that withdrawal of treatment for patients in remission is unlikely in practice, and the effects of this are difficult to quantify given the model structure and limited evidence over long-term maintenance of remission.

The company model applies a constant risk of relapse across each 8-week cycle of maintenance, with treatment stopping immediately when patients lose response. Thus, it assumes that maintenance treatment is stopped within 8 weeks of a loss of response. We consider this assumption to reflect UK practice. However, we have concerns that the costs of monitoring and follow-up in the company's model do not reflect the full cost of ensuring that treatment can be withdrawn within 8 weeks of a relapse. We address this by considering additional costs for outpatient visits to enable treatment cessation within 8 weeks of a relapse in our additional analyses.

Source of clinical effectiveness estimates

- *Choice of NMA models for economic analysis*

In general, we agree with company's approach to use locally-read clinical definitions of response and remission in economic model. The company states that their choice of NMA models was based on DIC measures of model fit, but they preferred the simpler fixed effect approach when DIC statistics were similar. In the case of the NMA for the TNFi-naïve population in the maintenance phase the fixed effect model was preferred because the

company thought the random effects results were implausibly imprecise with no treatment being predicted to be significantly better than placebo. The ERG has a general preference for the random effect NMA models, as we believe that the fixed effect models may underestimate uncertainty due to heterogeneity between the studies. We test the impact of different NMA models on cost-effectiveness results in our additional analyses.

- *Combination of TNFi-failed and TNFi-exposed subgroups*

The base case NMAs combine outcomes for subgroups defined as TNFi-failed for vedolizumab with TNFi-exposed subgroups for tofacitinib and adalimumab. We consider that combining results for TNFi-failed and TNFi-exposed subgroups is a potential source of bias in favour of tofacitinib. We conduct a scenario analysis using a more like-for-like comparison between tofacitinib and vedolizumab, using data for the TNFi-failed subgroups from the OCTAVE and GEMINI trials.

- *Transformation of NMA results to transition probabilities*

The company transformed the results of the clinical response/remission NMAs from the probit scale to the natural scale and converted to absolute probabilities for use in the model. They take a simpler approach by assuming constant ratio of patients in remission and response throughout maintenance phase and beyond in extrapolation. Clinical advice to the ERG is that these assumptions might not be realistic as clinical -experience indicates the risk is greatest in the first 6-12 months; and falls thereafter. The proportion of patients with response and remission is likely to increase over time as per our clinical advice. This is because responders (without remission) are more likely to stop or switch therapy whereas those in remission would continue with treatment. However, in the absence of evidence it is difficult to adapt the model. Therefore, we conclude that the model assumption of constant risk of loss of response for patients on maintenance treatment does not reflect clinical experience. Extrapolation of relapse and discontinuation rates from the maintenance trials is likely to underestimate the average duration of treatment and hence both the costs and QALYs of active treatments. However, it is not possible to estimate the net direction of bias in ICERs between comparators, because trends in long-term risks may vary between TNFi drugs, vedolizumab and tofacitinib.

- *Exclusion of other serious adverse events*

The company excluded adverse events other than serious infections. We agree that there would have been a risk of double-counting the costs and effects of ulcerative colitis exacerbations had

Table 111 ERG base case: drug sequencing scenarios (tofacitinib PAS, others at list price)

	Treatments	Total costs	Total QALYs	Fully incremental analysis ICER (£ per QALY)
Treatment sequencing	***TNFi- naive***			
	Conventional	[redacted];	[redacted];	[redacted];
	Gol-Ada-CT	[redacted];	[redacted];	[redacted];
	Inf-Ada-CT	[redacted];	[redacted];	[redacted];
	Ada-Ved-CT	[redacted];	[redacted];	[redacted];
	Ada-Tof-CT	[redacted];	[redacted];	[redacted];
	Gol-Ved-CT	[redacted];	[redacted];	[redacted];
	Gol-Tof-CT	[redacted];	[redacted];	[redacted];
	Inf-Ved-CT	[redacted];	[redacted];	[redacted];
	Inf-Tof-CT	[redacted];	[redacted];	[redacted];
	Tof-Ada-CT	[redacted];	[redacted];	[redacted];
	Ved-Ada-CT	[redacted];	[redacted];	[redacted];
	TNFi-exposed			
	Conventional	[redacted];	[redacted];	[redacted];
	Ada-Ved-CT	[redacted];	[redacted];	[redacted];
	Ved-Ada-CT	[redacted];	[redacted];	[redacted];
	Ada-Tof-CT	[redacted];	[redacted];	[redacted];
	Tof-Ada-CT	[redacted];	[redacted];	[redacted];