NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ozanimod for treating moderate to severe ulcerative colitis

Committee Papers

© National Institute for Health and Care Excellence [2022]. All rights reserved. See Notice of Rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Ozanimod for treating moderate to severe ulcerative colitis [ID3841]

Contents:

The following documents are made available to consultees and commentators:

The **final scope** and **final stakeholder list** are available on the NICE website.

1. Company submission from Celgene Ltd – A BMS Company

2. Clarification questions and company responses

3. Patient group, professional group and NHS organisation submissions from:

a. Crohn's & Colitis UK
b. British Society of Gastroenterology
c. UK Clinical Pharmacy Association

4. Evidence Review Group report prepared by Peninsula Technology Assessment Group (PenTAG)

5. Evidence Review Group report – factual accuracy check

6. Technical engagement response from company

7. Technical engagement responses and statements from experts:

a. Nancy Greig – Patient Expert, nominated by Crohn's & Colitis UK

8. Technical engagement responses from consultees and commentators:

a. Janssen

9. Evidence Review Group critique of company response to technical engagement prepared by PenTAG

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

© National Institute for Health and Care Excellence [2022]. All rights reserved. See Notice of Rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Ozanimod for treating moderately to severe ulcerative colitis

[ID3841]

Document B

Company evidence submission

17th January 2022

File name	Version	Contains confidential information	Date
ID3841_Ozanimod_UC_Document B_ACIC	Final	Yes	17th January 2022

©Celgene Ltd, a BMS Company (2022). All rights reserved

Contents

©Celgene Ltd, a BMS Company (2022). All rights reserved

©Celgene Ltd, a BMS Company (2022). All rights reserved

List of tables

©Celgene Ltd, a BMS Company (2022). All rights reserved

©Celgene Ltd, a BMS Company (2022). All rights reserved

List of figures

©Celgene Ltd, a BMS Company (2022). All rights reserved

©Celgene Ltd, a BMS Company (2022). All rights reserved

Abbreviations

Abbreviation	Definition
5-ASA	5-aminiosalicyclic acid
ADA	Adalimumab
AE	Adverse event
AESI	Adverse events of special interest
AGC	American College of Gastroenterology
ANCOVA	Analysis of covariance
ASCA	Anti-saccharomyces cerevisiae antibodies
AZA	Azathioprine
BID	Twice a day
BNF	British National Formulary
BSC	Best supportive care
BSG	British Society of Gastroenterologists
CcT	Concomitant Therapy
CD4	Cluster of differentiation 4
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
CODA	Convergence diagnosis and output analysis
CRC	Colorectal cancer
CRF	Case report form
CrI	Credible intervals
CRP	C-reactive protein
CSR	Clinical study report
CUA	Cost utility analysis
CvT	Conventional therapy
DIC	Deviance information criteria
DSA	Deterministic sensitivity analyses
DSU	NICE Evidence Synthesis Decision Support Unit
ECCO	European Crohn's and Colitis Organisation
ECG	Electrocardiogram
EIM	Extra-intestinal manifestations
EMA	European Medicine Agency
eMIT	Electronic Market Information Tool
EPAR	European public assessment report
EQ-5D	European quality of life-5 dimensions
ERG	Expert review group
ESR	Erythrocyte sedimentation rate
ESS	Effective sample size
EU	European Union
FDA	Food and Drugs Administration

©Celgene Ltd, a BMS Company (2022). All rights reserved

GGT	Gamma-glutamyl transferase
GI	Gastrointestinal
GOL	Golimumab
HBV	Hepatitis B virus
HCl	Hydrochloride
HCRU	Healthcare resource use
HRQoL	Health-related quality of life
HRU	Health resource utilisation
HTA	Health Technology Assessment
IBD	Inflammatory bowel disease
IBDQ	Inflammatory Bowel Disease Questionnaire
ICER	Incremental cost-effectiveness ratio
IFX	Infliximab
IgG	Immunoglobulin G
ITT	Intention to treat
IV	Intravenous administration
IVRS/IWRS	Interactive voice/web-based activated response system
JAK	Janus kinase
LSMD	Least squares mean difference
LYG	Life years gained
MAA	Marketing authorisation application
MCID	Minimum clinically important difference
MCS	Mental composite summary score
MCSE	Monte Carlo standard error
MedDRA	Medical Dictionary for Regulatory Activities
MIMS	Monthly Index of Medical Specialities
NA	Not available
NCT	National Clinical Trial
NHB	Net-health benefit
NHS	UK National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NRI	Non-responder imputation
NSAID	Nonsteroidal anti-inflammatory drug
OLE	Open-label extension
OR	Odds Ratio
OZA	Ozanimod
P-ANCA	Perinuclear anti-neutrophil cytoplasmic antibodies
PAS	Patient access scheme
PBO	Placebo
PCS	Physical component summary
PGA	Physician global assessment
PP	Per protocol

©Celgene Ltd, a BMS Company (2022). All rights reserved

PRO	Patient reported outcome
PSA	Probabilistic sensitivity analyses
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life year
RBS	Rectal bleeding subscore
RCT	Randomised controlled trial
RNR	Response no remission
RPLS	Reversible posterior leukoencephalopathy syndrome
RRMS	With relapsing remitting multiple sclerosis
RWE	Real world evidence
SC	Subcutaneous injection
SD	Standard deviation
SFS	Stool frequency subscore
SLR	Systematic literature review
SMD	Standardised mean difference
SmPC	Summary of product characteristics
SMR	Standard mortality ratios
SoC	Standard of care
SW	South-west
TAs	Technology appraisal
TBC	To be confirmed
TEAE	Treatment emergent adverse event
TEAESI	Treatment emergent adverse event of serious importance
TNFi	Tumour necrosis factor-alpha inhibitor
TOF	Tofacitinib
TSD	Technical Support Document
UC	Ulcerative colitis
UK	United Kingdom
UST	Ustekinumab
VAS	Visual analogue scale
VAT	Value added tax
VDZ	Vedolizumab
VZV	Varicella zoster virus
WTP	Willingness-to-pay threshold

©Celgene Ltd, a BMS Company (2022). All rights reserved

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

B.1.1 *Decision problem*

The objective of this appraisal is to determine the clinical and cost-effectiveness of ozanimod within its full marketing authorisation 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 (CvT) or a biologic agent. The decision problem addressed within this submission is consistent with the NICE final scope for this appraisal and any differences between the decision problem addressed within this submission and the NICE final scope are outlined in Table 1.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with moderately to severely active UC who are intolerant of, or whose disease has had an inadequate response, or loss of response to previous biologic therapy (a tumour necrosis factor-alpha inhibitor [TNFi], ustekinumab or vedolizumab), a JAK inhibitor (tofacitinib), or CvT (oral corticosteroids and/or immunomodulators)	Adults with moderately to severely active UC who have had an inadequate response, lost response, or were intolerant to either CvT or a biologic agent This comprises two mutually exclusive sub-populations: • TNFi-naïve: patients who have not previously received a TNFi • TNFi-experienced: patients who have previously received a TNFi and experienced treatment failure due to intolerance, lack of treatment efficacy or loss of response	The population addressed in the submission is in line with the final scope. TNFis are typically used as the first biologic treatment in patients who are intolerant or have had an inadequate response, or loss of response to CvT.[1] As a result, exposure to TNFis forms the basis for clinical decision-making, with treatment options differing in two distinct sub-populations: TNFi-naïve and TNFi-experienced. This is reflected in the NICE restriction on the use of ustekinumab and is in line with the current use of other biologic treatments in UK clinical practice.[1]
Intervention	Ozanimod	Ozanimod[a]	As per the final NICE scope
Comparator(s)	Current clinical management including: • TNFi (infliximab, adalimumab and golimumab) • Vedolizumab • Ustekinumab • Tofacitinib • Conventional therapies (aminosalicylates, oral corticosteroids and/or immunomodulators), without biological treatments	The submission population has been split into two distinct subpopulations: TNFi-naïve and TNFi-experienced. The relevant comparators differ in these two populations: • TNFi-naive: o Infliximab (and associated biosimilars) o Adalimumab (and associated biosimilars) o Vedolizumab • TNFi-experienced: o Vedolizumab o Ustekinumab	The SmPC for ozanimod states that patients must have failed CvT or a biologic. As biologics are only offered after failure on CvT in clinical practice, CvT is not viewed as a relevant comparator to ozanimod in either population. **TNFi-naïve:** • Following failure with CvT the majority of patients are initially treated with TNFis • As a result, whilst the NICE recommendation for vedolizumab and tofacitinib do not restrict their use in patients who have failed, cannot tolerate or are unsuitable for TNFis, neither tofacitinib nor vedolizumab are typically used first line in TNFi-naïve patients. This was supported by feedback received from clinical consultation conducted as part of this appraisal • TNFis are not suitable for all patients and

©Celgene Ltd, a BMS Company (2022). All rights reserved

			vedolizumab may be used in a small proportion of TNFi-naïve patients who are contraindicated to TNFis or have specific safety concerns surrounding their use • TNFis and vedolizumab have therefore been considered as relevant comparators in the TNFi-naïve population **TNFi-experienced:** • In line with the NICE final scope both ustekinumab and vedolizumab were considered relevant comparators in the TNFi-experienced populations • Neither tofacitinib or TNFis were considered relevant comparators in the TNFi-experienced population • Tofacitinib was not viewed as a relevant comparator as, in line with the opinion of clinicians consulted in TA633, feedback from clinical consultation received as part of this appraisal noted that whilst tofacitinib may be effective for some patients, concerns regarding its safety profile mean it is not typically used as a first line treatment option in TNFi-experienced patients. There has been no downgrading in the EMA warnings and restrictions associated with tofacitinib since the ustekinumab submission.[2] The restricted use of tofacitinib combined with concerns of its safety profile negates it as a standard comparator to ozanimod in this population (Section B.1.3.4) • TNFis were not considered relevant comparators in the TNFi-experienced population as TNFi switching is no longer routine clinical practice. As a result, receiving a second TNFi is only clinically relevant in a small proportion of TNFi-experienced patients. The exclusion of TNFis is in line with the accepted assumption in TA633
Outcomes	Outcome measures include: • Mortality	Outcome measures include: • Measures of disease activity; change	Mortality, rates of hospitalisation and rates of surgical intervention were not primary or secondary endpoints

©Celgene Ltd, a BMS Company (2022). All rights reserved

	• Measures of disease activity • Rates of and duration of response, relapse and remission • Rates of hospitalisation • Rates of surgical intervention • Endoscopic healing • Mucosal healing (combined endoscopic and histological healing) • Corticosteroid-free remission • Adverse effects of treatment • Health-related quality of life	in the 3-component Mayo score • Rates of and duration of response, relapse and remission • Endoscopic healing • Mucosal healing (combined endoscopic and histological healing) • Corticosteroid-free remission • Adverse effects of treatment • Health-related quality of life	in TRUENORTH. Data were therefore only collected on these events when assessing adverse events.
Economic analysis	• The cost-effectiveness of treatments is expressed in terms of incremental cost per quality-adjusted life year • The time horizon for estimating cost-effectiveness was set at a lifetime horizon to sufficiently reflect any differences in costs or outcomes between the technologies being compared • Costs are considered from a NHS and Personal Social Services perspective • The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account	As per final scope and NICE reference case	In line with the NICE final scope
Subgroups to be considered	If the evidence allows the following subgroups will be considered: • People who have been previously treated with one or more biologic • People who have not received prior biologic therapy	Clinical consultation conducted as part of this appraisal indicated that exposure to TNFis forms the basis for clinical decision-making, with treatment options differing in two distinct sub-populations: • TNFi-naïve • TNFi-experienced	• Economic analyses were conducted for ozanimod for sub-populations based on prior TNFi exposure owing to the relevant comparators differing between these sub-populations. These analyses informed the base case cost-effectiveness analysis for comparisons versus infliximab, adalimumab, golimumab and vedolizumab (in TNFi-naïve patents) and vedolizumab and ustekinumab (in TNFi-experienced patients)

©Celgene Ltd, a BMS Company (2022). All rights reserved

			• Subgroup analyses were informed by the Network Meta-analysis (NMA). The efficacy of ozanimod in the NMA was based on the subgroups of TRUENORTH stratified by TNFi exposure.

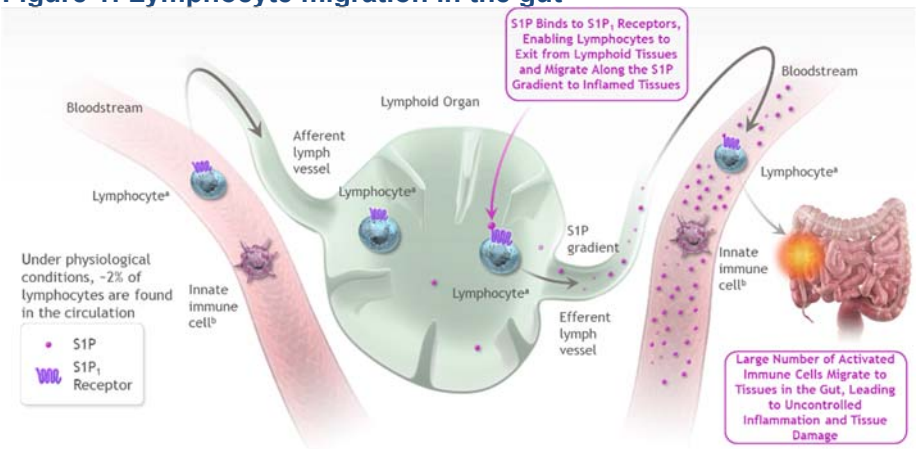
[a]Ozanimod presents in three distinct capsule strengths each with two reportable weights (ozanimod hydrochloride 0.25 mg,0.50 mg, and 1.0 mg, which are equivalent to ozanimod 0.23 mg, 0.46 mg, and 0.92 mg, respectively).
Abbreviations: NHS: National Health Service; NICE: National Institute for Health and Care Excellence; TNFi: tumour necrosis factor-alpha inhibitor; UC: ulcerative colitis; UK: United Kingdom.
Sources: Ozanimod NICE final scope [ID3841][3]; Sanborn *et al.* 2021.[4] TRUENORTH CSR.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

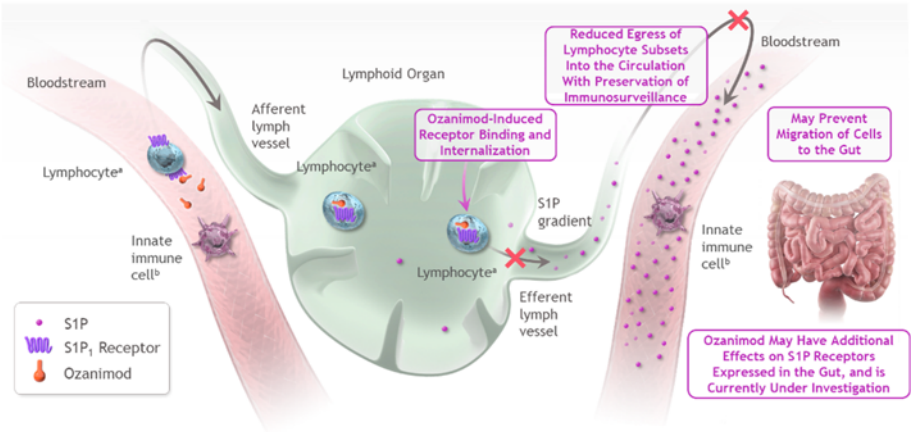
B.1.2 *Description of the technology being appraised*

A summary of the mechanism of action, marketing authorisation status, costs and administration requirements associated with ozanimod are presented in Table 2.

Table 2: Technology being appraised

UK approved name and brand name	Ozanimod (Zeposia®)
Mechanism of action	Ozanimod is an oral small-molecule, which acts as a selective sphingosine 1-phosphate (S1P) receptor modulator with specificity for receptor subtypes 1 and 5.[6] S1P1 signalling regulates the trafficking of lymphocytes out of secondary lymphoid tissue into the blood and plays an essential role in lymphocyte trafficking to sites of inflammation such as the gastrointestinal (GI) tract.[7, 8] In UC, the pathogenetic feature is inflammation of the colon which is caused by large numbers of activated inflammatory cells migrating to the gut.[8, 9] Inflammation and subsequent tissue ulceration result from failure to control the inflammatory response leading to symptoms of disease. **Figure 1: Lymphocyte migration in the gut**  [a]Including T cells and B cells. [b]Innate immune cells, which are responsible for antigen presentation and immunosurveillance, include macrophages, monocytes and natural killers cells, among others. **Abbreviations:** CD4: cluster of differentiation 4; HEV: high endothelial venules; NK: natural killer cells; S1P: sphingosine 1-phosphate. **Source:** Celgene (a BMS company). Data on file. When ozanimod or its active metabolites bind to the S1P1 receptors on lymphocytes, this causes them to become less responsive to the S1P gradient along which lymphocytes migrate thus interrupting lymphocyte trafficking. (Figure 2).[6] Reducing migration of lymphocytes to the intestinal mucosa to reduce inflammation represents a mechanism by which ozanimod exerts its therapeutic effect in UC.[8] Ozanimod may have additional effects on S1P receptors expressed in the gut.[10] Ozanimod offers a novel mechanism of action by targeting lymphocyte trafficking upstream. Currently licensed therapies such as anti-integrins, biologics and JAK inhibitors target inflammatory molecules downstream.

©Celgene Ltd, a BMS Company (2022). All rights reserved

	Figure 2: Effect of ozanimod on lymphocyte migration  [a]Including T cells and B cells. [b]Innate immune cells, which are responsible for antigen presentation and immunosurveillance, include macrophages, monocytes and natural killers cells, among others. **Abbreviations:** CD4: cluster of differentiation 4; HEV: high endothelial venules; NK: natural killer cells; S1P: sphingosine 1-phosphate. **Source:** Celgene (a BMS company). Data on file.
Marketing authorisation/CE mark status	A marketing authorisation application (MAA) for ozanimod for the treatment of adult patients with moderately to severely active UC who have inadequate response, lost response, or were intolerant to either CvT or a biologic agent was submitted to the European Medicines Agency (EMA) via the centralised procedure on 15th November 2020 and a positive opinion from the Committee for Medicinal Products for Human Use (CHMP) was received on 14th October 2021. Marketing authorisation approval for ozanimod in this indication was received from EMA on 18th November 2021.[11] Marketing authorisation approval for ozanimod in this indication from the MHRA is expected in [redacted] 2022.
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)[12]	• The EU and UK marketing authorisation wording for ozanimod in the indication of interest for this submission is: "For the treatment of adult patients with moderately to severely active UC who have had an inadequate response, lost response, or were intolerant to either CvT or a biologic agent" • The Summary of Product Characteristics (SmPC) for ozanimod in this indication are provided in Appendix C **Existing indications for ozanimod:** Ozanimod has existing marketing authorisations from the EMA in the following indication: • "Adult patients with relapsing remitting multiple sclerosis (RRMS) with active disease as defined by clinical or imaging features" **Contraindications of ozanimod:** • Contraindications for ozanimod are in line with existing therapies in this indication. Full details of contraindications can be found in the SmPC[12]
Method of administration and dosage[12]	**Induction period:** • Dose escalation (1 week): 0.25 mg[a] once daily (OD) taken orally on Days 1–4, increased to 0.50 mg[a] OD taken orally on Days 5–7 • 1.0 mg[a] OD taken orally starting on Day 8 **Maintenance period:** • 1.0 mg[a] OD taken orally

©Celgene Ltd, a BMS Company (2022). All rights reserved

<table>
<tr><td>Additional tests or investigations[12]</td><td>
Before first dose:

Perform baseline electrocardiogram (ECG)
If clinically indicated, arrange ophthalmological assessment before starting treatment in patients with diabetes mellitus, uveitis, or a history of retinal disease
A negative pregnancy test result in women of childbearing potential must be confirmed prior to starting treatment
As per routine practice for starting advanced treatments, check varicella zoster virus (VZV) antibody status in patients without a healthcare professional confirmed history of varicella or documentation of a full course of varicella vaccination. If negative, VZV vaccination is recommended at least 1 month prior to treatment initiation

Until 6 hours after first dose for patients requiring first dose observation:

In patients with certain pre-existing cardiac conditions (resting heart rate <55 bpm, second-degree [Mobitz type I] AV block or a history of myocardial infarction or heart failure):

Monitor for 6 hours after the first dose for signs and symptoms of symptomatic bradycardia, with hourly pulse and blood pressure measurement
Perform an ECG prior to and at the end of the 6-hour monitoring period

Extended monitoring may be required in the following situations if at hour 6 post-dose

Heart rate less than 45 bpm
The recorded heart rate is the lowest value in the total 6-hour post-dose period, suggesting that the maximum decrease in heart rate may not have occurred yet
Evidence of a new onset second-degree or higher AV block at the 6-hour post dose ECG
QTc interval ≥500 msec

</td></tr>
<tr><td>List price and average cost of a course of treatment</td><td>
The list prices of ozanimod are as follows (excluding VAT):
<table>
<tr><th>Pack</th><th>List price (£)</th></tr>
<tr><td>Initiation pack (4 x 0.25[a] mg & 3 x 0.50[a] mg)</td><td>£343.25</td></tr>
<tr><td>Maintenance pack (28 x 1.0[a] mg)</td><td>£1,373.00</td></tr>
<tr><td>Maintenance pack (98 x 1.0[a] mg)</td><td>£4,805.50</td></tr>
<tr><td>Annual cost[b] – induction (Year 1)</td><td>£17,910.29</td></tr>
<tr><td>Annal cost[b] – maintenance (Year 2 and onwards)</td><td>£17,910.29</td></tr>
</table>
</td></tr>
<tr><td>Patient access scheme (if applicable)</td><td>
[redacted]
<table>
<tr><th>Pack</th><th>PAS price (£)</th></tr>
<tr><td>Initiation pack (4 x 0.25[a] mg & 3 x 0.50[a] mg)</td><td>[redacted]</td></tr>
<tr><td>Maintenance pack (28 x 1.0[a] mg)</td><td>[redacted]</td></tr>
<tr><td>Maintenance pack (98 x 1.0[a] mg)</td><td>[redacted]</td></tr>
</table>
</td></tr>
</table>

©Celgene Ltd, a BMS Company (2022). All rights reserved

	Annual cost[b] – induction (Year 1)	[redacted]
	Annual cost[b] – maintenance (Year 2 and onwards)	[redacted]
	A confidential PAS is also available for the following relevant comparators to ozanimod: golimumab, vedolizumab, ustekinumab and tofacitinib. Since the PAS prices are not available to BMS, all results presented in the submission include the relevant comparators at list price.	

[a]Dosing based on dose of ozanimod hydrochloride (HCl). 0.25 mg, 0.50 mg and 1.0 mg of ozanimod HCl equivalates to 0.23 mg, 0.46 mg and 0.92 mg of ozanimod, respectively.
[b]Based on 365.25 days.
Source: Celgene (a BMS Company). Data on File. Draft SmPC for Zeposia®.[12]
Abbreviations: AV: atrioventricular block**;** CHMP: Committee for Medicinal Products for Humans Use; ECG: electrocardiogram; EMA: European Medicines Agency; EPAR: European public assessment report; EU: Europe; MAA: marketing authorization application; PAS: patient access scheme; QD: once daily; RRMS: relapsing remitting multiple sclerosis; S1P receptor: sphingosine 1-phosphate receptor; SmPC; summary of product characteristics; UC: ulcerative colitis.

©Celgene Ltd, a BMS Company (2022). All rights reserved

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

Disease overview

- UC, a type of inflammatory bowel disease (IBD), is a chronic, relapsing condition characterised by diffuse inflammation of the rectal and colonic mucosa.[13, 14] It is the most common form of IBD, with a prevalence in the UK of 570 per 100,000 people, equating to approximately 390,000 people currently living with the disease. The condition can develop at any age, but peak incidence is between the ages of 15 and 35 years[15-17]
- UC is classified according to disease severity and extent. Disease severity is categorised as mild, moderate or severe.[16] Categorisation can be achieved through the use of scoring systems, the most widely used of which include The Mayo Scoring system and The Truelove and Witts' severity index system.[18-20] Patients on the boundary between two categories are classed as 'mild to moderate' or 'moderate to severe'[1]
- Of the estimated 390,000 people in the UK living with UC, approximately 52% are categorised as having moderately to severely active UC[1]
- Patients with UC typically present with rectal bleeding, faecal urgency, diarrhoea and lower abdominal pain.[21, 22] Some patients may have symptoms affecting other parts of the body such as the joints, eyes, skin and liver, known as extra-intestinal manifestations (EIMs). [23, 24]
- Symptoms have a detrimental impact on patients' lives, preventing them for leading a 'normal' life when compared to people of a similar age, socioeconomic status and geographical region. Indeed, living with moderately to severely active UC has been found to be both physically and emotionally disabling[25, 26]
- UC also impacts the lives of patients friends and family who frequently serve as informal carers, particularly in the more severe stages of disease.[27, 28] Carers of patients with chronic diseases are placed under significant financial, psychological, and physical strain, which increases with disease severity[28, 29]

Current treatment pathway and position of the technology

- The treatment pathway followed by patients with UC is individualised, depending on the severity and extent of the disease, response to previous medication and patient preference.[16, 30] Guidance on the typical clinical pathway of care for patients with UC in the UK is provided in NICE Guideline 130.[16]
- Mild to moderate disease is primarily managed with CvT such as aminosalicylates, corticosteroids and thiopurines
- The management of moderately to severely active UC is more challenging and non-CvT including biologics, such as TNFis, may be used in patients who have had an inadequate response with, lost response to, or were intolerant of conventional therapies[16]
- Patients who ultimately fail to respond to medical therapies may undergo surgery (colectomy). Surgery is viewed as an undesirable treatment option by the majority of patients owing to the associated risk of life-long, irreversible consequences including risk of permanent stoma formation and decreased fertitlity.[31, 32] Despite this, approximately 10–15% of patients have to undergo surgical resection in their lifetime[33]
- The availability of biologics has resulted in improvements in disease management and quality of life for patients with moderately to severely active UC. However, given the life-long relapsing and remitting nature of UC there is still a substantial unmet need among patients with moderately to severely active UC for additional, safe, effective and easily administrable therapies, that can not only induce remission but maintain patients in response over the long term so they do not have to resort to surgery.
- Ozanimod satisfies this need, with a novel mechanism of action, convenient oral method of administration and tolerable safety profile

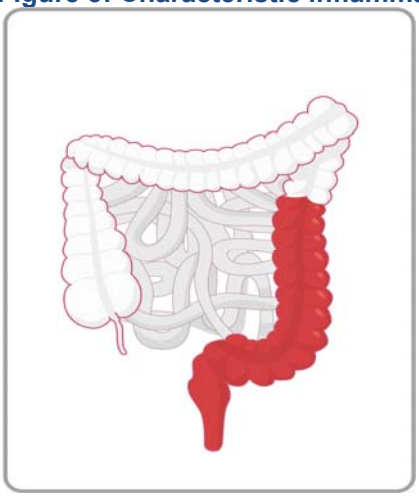
©Celgene Ltd, a BMS Company (2022). All rights reserved

B.1.3.1 Disease overview

UC is a type of inflammatory bowel disease (IBD) and is a chronic, lifelong condition characterised by diffuse inflammation of the rectal and colonic mucosa. Inflammation typically starts in the rectum and lower colon and may then extend to involve the entire colon (Figure 3).[13, 14] As a result, ulcers develop on the surface of the lining of the colon which may bleed and produce pus.[34]

UC follows a relapsing-remitting course, whereby symptoms are apparent when the disease is active and abate in periods of remission. Approximately 30–60% of patients with UC have at least one relapse per year.[16, 35] Clinicians are currently unable to predict which patients will flare and it is only with passage of time that the pattern of a patient's disease becomes apparent. Despite this, patients with a younger age of disease onset typically have a higher relapse rate.[36] More frequent or severe flares are indicative of a more difficult to control disease and therefore such patients receive more frequent treatment escalation or greater use of steroids.[37] Patients with UC symptoms are said to have 'active' UC, while patients in remission periods (no symptoms) are typically referred to as having 'inactive' UC.[38]

Figure 3: Characteristic inflammation in UC



Cause of UC

The exact aetiology of UC is not completely understood, meaning curative medical therapies are not currently available, with treatment instead focussing on symptom management.[23, 39] The pathogenesis of UC is complex and is thought to result from an interplay of several factors including genetic pre-disposition, environmental factors, a dysregulated immune response and defects in the colonic epithelial barrier (protective lining).[40]

It is accepted that UC occurs in genetically susceptible individuals with environmental influences that result in a dysregulated immune response to commensal intestinal microbiota. This leads to a release of cytokines and chemokines leading to colonic inflammation.[1, 41] The dysregulated immune response results in a retention of circulating leukocytes in the inflamed intestinal mucosa leading to the development of chronic inflammation. The complexity in disease pathogenies contributes to variation in disease presentation and response to treatment.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Given the differences in individual responses to inflammation it is important to have novel treatments, especially those that move beyond targeting individual molecules and can lead to a reduction in inflammation through other mechanisms such as lymphocyte trafficking.

Epidemiology

UC is the most common form of IBD in the UK, with an incidence of 10–14 per 100,000 people and a prevalence of approximately 570 per 100,000 people.[17, 42] Based on the latest UK population estimates this equates to approximately 4,450–6,220 adult patients being diagnosed with UC every year and 390,000 people currently living with UC in the UK.[16, 43]

The condition can develop at any age, but peak incidence is between the ages of 15 and 35 years (with a second, smaller peak between the ages of 55 and 65 years).[16] Thus the burden of disease and associated disability can impact patients from a young age and continue throughout their working life. The disease affects men and women at similar rates, however men are more likely to be diagnosed with UC in their 50s and 60s than women.[44-46]

B.1.3.2 Disease burden

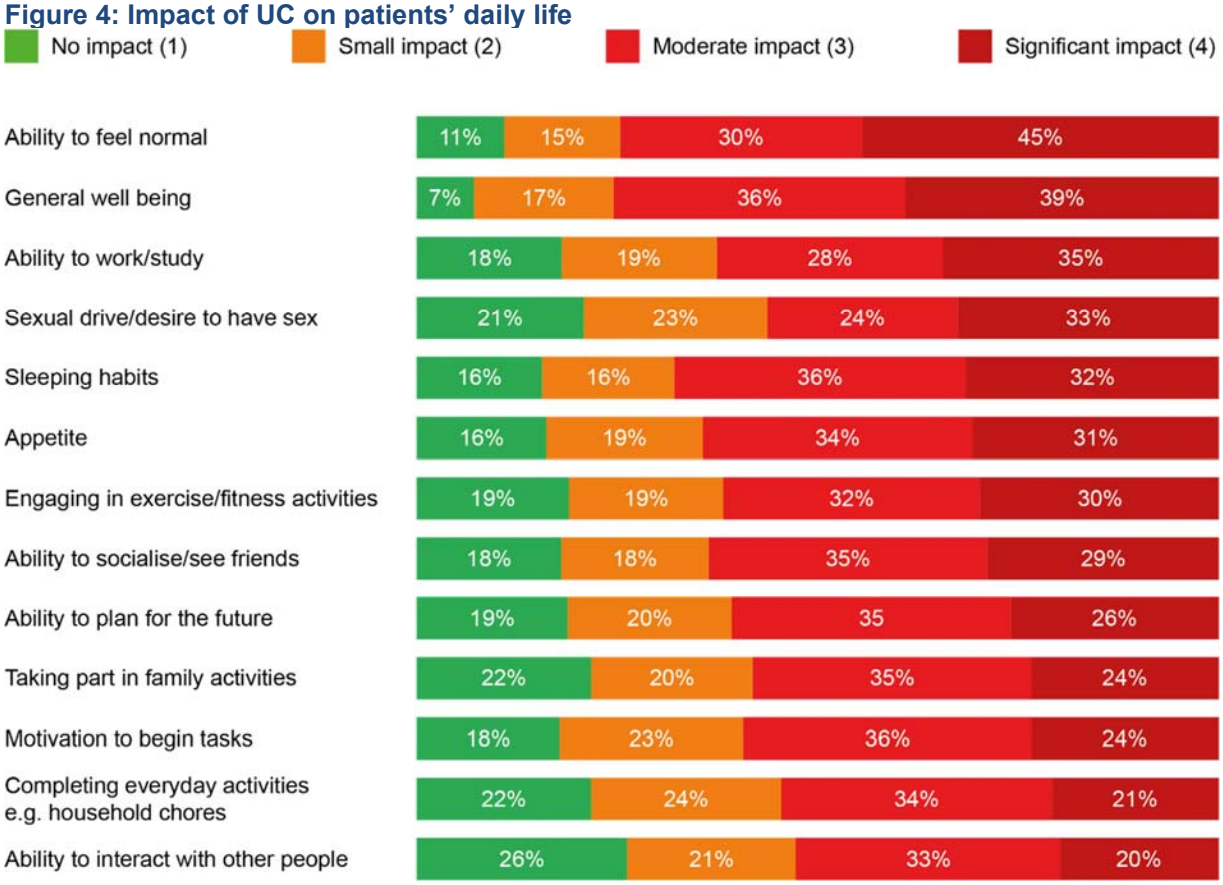
Symptoms of UC and impact on quality of life

Patients with UC typically present with rectal bleeding, faecal urgency, tenesmus (recurrent inclination to evacuate the bowl), diarrhoea, mucous discharge from the rectum and lower abdominal pain.[21, 22] In a UK-based survey of patients with IBD aged ≤29 years (N=1,081; 36.3% UC) extreme fatigue was the most frequently reported symptom (75.9%) followed by pain (49.4%) and diarrhoea (44%).[47] The survey also found that 78% of participants had experienced a disease flare that year with almost half being hospitalised in the same period.[47] Increased disease activity is associated with increased disease symptoms, and reduced patient health-related quality of life (HRQoL).[48] A cross-sectional UK study reporting outcomes for the Inflammatory Bowel Disease Questionnaire (IBDQ), which reports on four domains (bowel symptoms, systematic symptoms, emotional function and social function), found that the baseline score for patients with moderate to severe disease was significantly ($p<0.001$) reduced compared to patients with mild disease (116.41 versus 148.81, respectively), signifying worsening symptoms and reduced HRQoL.[32, 49, 50]

UC symptoms have a detrimental impact on patients' lives, preventing them for leading a 'normal' life when compared to people of a similar age, socioeconomic status and geographical region.[51] A study in patients aged 18–54 years with UC found that 89% of patients felt UC affected their ability to feel normal, with 45% reporting a significant impact (Figure 4). Additionally, almost 2/3 of patients felt the disease moderately or significantly impacted their ability to work/study.[52] Another study reported that 73% of UC patients felt their symptoms limited their ability to enjoy leisure activites.[53]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 4: Impact of UC on patients' daily life



Abbreviations: UC: ulcerative colitis
Source: Figure adapted from Research Partnership 2017: Living with Ulcerative Colitis[52]

The burden of UC also extends beyond the physical symptoms. The relapsing-remitting course of the disease means patients live in fear of flare-ups and the unpredictable nature of symptoms can lead to depression and anxiety. Indeed depression and anxiety are two of the most commonly reported concomitant conditions in patients with UC.[52] Further to this, living with the disease can cause patients increased stress with 72% of patients (N=256) in a European study saying the disease made their lives more stressful.[54] Patients also struggle to adapt to body image changes caused by adverse effects of medications and faecal incontinence and soiling can lead to a loss of self-worth.[55, 56] Furthermore, patients' constant fear of losing control of bowel movements means they struggle to attend social events, leading to an increased risk of self-isolation.[57] A large international survey (N=4670; 33% with UC) found that 35% of patients with IBD felt their disease prohibited them from pursuing an intimate relationship and 26% felt it prevented them from making or keeping friends.[58]

UC also impacts the lives of patients friends and family who frequently serve as informal carers, particularly in the more severe stages of disease.[27, 28] Carers of patients with chronic diseases are placed under significant financial, psychological, and physical strain.[28, 29] A study of carers of adult patients with IBD found carers experienced a similar level of burden to carers of patients with acute brain injury.[29] In addition, carers of patients with IBD can suffer from burnout symptoms related to

©Celgene Ltd, a BMS Company (2022). All rights reserved

persistent physical fatigue syndrome, energy loss and high physical stress.[28] Severe disease as well as symptom flares were found to be predictors for high caregiver burden and reduced HRQoL.[29, 59]

Collectively, these factors demonstrate the significant disease burden of UC on both patients and carers, which increases with disease severity.

Complications of UC

UC can be associated with a range of complications, including acute severe colitis, which typically requires hospitalisation and may progress to toxic megacolon (colonic dilatation associated with systemic toxicity). Patients experiencing toxic megacolon are at risk of colonic perforation and if they do not respond to medical therapies require an emergent colectomy.

Longer-term complications associated with UC include the development of colorectal cancer (CRC), the formation of strictures and a requirement for surgery (colectomy) in cases where the disease is refractory to medical therapy.[22, 60, 61] Patients with UC are twice as likely to develop CRC compared with the general population.[1, 62] The risk of developing CRC increases with both disease extent and duration and therefore it is thought the presence of chronic inflammation promotes carcinogensis.[62, 63] UC patients may also develop symptoms beyond the colon; these are known as extra intestinal manifestations (EIMs). EIMs can involve organ systems including musculoskeletal, skin, hepato-pancreato-biliary and eyes.[61] Examples include peripheral arthritis (joints), uveitis (eyes), erythema nodosum (skin) and primary sclerosing cholangitis (liver).[23, 64] Over 50% of patients with UC experience at least one EIM 30 years after diagnosis, with up to 25% experiencing more than one.[60, 61]

Surgery

As discussed above, complications associated with UC can result in patients having to receive surgery. Short and long-term complications of surgery are common and can have a profound impact on patients' lives as well as being a considerable expense for the healthcare system.[65, 66] Short-term complications, occurring within 30 days of a procedure, include infections (20%), ileus (18%), pouch-related complications (8%), small bowel obstructions (8%), anastomotic leakage (2%)[65, 67] Longer-term complications, occurring more than 30 days post-procedure, include pouchitis (29%), faecal incontinence (21%), small bowel obstruction (17%), ileus (11%), fistula (6%), and pouch failure (5%).[65] In patients who have restorative surgery following a colectomy short-term gains in HRQoL have shown to decrease over time as patients experience pouch failures, pouchitis, cuffitis and irritable pouch syndrome[62] Further to this, short-term improvements in HRQoL in 80% of patients were not sustained over the long term due to depression, body image, greater eating restrictions, sexual disfunction and reduced productivity.[32] Finally, surgery carries the associated risk of significant changes in sexual and reproductive function, in some cases leading to infertility.[31, 32] This is of particular concern for UC patients, who due to the early age of onset of UC, are typically sexually active and of child-bearing age.[68]

Economic burden

The early age of onset of UC, and the intractable chronic relapsing-remitting nature of the disease, which results in the requirement for ongoing follow-up and monitoring as well as patients having to cycle through various therapies, means that it poses a substantial burden to resource utilisation in

©Celgene Ltd, a BMS Company (2022). All rights reserved

the healthcare system.[42] In addition, owing to the early age of onset, UC negatively impacts patients' work life during the most productive years of work (18–64).[58, 69] Patients with UC may be forced to take significant periods of absence from work, as well as reduced working hours, due to the symptoms of the disease.[58, 70] Findings from a large international survey (N=4670; 33% with UC) showed that approximately 3 out of 4 patients with IBD missed work due to their illness and one-third of patients lost or quit a job. In particular, work productivity was found to be seriously impaired in patients with moderately to severely active UC, with productivity impairment increasing with disease severity.[58, 69] A US study incorporating chart reviews and patient questionnaires from 2015 to 2017 showed patients with UC categorised as 'moderate-to-severe' incurred 30 times the mean annual absenteeism costs as patients with 'mild' UC.[71] As a result, indirect costs from lost work productivity and presenteeism can account for up to 68% of total disease costs.[72]

Treatment with biologics often requires dose escalation to maintain treatment effect which leads to increased health expenditure. A UK retrospective cohort study in 2016 found approximately 43% of patients treated with adalimumab had their doses escalated by ≥100% during the maintenance period and incurred greater mean healthcare costs than those who did not have their dose escalated (£14,596 versus £13,351).[73] Additionally, a systematic literature review (SLR) collecting real-world evidence (RWE) from 48 studies investigating interventions for the treatment of moderately to severely active UC found that 35% and 33% of patients receiving vedolizumab and tofacitinib, respectively, receive dose escalation.[74]

Collectively, the direct medical costs and indirect costs associated with lost work productivity associated with UC represent a significant economic burden to society. The total economic burden of UC has been estimated at €12.5–29.1 billion in Europe, with direct medical costs ranging from €2,210 to €10,395 per patient per year.[72] Symptom flares have been found to lead to a 2–3-fold increase in healthcare costs for patients compared to those with stable disease. If hospitalisation is needed to control a flare this leads to a greater than 20-fold increase in costs; indeed, hospitalisation costs have been found to account for 41–55% of direct medical costs.[72] This reinforces the need for novel therapies that are capable of maintaining remission, thus limiting expensive flare-up and subsequent hospitalisation costs.

B.1.3.3 Disease diagnosis and classification

The onset of UC can be insidious and symptoms can often occur for weeks or months before medical advice is sought.[21] As UC symptoms can be non-specific, medical history and clinical evaluation alone are often not sufficient to form the basis of UC diagnosis. Typically, a combination of history of clinical symptoms, clinical evaluation, endoscopic and histological findings and the exclusion of other causes of colonic inflammation, such as infection, form the basis of diagnosis.[1, 75, 76]

Endoscopic findings aid disease diagnosis and may reveal mucosal changes characteristic of UC, including loss of typical vascular pattern, granularity and ulceration.[77-79] These findings allow for an assessment of disease extent, the severity of mucosal damage and hence disease severity.[23, 53, 80] Endoscopies also allow for biopsies and histological analysis which can be used to form an accurate diagnosis and assess level of histological inflammation.[81] Clinical investigations such as blood tests to check for biomarkers of inflammation such as C-reactive protein (CRP), and erythrocyte sedimentation rate (ESR) can be used to aid diagnosis and inform disease activity.[23, 79] Further

©Celgene Ltd, a BMS Company (2022). All rights reserved

details of clinical findings and investigative techniques used to assist disease diagnosis can be found in Table 3.

Table 3: Diagnostic features of UC

Feature	Description for UC
Clinical	Diarrhoea, bloody stools, abdominal pain, faecal urgency, tenesmus
Lab findings	Inflammatory markers: ESR, CRP, faecal calprotectin Full blood count, iron studies, albumin levels Serology: p-ANCA (+)
Endoscopic findings	Inflammation begins in the rectum and extends proximally in a circumferential and continuous fashion. Findings include erythema, oedema/loss of the fine vascular pattern, increased granularity of the mucosa, friability/spontaneous bleeding, pseudopolyps, erosions, and ulcers
Histological findings	Crypt abscesses, crypt branching, crypt shortening, crypt disarray, crypt atrophy, mucin depletion, paneth cell metaplasia, increased lamina propria cellularity, basal plasmacytosis, basal lymphoid aggregates, and lamina propria eosinophils

Abbreviations: ASCA: Anti-Saccaromyces cerevisiae antibodies; CRP: c-reactive protein; ESR: erythrocyte sedimentation rate; FC: faecal calprotectin; p-ANCA: perinuclear anti-neutrophil cytoplasmic antibodies.
Source: Gejendran 2019.[82]

Disease classification: extent

UC is classified according to disease extent and disease severity. The classification determines which treatment pathway a patient follows and therefore accurate disease classification is key to ensuring patients receive optimal treatment and are able to achieve the best outcomes.[16, 83]

The disease extent of UC is typically defined by the maximal macroscopic extent of the disease observed at colonoscopy. There are three main subgroups associated with disease extent: proctitis, left-sided colitis and extensive colitis (Table 4).[23] At disease presentation, typically 30–60% of patients with UC have proctitis, 16–45% have left-sided colitis and 14–35% have extensive colitis.[40] Due to the progressive nature of the disease, the extent of disease often increases with time. Studies have shown that UC progresses proximally in 10–19% of patients after 5 years and 28% of patients after 10 years.[14, 84] Disease progression in UC generally follows a gradual course with <20% of patients going from proctitis directly to extensive colitis.[82]

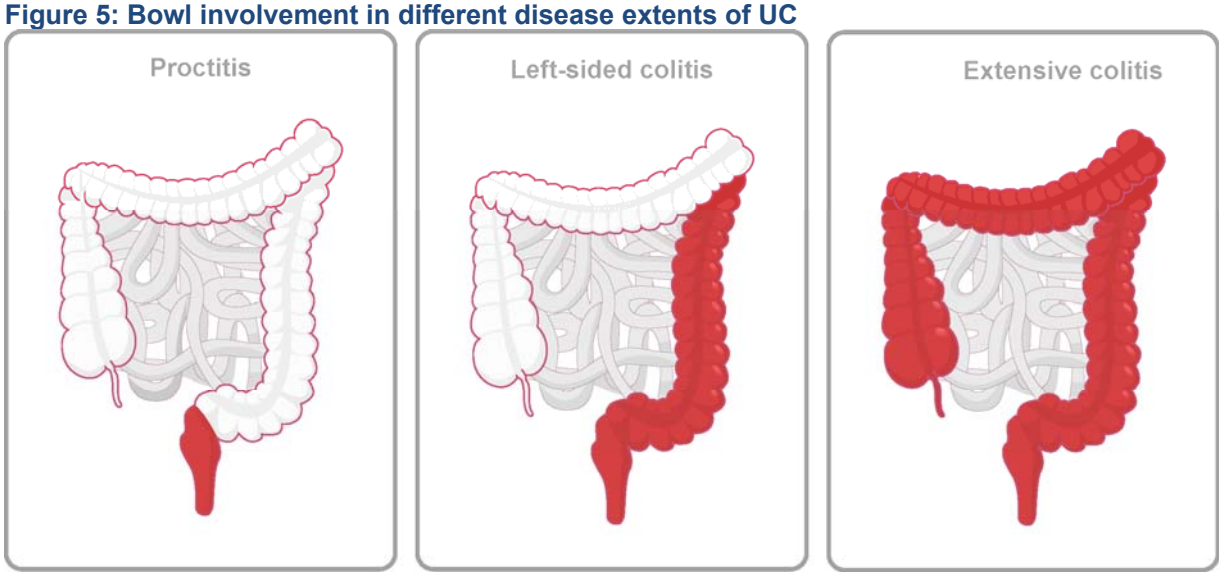
Table 4: Disease extent definition

Disease extent	Area of bowl involved
Proctitis	Involvement limited to the rectum (the distal 15 cm of the large intestine).
Left-sided colitis	Involvement limited to the left portion of the colon; inflammation extends from the rectum up the colon and stops at the splenic flexure
Extensive colitis	Inflammation extends from the rectum up the colon and beyond the splenic fixture and involves most or all of the colon.

Sources: Magro et al. (2017).[23] Unagro 2017.[40]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 5: Bowl involvement in different disease extents of UC



Disease classification: severity

UC is further categorised according to disease severity as either mild, moderate or severe. The British Society of Gastroenterology and the European Crohn's and Colitis Organisation (ECCO) categorise severity of disease activity based on clinical presentation.

Scoring systems

A number of scoring systems have been developed to measure disease severity, with most only being used in a clinical trial setting. The most widely used of these systems are the Mayo Scoring system and the Truelove and Witts' severity index system.[18, 20].

The Mayo Scoring system is the most widely used scoring system in both the clinical trial setting and in clinical practice.[1, 19, 85] The Mayo Scoring System was used to assess clinical response and clinical remission in the TRUENORTH trial, which forms the principal evidence base for this submission (Section B.2.3.1). There is a 4-component Mayo score, a 3-component Mayo score and a partial Mayo score (Table 5). The 4-component Mayo score ranges from 0–12 points and consists of the sum of four sub-scores: stool frequency, rectal bleeding, endoscopic findings and physician global assessment (PGA). Each subgroup is scored on a scale of 0–3 with a higher score indicating more severe disease (0: normal, 1: mild, 2: moderate and 3: severe).[5] A total score of ≥6 signifies moderately to severely active disease.

The 3-component score ranges from 0–9 points and does not include the PGA subscore.[86] The partial Mayo score also ranges from 0–9 points and is the sum of the rectal bleeding sub-score, stool frequency sub-score and the PGA subscore.[87]

The Mayo scoring system has been found to be clinically relevant owing to its correlation with both disease-specific health-related quality of life (HRQoL), as measured by the Inflammatory Bowel Disease Questionnaire (IBDQ) and generic quality of life scores, such as 36-item Short Form survey (SF-36).[88, 89] The 3-component Mayo score and the partial Mayo score have been found to be just as effective as the 4-component mayo score in identifying patient responses to therapies.[87]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 5: Mayo scoring system (4-component)

	Normal (0)	Mild (1)	Moderate (2)	Severe (3)
Number of bowel movements per day	Normal	1–2 stools more than normal for the patient	3–4 stools more than normal for the patient	≥5 stools more than normal for the patient
Rectal bleeding	None	Streaks of blood stool occurs less than half the time	Obvious blood with stool most of the time	Blood alone passes
Findings on endoscopy	Normal	Erythema, decreased vascular pattern, mild friability	Marked erythema, lack of vascular pattern, friability, erosions	Spontaneous bleeding, ulceration
Physician's global assessment	Normal	Mild	Moderate	Severe

Source: Lewis et al. (2008).[87]

The focus of UC management has begun to shift from just symptom alleviation to the longer term goal of clinical remission as evidenced by mucosal healing, which is the absence of endoscopic and/or histologic activity.[90] Both the American College of Gastroenterology (AGC) and The International Organisation for the Study of IBD (IOIBD) recognise mucosal healing as a treatment goal for patients with UC and studies have shown that mucosal healing is associated with lower rates of relapse, hospitalisation, colectomy, and colorectal cancer.[90-92] The Mayo endoscopy subscore is frequently used in a clinical trial setting to assess mucosal healing, where an endoscopy subscore of 0 symbolises normal mucosa.[90] Although mucosal healing is a goal for UC treatment, relapse can still occur in patients who reach this endpoint.[90] Moreover, microscopic inflammation can persist in the absence of endoscopically visible disease as histologic activity is present in approximately 25% of patients with a normal-appearing mucosa.[90] As a result, histologic activity should also be accounted for when assessing mucosal healing. The Geboes scoring system is used to measure histologic activity and has demonstrated good reproducibility.[93 94] The score ranges from 0–6 with a lower score representing reduced inflammation (Table 6).[94] In the TRUENORTH trial mucosal healing was assessed using a combination of both the Mayo endoscopic subscore and the Geboes scoring system (Section B.2.3.1).

Table 6: Geboes scoring system

Grade	Description
0	Structural change only
1	Chronic inflammation
2	Lamina propria neutrophils
3	Neutrophils in epithelium
4	Crypt destruction
5	Erosions and ulcers

Source: Geboes 2000.[94]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.1.3.4 Current treatment pathway

The most recent guidelines and treatment pathway for UC in the UK are available from NICE (NG130 [2019]) and The British Society of Gastroenterologists (BSG) [2019].[16, 37]

Treatment aims

Current UC therapies are not curative and treatment focuses on reducing colonic inflammation to alleviate symptoms and to drive disease into and maintain remission. The main aims of treating patients with UC are:[16, 95-97]

- To relieve symptoms and induce remission in an active period of disease
- To maintain remission in order to minimise subsequent flares and prevent development of long-term complications
- To improve patient HRQoL

Additional treatment aims include achieving corticosteroid-free remission, inducing mucosal healing (defined as endoscopic improvement +/- histologic remission) and preventing the need for a colectomy. Long term use of corticosteroids is associated with a multitude of significant adverse events and comorbidities including osteoporosis, pre-disposition to diabetes, weight gain, hypertension, risk of infection, cataracts and psychological effects.[98] Long-term corticosteroid free remission is associated with improved HRQoL in UC patients and a decreased risk of surgery.[23, 53] As mentioned in Section B.1.3.1, mucosal healing has been associated with lower rates of disease relapse, hospitalisation, colectomy, and colorectal cancer.[90, 92]

The treatment pathway followed by patients with UC is individualised and depends on the severity and extent of disease, patients' response to previous medication and patient preference.[16, 30] Guidance on the typical clinical pathway of care for patients with UC in the UK is provided in NG130 and is summarised below.[16]

Step 1: CvT

The first line of treatment for UC is CvT, which is typically used to manage patients with mild to moderate disease. CvT includes aminosalicylates, corticosteroids and thiopurines.[16] Whilst these therapies are effective for the majority of patients with mild UC, a proportion of patients will relapse and/or progress to moderately to severely active UC. Given the challenges associated with diagnosis of UC (see Section B.1.3.1), some patients may also have moderately to severely active UC at diagnosis. The management of moderately to severely active UC is more challenging and may require the use of advanced therapies, such as biologics.[16]

Step 2: Advanced/biologic therapy

When CvT cannot be tolerated or is contraindicated, or when patients have responded inadequately or lost response to treatment (see Table 7 for definitions) patients are treated with biological (sometimes referred to as advanced therapies). Treatment options at this stage recommended by NICE include: TNFis, (infliximab, adalimumab, and golimumab [TA329][35]), an interleukin 12/23 inhibitor (ustekinumab [TA633]), an anti-integrin (vedolizumab [TA342][85]) and a JAK inhibitor (tofacitinib [TA547]).[88] It is anticipated that ozanimod will be positioned at this point in the treatment pathway.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 7: Definitions of terms used in treatment pathway for patients with moderately to severely active UC

Term	Definition
Intolerant	Patient is unable to achieve therapeutic doses, or persist on treatment due to treatment-related side effects and/or laboratory abnormalities
Inadequate response	Patient does not achieve sufficient reduction in UC disease activity (e.g. defined by reductions in Mayo score and rectal bleeding subscore)
Lost response	Patient has recurrence of symptoms during maintenance dosing following prior clinical benefit
Contraindication	A condition or circumstance that suggests that a particular technology or drug should not be used for a given patient

Source: Adapted from TRUENORTH CSR.[5]

Treatment in Step 2 is individualised and is largely based on clinician assessment of need, cost-effectiveness and patient preference.[99, 100] The following describes the typical treatment pathway taken by a moderately to severely active UC patient in Step 2:

TNFis: The majority of patients with moderately to severely active UC receive TNFis as a first choice Step 2 treatment, due to the wealth of experience and familiarity with their use in practice, as well as the availability and affordability of biosimilar products for both infliximab and adalimumab.[1, 101].[68 35] Clinical consultation conducted as part of this appraisal noted that use of golimumab in UK clinical practice is limited.

Switching to a second TNFi following failure of a first is no longer routine clinical practice. This is due to the development and availability of TNFi therapeutic drug monitoring to rationalise clinical decision making as well as the availability of alternative drugs with different modes of action. As a result, receiving a second TNFi is only clinically relevant in a small subgroup of TNFi-experienced patients. Guidelines on therapeutic drug monitoring by experts in IBD provide algorithms with recommendations if there is loss of response to TNFi therapy.[102, 103] Where the drug and antibody levels indicate patients have failed a TNFi due to the mechanism of action, patients are offered advanced therapies with other modes of action.

Non-TNFi biologics and small molecules: Non-TNFi biologics recommended by NICE to treat to treat moderately to severely active UC include ustekinumab and vedolizumab. Tofacitinib is currently the only small molecule recommended by NICE to treat moderately to severely active UC in the UK.

Ustekinumab is recommended by NICE for the treatment of patients who have failed a TNFi or for whom a TNFi is not tolerated or is unsuitable. Whilst NICE recommendations for vedolizumab and tofacitinib do not restrict use to patients who have failed or cannot tolerate TNFis, these treatments are most commonly used in clinical practice following treatment failure on TNFis or in patients who are contraindicated to TNFis:[1]

- UK clinical expert feedback received as part of this appraisal noted that patients contraindicated to TNFis or those with safety concerns in the first line setting, typically receive vedolizumab

©Celgene Ltd, a BMS Company (2022). All rights reserved

- For patients that experience treatment failure on TNFis, treatment options include vedolizumab, ustekinumab or tofacitinib.[1] Based on UK clinical feedback received as part of this appraisal, and as confirmed in TA633, treatment choice for these patients is influenced by the safety profile of available treatment options
- Vedolizumab is most frequently prescribed next owing to the strong clinical experience of its use in practice, as well as its tolerable safety profile.[1, 75] For the same reason, ustekinumab may be offered
- As highlighted in TA633 and validated by a clinical expert consulted as part of this submission, due to the safety concerns associated with tofacitinib, it is not routinely used in UK clinical practice and when used is typically reserved for later treatment lines[1]
    - Current guidance from EMA advises that tofacitinib should be used with caution in patients with UC who are at high risk of blood clots. This includes patients with previous blood clots, inherited blood clotting, those taking combined hormonal contraceptives, receiving hormone replacement therapy or patients with heart failure or cancer[2]
    - In addition, due to the increased risk of infection, it is recommended that tofacitinib is not used in patients over the age of 65 years unless there is no alternative[2]
    - Further to this, tofacitinib is associated with multiple black box warnings issued by the US Food and Drug Administration (FDA) for: serious infections, thrombosis, malignancies and mortality in over 50 with risk factor for cardiovascular disease.[104]

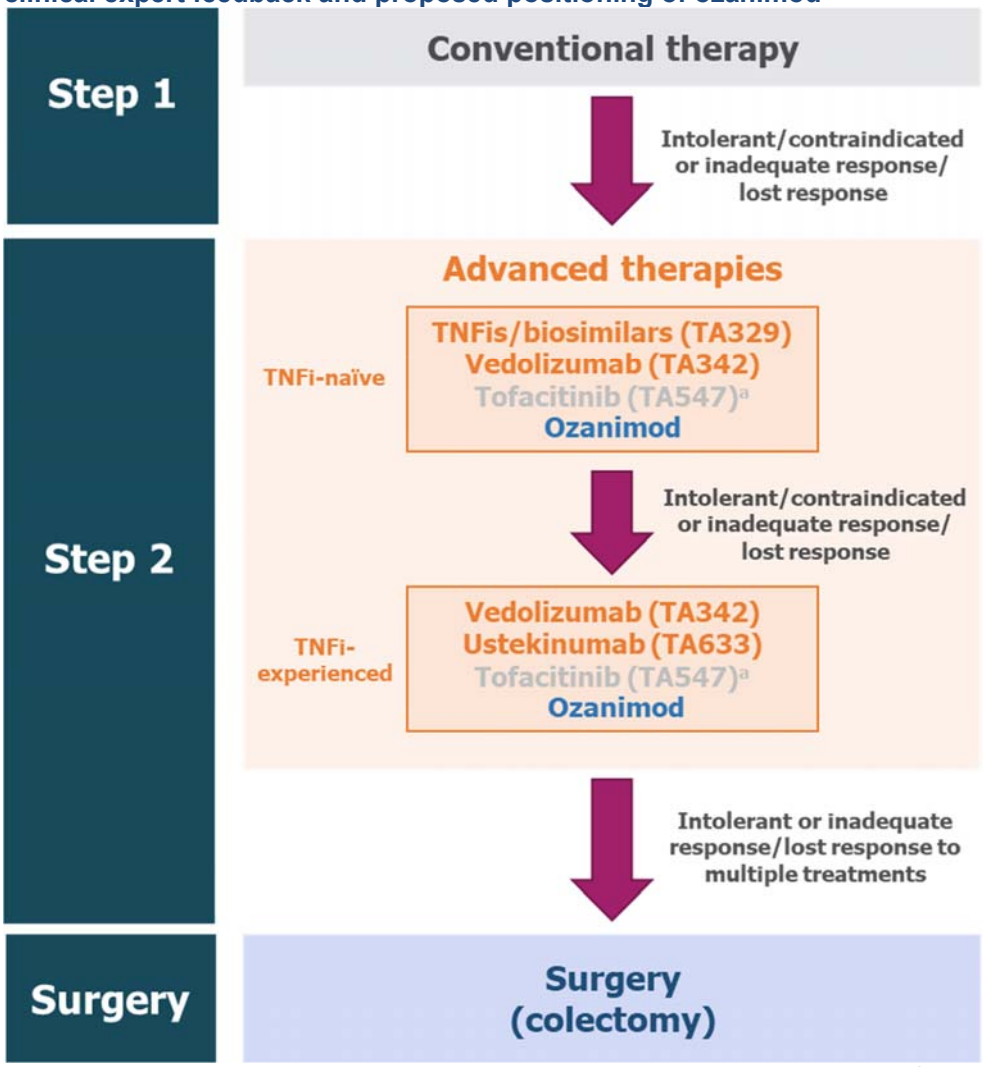
Surgery: Surgery is considered for UC patients whose disease has not responded or is refractory to medical treatment. However, owing to the irreversible consequences of surgery and associated short and long term risks (Section B.1.3.2) patients remain motivated to try further treatment options and surgery is typically viewed as a last resort.[68] A European survey of 2,333 UC patients, including patients from the UK, found that 86.4% of patients would rather try a new UC drug than undergo surgery.[53] A small number of patients may choose surgery at any stage due to personal preferences, but it remains an undesirable option for most patients.[68]

B.1.3.4.1 Proposed positioning of ozanimod in clinical practice and relevant comparators

A summary of the UK clinical pathway of care for patients with moderate to severely active UC, including the anticipated positioning of ozanimod, is presented in Figure 6. This pathway has been adapted from the NICE Pathway for UC,[16] based on feedback from clinical consultation conducted as part of this appraisal and reflects the differential usage of currently available biologic treatment options as described above.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 6: Typical treatment pathway for moderately to severely active UC based on UK clinical expert feedback and proposed positioning of ozanimod



[a]Not typically used at this point in the treatment pathway in UK clinical practice due to safety concerns regarding use
Abbreviations: TA: technology appraisal; TNFi: tumour necrosis factor alpha inhibitor.

In this submission, ozanimod is positioned as a treatment option for patients with moderately to severely active ulcerative colitis (UC) who have had an inadequate response, lost response, or were intolerant to either conventional therapy (CvT) or a biologic agent, in line with the full marketing authorisation received on 18th November 2021.[11] The relevant comparators differ in the TNFi-naïve and TNFi-experienced populations, as outlined below.

TNFi-naïve

TNFis or vedolizumab are considered the only relevant comparators to ozanimod in the TNFi-naive population (Section B.1.3.4).

- TNFis are the most frequently used first choice treatments following CvT

©Celgene Ltd, a BMS Company (2022). All rights reserved

- However, clinical consultation received as part of this appraisal noted that some patients who are contraindicated to TNFis or have significant safety concerns regarding their use most commonly receive vedolizumab due to its tolerable safety profile and experience of use in clinical practice

TNFi-experienced

Vedolizumab or ustekinumab are considered the only relevant comparators to ozanimod in the TNFi-experienced population, in line with the relevant comparators considered in the NICE submission for ustekinumab (TA633).[1]

- TNFis and tofacitinib were not considered relevant comparators in the TNFi-experienced population
- Receipt of a second TNFi following failure with an initial TNFi treatment is only clinically relevant in a small subgroup of TNFi-experienced patients and therefore TNFis are not considered a relevant comparator to the general TNFi-experienced population, in line with the accepted assumption in TA633[1]
- Tofacitinib is typically reserved for later treatment lines due to concerns relating to its safety profile and the resulting EMA restrictions on its use and therefore does not represent a standard comparator to ozanimod in this population. This is in line with the accepted assumption in TA633[1]

B.1.3.5 Unmet need

The availability of biologics has resulted in improvements in disease management and HRQoL for patients with moderately to severely active UC. However, there are still considerable limitations associated with the available treatments for patients in both the TNFi-naïve and TNFi-experienced populations. These limitations are summarised below.

- Owing to the lifelong, chronic nature of the disease patients often require long-term treatments. Patients may fail to respond to existing treatments or subsequently lose their response over time; approximately 33−55% of patients treated with a TNFi do not respond and 50% of patients who initially respond lose their response within a year.[1, 105, 106] In addition, a long-term extension trial of vedolizumab in TNFi failure patients found approximately 80% of patients do not achieve remission at two years.[107] In clinical practice there remains a high rate of incomplete or non-response to biologics, resulting in disappointment for patients as fewer treatment options remain to control the disease before resorting to surgery. This indicates a need for novel therapeutic options, particularly given the differences in individual responses to inflammation with differing cytokine profiles (Section B.1.3.1).[108] As a result, the availability of new therapeutics options with novel mechanisms of action were identified as an area of unmet need by an expert group consensus in UC published in 2019.[108]
- The availability of novel treatment options would offer patients additional effective treatments options before surgery. Surgery does return bowel function and carries multiple short and long term risks including impaired sexual and reproductive function, which in some cases can lead to infertility.[31, 32,109] This is of particular concern for UC patients, who due to the early age of onset of UC, are typically sexually active and of child-bearing age.[68]

©Celgene Ltd, a BMS Company (2022). All rights reserved

- Biologic treatments are limited due to immunogenicity which may require co-administration of immune-modulators or often dose escalation, which is both expensive and may result in an increased risk of AEs.[110-112] Approximately 30% of patients taking TNFis receive dose escalation after 12 months due to loss of response, rising to 50% after 3 years.[113-116] Further to this, an SLR collecting RWE from 48 studies investigating interventions for the treatment of moderately to severely active UC found 35% of patients receiving vedolizumab received dose escalation.[74] There is therefore an unmet need for new small molecule treatment option which are not associated with immunogenicity issues.
- The vast majority of moderately to severely active UC patients requiring more advanced treatments are currently treated with biologic therapies, which are administered either IV or by SC injection, which can be viewed by patients as inconvenient and intrusive methods of administration.[117, 118]
- Tofacitinib, the only small molecule and oral treatment option currently licenced and approved by NICE to treat patients with moderately to severely active UC, is associated with serious safety concerns and therefore, as confirmed by clinical feedback received as part of this appraisal, its use in clinical practice is limited (Section B.1.3.4).[1] The availability of treatments options with convenient methods of administration are particularly important in chronic diseases such as UC where treatment may be required for the majority of the patients life. A study in 298 patients with IBD showed that patients preferred oral administration over IV or SC injection (91% versus 33% and 34%, respectively).[119] There is therefore an unmet need for new treatment options with a convenient oral method of administration.
- In addition, during the COVID-19 era there is a drive to keep vulnerable people (such as those with moderately to severely active UC) away from the hospital to reduce risk of infection. As a result, there is an increased need for the availability of oral treatments which can be easily administered at home to prevent delays in patients receiving treatments they need and reduce the burden on overstretched infusion clinics.[120]
- Ozanimod, a small molecule with a novel mechanism of action has the potential to address the unmet need amongst patients with moderately to severely active UC in both TNFi-naïve and TNFi-experienced populations by providing a novel treatment option with a convenient oral method of administration alongside a tolerable safety profile.

B.1.4 *Equality considerations*

It is not anticipated that the provision (or non-provision) of ozanimod would exclude from consideration any people protected by 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 an adverse impact on people with a particular disability or disabilities.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2 Clinical effectiveness

Clinical effectiveness summary

Study identification

- An SLR identified two clinical trials for ozanimod in patients with moderately to severely active UC; one Phase III; TRUENORTH (NCT02435992), and one Phase II; TOUCHSTONE (NCT01647516)
- The Phase III TRUENORTH trial is the pivotal clinical trial for this indication and is presented in full in this section. TOUCHSTONE is the key supporting trial and is presented in Appendix J

TRUENORTH

- TRUENORTH evaluated the safety and efficacy of ozanimod compared to placebo in adults with moderately to severely active UC
- The trial was composed of 2 periods, a 10-week induction period followed by a 42-week maintenance period (total 52 weeks).
- The induction period was composed of 2 cohorts, a randomised cohort; Cohort 1 and an open-label enrichment cohort; Cohort 2. Patients with a clinical response at the end of the induction period (from either cohort) proceeded into the maintenance period
- The primary endpoint of the induction and maintenance periods of TRUENORTH were the proportion of patients in clinical remission at Week 10 and Week 52, respectively. Remission was assessed using the 3-component Mayo score
- Key secondary endpoints of TRUENORTH induction and maintenance periods, measured at Weeks 10 and 52, respectively, were the proportion of patients with clinical response (based on the 3-component Mayo definition) the proportion of patients with endoscopic improvement and the proportion of patients who reported mucosal healing
- The proportion of patients who reported sustained clinical remission, corticosteroid-free remission and durable clinical remission at Week 52 were also assessed as key secondary endpoints of maintenance period
- Patients were stratified according to prior corticosteroid use (yes or no) as well as prior TNFi exposure (yes or no)

Efficacy

- In the induction period, treatment with ozanimod versus placebo resulted in statistically significantly greater proportions of patients achieving clinical remission (18.4% vs 6.0%, respectively; p<0.0001), clinical response (47.8% vs 25.9%, respectively; p<0.0001), endoscopic improvement (27.3% vs 12.0%, respectively; p<0.0001), and mucosal healing (12.6% vs 3.7%, respectively; p<0.001) at Week 10
- Maintenance therapy with ozanimod versus placebo resulted in statistically significantly greater proportions of patients achieving clinical remission (37.0% vs 18.5%, respectively; p<0.0001), clinical response (60.0% vs 41.0%, respectively; p<0.0001), endoscopic improvement (45.7% vs 26.4%, respectively; p<0.0001), and mucosal healing (29.6% vs 14.1%, respectively; p<0.001) at Week 52[4]
- Clinical remission rates were also greater in both TNFi-naïve ([redacted] vs [redacted] respectively) and TNFi-experienced patients ([redacted] vs [redacted]) treated with ozanimod compared with placebo at Week 10
- Similarly, clinical remission rates were significantly greater in both TNFi-naïve ([redacted] vs [redacted]; [redacted]) and TNFi-experienced ([redacted] vs [redacted]; [redacted]) patients treated with ozanimod versus placebo at Week 52

Safety

- The safety profile of ozanimod was consistent with the known tolerable safety profile for ozanimod in Multiple Sclerosis (MS)
- Treatment with ozanimod was well tolerated, with similar overall incidence of treatment-emergent adverse events (TEAEs) in both the ozanimod and placebo arms in the induction period.
- Severe TEAEs were comparable between the ozanimod and placebo arms in in the induction period (Cohort 1: 3.3% versus 1.9%, respectively) and a similar incidence of severe TEAEs in the ozanimod and placebo arms was observed in the maintenance period ([redacted] versus [redacted], respectively)[4]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.1 *Identification and selection of relevant studies*

An SLR was conducted in October 2020 and updated in October 2021 to identify relevant clinical evidence for the efficacy and safety of ozanimod and relevant comparators for the treatment of moderate to severely active UC in the form of randomised controlled trials (RCTs). The SLR identified 157 relevant publications, reporting on 28 unique studies. Full details of the SLR, including search strategy, study selection process and detailed results are presented in Appendix B.

B.2.2 *List of relevant clinical effectiveness evidence*

The SLR identified two RCTs investigating the efficacy and safety of ozanimod for the treatment of moderate to severely active UC: TRUENORTH (NCT02435992) and TOUCHSTONE (NCT01647516) (Table 8).

The principal clinical evidence base for the use of ozanimod in this indication is TRUENORTH, a Phase III double-blind, randomised, placebo-controlled, multi-centre trial, which studied the efficacy and safety of ozanimod in the treatment of moderately severely active UC in adults. Supportive evidence is provided by the TOUCHSTONE trial, a Phase II, double-blind, randomised, placebo-controlled, multi-centre trial.

The information presented in this submission has been derived from the TRUENORTH CSR and Sandborn *et al.* 2021 publication.[4, 5] The publication provides details of the trial design and presents the key efficacy and safety results of the induction and maintenance periods in the ITT population.

Table 8: Clinical effectiveness evidence

Study	TRUENORTH (NCT02435992)	TOUCHSTONE (NCT01647516)
Study design	Phase III, multicentre, randomised, double-blind placebo-controlled study consisting of a 10-week induction period (including a 1-week dose titration) with responders to ozanimod re-randomised to a 42-week maintenance period	Phase II, multicentre, randomised, double-blind, placebo-controlled study consisting of an 9-week induction period (including a 1-week dose titration) with responders to ozanimod re-randomised to a 24-week maintenance period
Population	Adults with moderate to severely active UC	Adults with moderate to severely active UC
Intervention(s)	1 mg/day of ozanimod HCl[a] administered orally during both the maintenance and induction period	Patients were randomised in a 1:1 ratio to receive either 0.5 or 1 mg/day of ozanimod[b] administered orally during both the maintenance and induction period
Comparator(s)	Placebo administered orally	Placebo administered orally
Trial supports application for marketing authorisation?	Yes	Yes
Trial used in the economic model?	Yes	No

©Celgene Ltd, a BMS Company (2022). All rights reserved

Rationale for use/non-use in the model	TRUENORTH provides the primary source of evidence for the clinical efficacy and safety of ozanimod, is relevant to the decision problem and informed the marketing authorisation application. The trial data were therefore used to inform relative efficacy in the NMA which fed into the economic model	As discussed in Section B.1.1, due to TNFis typically being prescribed first-line in patients who are intolerant or have failed CvT, exposure to TNFis forms the basis for clinical decision-making. The decision problem has therefore been split into two distinct populations based on prior TNFi exposure (Section B.1.1). No data for TNFi subgroups were available from TOUCHSTONE, and therefore results from TOUCHSTONE were not incorporated in the economic model. As 82% of the participants of TOUCHSTONE were naïve to any biologic treatment, a sensitivity analysis was conducted to explore the effect on the results of the NMA of including the trial in the biologic-naïve induction analyses. The results were not found to differ significantly and therefore the analysis was not used to inform the model (Section B.3.8)
Reported outcomes specified in the decision problem[c]	• Measures of disease activity; change in Mayo score • **Rates of clinical response** • **Rates of clinical remission** • Rates of durable remission • Endoscopic improvement • Mucosal healing (combined endoscopic and histological healing) defined as an endoscopy sub-score of ≤1 point, without friability and Geboes index score <2.0 • Corticosteroid-free remission • **Adverse effects of treatment** • Rates of hospitalisation • Health-related quality of life	• Measures of disease activity; change in the 3-component Mayo score • Rates of clinical response • Rates of clinical remission • Mucosal healing • Adverse effects of treatment

[a]Equivalent to 0.92 mg of ozanimod.
[b]Equivalent to 0.92 mg and 0.46 mg of ozanimod respectively
[a]Outcomes in bold indicate those used in the cost effectiveness analysis.
Abbreviations: CRP: C-reactive protein; EQ-5D: European quality of life-5 dimensions; HCl: hydrochloric acid; HRQoL; health-related quality of life; IBDQ; inflammatory bowel disease questionnaire; SF-36: 36-item Short Form Survey; UC: ulcerative colitis.
Sources: Sanborn *et al.* 2021.[4]

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

As the pivotal trial supporting this submission, the methodology and results of TRUENORTH are presented within this section. The methodology and results of TOUCHSTONE are summarised in Appendix J.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.3.1 Trial methodology

The methodology of TRUENORTH, for both the induction period and the maintenance period, is summarised in Table 9.

Table 9: Summary of methodology for TRUENORTH

	Induction period (N=1,012)	Maintenance period (N=526)
Location	Multinational study, including 285 study sites in North America, Europe, Asia Pacific, South America, and South Africa.	
Study objective	To evaluate the safety and efficacy of ozanimod induction therapy in patients with moderate to severely active UC	To evaluate the safety and efficacy of ozanimod maintenance therapy in patients with moderate to severely active UC
Trial design	Phase III, randomised, double-blind, placebo-controlled, multicentre study consisting of a 10-week induction period followed by a 42-week maintenance period. The induction period consisted of 2 cohorts, one randomised double-blinded; Cohort 1 and one open-label enrichment cohort; Cohort 2	
Population	Adult patients aged 18 years or older with moderate to severely active ulcerative colitis (N = 1,012)	
Eligibility criteria for participants	The key eligibility criteria for the TRUENORTH trial are summarised below. The full eligibility criteria can be found in Appendix O. • Patients aged 18–75 years • Diagnosed with UC ≥3 months before first study drug administration, with disease extending ≥15 cm from the anal verge as determined by baseline endoscopy (flexible sigmoidoscopy or colonoscopy) • Active UC, defined as a 4-component Mayo score 6–12, with an endoscopic sub-score of ≥ 2, a RBS of ≥ 1, and a stool frequency score ≥ 1 • A document presence of the varicella-zoster virus IgG antibody or complete varicella-zoster vaccination at least 30 days before randomisation	
Randomisation	• Patients in Cohort 1 were randomised in a 2:1 ratio to receive either ozanimod 1 mg or placebo OD. • Randomisation was performed using the interactive voice/web-based activated response system (IVRS/IWRS) • Randomisation was stratified by corticosteroid use at screening and prior TNFi therapy. This stratified randomisation was centrally allocated across all centres via IVRS/IWRS	• Patients in Cohort 1 who were randomised to ozanimod during the induction period and patients receiving open-label ozanimod in Cohort 2 who achieved at least clinical response at the end of the Week 10 were re-randomised in a 1:1 ratio to receive either ozanimod or placebo until Week 52 • Patients randomised to placebo at induction and achieved at least clinical response at Week 10 continued to receive placebo in the maintenance period • Randomisation was stratified

©Celgene Ltd, a BMS Company (2022). All rights reserved

	Induction period (N=1,012)	Maintenance period (N=526)
		by clinical remission status and corticosteroid use at Week 10
Interventional trial drugs	**Ozanimod** • Dose titration: Ozanimod 0.25 mg[a] OD, taken orally on Days 1–4 increasing to 0.50 mg[b] OD on Days 5–7 (Week 1) • Ozanimod 1 mg[c] OD, oral (Weeks 2–10) **Placebo** • Placebo OD, oral	**Ozanimod** • Ozanimod 1 mg[c] OD, oral **Placebo** • Placebo OD, oral
Method of blinding	The Investigator site personnel, the Sponsor, and the Sponsor's representatives involved in the monitoring or conducting of the trial, and the patients were all blinded to the treatment allocation in Cohort 1 during the induction period and during the maintenance period. Endpoints were assessed via the investigator. Ozanimod and placebo capsules were identical in physical appearance. Cohort 2 in the induction period was open-label.	
Permitted and disallowed concomitant medications	**Permitted:** • Patients receiving oral 5-ASA or oral corticosteroids at Screening were required to keep their prescribed dosage steady through to Week 10 o Patients receiving oral corticosteroid were to have a steroid tapering regimen introduced after Week 10 o Patients receiving oral 5-ASA were to maintain a stable dose through to Week 52 of the maintenance period **Prohibited:** • Oral corticosteroids or 5-ASA in patients not already receiving a stable dose at Screening • Marketed biologic therapies (e.g. adalimumab, vedolizumab) • Immunosuppressive agents (e.g. AZA/6-MP) • Any investigational drug other than the investigational drug specified in this trial	
Primary outcome (including scoring methods and timings of assessments)	The proportion of patients in clinical remission at Week 10 measured using the 3-component Mayo scoring system	The proportion of patients in clinical remission at Week 52 measured using the 3-component Mayo scoring system
Key secondary outcomes (including scoring methods and timings of assessments)	**Hierarchically ranked key secondary efficacy endpoints:** • Proportion of patients with a clinical response at Week 10 • Proportion of patients with endoscopic improvement at Week 10 • Proportion of patients with mucosal healing at Week 10 All key secondary efficacy endpoints were assessed using the 3-component	**Hierarchically ranked key secondary efficacy endpoints:** • Proportion of patients with a clinical response at Week 52 • Proportion of patients with endoscopic improvement at Week 52 • Proportion of patients with maintenance of remission (clinical remission at Week 52 among subset of patients in remission at Week 10)

©Celgene Ltd, a BMS Company (2022). All rights reserved

	Induction period (N=1,012)	Maintenance period (N=526)
	Mayo scoring system, or relevant subscore thereof (see Table 10)	• Proportion of patients with corticosteroid-free remission (clinical remission at Week 52 after ≥12 weeks without corticosteroids) • Proportion of patients with mucosal healing at Week 52 • Proportion of patients with durable clinical remission (remission at Weeks 10 and 52 among patients entering maintenance phase) All key secondary efficacy endpoints were assessed using the 3-component Mayo scoring system, or relevant subscore thereof (see Table 10)
Other secondary outcomes	• Changes from baseline to Week 10 in 3-component Mayo score, 4-component Mayo score, and partial Mayo score[d] • Proportion of patients with histologic remission at Week 10 • Proportion of patients in clinical remission (4-component Mayo definition) at Week 10 • Proportion of patients with a clinical response (4-component Mayo definition) at Week 10 • Proportion of patients with clinical response, clinical remission (3-component Mayo score), or endoscopic improvement at Week 10 in patients who previously received TNFi therapy • Changes from baseline to Week 10 in the 36-Item Short Form Health Survey (SF-36) and the EuroQol-5 Dimension Questionnaire (EQ-5D)	• Change from Baseline to Week 52 in 3-component Mayo score, 4-component Mayo score, and partial Mayo score[d] • The proportion of patients with histological remission at Week 52[c] • The proportion of patients in clinical remission (4-component Mayo) at Week 52 • The proportion of patients with a clinical response (4-component Mayo) at Week 52 • The proportion of patients with clinical response, remission (3-component Mayo score), or endoscopic improvement at Week 52 in patients who previously received • Change in the SF-36 and the EQ-5D from Baseline to Week 52 • Health resource utilisation (capturing the costs associated with treatment of UC, including hospitalisation for UC and colectomy) at 28 weeks, 40 weeks, and at Week 52[e] • Work productivity at 28 weeks, 40 weeks, and at Week 52[e]
Pre-planned subgroup analyses in the overall population	Subgroup analyses were also performed for the endpoints of clinical remission and clinical response at Week 10 and Week 52 for induction and maintenance, respectively:	

©Celgene Ltd, a BMS Company (2022). All rights reserved

	Induction period (N=1,012)	Maintenance period (N=526)
	• UC medication history: Corticosteroid use at screening (yes vs no); Prior TNFi use (yes vs no) • Baseline disease characteristics: Baseline 4-component Mayo score (≤ 9 vs > 9); Extent of colitis (left-sided vs extensive); Years since initial UC diagnosis (≤ 4 vs > 4 years); Baseline partial Mayo score (≤ median vs > median); Baseline partial Mayo score (≤ 7 vs > 7); Baseline endoscopy subscore (2 vs 3) • Moderate UC status at Baseline (4-component Mayo score 6 to 10; yes versus no) • Baseline demographics: Sex (female vs male); Age at screening (≤ median vs > median); Baseline faecal calprotectin (≤ 250 vs > 250 mg/kg); Baseline ALC (≤ 1,500 vs > 1,500 10^6/L); Region (North America, Eastern Europe, Western Europe, Asia Pacific)	

[a]Equivalent to 0.23 mg ozanimod. [b]Equivalent to 0.46 mg ozanimod. [c]Equivalent to 0.92 mg ozanimod. [d]Only change in 3-component Mayo score from baseline presented in this submission. [e]Outcome not presented in this submission, full details are presented in the TRUENORTH CSR.[5]
Abbreviations: 5-ASAs: 5-aminiosalicyclic acid AZA: azathioprine; IV: intravenous; 6-MP: 6-mercaptopurine; RBS: rectal bleeding subscore; TNFi: tumour necrosis factor inhibitor; UC: ulcerative colitis; VZV: varicella zoster virus; NSAID: nonsteroidal anti-inflammatory drug.
Sources: Sanborn *et al.* 2021.[4] TRUENORTH CSR.[5]

Outcomes

Outcomes were measured for disease activity and health utility using different instruments and scoring systems, full details of these are presented in Table 10. To evaluate the primary and key secondary efficacy endpoints (clinical remission and clinical response) the 7-day scoring algorithm was used. The algorithm uses 7 days prior to each visit to calculate the rectal bleeding subscore (RBS) and stool frequency subscore (SFS).

The Mayo scoring system was used to assess clinical remission and response in TRUENORTH. Full details of the Mayo Scoring system are presented in Section B.1.3.1. The primary analyses were based on the 3-component Mayo definition; all results presented for clinical remission and response are based on the 3-component Mayo unless otherwise stated.[20 87]

Mucosal healing was a key secondary outcome of TRUENORTH owing to its association with long-term remission of disease activity, decreased risk of surgery, and improved health-related quality of life (HRQoL) in UC patients.[23, 53, 80] Pivotal trials for other treatments in UC (TNFis, vedolizumab and tofacitinib) defined mucosal healing as a Mayo endoscopic subscore of ≤1.[24, 26, 121-123] However, defining mucosal healing solely on endoscopic assessment overlooks that approximately 25% of patients with an endoscopically normal-appearing mucosa have persistent histologic inflammation.[90] For this reason, TRUENORTH used both the endoscopic subscore ≤1 (without friability) AND a Geboes index score <2.0 to define mucosal healing. Details of the Geboes score are presented in Section B.1.3.1. The pivotal trial for ustekinumab also included histologic improvement as part of the definition of mucosal healing however it used a less strict definition than TRUENORTH (Geboes score <3.1 versus 2).

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 10: Outcome measured used in TRUENORTH

Outcome	Definition
Clinical remission	**3-component Mayo score:**[a] • RBS= 0 **AND** • SFS ≤ 1 (and a decrease of ≥ 1 point from the baseline stool frequency subscore) **AND** • Endoscopy subscore ≤ 1 **4-component Mayo score:**[b] • 4-componet Mayo score ≤2, **AND** • RBS ≤ 1 • SFS ≤ 1 • Endoscopy subscore ≤ 1 • PGA ≤ 1
Clinical response	**3-component Mayo score:**[a] • A reduction from Baseline in the 3-component Mayo score of ≥ 2 points and ≥ 35% **AND** • A reduction from Baseline in the RBS of ≥ 1 point **OR** • An absolute RBS of ≤ 1 point **4-component Mayo score:**[b] • A reduction from Baseline in the 4-component Mayo score of ≥ 3 points and ≥ 30% **AND** • A reduction from Baseline in the RBS of ≥ 1 point **OR** • An absolute RBS of ≤ 1 point
Endoscopic healing	Mayo endoscopy subscore of ≤ 1 point.
Histologic healing	Geboes index score < 2.0, indicating no neutrophils in the epithelial crypts or lamina propria and no increase in eosinophils, no crypt destruction, and no erosions ulcerations or granulation tissue.
Mucosal healing	Endoscopic improvement with histologic remission (Mayo endoscopy subscore ≤ 1 point AND Geboes index score < 2.0).
Corticosteroid-free remission	Clinical remission at 52 weeks while off corticosteroids for ≥ 12 weeks. The ≥12 week time period was clinically meaningful, owing to relapse within 12 weeks of discontinuation of corticosteroids being a defining characteristic of patients with steroid-dependent UC.[21]
Maintenance of clinical remission	The proportion of patients in clinical remission (defined above) at Week 52 in the subset of patients in clinical remission at Week 10
Durable clinical remission	The proportion of patients achieving clinical remission at Week 10 and at the end of the 42-week maintenance period.
36-item short-form health survey (SF-36)	SF-36 evaluates 36 measures, providing 2 scales on patient's HRQoL: the physical composite summary (PCS) score and the mental composite summary (MCS) score.[68] The PCS score is based on the answers to questions from 4 domains; general health, physical function, physical role functioning (e.g. being physically able to perform work and other activities), and bodily pain. The MCS score is also based on the answers to questions from 4 domains; vitality (e.g. energy level), emotional role functioning (e.g. being emotionally able to perform work and other activities), mental health, and social functioning.

©Celgene Ltd, a BMS Company (2022). All rights reserved

	Both the PCS score and the MCS score were assessed in the trial. An improvement in score of ≥5 points was defined as the minimum clinically important difference (MCID).
EuroQoL 5-dimensions questionnaire (EQ-5D)	The EQ-5D is composed of 5 dimensions (mobility, self-care, usual activities, pain and discomfort, anxiety and depression), each of which have 5 severity levels. The summary index and Visual Analogue Scale (VAS) were assessed.

[a]The 3-component Mayo score is the sum of the rectal bleeding subscore, stool frequency subscore, and the endoscopy subscore. The 3-component Mayo score has a range of 0 to 9 points.
[b]The 4-component Mayo score is the sum of the rectal bleeding subscore, stool frequency subscore, the endoscopy subscore and the PGA subscore. The 4-component Mayo score has a range of 0 to 12 points.
Abbreviations: MCS: mental composite summary score; MCID: minimum clinically important difference; PCS: physical composite summary score; PGA: Physician Global Assessment; QoL: quality of life; RBS: rectal bleeding subscore; SFS: stool frequency subscore; VAS: visual analogue scale.
Sources: Sanborn *et al.* 2021[4]; TRUENORTH CSR.[5]

B.2.3.2 Trial design

A schematic of the trial design for TRUENORTH is presented in Figure 7. The trial was composed of two periods; a 10-week induction period followed by a 42-week maintenance period.

The induction period was composed of 2 cohorts: a randomised cohort; Cohort 1 and an open label enrichment cohort; Cohort 2. Only data from patients in Cohort 1 were used to inform the efficacy endpoints of the induction period of the trial. In Cohort 1 patients were randomised in a 2:1 ratio to receive either ozanimod 1 mg OD or placebo OD in a double-blinded fashion. Randomisation was conducted using an interactive voice/web-based activated response system (IVRS/IWRS). In Cohort 2 all patients received open-label ozanimod 1 mg OD. The purpose of Cohort 2 was to allow for increased number of responders to ozanimod at the end of the induction period to enter re-randomisation for the maintenance period; further details are presented in Table 12. A total of 1,012 patients were enrolled into the induction period (645 in Cohort 1, 367 in Cohort 2). Results from Cohort 2 were not used to assess the efficacy endpoints of the induction period as this was a non-controlled, non-randomised open-label cohort. Efficacy results from Cohort 2 are therefore not reported as part of this submission, however safety results from this cohort have been included for completeness.

Patients in Cohort 1 were stratified by corticosteroid use and previous TNFi use before randomisation. In Cohort 1 the proportion of TNFi-experienced patients was limited to ≤30%. Once the quota for TNFi-experienced patients in Cohort 1 was reached, TNFi-experienced patients were automatically enrolled into Cohort 2. TNFi-naïve patients continued to be enrolled into Cohort 1 until enrolment was complete, upon which they could be enrolled into Cohort 2. In Cohort 2, the proportion of TNFi-experienced patients was limited to <50% of the total cohort. In total, 30.2% (195/645) and 43.3% (159/367) of patients enrolled into Cohort 1 and Cohort 2, respectively, were TNFi-experienced.

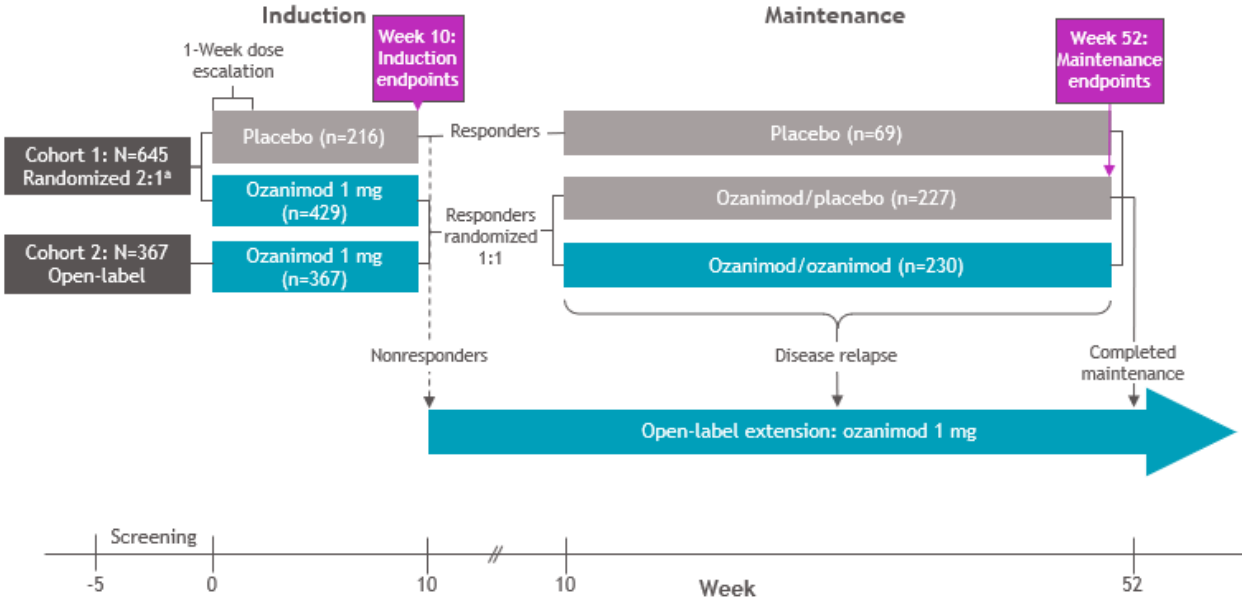
Only patients who reported a clinical response (according to the 3-component Mayo score) whilst receiving ozanimod (n=457), in either cohort, were eligible to enter the maintenance period. These patients were re-randomised in a 1:1 ratio in a double-blind manner to either maintenance with ozanimod [1 mg] (n=230) or placebo (n=227). Randomisation to maintenance was stratified by clinical remission status (according to the 3-or 4-component Mayo definition), at Week 10 and

©Celgene Ltd, a BMS Company (2022). All rights reserved

corticosteroids use at Week 10. Patients who reported a clinical response whilst receiving placebo in the Cohort 1 (n=69) continued to receive placebo during the maintenance period.

Patients who did not report a clinical response at the end of induction or relapsed during maintenance treatment as well as those who completed the 42-week maintenance period were eligible to enter an open-label extension (OLE) trial (RPC01-3102). Further details of the OLE are presented in Section B.2.10 and full details of patient disposition during the TRUENORTH trials are presented in Section B.2.3.5.

Figure 7: Trial schematic for TRUENORTH



[a]Responders re-randomised into the maintenance period were stratified by clinical remission status at Week 10 (yes/no) and corticosteroids use at Week 10 (yes/no).
Abbreviations: TNF: tumour necrosis factor.
Source: TRUENORTH CSR: Figure 1.[5]

B.2.3.3 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

The data sets analysed in TRUENORTH, including the number of patients in each set, are presented in Table 11. For the purposes of the statistical analyses, the induction period and the maintenance period were treated as two independent studies. All safety and efficacy endpoints were summarised by trial arm.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 11: Trial populations used in the analysis of outcomes in TRUENORTH

Analysis set	Description
Intention-to-treat (ITT) population	**Induction period** • Cohort 1: all randomised patients from Cohort 1 of the induction period who received at least 1 dose of investigational drug (ozanimod [n=429] or placebo [n=216]) **Maintenance period** • All re-randomised patients who received at least 1 dose of investigational drug (ozanimod or placebo) in the maintenance period (ozanimod [n=230] or placebo [n=227])
Safety population	All patients who received at least 1 dose of investigational drug. Patients randomised to placebo who received any amount of ozanimod were included in the safety population. All patients randomised to ozanimod received ozanimod, and thus patient numbers in the ITT and safety populations were identical (Induction period [n=645]; Maintenance period [n=526])

Abbreviations: ITT: intention-to-treat; PP: per protocol.
Source: TRUENORTH CSR.[5] Sanborn *et al.* 2021.[4]

Primary efficacy analysis

The primary efficacy endpoint for TRUENORTH was the proportion of patients in clinical remission, measured via the 3-component Mayo score, at Week 10 (induction period) and Week 52 (maintenance period). The statistical methods for the primary analysis of the primary endpoint in TRUENORTH are presented in Table 12. The primary analysis of the primary and key secondary endpoints was conducted on the ITT population (see Section B.2.5) and repeated on each subgroup listed in Table 9, including patients with and without prior TNFi experience. A forest plot showing the weighted treatment differences with associated 95% CI and p-values was produced and is presented in Appendix E.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 12: Statistical methods for the primary analyses of TRUENORTH

	TRUENORTH	
	Induction period	**Maintenance period**
Hypothesis	**Cohort 1** • Null hypothesis: an equal proportion of patients will be in clinical remission at Week 10 (ozanimod vs placebo) **Cohort 2** • Cohort 2 was open-label and non-randomised; no tests of hypotheses were pre-specified	• Null hypothesis: there will be no difference in the proportion of patients re-randomised to ozanimod and placebo in clinical remission at Week 52
Statistical analysis	**Induction** • Only data collected from Cohort 1 (double-blind ozanimod OD or placebo OD) were used to assess the efficacy endpoints for the induction period of the study • Statistical analysis of the primary endpoint (proportion of patients in clinical remission at Week 10) expressed as proportions of patients was carried out on the ITT population using a two-sided Cochran-Mantel-Haenszel (CMH) test at the 5% level of significance o Patients were stratified by corticosteroid use at screening (yes or no), and prior TNFi use (yes or no) o The odds ratio (OR), 95% confidence intervals (CIs) and statistical significance (p-values) were calculated o Patients who met the criteria for treatment failure were imputed using non-responder imputation (NRI)[124] • An ANCOVA model was used to analyse all secondary efficacy endpoints that were expressed as changes from baseline, with corticosteroid use at screening, previous use of TNFi therapy, and baseline value of the corresponding outcome included as covariates • A closed hierarchical testing procedure was used to control the overall Type I error rate for multiplicity, starting with the primary endpoint, followed by the key secondary efficacy endpoints in the order provided in Table 9 o Formal testing only proceeded to the key secondary outcomes if the result from the primary analysis were significant (2-sided Cochran-Mantel-Haenszel (CMH) test P<0.05); if the result was not significant, all subsequent analyses were considered exploratory, with the corresponding P-values being nominal. The same approach was taken with the key secondary endpoints; analysis only proceeded to the next secondary endpoint if the results from the previous analysis were significant • Other secondary efficacy endpoints were tested in a non-hierarchical manner without multiplicity adjustments • As Cohort 2 (open-label ozanimod OD) did not have a placebo control, no inferential statistics were conducted for data collected from Cohort 2 during the induction period. Descriptive statistics for the efficacy endpoints for Cohort 2 are available instead and can be found in the CSR[5] **Maintenance** • Statistical analysis of the primary endpoint (proportion of patients in clinical remission at Week 52) was carried out on the ITT population using a two-sided Cochran-Mantel-Haenszel (CMH) test at the 5% level of significance	

©Celgene Ltd, a BMS Company (2022). All rights reserved

	o The test accounted for stratification by corticosteroid use at Week 10 of the Induction Period (yes or no) and previous TNFi use (yes or no) o The OR, 95% confidence intervals (CIs) and statistical significance (p-values) were calculated o Patients who met the criteria for treatment failure were imputed using NRI, described below[124] • An ANCOVA model was used to analyse all secondary efficacy endpoints that were expressed as changes from baseline, with clinical remission status at Week 10 of the induction period, corticosteroid use at Week 10 of the induction period, and baseline value of the corresponding outcome included as covariates • As in induction, a hierarchical testing procedure was used to control the overall Type I error rate for multiplicity, starting with the primary endpoint, followed by the 6 key secondary efficacy endpoints in the order provided in Table 9 o Formal testing proceeded to the next outcome analysis if results from the previous analysis were significant (2-sided $P<0.05$); if the results were not significant, all subsequent analyses were exploratory, with the corresponding P-values being nominal • Other secondary efficacy endpoints were tested in a non-hierarchical manner without multiplicity adjustments **Subgroup/sensitivity analyses** • The analysis of primary and key secondary endpoints was repeated in the per-protocol population and pre-specified subgroups (Section B.2.5) and were not subject to family-wise Type I error control. In addition to reporting the subgroup analyses in full, a forest plot showing the weighted treatment differences with associated 95% CI and p-values was produced • In addition to the primary analysis of clinical remission assessed using the 3-component Mayo score (Table 10) at Week 10 and Week 52 in induction and maintenance, respectively, sensitivity analyses were explored which utilised alternative definitions of remission including the 4-component Mayo score	
Sample size, power calculation	**Cohort 1 (power calculation)** • Power calculations were based on the remission results from a previous Phase II induction study of ozanimod 1 mg (TOUCHSTONE); 16% for ozanimod 1 mg OD and 6% for placebo (rate difference = 10%) • Based on a 2:1 randomisation ratio to ozanimod or placebo, 600 patients (400 ozanimod and 200 placebo) were required to achieve 90% power to detect a 10% difference, using a 2-sided Fisher's exact test (α = 0.05) **Cohort 2 (sample size calculation)** • Based on the same Phase II study, it was anticipated that at least 60% of patients treated with ozanimod would have a clinical response at the end of the induction period • In order to ensure there were approximately 420 patients with a clinical response to ozanimod for potential enrolment of approximately 400 patients into the maintenance period	• A 16% remission rate (Week 52) with placebo was used, from a randomised withdrawal study in UC patients who had previously had a clinical response to induction therapy (Feagan *et al.* 2013).[24] For ozanimod, a rate difference of 14% (30% remission rate with ozanimod) was assumed. A sample size of 400 patients (200 in each trial arm) was required to achieve 90% power, using a 2-sided Fisher's exact test (α = 0.05)

©Celgene Ltd, a BMS Company (2022). All rights reserved

	(assuming a 5% dropout rate), it was necessary to enrol approximately 900 patients overall into the induction period, of which 700 would receive treatment with ozanimod • Therefore, approximately an additional 300 patients receiving ozanimod were enrolled into Cohort 2, the non-randomised enrichment cohort	
Data management, patient withdrawals	• The primary analysis utilised the ITT principle and included all patients who were randomised to treatment o Treatment failure rules were applied to the primary analyses of all efficacy endpoints, meaning patients meeting the criteria for treatment failure were treated as non-responders using NRI for statistical analyses of efficacy. Patients with insufficient data to determine endpoint status at Week 10 and/or Week 52 were classified as non-responders. Patients with any of RBS, SFS and endoscopic subscores missing at Week 10 or Week 52 were classed a non-remitters. In addition, patients meeting criteria for treatment failure were considered non-responders using NRI for efficacy analyses. NRI is considered to be a more conservative approach for managing missing data than last observation carried forward[124] • Patients were considered to have failed treatment if any of the following occurred: o Any protocol-prohibited change in medications o A colectomy (partial or total) or an ostomy o Discontinuation of investigational drug due to lack of therapeutic effect before the Week 10 or Week 52 efficacy evaluations o Patients with missing Week 10 efficacy data for the induction period and/or Week 52 efficacy data for the maintenance period were also considered non-responders using NRI	

Abbreviations: CI: confidence interval: CMH: Cochran-Mantel-Haenszel: ITT: intention-to-treat: NRI: non-responder imputation; OR: odds ratio; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.3.4 Baseline characteristics

Baseline for the induction period was defined as the last observed measurement prior to the Day 1 receipt of investigational drug (ozanimod or placebo). Baseline for the maintenance period was defined as the last observed measurement prior to the first dose of investigational drug (ozanimod or placebo) following re-randomisation into the maintenance period. As discussed in Section B.2.3.1, the primary analysis was conducted on the overall ITT population as well as key subgroups of the ITT population, including patients with and without prior TNFi experience. The subgroups by prior TNFi experience inform the base case economic analyses for this submission. Baseline characteristics for the overall ITT population are presented in Section B.2.3.4.1, and baseline characteristics for the subgroups by prior TNFi experience are presented in Section B.2.3.4.2.

B.2.3.4.1 Overall population

Baseline demographic and disease characteristics

The baseline demographic and disease characteristics of patients in the induction period of TRUENORTH were generally well balanced between trial arms (Table 13). The mean age across Cohort 1 was approximately 42.[4] A high proportion of patients enrolled were white ([redacted]) and over [redacted] of patients across both Cohort 1 and 2 were from Europe. The baseline demographic characteristics of patients entering the maintenance period were similar to those for the induction period (Table 13).

The baseline disease characteristics of patients in the induction period were also well balanced across Cohort 1 (Table 13). The mean duration of UC symptoms at the beginning of the induction period was approximately [redacted]. Across all patients, the mean 3-component Mayo score at baseline was 6.6.[4] The disease characteristics for patients who entered the maintenance period were also generally were well balanced between trial arms (Table 13). The mean duration of UC symptoms was slightly longer in patients randomised to maintenance with ozanimod (approximately [redacted]) compared to placebo (approximately [redacted]).

Full details of baseline demographic and disease characteristics can be found in Appendix K.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 13: Patient baseline demographics and disease characteristics in TRUENORTH (ITT population)

	Induction period			Maintenance period		
	Cohort 1		Cohort 2		Re-randomised patients	
	Ozanimod (N = 429)	Placebo (N = 216)	Ozanimod (N = 367)	Placebo (N = 69)	Ozanimod – Placebo (N = 227)	Ozanimod – Ozanimod (N = 230)
Male, n (%)	245 (57.1)	143 (66.2)	214 (58.3)	[redacted]	[redacted]	[redacted]
Age, year, mean (SD)	41.4 (13.54)	41.9 (13.64)	42.1 (13.72)	[redacted]	[redacted]	[redacted]
Body mass index (kg/m²), n (SD)	25.40 (5.492)	25.11 (4.477)	25.88 (5.796)	[redacted]	[redacted]	[redacted]
Region, n (%)						
North America	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Europe	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other (Asia pacific, South America, South Africa)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease characteristics						
Age at UC symptoms onset (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Years since UC diagnosis, mean (SD)	6.9 (6.61)	6.8 (7.04)	7.91 (7.365)	[redacted]	[redacted]	[redacted]
3-component Mayo score at Baseline, mean (SD)	6.6 (1.21)	6.6 (1.15)	6.8 (1.26)	[redacted]	[redacted]	[redacted]
4-component Mayo score at Baseline, mean (SD)	8.9 (1.47)	8.9 (1.35)	9.1 (1.49)	[redacted]	[redacted]	[redacted]
Faecal calprotectin (ug/g), median	1080	1350	1260	[redacted]	[redacted]	[redacted]
C-reactive protein (mg/L), median	4.0	5.0	5.0	[redacted]	[redacted]	[redacted]
Extent of disease, n (%)						
Limited to left side of colon	268 (62.5)	134 (62.0)	237 (64.6)	[redacted]	[redacted]	[redacted]
Extensive	161 (37.5)	82 (38.0)	130 (35.4)	[redacted]	[redacted]	[redacted]

[a]Median is derived based on the ITT population from Cohort 1.
Abbreviations: ITT: Intent-to-Treat; N: number of patients in trial arm; n: number of patients in specific category; SD: standard deviation.
Sources: Sanborn *et al.* 2021.[4] TRUENORTH CSR: Tables 12–15.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Prior UC medication

The trial arms of Cohort 1 were generally well balanced with regard to prior UC medication (Table 14). Approximately 30% of patients in Cohort 1 had experienced an inadequate response, loss of response, or intolerance to prior TNFi use.[4] Of these patients, 35.0% failed to respond initially to at least one TNFi (primary non-response) and 64.6% lost response to a TNFi with time (secondary nonresponse).[4] Amongst these secondary non-responder patients approximately [redacted] had received two or more biologics and approximately [redacted] had received vedolizumab.

Of those patients who had received a prior biologic therapy the most commonly used biologic medications (used by [redacted] of patients) for patients in Cohort 1 in the induction period were infliximab, adalimumab and vedolizumab. This is reflective of current UK clinical practice, as described in Section B.1.3.4.[1, 75]

Within the maintenance period the trial arms were also generally well balanced with regards to prior UC medication use (Table 14). Prior treatments for UC included TNFis (as randomised) in approximately [redacted] of patients, and other biologics in approximately [redacted]. A slightly higher proportion of patients randomised to maintenance with ozanimod had been treated with a prior TNFi ([redacted]) compared to those re-randomised to placebo ([redacted]).

Full details of patients prior UC medication are reported in Appendix M.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 14: Prior UC medication (safety population)

Prior UC medication and response category	Induction period			Maintenance period		
	Cohort 1		Cohort 2	Placebo (N = 69) n (%)	Re-randomised	
	Ozanimod (N = 429) n (%)	Placebo (N = 216) n (%)	Ozanimod (N = 367) n (%)		Ozanimod – Placebo (N = 227) n (%)	Ozanimod – Ozanimod (N = 230) n (%)
Glucocorticoids	322 (75.1)	162 (75.0)	286 (77.9)	[redacted]	[redacted]	[redacted]
Immunomodulators	174 (40.6)	93 (43.1)	166 (45.2)	[redacted]	[redacted]	[redacted]
Oral aminosalicylates	418 (97.4)	210 (97.2)	362 (98.6)	[redacted]	[redacted]	[redacted]
TNFi[a]	130 (30.3)	65 (30.09)	159 (43.3)	[redacted]	[redacted]	[redacted]
Primary non-responder[a]	49 (37.7)	21 (32.3)	60 (37.7)	[redacted]	[redacted]	[redacted]
Secondary non-responder[a]	84 (64.6)	42 (64.6)	109 (68.6)	[redacted]	[redacted]	[redacted]
Intolerant[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-TNFi biologics	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Primary non-responder	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Secondary non-responder	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intolerant	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab[b]	71 (16.6)	38 (17.6)	93 (25.3)	[redacted]	[redacted]	[redacted]
Tofacitinib[b]	3 (0.7)	4 (1.9)	13 (3.5)	[redacted]	[redacted]	[redacted]

[a]Percentages for TNFi non-responders and intolerant patients were calculated as a percent of the number of patients who received prior TNFi treatment, rather than the total Safety Population for each trial arm. Primary non-response was defined as signs and symptoms of persistently active disease despite an adequate trial of induction treatment with an TNFi agent (per country's approved label). Secondary non-response was defined as recurrence of symptoms during maintenance dosing following prior clinical benefit. Intolerance included inability to achieve doses, dose levels, or treatment durations because of treatment-related side effects and/or laboratory abnormalities. Patients may be classified under more than 1 response category if they received more than 1 prior TNFi and experienced a different response to each therapy.
[b]Only selected prior non-TNFi biologics are presented here, hence values presented do not always add to 100%. For a full breakdown of previous non-TNFi biologics use please refer to the TRUENORTH CSR.
Abbreviations: N: number of patients in trial arm; n: number of patients in response category; TNFi: tumour necrosis factor inhibitor; UC: ulcerative colitis.
Sources: Sandborn *et al.* 2021.[4] TRUENORTH CSR: Tables 16 and 17.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.3.4.2 TNFi experience

The baseline characterises of patients in the induction and maintenance period of TRUENORTH by prior TNFi use are presented below.

Baseline demographic and disease characteristics

The baseline demographic characteristics of patients categorised according to TNFi exposure were well balanced across treatment arms in both the induction and maintenance period of TRUENORTH (Table 15 and Table 16, respectively). On average, patients with prior TNFi exposure had a longer time since UC diagnosis, reflecting time spent with the disease, number of treatments received and disease stage (Table 15 and Table 16). Overall, baseline characteristics in the TNFi-naïve and experienced subgroups were similar to the ITT population.

Prior UC medication

A small proportion of TNFi-naïve patients had prior exposure to vedolizumab, ustekinumab or tofacitinib (Table 17 and Table 18). Vedolizumab was the most common non-TNFi-biologic received, which is reflective of UK clinical practice, where vedolizumab may be used in a small proportion of patients who are contraindicated to TNFis or have specific safety concerns regarding their use (Section B.1.3.4).

Unsurprisingly, a far greater proportion of TNFi-experienced patients had received prior UC treatment with CvTs (see Appendix K) and non-TNFi biologics than TNFi-naïve patients. The most commonly used non-TNFi biologic was vedolizumab, reflecting current UK clinical practice (see Section B.1.3.4).

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 15: Patient baseline demographics and disease characteristics in TRUENORTH by TNFi exposure (Induction period)

	TNFi-naïve			TNFi-experienced		
	Cohort 1		Cohort 2	Cohort 1		Cohort 2 ([redacted])
	Ozanimod ([redacted])	Placebo ([redacted])	Ozanimod ([redacted])	Ozanimod ([redacted])	Placebo ([redacted])	
Male, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age, year, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Body mass index (kg/m²), n (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Region, n (%)						
North America	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Europe	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease characteristics						
Age at UC symptoms onset (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Years since UC diagnosis, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Extent of disease, n (%)						
Limited to left side of colon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Extensive	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: N: number of patients in trial arm, n: number of patients in specific category.
Source: TRUENORTH CSR: Table 14.1.5.1.2.1A –2A, Tables 14.1.5.2.1.1A–2A .[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 16: Patient baseline demographics and disease characteristics in TRUENORTH by TNFi exposure (Maintenance period)

	TNFi-naïve			TNFi-experienced		
		Re-randomised patients			Re-randomised patients	
	Placebo ([redacted])	Ozanimod – Placebo ([redacted]) n (%)	Ozanimod – Ozanimod ([redacted]) n (%)	Placebo ([redacted])	Ozanimod – Placebo ([redacted]) n (%)	Ozanimod – Ozanimod ([redacted]76) n (%)
Male, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Age, year, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Body mass index (kg/m²), n (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Region, n (%)						
North America	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Europe	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other (Asia pacific, South America, South Africa)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease characteristics						
Age at UC symptoms onset (years), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Years since UC diagnosis, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Extent of disease, n (%)						
Limited to left side of colon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Extensive	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: N: number of patients in trial arm, n: number of patients in specific category.
Source: TRUENORTH CSR: Tables 14.1.5.1.1B –2B, Tables 14.1.4.2.1B1B–2B.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 17: Prior UC medication by TNFi exposure (Induction period)

Prior UC medication and response category	TNFi-naïve			TNFi-experienced		
	Cohort 1		Cohort 2	Cohort 1		Cohort 2 ([redacted])
	Ozanimod ([redacted])	Placebo ([redacted])	Ozanimod ([redacted])	Ozanimod ([redacted])	Placebo ([redacted])	
Glucocorticoids	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Immunomodulators	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oral aminosalicylates	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-TNFi biologics	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Primary non-responder	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Secondary non-responder	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intolerant	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ustekinumab[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a]Only selected prior non-TNFi biologics are presented here, hence values presented do not always add to 100%. For a full breakdown of previous non-TNFi biologics use please refer to the TRUENORTH CSR.
Abbreviations: N: number of patients in trial arm; n: number of patients in response category; TNFi: tumour necrosis factor inhibitor; UC: ulcerative colitis.
Source: TRUENORTH CSR: Tables 14.1.8.2.1.1A–2Aa.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 18: Prior UC medication by TNFi exposure (Maintenance period)

Prior UC medication and response category	TNFi-naïve			TNFi-experienced		
		Re-randomised patients			Re-randomised patients	
	Placebo ([redacted])	Ozanimod – Placebo ([redacted]) n (%)	Ozanimod – Ozanimod ([redacted]) n (%)	Placebo ([redacted])	Ozanimod – Placebo ([redacted]) n (%)	Ozanimod – Ozanimod ([redacted]) n (%)
Glucocorticoids	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Immunomodulators	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oral aminosalicylates	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-TNFi biologics	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Primary non-responder	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Secondary non-responder	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intolerant	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ustekinumab[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a]Only selected prior non-TNFi biologics are presented here, hence values presented do not always add to 100%. In certain cases, patients may have received more than one previous non-TNFI biologic, meaning the total exceeds 100%. For a full breakdown of previous non-TNFi biologics use please refer to the TRUENORTH CSR.
Abbreviations: N: number of patients in trial arm, n: number of patients in specific category.
Source: TRUENORTH CSR: Tables 14.1.8.2.2A –2B.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.3.5 Participant flow

A CONSORT diagram showing patient flow through TRUENORTH is shown in Figure 8.

Induction period

A total of 1,012 patients were enrolled in the induction period of the study; 645 patients were enrolled into Cohort 1 (429 randomised to ozanimod and 216 to placebo) and 367 patients were enrolled into Cohort 2 (all treated with ozanimod) (Figure 8).[4] The majority of ozanimod-treated patients completed the 10-week induction period (93.5% in Cohort 1 and 88.9% in Cohort 2). A high proportion of patients receiving placebo in Cohort 1 also completed the induction period (88.9%). Of those receiving ozanimod, 54.3% and 61.0% in Cohort 1 and Cohort 2, respectively achieved clinical response and proceeded into the maintenance period. In comparison, only 31.9% of patients receiving placebo in Cohort 1 achieved clinical response and continued into the maintenance period.

Almost half as many patients receiving ozanimod in Cohort 1 discontinued treatment during the induction period compared to those receiving placebo (6.5% versus 11.1%, respectively). The primary reason for discontinuation in the placebo arm was lack of efficacy (41.7% of discontinuations) and patients withdrawing consent (33.3% of discontinuations). The primary reason for study discontinuation in the ozanimod arm was due to an AE, however the proportion of patients withdrawing due to AEs was comparable between the placebo and ozanimod arm in Cohort 1 (2.8% and 2.6% of those randomised to the placebo and ozanimod arms, respectively).[4]

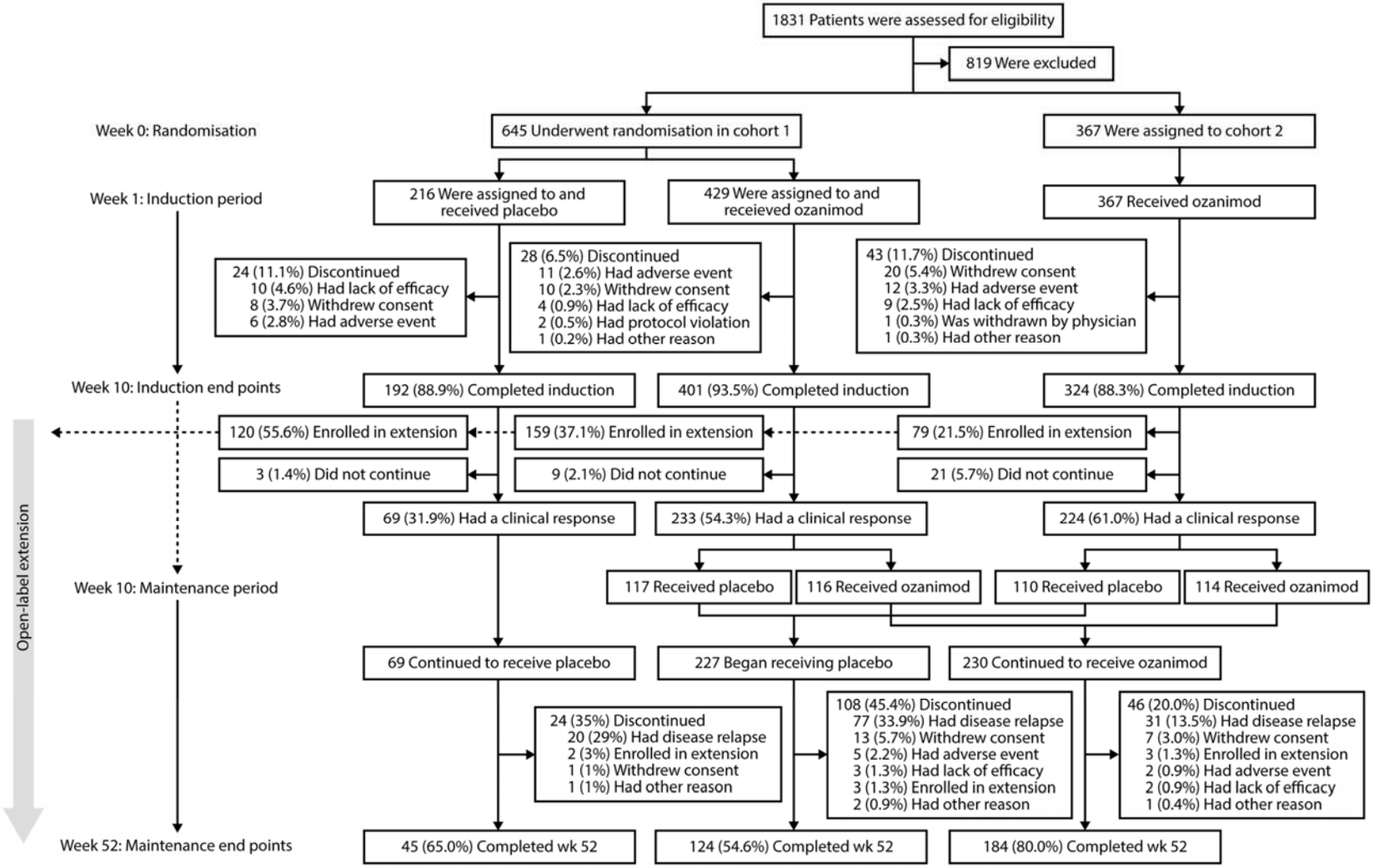
Maintenance period

Of the 457 patients who responded to ozanimod during the induction period, and were re-randomised into the maintenance period, 230 were randomised to maintenance with ozanimod and 227 were randomised to placebo. Patients receiving placebo in the induction period who achieved clinical response continued to receive placebo in the maintenance period.[4] The completion rate for the maintenance period was 80% in patients continuously treated with ozanimod, 55% in patients re-randomised from ozanimod to placebo and 65% in patients continuously treated with placebo.

The most frequently reported reason for study discontinuation from the maintenance period was disease relapse, which occurred in a higher proportion of patients re-randomised to placebo than ozanimod (33.9% and 13.5%, respectively).[4]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 8: CONSORT diagram for TRUENORTH



Abbreviations: OLE: open-label extension
Sources: Adapted from Sandborn *et al.* 2021.[4]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.4 *Quality assessment of the relevant clinical effectiveness evidence*

A summary of the quality assessment for TRUENORTH is presented in Table 19, with full details given in Appendix B.4.4.

Table 19: Overview of quality assessment for TRUENORTH

Criteria	Outcome	Justification	Relevant section in Document B
Was randomisation carried out appropriately?	Yes	Patients were randomised using the interactive voice/web-based activated response system (IVRS/IWRS) and stratified randomisation was centrally allocated across all centres via the IVRS/IWRS	Table 9
Was the concealment of treatment allocation adequate?	Yes	Patients were assigned to treatment/randomised using the IVRS/IWRS	Table 9
Were the trial arms similar at the outset of the study in terms of prognostic factors	Yes	Baseline characteristics were balanced between treatment arms	B.2.3.4
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Ozanimod and placebo capsules were identical in physical appearance and the treatment each patient received was not disclosed to the Investigator, site staff, patient, Sponsor, or the clinical staff at the Contract Research Organisation involved with trial conduct or data collection/analysis	Table 9
Were there any unexpected imbalances in drop-outs between trial arms?	No	There were imbalances in drop-out with the placebo arm having twice as many drop-outs as ozanimod in the induction and maintenance period. However, as the main reason for dropout was lack of efficacy this imbalance was not unexpected	B.2.3.5
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Data for all reported endpoints were available	N/A
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	ITT population was used. NRI was used to account for missing data where patients with missing Week 10 efficacy data for the induction period and/or patients with missing Week 52 efficacy data for the maintenance period were considered non-responders	Section B.2.3.3

Abbreviations: ITT: intention to treat; N/A: not applicable; NRI: non-responder imputation; IVRS/IWRS: web-based activated response system.
Sources: Sandborn *et al*. 2021.[4] TRUENORTH CSR.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.5 *Clinical effectiveness results of the relevant trials*

As clinical decision making in UK clinical practice for moderately to severely active UC patients following failure on CvT is typically influenced by patients' suitability for or prior exposure to TNFis (Section B.1.1), the population in this submission is split into two distinct subpopulations: TNFi-naïve and TNFi-experienced. Clinical effectiveness results for the overall ITT population of TRUENORTH are presented for below for completeness. Clinical effectiveness results for the TNFi-naïve and TNFi-experienced populations are presented in Section B.2.6.1.

Summary – Overall population

- At Week 10 a statistically significantly greater proportion of patients treated with ozanimod achieved clinical remission compared with the placebo arm in Cohort 1 (18.4% vs 6.0%, p<0.001).
- In addition, each of the key secondary efficacy endpoints (clinical response, endoscopic improvement, and mucosal healing) were achieved by a statistically significantly greater proportions of patients in the ozanimod arm versus the placebo arm (each p<0.001) in Cohort 1 at Week 10
- In the maintenance period (Week 52), a statistically significantly greater proportion of patients randomised to ozanimod achieved clinical remission compared to placebo (37.0% vs 18.5%, p<0.0001)
- Similarly, each of the key secondary efficacy endpoints (clinical response, endoscopic improvement, maintenance of remission, corticosteroid-free remission, mucosal healing, and durable clinical remission) were achieved by a statistically significantly greater proportions of patients randomised to maintenance with ozanimod compared to placebo (each p<0.001) at Week 52
- Ozanimod-treated patients in clinical response at Week 10 who were randomised to maintenance with ozanimod where 2.3-times more likely to still be in clinical response at Week 52 compared to placebo. (Odds ratio [OR]: 2.27; p<0.0001)

B.2.5.1 Induction period – Overall population

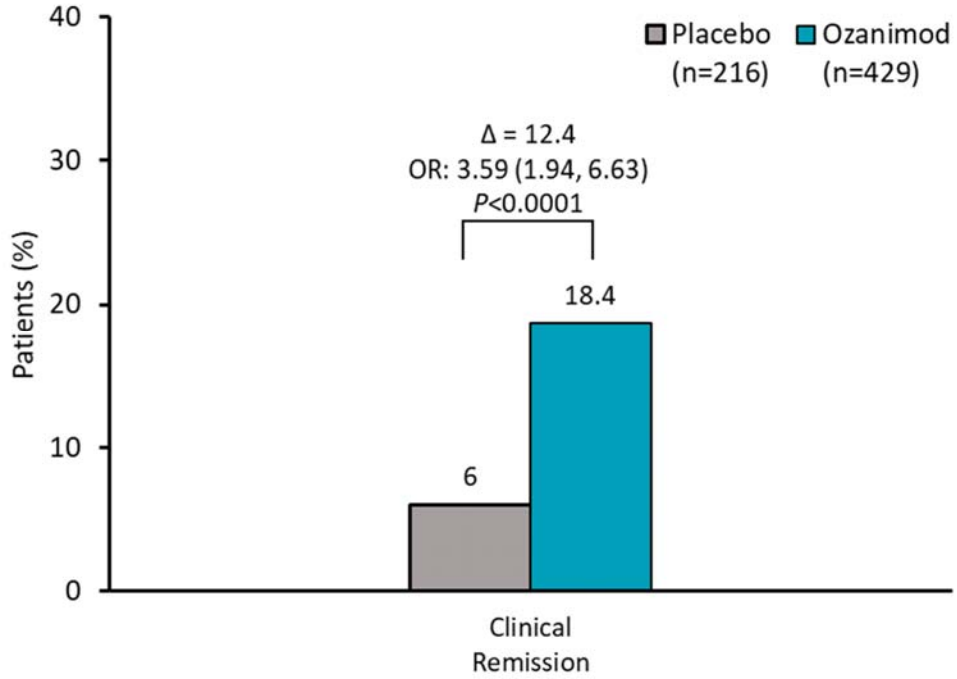
B.2.5.1.1 Primary endpoint

Clinical remission at Week 10

At Week 10, a statistically significantly greater proportion of patients achieved clinical remission in the ozanimod arm compared with the placebo arm in Cohort 1 (18.4% vs 6.0%, p<0.0001; Figure 9).[4] Clinical remission is associated with not only symptomatic but also endoscopic improvement and is therefore an important and clinically relevant outcome measure reflecting real aims of clinical practice.[48] These results provide strong evidence that ozanimod is able to induce clinical remission in patients, improving patients HRQoL and reducing the disease burden.[48]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 9: Proportion of patients in clinical remission at Week 10 (induction period; ITT population)



Notes: Odds ratio (ozanimod/placebo), treatment difference, and 2-side 95% Wald CI and p-value for comparison between Cohort 1 ozanimod and placebo arm are based on the CMH test, stratified by corticosteroid use at Screening and prior TNFi use (yes or no). Patients with any of RBS, SFS, and endoscopy subscores missing at Week 10 were classified as non-remitters.
Abbreviations: CI: Confidence interval; CMH: Cochran-Mantel-Haenszel; ITT: Intent-to-Treat; N: number of patients in trial arm, n: number of patients that achieved the endpoint; NRI; non-responder imputation; RBS: rectal bleeding subscore; OR: odds ratio; SFS: stool frequency subscore; TNFi: tumour necrosis factor alpha-inhibitor.
Source: TRUENORTH CSR: Table 18;[5] Sandborn *et al.* 2021.[4]

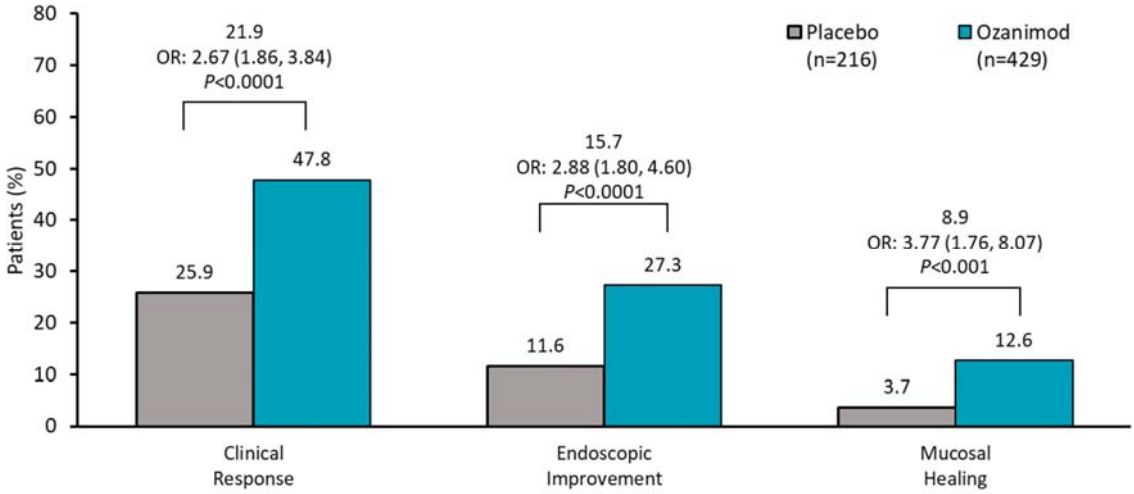
B.2.5.1.2 Key secondary endpoints

Key secondary endpoints for the induction period of TRUENORTH, as presented in Figure 10, were: the proportion of patients in clinical response, the proportion of patients with endoscopic improvement and the proportion of patients with mucosal healing at Week 10 (as discussed in Section B.2.3.1, mucosal healing during TRUENORTH had a stricter definition than trials for existing therapies in UC).

At Week 10, a statistically significantly greater proportion of patients in the ozanimod arm achieved clinical response (3-component Mayo; 47.8% vs 25.9%, p<0.0001), endoscopic improvement (27.3% vs 12.0%, p<0.0001) and mucosal healing (12.6% versus 3.7%, p<0.001) compared with the placebo arm in Cohort 1 (Figure 10).[4] Improvements in clinical response and mucosal healing have been linked to lower rates of disease relapse, hospitalisation, colectomy, as well as improved HRQoL in patients with UC.[90, 92, 125] These therefore represent key outcomes for patients, and have all demonstrated significant improvements with treatment with ozanimod compared to placebo.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 10: Proportion of patients with clinical response, endoscopic improvement and mucosal healing (induction period; ITT population)



Notes: Odds ratio (ozanimod/placebo), treatment difference, and 2-side 95% Wald CI and p-value for comparison between Cohort 1 ozanimod and placebo arm are based on the CMH test, stratified by corticosteroid use at Screening and prior TNFi use (yes or no). Patients with any of RBS, SFS, and endoscopy subscores missing at Week 10 were classified as non-remitters.
Abbreviations: CI: Confidence interval; CMH: Cochran-Mantel-Haenszel; ITT: Intent-to-Treat; N: number of patients in trial arm, n: number of patients that achieved the endpoint; NRI; non-responder imputation; RBS: rectal bleeding subscore; OR: odds ratio; SFS: stool frequency subscore.
Source: TRUENORTH CSR: Table 20, Table 21, Table 22[5]; Sandborn *et al.* 2021.[4]

B.2.5.1.3 Other efficacy endpoints

Change from baseline in 3-component Mayo score at Week 10

Following the trial induction period, patients treated with ozanimod had a significantly greater reduction in their 3-component mayo score from baseline compared those in the placebo arm. ([redacted]). (Table 20).

Table 20: Change from baseline in 3-component Mayo Score at Week 10 (induction period; ITT population, observed cases)

	Cohort 1	
	Ozanimod ([redacted])	Placebo ([redacted])
Baseline[a], mean (SD)	[redacted]	[redacted]
Change from baseline, LS mean (SE)[b]	[redacted]	[redacted]
LSMD (95% CI)[b]	[redacted]	
P-value[b]	[redacted]	

[a]Baseline is derived from the latest subscores on or prior to the date of initial dose.[b]Based on ANCOVA for change from baseline adjusted for corticosteroid use at Screening (yes or no), prior TNFi use (yes or no), and the Baseline 4-component, partial, or 3-component Mayo score.
Abbreviations: ANCOVA: analysis of covariance; CI: confidence interval; ITT: Intent-to-Treat; LS: least squares; LSMD: least squares mean difference; SD: standard deviation; SE: standard error; TNF: tumour necrosis factor.
Source: TRUENORTH CSR: Table 29.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Health-related quality of life (HRQoL)

In the induction period of TRUENORTH, HRQoL was assessed using the Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) and the EuroQol-5 Dimension Questionnaire (EQ-5D), using the latest five-level (5L) version, from Baseline to Week 10. Further information on these two surveys is provided in Table 10.

SF-36

The SF-36 PCS score was significantly improved ([redacted]) in patients treated with ozanimod compared to placebo in Cohort 1 during the induction period. The MCS score showed a numerical (but not statically significant) improvement ([redacted]) in patients randomised to ozanimod ([redacted]) versus placebo ([redacted]) in Cohort, however individual domain scores for the MCS such as vitality ([redacted]), social functioning ([redacted]), and mental health ([redacted]) were significantly improved during the induction period. Scores were also found to be improved on the SF-36 global health ([redacted]) and health utility score ([redacted]) for patients treated with ozanimod versus placebo in Cohort 1, suggesting treatment with ozanimod led to improved patient HRQoL compared to placebo.

The minimally clinically important difference (MCID) for the SF-36 PCS and MCS scores was defined as a ≥ 5 point improvement in each score from Baseline to Week 10. For the PCS score, a nominally significantly ([redacted]) greater proportion of patients treated with to ozanimod ([redacted]) achieved an MCID compared to those randomised to placebo ([redacted]). (Figure 11). No difference in the proportion of patients who achieved an MCID in the SF-36 MCS was observed.

Figure 11: MCID for SF-36 PCS at Week 10 (induction period; ITT population; observed cases)

[redacted]

[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied
Abbreviations: ITT: intention-to-treat; n: number of patients in subgroup; NRI; non-responder imputation; OR: odd ratio.
Source: TRUENORTH CSR.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

EQ-5D-5L

At Week 10 the mean change from baseline in EQ-5D summary index score for patients treated with ozanimod was statistically significantly greater than the mean change from baseline in patients treated with placebo in Cohort 1 (nominal [redacted]; Figure 12). The EQ-5D self-reported questionnaire includes a visual analogue scale (VAS), which records the respondent's self-rated health status on a graduated (0–100) scale. The mean change from baseline in EQ-5D VAS score for ozanimod was statistically significantly greater (nominal [redacted]; Figure 13) than the mean change from baseline for placebo.

These HRQoL results demonstrate that the improvements in efficacy results in patients treated with ozanimod compared to placebo translate to improvements in patients' HRQoL.

Figure 12: Mean change from Baseline to Week 10 in EQ-5D summary index scores (induction period; ITT population, observed cases)

Abbreviations: EQ-5D: EuroQoL-5 Dimension: ITT: intention-to-treat.
Source: TRUENORTH CSR: Table 14.2.8.1A.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 13: EQ-5D Visual Analogue Scale change from Baseline at Week 10 (induction period; ITT population, observed cases)

Abbreviations: EQ-5D: EuroQoL-5 Dimension: ITT: intention-to-treat.
Source: TRUENORTH CSR: Table 14.2.8.2A.[5]

Health resource utilisation

A health resource utilisation (HRU) questionnaire was used to collect data on reported doctor visits, emergency room visits, and hospitalisations in the induction period. However, a low overall number of doctor visits, emergency room visits, and hospitalisations in the induction period made inferring trial arm differences in HRU difficult. Overall, a low rate of hospitalisations occurred in the induction period and was consistent across both the ozanimod 1 mg and placebo arms of Cohort 1 (and , respectively).[5]

B.2.5.2 Maintenance period

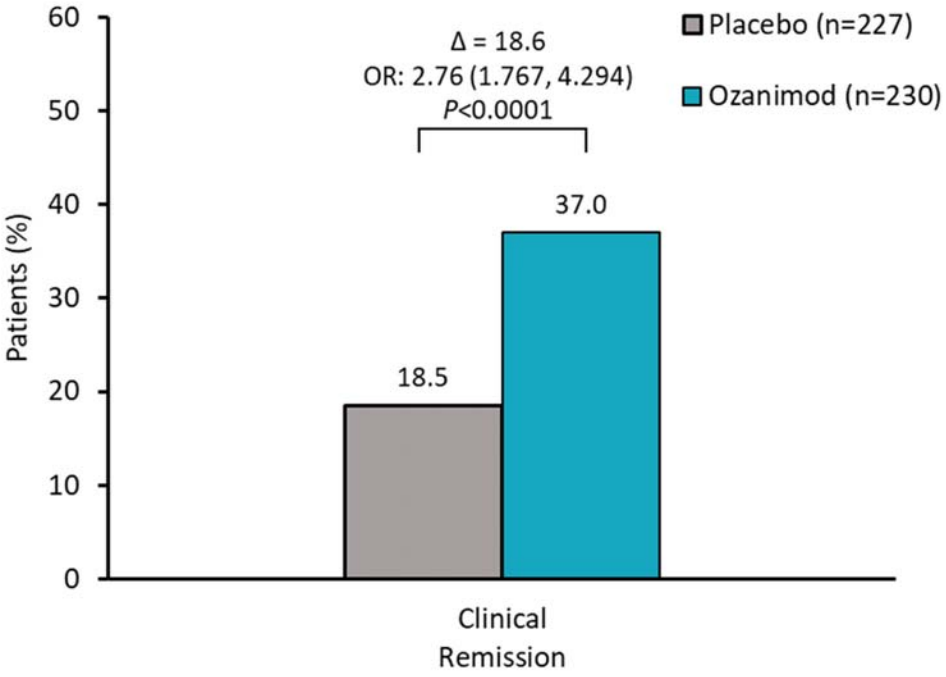
B.2.5.2.1 Primary endpoint

Clinical remission at Week 52

Clinical remission at Week 52, was achieved in a greater proportion of patients randomised to ozanimod maintenance treatment compared with placebo (37.0% vs 18.5%, p<0.0001; Figure 14).[4, 126] The results for clinical remission at Week 52 provide strong evidence for ozanimod's ability to maintain clinical remission in the longer term following clinical response at induction.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 14: Proportion of patients in clinical remission at Week 52 (maintenance period; ITT population)



Notes: Odds ratio (ozanimod/placebo), treatment difference, and 2-side 95% Wald CI and p-value for comparison between patients randomised to ozanimod or placebo are based on the CMH test, stratified by corticosteroid use at Screening and prior TNFi use (yes or no).
Abbreviations: CI: Confidence interval; CMH: Cochran-Mantel-Haenszel; ITT: Intent-to-Treat; N: number of patients in trial arm, n: number of patients that achieved the endpoint; NRI; non-responder imputation; RBS: rectal bleeding subscore; OR: odds ratio; SFS: stool frequency subscore; TNFi: tumour necrosis factor alpha-inhibitor.
Source: TRUENORTH CSR: Table 19.[5] Sandborn *et al.* 2021.[4]

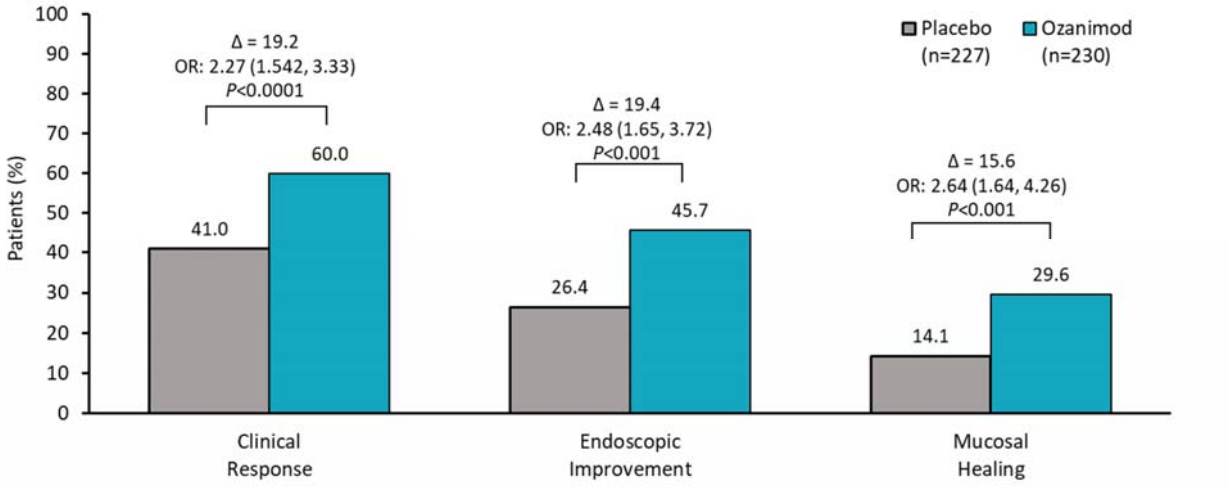
B.2.5.2.2 Key secondary endpoints

Key secondary endpoints for the maintenance period of TRUENORTH in ascending hierarchal order were: clinical response, endoscopic improvement, maintenance of remission, corticosteroid-free remission, mucosal healing and durable clinical remission.

At the end of Week 52, a statistically significantly greater proportion of patients randomised to maintenance treatment with ozanimod compared to placebo achieved clinical response (60.0% versus 41.0%, p<0.0001), endoscopic improvement (45.7% versus 26.4%, p<0.0001) and mucosal healing (29.6% versus 14.1%, p<0.001) (Figure 15).[4]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 15: Proportion of patients with clinical response, endoscopic improvement and mucosal healing (maintenance period; ITT population)



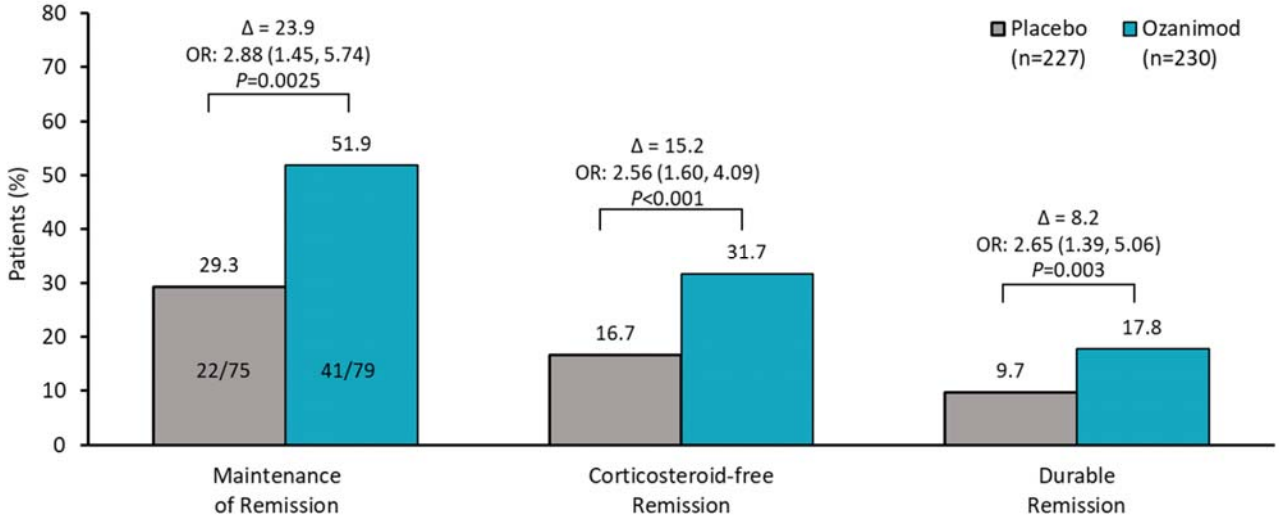
Notes: Odds ratio (ozanimod/placebo), treatment difference, and 2-side 95% Wald CI and p-value for comparison between patients re-randomised to ozanimod and placebo are based on the CMH test, stratified by corticosteroid use at Screening and prior TNFi use (yes or no).
Abbreviations: CI: Confidence interval; CMH: Cochran-Mantel-Haenszel; ITT: Intent-to-Treat; N: number of patients in trial arm, n: number of patients that achieved the endpoint; NRI; non-responder imputation; RBS: rectal bleeding subscore; OR: odds ratio; SFS: stool frequency subscore; TNFi: tumour necrosis factor alpha-inhibitor.
Source: TRUENORTH CSR: Tables 23, 24 and 27.[5] Sandborn *et al.* 2021.[4]

Patients treated with ozanimod in the maintenance period compared to placebo achieved statistically greater rates of corticosteroid-free remission (31.7% vs 16.7%, p<0.001), maintenance of clinical remission (51.9% versus 29.3%, p=0.0025) and durable clinical remission (17.8% versus 9.7%, p=0.003).[4]

Long-term corticosteroid free remission is an important marker and target of UC treatment and is associated with improved HRQoL in UC patients and a decreased risk of surgery.[23, 53] Meanwhile, maintenance and durability of disease remission are key measures indicating patients do not relapse and therefore maintain reduced symptom burden achieved during induction.[127] Overall, these results demonstrate that treatment with ozanimod leads to a prolonged reduction in symptom burden and hence prolonged improvements in patient HRQoL compared to treatment with placebo.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 16: Proportion of patients with maintenance of remission, corticosteroid-free remission and durable remission (maintenance period; ITT population)



Notes: Odds ratio (ozanimod/placebo), treatment difference, and 2-side 95% Wald CI and p-value for comparison between patients re-randomised to ozanimod and placebo are based on the CMH test, stratified by corticosteroid use at Screening and prior TNFi use (yes or no).
Abbreviations: CI: Confidence interval; CMH: Cochran-Mantel-Haenszel; ITT: Intent-to-Treat; N: number of patients in trial arm, n: number of patients that achieved the endpoint; NRI; non-responder imputation; RBS: rectal bleeding subscore; OR: odds ratio; SFS: stool frequency subscore; TNFi: tumour necrosis factor alpha-inhibitor.
Source: TRENORTH CSR: Tables 25, 26 and 27;[5] Sandborn *et al* 2021.[4]

B.2.5.2.3 Other efficacy endpoints

Change from baseline in 3-component Mayo score at Week 52

Patients treated with Ozanimod during the maintenance period had a significantly greater reduction in their 3-component mayo score from baseline compared those in the placebo arm ([redacted]; Table 21).

Table 21: Change from baseline in 3-componenet Mayo Score at Week 52 (maintenance period; ITT population, observed cases)

	Re-randomised patients	
	Ozanimod – Placebo ([redacted])	Ozanimod – Ozanimod ([redacted])
Baseline[a], mean (SD)	[redacted]	[redacted]
Change from baseline, LS mean (SE)[b]	[redacted]	[redacted]
LSMD (95% CI)[b]	[redacted]	
P-value[b]	[redacted]	

[a]Baseline is derived from the latest subscores on or prior to the date of initial dose.[b]Based on ANCOVA for change from baseline adjusted for corticosteroid use at Screening (yes or no), prior TNFi use (yes or no), and the Baseline 4-component, partial, or 3-component Mayo score.
Abbreviations: ANCOVA: analysis of covariance; CI: confidence interval; ITT: Intent-to-Treat; LS: least squares; LSMD: least squares mean difference; SD: standard deviation; SE: standard error; TNF: tumour necrosis factor.
Source: TRUENORTH CSR: Table 34.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

HRQoL

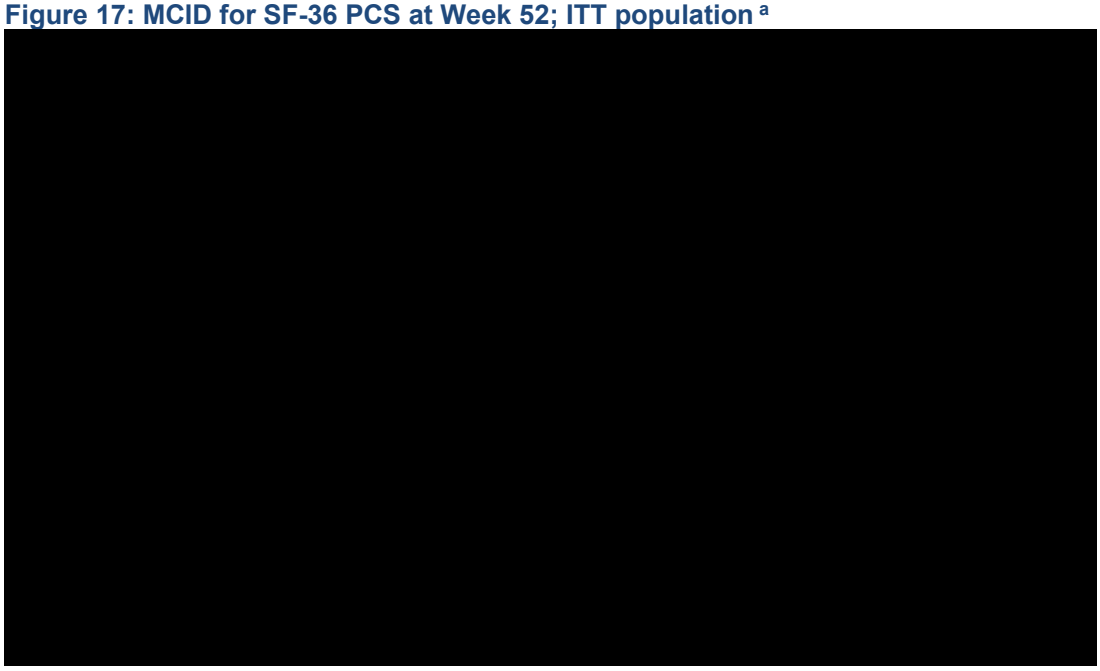
In the maintenance period of TRUENORTH, HRQoL was assessed again via SF-36 and the EQ-5D (Table 10).

SF-36

Scores on SF-36 generally improved during the maintenance period for both patients randomised to ozanimod or placebo. However, the PCS score was more significantly improved (nominal [redacted]) in patients randomised to maintenance with ozanimod compared to placebo at Week 52.

The MCID (improvement of ≥ 5 points) for the SF-36 PCS score was achieved in a nominally significantly ([redacted]) greater proportion of patients treated with ozanimod ([redacted] versus placebo ([redacted]) at Week 52 (Figure 17). No difference in the proportion of patients achieving an MCID in the SF-36 MCS score was observed.

Figure 17: MCID for SF-36 PCS at Week 52; ITT population [a]



[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied
Abbreviations: ITT: intention-to-treat; n: number of patients in subgroup; NRI; non-responder imputation; OR: odd ratio.
Source: TRUENORTH CSR.[5]

EQ-5D-5L

At Week 52, patients randomised to maintenance with ozanimod had nominally significantly ([redacted]) improved scores on the EQ-5D VAS (Figure 18)**Error! Reference source not found.** but not for the EQ-5D Summary Index ([redacted]), relative to placebo at Week 52. These results demonstrate the long-term improvements in patients' HRQoL associated with ozanimod, as compared to placebo.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 18: EQ-5D VAS mean change from Baseline at Week 52 (maintenance period; ITT population, observed cases)

[a]Based on analysis of covariance (ANCOVA) for change from Baseline adjusted for corticosteroid use at screening (yes or no), prior TNFi use (yes or no), and the baseline EQ-5D visual analogue scale.
Abbreviations: EQ-5D: EuroQoL-5 Dimension: ITT: intention-to-treat.
Source: TRUENORTH CSR: Table 14.2.8.2A.[5]

Figure 19: EQ-5D summary index mean change from Baseline at Week 52 (maintenance period; ITT population, observed cases)

Abbreviations: EQ-5D: EuroQoL-5 Dimension: ITT: intention-to-treat.
Source: TRUENORTH CSR: Table 14.2.12.1B.[5]

Health resource utilisation

©Celgene Ltd, a BMS Company (2022). All rights reserved

As in the induction period a HRU questionnaire was used to collect data on reported doctor visits, emergency room visits, and hospitalisations in the maintenance period. The low overall number of doctor visits, emergency room visits, and hospitalisations however made inferring trial arm differences in health resource utilisation difficult. Overall, a very low rate of hospitalisations occurred in both patients randomised to maintenance with ozanimod or placebo during the maintenance period. ([redacted] and [redacted], respectively).[5]

B.2.6 *Subgroup analysis*

B.2.6.1 TNFi exposure

Subgroup analyses were performed on all the primary and key secondary endpoints in the TNFi-naïve and experienced population at Week 10 and Week 52 for induction and maintenance. Results for these subgroups have been presented as these were the key populations in the economic analysis, nominal P-values are reported for all analyses.

Summary of results

- In the induction period in both the TNFi-naïve and TNFi-experienced subgroups, patients treated with ozanimod had greater rates of clinical remission compared with placebo ([redacted] versus [redacted] and [redacted] versus [redacted] for the TNFi-naïve and experienced populations, respectively).
- Similarly, in both the TNFi-naïve and experienced subgroups treatment with ozanimod resulted in a greater proportion of patients of patients achieving clinical response, endoscopic improvement and mucosal healing compared to placebo.
- In the maintenance period treatment with ozanimod in both the TNFi-naïve and experienced patients results in a greater proportion of patients achieving clinical remission and clinical response compared with placebo ([redacted] versus [redacted] and 55.3% vs 24.6% for the TNFi-naïve and experienced population, respectively).[126]
- Similarly, a greater proportion of both TNFi-naïve and experienced patients randomised to maintenance with ozanimod met the key secondary endpoints of the maintenance period compared to placebo including endoscopic improvement, mucosal healing, maintenance of remission, corticosteroid-free remission and durable clinical remission at Week 52.
- Overall, a greater proportion of TNFi-naïve patients achieved the primary and key secondary endpoints compared to the TNFi-experienced population. This was in line with clinical expectation as patients with prior TNFi exposure typically have more advanced disease
- The results were broadly in line with the ITT population.

B.2.6.1.1 Induction

Clinical remission at Week 10

In line with the overall population, at Week 10, more patients achieved clinical remission with ozanimod compared to placebo in both the TNFi-naïve ([redacted] vs [redacted]; Figure 20) and TNFi-experienced subgroups ([redacted] versus [redacted]; Figure 21). The lower rates of remission and smaller difference observed in the TNFi-experienced group is clinically reasonable, as patients with prior exposure to TNFis typically have disease which is more difficult to treat and are therefore less likely to enter remission.

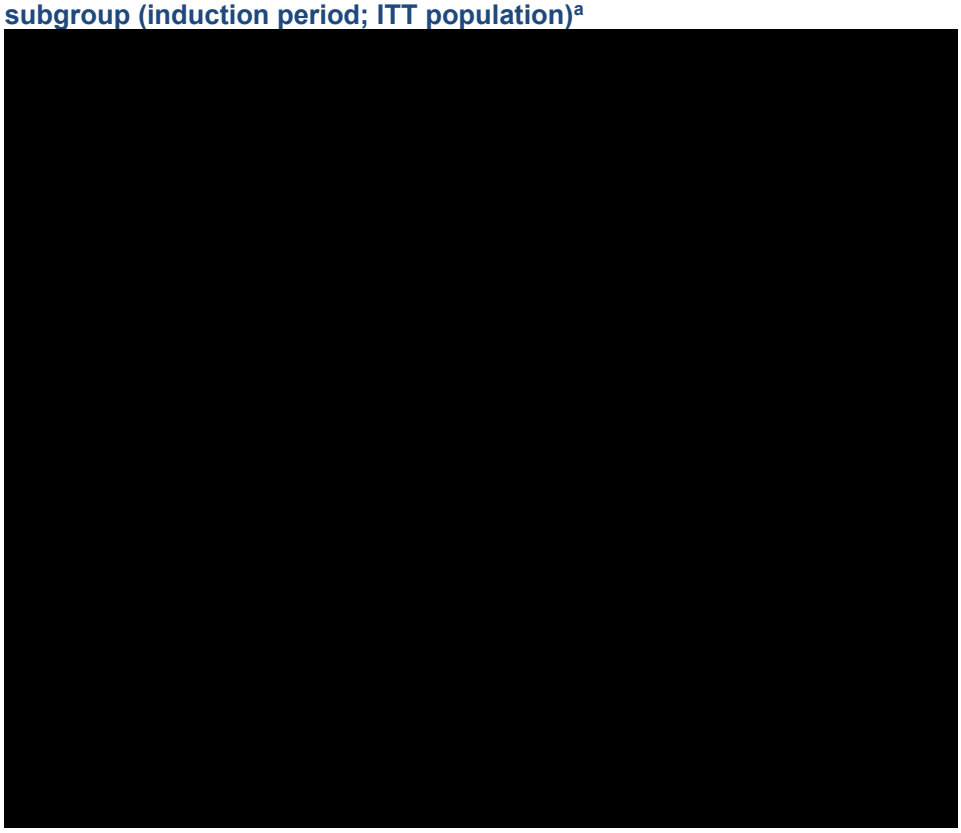
©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 20: Proportion of patients with clinical remission at Week 10, TNFi-naïve subgroup (induction period)[a]

[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: CI: Confidence interval; ITT: intent-to-treat; N: number of patients in subgroup; n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 33.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 21: Proportion of patients with clinical remission at Week 10, TNFi-experienced subgroup (induction period; ITT population)[a]



[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: CI: Confidence interval; ITT: intent-to-treat; N: number of patients in subgroup; n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 33.[5]

Secondary outcomes

In line with the overall population, clinical response at Week 10 was higher in patients randomised to ozanimod compared with placebo in both the TNFi-naïve (vs ; Figure 22) and TNFi-experienced subgroups (% vs %; Figure 23).

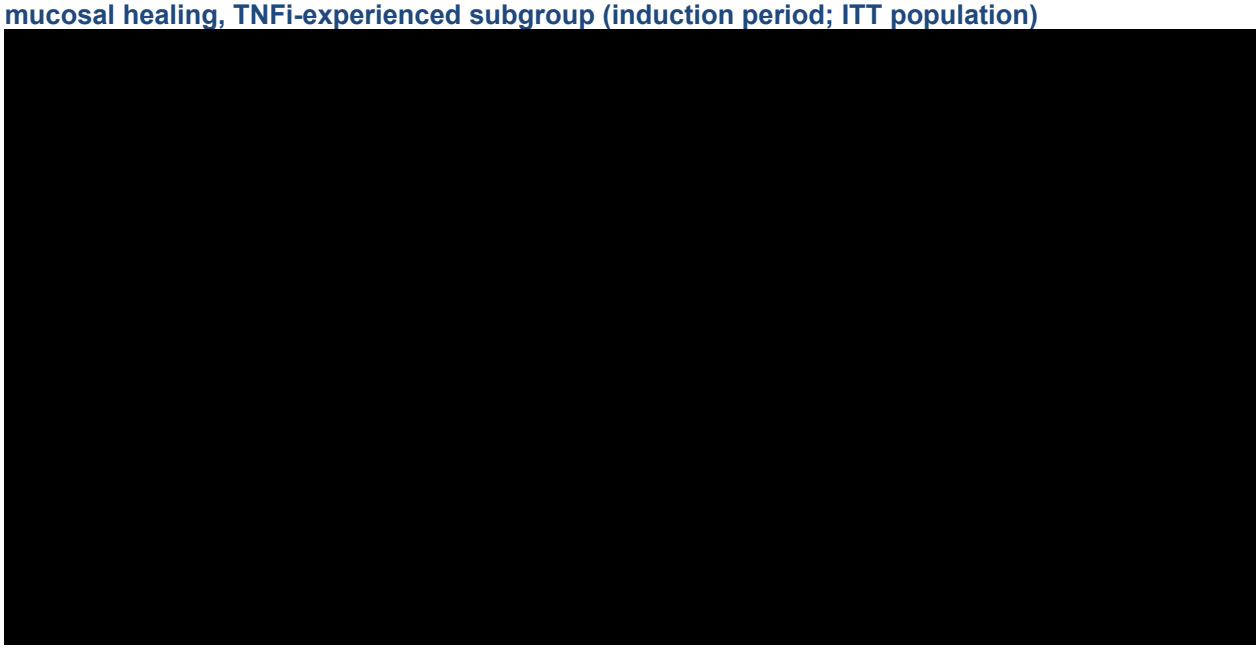
Further to this, more patients randomised to ozanimod achieved endoscopic improvement (versus ; Figure 22) and mucosal healing (vs ; Figure 22) compared to placebo in the TNFi-naïve subgroup. Similarly, more patients randomised to ozanimod compared with placebo achieved mucosal healing (versus ; Figure 23) and endoscopic improvement (versus ; Figure 23) in the TNFi-experienced subgroup.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 22 : Proportion of patients with clinical response, endoscopic improvement and mucosal healing, TNFi-naïve subgroup (induction period; ITT population)

[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: ITT: intent-to-treat; N: number of patients in subgroup, n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 33.[5]

Figure 23: Proportion of patients with clinical response, endoscopic improvement and mucosal healing, TNFi-experienced subgroup (induction period; ITT population)



[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: ITT: intent-to-treat; N: number of patients in subgroup, n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 33.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.6.1.2 Maintenance

Clinical remission at Week 52

At Week 52 a greater proportion of patients randomised to maintenance with ozanimod achieved clinical remission compared to placebo in the TNFi-naive ([redacted] versus [redacted]; Figure 24). And TNFi-experienced subgroups ([redacted] versus [redacted]; Figure 24). These results were approximately aligned with the results for the overall population and provide evidence that ozanimod is able to maintain reduced disease symptoms (Section B.2.5.1), improving patients HRQoL and reducing the disease burden.[48]

Figure 24: Proportion of patients in clinical remission at Week 52, TNFi-naïve subgroup (maintenance period; ITT population)[a]

[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: CI: Confidence interval; ITT: intent-to-treat; N: number of patients subgroup; n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 38.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 25: Proportion of patients in clinical remission at Week 52, TNFi-experienced subgroup (maintenance period; ITT population)[a]

[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: CI: Confidence interval; ITT: intent-to-treat; N: number of patients subgroup; n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 38.[5]

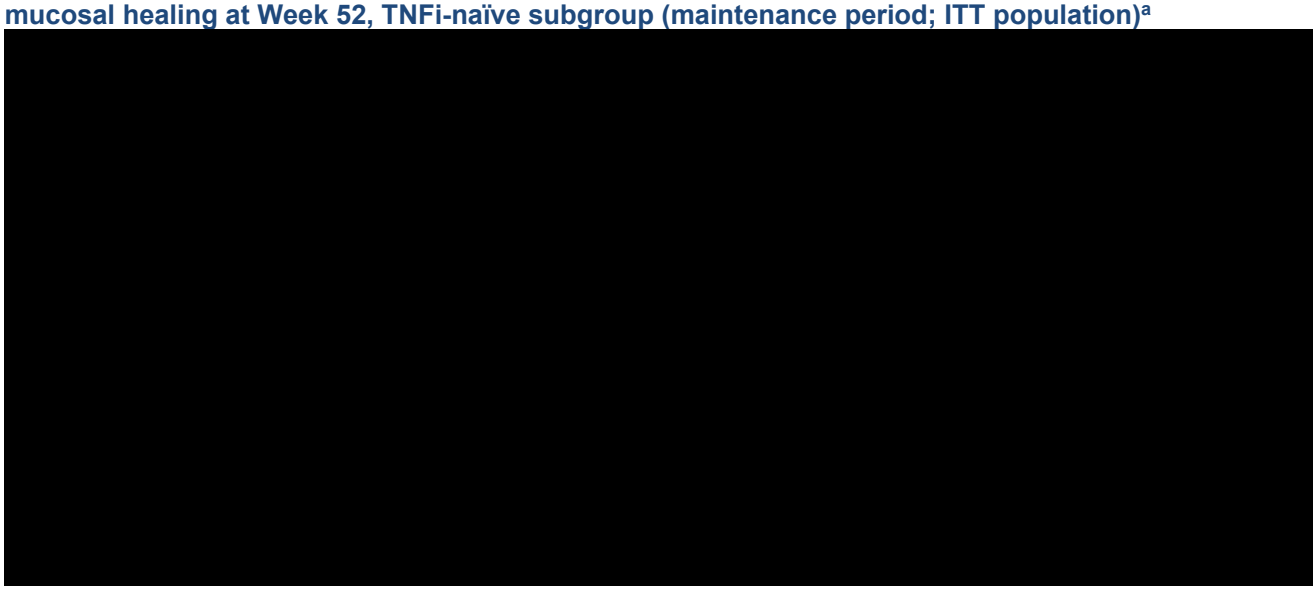
Secondary outcomes

In line with the overall population, at Week 52, a larger proportion of patients randomised to maintenance with ozanimod achieved clinical response compared to placebo in the TNFi-naïve (versus ; Figure 26) and TNFi-experienced subgroups (% vs %; Figure 27).[126]

Additionally, more patients randomised to maintenance with ozanimod achieved endoscopic improvement (versus) and mucosal healing (versus ,) compared to placebo in the TNFi-naïve subgroup (Figure 26). Similarly, a greater number of patients randomised to maintenance with ozanimod achieved endoscopic improvement (vs) and mucosal healing (vs) in the TNFi-experienced subgroup at Week 52 (Figure 27). These results demonstrate the ability of ozanimod to reduce the disease burden on patients and hence improve patients HRQoL (Section B.2.3.4.1).

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 26: Proportion of patients with clinical response, endoscopic improvement and mucosal healing at Week 52, TNFi-naïve subgroup (maintenance period; ITT population)[a]



[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: CI: Confidence interval; ITT: intent-to-treat; N: number of patients subgroup; n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 38.[5]

Figure 27: Proportion of patients with clinical response, endoscopic improvement and mucosal healing at Week 52, TNFi-experienced subgroup (maintenance period; ITT population)[a]

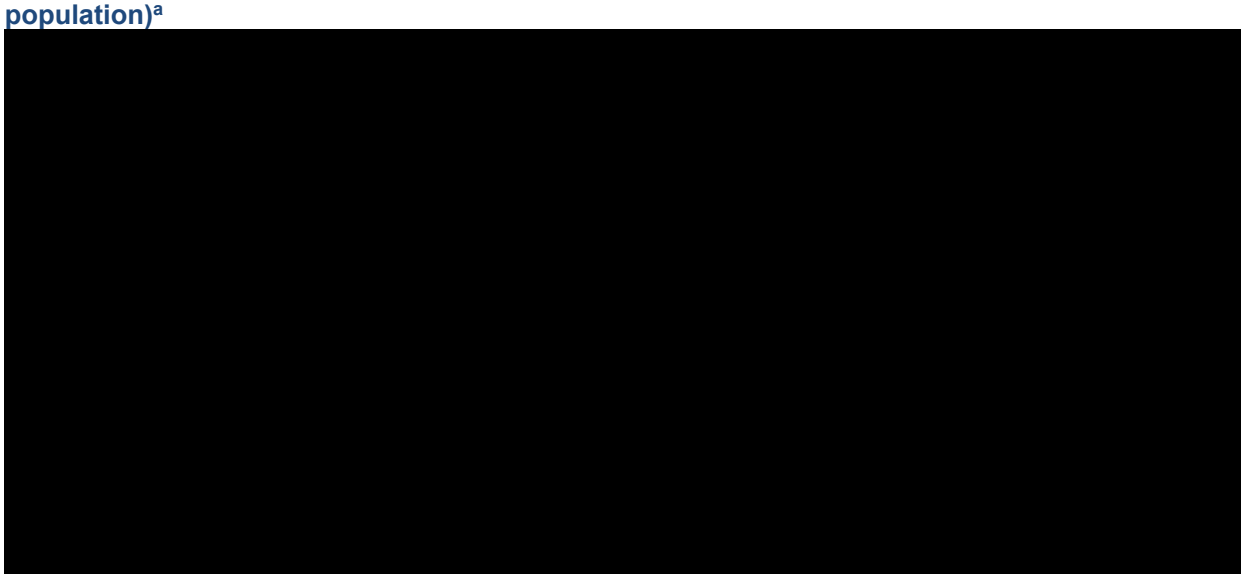
[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: CI: Confidence interval; ITT: intent-to-treat; N: number of patients subgroup; n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 38.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

At Week 52 a greater number of patients randomised to maintenance with ozanimod achieved maintenance of remission ([redacted] versus [redacted]), corticosteroid-free remission ([redacted] versus [redacted]) and durable clinical remission ([redacted] versus [redacted]) compared to placebo in the TNFi-naïve subgroup (Figure 28). Similarly, more patients randomised to maintenance with ozanimod achieved maintenance of remission ([redacted] vs [redacted]), corticosteroid-free remission ([redacted] vs [redacted]) and durable clinical remission ([redacted] versus [redacted]) when compared to placebo in the TNFi-experienced subgroup at Week 52 (Figure 29). These results were approximately in line with the overall population (Section B.2.3.4.1) and again highlight the ability of ozanimod to reduce the disease burden on patients and hence lead to improvements in patient HRQoL (Section B.2.3.4.1).

Therefore, ozanimod proved more effective than placebo in all primary and secondary endpoints in both TNFi-naïve and experienced subgroups, as well as in the overall population. This demonstrates that ozanimod is effective in treating UC in patients, regardless of previous exposure to TNFis.

Figure 28: Proportion of patients with maintenance of remission, corticosteroid-free remission and durable remission at Week 52, TNFi-naïve subgroup (maintenance period; ITT population)[a]



[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: CI: Confidence interval; ITT: intent-to-treat; N: number of patients subgroup; n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 38.[5]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 29: Proportion of patients with maintenance of remission, corticosteroid-free remission and durable remission at Week 52, TNFi-experienced subgroup (maintenance period; ITT population)[a]

[a]P-values < 0.05 are considered nominally significant, because no multiplicity adjustment was applied.
Abbreviations: CI: Confidence interval; ITT: intent-to-treat; N: number of patients subgroup; n: number of patients that achieved the endpoint; TNFi: tumour necrosis factor inhibitor.
Source: TRUENORTH CSR: Table 38.[5]

B.2.6.1.3 Sensitivity analysis: 4-component Mayo score

Sensitivity analyses were conducted using the 4-component Mayo score as opposed to the 3-component score used in the primary analysis.

Results based off the 4-component Mayo score were incorporated into the network meta-analysis (NMA) (Section B.2.8) which was used to inform the base-case economic analysis. As the decision problem in this submission has been split into two distinct populations based on TNFi experience, results for the TNFi subgroups were used in the NMA and subsequently incorporated into the economic model. Overall, the results for clinical remission and response at induction and maintenance assessed using the 4-component Mayo score were found to be comparable to the results assessed using the 3-component Mayo score. A summary of the sensitivity analysis results for the TNF exposure subgroups are presented in

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 22, full results are presented in full in Appendix E.1.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 22: Summary sensitivity analysis results: 4-component Mayo score at Week 10 and Week 52, by TNFi experience

	TNFi-naïve				TNFi-experienced			
	Ozanimod		Placebo		Ozanimod		Placebo	
	4-component	3-component	4-component	3-component	4-component	3-component	4-component	3-component
Induction period (N)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clinical remission[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clinical response[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Maintenance period (N)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clinical remission[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clinical response[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviaitons: TNFi: tumour necrosis factor alpha inhibitor.

[a]Clinical remission is defined using the 4-component Mayo definition: 4-component Mayo score of ≤ 2 points with no individual subscore of > 1 point. [b]Clinical response is defined using the 4-component Mayo definition: a reduction from Baseline in the 4-component Mayo score of ≥ 3 points and ≥ 30%, and a reduction from Baseline in the Rectal Bleeding subscore of ≥ 1 point or an absolute Rectal Bleeding subscore of ≤ 1 point

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.6.2 Other pre-planned subgroup analyses – ITT population

The pre-defined subgroups for the induction and maintenance period can be found in Table 9. Pre-specified subgroup analyses were performed for the endpoints of clinical remission, clinical response, endoscopic improvement and mucosal healing for the induction and maintenance period in the overall population.

The forest plots for the proportion of patients in clinical remission at Week 10 and Week 52 for the induction and maintenance period are presented in Appendix E and analyses for the secondary endpoints can be found in the CSR.[5] Across the majority of subgroups, a nominally significantly higher proportion of patients treated with ozanimod achieved clinical remission in both the induction and maintenance period compared to placebo, and a consistently favourable response for ozanimod was observed versus placebo across all subgroups.

B.2.7 *Meta-analysis*

Given the lack of head-to-head RCT data for ozanimod versus the relevant comparators in UK clinical practice, a network meta-analysis was performed, and is presented in Section B.2.8.

B.2.8 *Indirect and mixed treatment comparisons*

Comparative effectiveness

- Following the identification of relevant studies from the clinical SLR, an NMA was performed to assess the efficacy and safety of ozanimod versus the treatment options specified in the NICE final scope; TNFis, ustekinumab, vedolizumab and tofacitinib
- Of the 28 RCTs identified in the SLR, 22 were included in the base case NMA to assess the comparable efficacy of ozanimod at both induction and maintenance of clinical response and clinical remission
- The relative efficacy was assessed in the two subgroups of relevance to the decision problem: TNFi-naïve and TNFi-experienced patients
- Results of the NMA demonstrated ozanimod to be associated with comparable efficacy to all relevant comparators (TNFis, ustekinumab and vedolizumab) in terms of clinical remission and clinical response in the induction and maintenance period for both subgroups
    - In the TNFi-naïve subgroup, the NMA found no statistically significant differences between ozanimod and the comparators (infliximab, adalimumab, golimumab and vedolizumab) for induction of clinical response and remission.
    - For maintenance, ozanimod was found to have comparable results for clinical remission and response to all relevant comparators in the TNFi-naïve subgroup, with the exception of vedolizumab
    - In the TNFi-experienced subgroup, no statistically significant differences were found between ozanimod and the relevant comparators (vedolizumab and ustekinumab) for clinical response and remission in induction and maintenance
- Sensitivity analyses were conducted to test the key areas of uncertainty within the NMA and showed that results obtained were generally robust
- Full details of the NMA methodology are presented in Appendix B.4

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.8.1 Identification and selection of relevant studies for the clinical SLR

An SLR (Section B.2.1) was conducted in October 2020 and updated in October 2021 to identify relevant clinical evidence of the efficacy and safety of ozanimod and other therapies for the treatment of moderately to severely active UC. In total 5,824 publications were screened at the title and abstract stage, of which 816 publications were reviewed at the full-text stage. After exclusion of publications not meeting the eligibility criteria, 157 publications (reporting on 28 unique RCTs) were included in the SLR.[128] A full list of the 28 included RCTs is presented in Appendix B.3.1. A risk of bias assessment was performed on all eligible RCTs using the University of York Centre for Reviews and Dissemination criteria for assessment of risk of bias in RCT.[128] These criteria include questions on randomisation scheme, allocation concealment, balance of prognostic factors, blinding of patients, care providers, and outcome assessors, imbalances in dropouts between groups, selective outcome reporting, and intention to treat analysis/handling of missing data. Results of the quality assessment of the included trials can be found in Appendix B.6.

B.2.8.2 Eligibility for the NMA

Studies considered for inclusion in the NMA were informed by the clinical SLR. As the clinical SLR captured data from for all potentially relevant studies from a global perspective, including multiple outcomes, only those reporting on relevant clinical outcomes, namely, clinical remission and clinical response, as defined by the Mayo clinic score, were included in the NMAs. Studies investigating treatments that had not been FDA or EMA approved for moderately to severely active UC at the time of the report were also not included in the NMA. This included studies in etrasimod (OASIS study), intravenous golimumab (PURSUIT-IV) and etrolizumab (HICKORY and EUCALYPTUS studies), in line with previous NMAs evaluating studies in this indication.[1, 88] Further to this, only studies investigating ozanimod and the comparators specified in the NICE final scope (TNFis [adalimumab, infliximab and golimumab], ustekinumab, vedolizumab and tofacitinib) were included. A full list of the eligibility criteria for inclusion in the NMA is provided in Table 23 and is in alignment with previous NMAs evaluating studies in this indication.[1, 88]

Of the 28 unique RCTs included in the SLR, 47 publications reporting on 25 unique RCTs were considered for inclusion in the NMA; a complete list of included/excluded studies is provided in Table 24. Twenty two unique RCTs were ultimately included in the NMA following the feasibility assessment in which a further 3 studies were excluded due to comparator choice (SERENE UC)[129] and endpoint definitions (Probert 2003 and Sands 2001)[130, 131] (see Section B.2.8.3).

Table 23: Eligibility criteria for inclusion of data in NMA

Criteria	Inclusion Criteria
Population	Patients with moderately to severely active UC
Comparators	ADA, IFX, GOL, TOF, VDZ, UST, and OZA according to licensed EMA and FDA doses in addition to infliximab 10 mg/kg and ustekinumab 130 mg[a]
Outcomes	Clinical response and clinical remission defined by the Mayo scale as well as occurrence of adverse events, serious adverse events, serious infections, and discontinuations due to adverse events

©Celgene Ltd, a BMS Company (2022). All rights reserved

Subgroups	TNFi-naïve, defined as patients who have not previously received a prior TNFi therapy, and TNFi-experienced, defined as patients who had received prior TNFi therapy

[a]Alternative doses of infliximab 10 mg/kg and ustekinumab 130 mg were included as they are used in clinical practice in some markets.
Abbreviations: ADA: adalimumab; EMA: European Medicines Agency; FDA: Food and Drug Administration; GOL: golimumab; IFX: infliximab; NMA: network meta-analysis; OZA: ozanimod; TNFi: tumour necrosis factor alpha inhibitor; TOF: tofacitinib; UST: ustekinumab; VDZ: vedolizumab.

Table 24: Trials included in NMAs

Trial name, NCT	Treatment
TRUENORTH[132], NCT02435992	PBO
	OZA 1 mg
TOUCHSTONE[133], NCT01647516	PBO
	OZA 1 mg
ULTRA 1[134], NCT00385736	PBO
	ADA 80/40 mg
	ADA 160/80/40 mg
ULTRA 2[33], NCT00408629	PBO
	ADA 160/80/40 mg
Suzuki 2014[135], NCT00853099	PBO
	ADA 80/40 mg
	ADA 160/80/40 mg
SERENE-UC[a],[129] NCT02065622	ADA 160/40 mg
	ADA 160/80/40 mg
	ADA 40 mg EW
	ADA 40 mg EOW
PURSUIT-SC[136], NCT00487539	PBO
	GOL 100/50 mg
	GOL 200/100 mg
	GOL 400/200 mg
PURSUIT M[26], NCT00488631	PBO
	GOL 50 mg
	GOL 100 mg
PURSUIT J[137], NCT01863771	PBO
	GOL 100 mg
	GOL 200/100 mg
ACT 1[123], NCT00036439	PBO
	IFX 5 mg
	IFX 10 mg
ACT 2[123],	PBO

©Celgene Ltd, a BMS Company (2022). All rights reserved

NCT00096655	IFX 5 mg
	IFX 10 mg
UC-SUCCESS[138], NCT00537316	AZA 2.5 mg/kg
	IFX 5 mg/kg
	AZA 2.5 mg/kg + IFX 5 mg/kg
Jiang 2015[139]	PBO
	IFX 3.5 mg/kg
	IFX 5 mg/kg
Kobayashi 2016[140]	PBO
	IFX 5 mg/kg
Probert 2003[a130]	PBO
	IFX 5 mg/kg
Sands 2001[a131]	PBO
	IFX 5 mg/kg
	IFX 10 mg/kg
GEMINI 1[24], NCT00783718	IFX 20 mg/kg
	Cohort 1 VDZ 300 mg
	Cohort 2 VDZ 300 mg
	Total VDZ 300 mg
	PBO
	VDZ 300 mg Q8W
	VDZ 300 mg Q4W
VARSITY[141], NCT02497469	VDZ 300 mg
	ADA 160/80/40 mg
VISIBLE 1[142], NCT02611830	VDZ 300 mg
	PBO
	VDZ 108 mg Q2W
	VDZ 300 mg Q8W
Motoya 2019[143], NCT02039505	PBO
	VDZ 300 mg
UNIFI[144], NCT02407236	PBO
	UST 130 mg
	UST 6 mg/kg
	PBO
	UST 90 mg Q12W
	UST 90 mg Q8W
Study A3921063[145], NCT00787202	PBO
	TOF 0.5 mg
	TOF 3 mg

©Celgene Ltd, a BMS Company (2022). All rights reserved

	TOF 10 mg
	TOF 15 mg
OCTAVE 1[121], NCT01465763	PBO
	TOF 10 mg
OCTAVE 2[121], NCT01458951	PBO
	TOF 10 mg
OCTAVE SUSTAIN[121], NCT01458574	PBO
	TOF 5 mg
	TOF 10 mg

*With the exception of OCTAVE SUSTAIN and UNIFI, all trials only re-randomised responders receiving active treatment in the induction phase. OCTAVE SUSTAIN also re-randomised placebo responders in the induction phase and UNFI re-randomised delayed responders, defined as patients who did not respond to placebo during the 8-week induction phase but who then received ustekinumab at week 8 and were responders at week 16. Across all trials, patients who were non-responders during the induction phase (excluding delayed responders from UNIFI) received open label active treatment during the maintenance phase.
[a]Excluded from the NMA following feasibility assessment (Section B.2.8.3).
Abbreviations: ADA: adalimumab; AZA: azathioprine; GOL: golimumab; IFX: infliximab; NA: not applicable; NR: not reported; OZA: ozanimod; PBO: placebo; TOF: tofacitinib; UST: ustekinumab VDZ: vedolizumab.

B.2.8.3 Feasibility assessment

In line with all major previous HTA-focussed NMAs in UC, separate analyses were conducted in the two sub-populations of relevance to this submission: TNFi-naïve and TNFi-experienced patients. Subgroup analyses in the TRUENORTH trial indicated that there may be an interaction between prior TNFi use and treatment effect (see section B.1.1.1). Several studies have shown lower efficacy with second-line biologics than with first biologics in UC (i.e., lower response rates, more patients requiring dose escalation).[146-149] As such, the available evidence was reviewed for the feasibility of conducting NMA analyses in subgroups by TNFi experience, in order to generate estimates of relative efficacy for ozanimod versus relevant comparators that could be subsequently used to inform the economic analysis. A rigorous qualitive and quantitative assessment of between-trial heterogeneity was conducted based on trial design, patient eligibility criteria, baseline patient characteristics, population definitions, as well as trial specific outcome availability and definitions.

Trial design

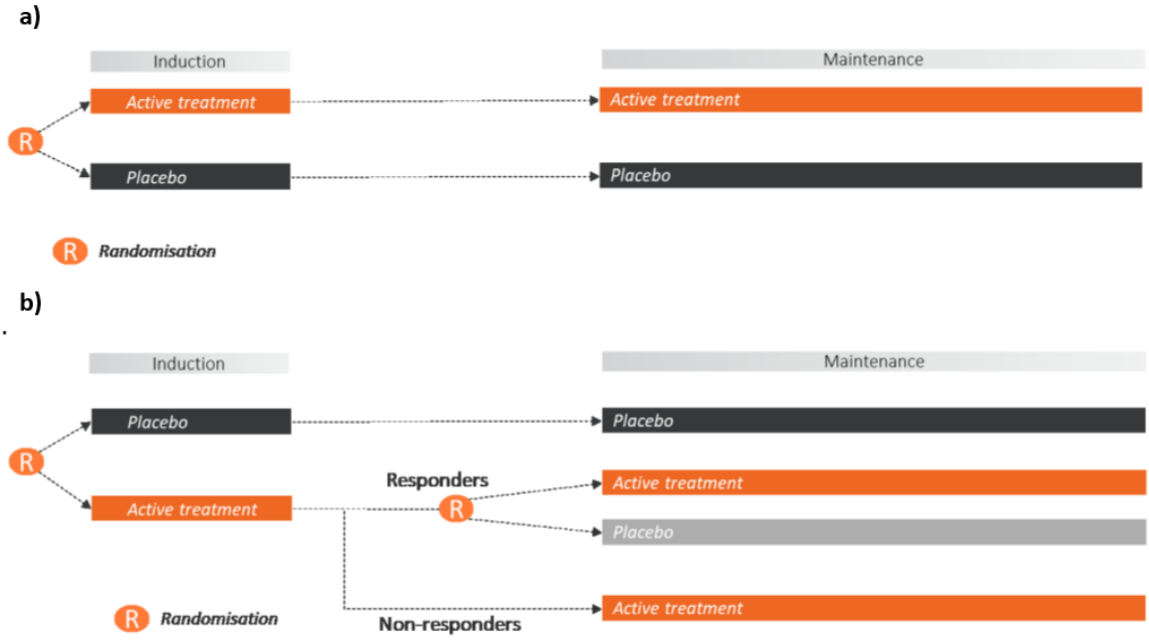
A summary of the key trial characteristics of all included studies in the NMA are presented in Appendix D.4.

Maintenance trial design

Trials that included a maintenance period were a combination of “treat-through” and “re-randomised” trial designs, the latter of which involved an additional randomisation period at the end of induction period, on top of the initial randomisation that usually occurs at baseline in treat-through trial designs (Figure 30). Earlier trials evaluating therapies like infliximab (ACT I) and adalimumab (ULTRA II and VARSITY [also included vedolizumab]) are based on treat-through trial designs whereas trials evaluating newer treatments including vedolizumab (GEMINI II), tofacitinib (OCTAVE), golimumab (PUSRUIT) and ustekinumab (UNIFI) are based on re-randomised trial designs.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 30: Treat-through versus re-randomised trial design



Trial schematics for a) conventional treat-through design that involves a single baseline randomisation step and b) response-based re-randomised design that involves an additional re-randomisation step for patients who are responders at induction.
Note: Not all non-responders in re-randomised trials will necessarily have received active treatment, e.g. if the trial design did not include an extended induction or mandatory extension phase
Source: Adapted from NICE TA633 committee papers.[1]

With the exception of OCTAVE SUSTAIN and UNIFI, trials with a re-randomised trial design only re-randomised responders receiving active treatment in the induction period. OCTAVE SUSTAIN also re-randomised placebo responders in the induction phase and UNIFI re-randomised delayed responders, defined as patients who did not respond to placebo during the 8-week induction phase but who then received ustekinumab at Week 8 and were responders at Week 16.

Comparisons across re-randomised and treat-through trials are difficult, since they include heterogeneous groups with respect to study drug exposure. In re-randomised trials, patients who have received active treatment during induction but are re-randomised to placebo in maintenance may “carry over” the effect of the induction therapy into the maintenance period, resulting in a heightened level of response at maintenance.

Heterogeneity generated by differences in trial design in UC may influence the maintenance NMA findings if maintenance results for induction responders were directly compared to those of non-induction responders. To account for these differences in trial design, statistical adjustments were made to treat-through trials to align with the data presented in re-randomised trials during the maintenance period. This approach is in in alignment with that used in TA547 and the ERG ‘maintenance only’ scenario in TA633.[1, 88] Full details of adjustments made are reported in Section B.2.8.4. In cases where sustained clinical response and remission among induction responders were available in treat-through trials, this data was directly inputted into the NMAs as this group of patients were conditioned on an induction responder status therefore aligning with patients from re-randomised trials during the maintenance period. A sensitivity analysis was performed which

©Celgene Ltd, a BMS Company (2022). All rights reserved

excluded treat-through trial designs to assess potential bias introduced from the adjustments to treat-through trials (Section B.2.8.5.3).

Timepoint of assessment

The timepoint of assessment in the induction and maintenance periods was a further source of heterogeneity between trials included in the NMA. The primary induction period assessment varied from 2–14 weeks across trials, while the maintenance period timepoint of assessment varied from 32–60 weeks (Table 25). In order to minimise heterogeneity associated with this issue, the timepoint of assessment deemed eligible for the NMA was restricted to 6–14 weeks for induction, and 52–60 weeks for maintenance. As a result, several trials were not included in the maintenance period NMA despite including a maintenance period. The approach to exclude trials that have significantly different timepoints of assessments is in alignment with previous UC NMAs.[150, 151]

Despite restricting the induction and maintenance period timepoint eligible for the NMA, some heterogeneity may remain within the allowable timeframe.

Table 25: Timepoint of assessment in induction and maintenance period in trials included in NMAs

Drug	Trials	Induction duration and timepoint of assessment	Maintenance timepoint of assessment
OZA	TRUE NORTH[132]	10 weeks	52 weeks
	TOUCHSTONE[133]	8 weeks[a]	32 weeks
ADA	ULTRA 1[134] ULTRA 2[33]	8 weeks	52 weeks
	SERENE-UC[129]	8 weeks	52 weeks
	Suzuki 2014[135]	8 weeks	52 weeks
GOL	PURSUIT-SC[136] PURSUIT M[26]; PURSUIT J[137]	6 weeks	60 weeks
IFX	ACT 1[123]	8 weeks	54 weeks
	ACT 2[123]	8 weeks	30 weeks
	UC-SUCCESS[138]	8 weeks[b]	NA
	Jiang 2015[139]	8 weeks	30 weeks
	Kobayashi 2016[140]	8 weeks	38 weeks
	Probert 2003[130]	6 weeks[c]	NA
	Sands 2001[131]	2 weeks[d]	NA
VDZ	GEMINI 1[24]	6 weeks	52 weeks
	VARSITY[141]	14 weeks	52 weeks
	VISIBLE 1[142]	6 weeks	52 weeks
	Motoya 2019[143]	10 weeks	60 weeks
UST	UNIFI[144]	8 weeks	52 weeks
TOF	Study A3921063[145]	8 weeks	NA

©Celgene Ltd, a BMS Company (2022). All rights reserved

	OCTAVE 1[121] OCTAVE 2[121] OCTAVE SUSTAIN[121]	8 weeks	60 weeks

[a]9 weeks including the 1-week dose escalation phase. [b]Extended induction data available at week 16. [c]Extended induction data available at week 8. [d]Extended induction data available at week 12.
Abbreviations: ADA: adalimumab; GOL: golimumab; IX: infliximab; NA: not applicable; NMA: network meta-analysis; OZA: ozanimod; TOF: tofacitinib; UST: ustekinumab; VDZ: vedolizumab.

Eligibility criteria

Trial eligibility criteria is an important source of potential heterogeneity as it defines the patient population of interest within each trial, which may vary between trials. A list of the key inclusion and exclusion criteria for the trials included in the NMAs are presented in Appendix D.4.1

In general, inclusion and exclusion criteria were similar across trials, often requiring a combination of:

- Adults aged ≥18 years
- UC diagnosis of ≥3 months
- Active UC based on Mayo score of 6 to 12, with endoscopic subscore of ≥2
- Inadequate response to, or had failed to tolerate, at least one of the conventional therapies: oral aminosalicylates, oral corticosteroids, azathioprine, and/or mercaptopurine

Several trials restricted recruitment to an entirely Asian population. A sensitivity analyses was performed in which Asian trials were excluded in line with previous NMAs in UC (Section B.3.8).

©Celgene Ltd, a BMS Company (2022). All rights reserved

Subgroup definitions

'TNFi' versus 'biologic'

NMA analyses were explored for subgroups of patients defined by TNFi experience, in line with previous NMAs in UC (TA547).[88] 12 trials included proportions of patients who had previously received treatment with TNFis, whilst the remaining 10 trials (typically investigating infliximab and golimumab) recruited entirely TNFi-naïve populations. Of those trials that included TNFi-naïve and experienced patients, there was heterogeneity in whether trials defined subgroups according to TNFi experience or experience of any biologic therapy (based on trial stratification factors):

- Subgroups of the UNIFI trials were defined as 'biologic failure' and 'non-biologic failure'. However, a total of 98.8% of patients in the 'biologic failure' subgroup had had treatment failure with at least one TNF antagonist.[144] As such, it was considered appropriate to include the 'biologic failure' and 'non-biologic failure' subgroups in the TNFi-experienced and TNFi-naïve analyses, respectively.
- In all other trials included in the NMA, including TRUENORTH and TOUCHSTONE, subgroups were defined based on TNFi experience.

All 12 trials except for TOUCHSTONE reported some form of TNFi-experienced and TNFi-naïve subgroup data (Table 26). However, as 82% of patients were TNFi-naïve in TOUCHSTONE, a sensitivity analysis was conducted to explore the effect of including the trial in the TNFi-naïve induction analyses (Section Section B.3.8).

TNFi 'Experience' versus 'failure'

In addition, the definition of "experience" varied across the trials. For example, certain trials reported data for patients who had failed previous TNFi therapies (i.e., were intolerant or inadequate responders), whilst others reported data for patients who were exposed to previous TNFi therapies (i.e., failed or exposed without failure). Likewise, the TNFi-naïve populations could have been defined as "no prior failure" (which may include those exposed without failure) or "no prior exposure" to a previous TNFi. Thus, there were two possible definitions for the TNFi-experienced and TNFi-naïve subgroups, one based on exposure and one on prior failure (intolerant or inadequate response). A summary of which trials reported experience based on exposure or failure is provided in Table 26 below.

Despite this heterogeneity around the definition of TNFi "experience", there were often only minor differences between the "exposed" and "failure" populations between subgroups, because treatment failure was the most common reason to stop treatment. This is especially apparent for ULTRA-2, where all TNFi-exposed patients were TNFi failures. Given published data were not reported for consistent subgroups across trials, no sensitivity analysis could be performed.

Prior biologics in TNFi-experienced subgroups

TNFi-experienced populations differed across trials according to the biologic agents available for treatment of UC at the time of study. Earlier trials often only reflect prior TNFis in the TNFi-experienced population while TNFi-experienced populations in more recent trials reflect previous exposure to newer biologic UC agents, namely vedolizumab and ustekinumab (Table 26). For

©Celgene Ltd, a BMS Company (2022). All rights reserved

example, in TRUENORTH, a large proportion of patients in the TNFi-experienced population had also received another non-TNFi biologic therapy; [redacted] of patients receiving ozanimod had prior exposure to vedolizumab.

Table 26: Summary of TNFi subgroup data in studies included in NMAs

Trials	Prior Biologics	TNFi-naïve patients		TNFi-experienced patients	
		TNFi non-exposed	TNFi non-failure	TNFi-exposed	TNFi-failure
TRUENORTH	UST, VEDO, TNFi	Available	NR	Available	NR
TOUCHSTONE	UST, VEDO, TNFi	NR	NR	NR	NR
ULTRA 2[a]	TNFi (excluding ADA)	Available		Available	
GEMINI 1	TNFi	Available	NR	NR	Available
Motoya 2019	TNFi	Available	NR	Available	NR
VARSITY	TNFi (excluding ADA)	Available	NR	Available	NR
VISIBLE 1	TNFi	Available	NR	Available	NR
UNIFI[b]	VEDO, TNFi	Available	Available	NR	Available
Study A3921063	TNFi	Available	NR	Available	NR
OCTAVE 1	TNFi	Available	Available	Available	Available
OCTAVE 2	TNFi	Available	Available	Available	Available
OCTAVE SUSTAIN	TNFi	NR	Available	NR	Available

[a]All TNFi-exposed patients were TNFi-failures. [b]Subgroups were defined according to treatment failure with any biologic agent.
Abbreviations: ADA: adalimumab; NR: not reported; TNFi: tumour necrosis factor alpha inhibitor; UST: ustekinumab; VEDO: vedolizumab.

Baseline characteristics

Baseline characteristics of patients in entering the induction and maintenance period of the studies included in the NMA are provided in Appendix B.4.1. Baseline characteristics were sparsely reported for the maintenance periods of trials however, where reported, characteristics were broadly similar across trials in terms of mean age, the proportion of males and the mean Mayo score. Of note, CRP levels were found to vary across trials ranging from 3.2–142.98 mg/L at induction and 0.7–35.8 mg/L at maintenance. Similarly, years since UC diagnosis varied across trials, ranging from 4.0–14.6 years at induction and 5.4–8.7 years at maintenance. Disease extent varied across trials, with the proportion of patients with left sided colitis ranging from 15.0–63.0% at induction and 30.6–69.2% at maintenance, whilst the proportion of patients with extensive disease varied by 6.6–80.8% and 11.2–68.3% at induction and maintenance, respectively. Use of concomitant steroids ranged from 25.0–100.0% at induction and 28.2–58% at maintenance. Likewise, the proportion of patients who were naïve to TNFis or other biologics was also variable, ranging from 19.0–100% at induction and 30.4–51.3% maintenance. This was partially due to the eligibility criteria of early UC trials restricting the patient population to those who are TNFi-naïve. Previous NMAs in UC have highlighted similar differences in patient baseline characteristics as a potential source of heterogeneity between trials.[150, 151]

©Celgene Ltd, a BMS Company (2022). All rights reserved

It is not clear whether CRP levels, years since diagnosis, disease extent, or use of concomitant steroids are treatment effect modifiers and thus whether these differences would impact the results. Differences in the proportion of patients with TNFi experience are not a source of bias, as separate analyses were explored by prior TNFi experience. Several baseline characteristics were not reported across trials, including moderate UC status at baseline, and thus the extent to which differences in these characteristics may impact the NMA results is unclear.

Outcomes

The availability of data for key outcomes of interest for the NMA (clinical response and clinical remission) in the TNFi-naïve and experienced population are summarised in Table 27 and Table 28 for the induction and maintenance period, respectively.

Several studies investigating TNFis, including infliximab and golimumab, included only TNFi-naïve populations. However, as outlined in Section B.1.1, TNFis are not relevant comparators in the TNFi-experienced population in this appraisal. Of the 17 induction trials, all trials reported some relevant data for the NMAs by TNFi subgroup status. Likewise, of the 11 maintenance trials, all trials reported some relevant data by TNFi subgroups. Of note, clinical response and remission data during the induction period were not available for TNFi subgroups for OCTAVE 1 and OCTAVE 2 trials. Instead, available subgroup data for pooled OCTAVE 1 + OCTAVE 2 was leveraged. Certain studies only reported on clinical remission data instead of both response and remission in the TNFi subgroups (Study A3921063, VARSITY, VISIBLE 1). By leveraging an ordinal response-remission NMA, these studies were retained in analyses.

It should be noted that, where possible, comparisons were explored against all comparators in both the TNFi-naïve and TNFi-subgroups. However, as noted in Section B.1.1 only TNFis and vedolizumab are relevant comparators in the TNFi-naïve population and only ustekinumab and vedolizumab are relevant comparators in the TNFi-experienced population and therefore relevant for the decision problem.

Table 27: Trials included in induction period NMAs of clinical response and clinical remission

Trial name	Clinical response		Clinical remission	
	TNFi-naïve	TNFi-experienced	TNFi-naïve	TNFi-experienced
TRUE NORTH[132]	✓	✓	✓	✓
TOUCHSTONE[133]	×	×	×	×
ULTRA 1[134]	✓	×	✓	×
ULTRA 2[33]	✓	✓	✓	✓
Suzuki 2014[135]	✓	×	✓	×
PURSUIT-SC [136]	✓	×	✓	×
ACT 1[123]	✓	×	✓	×
ACT 2[123]	✓	×	✓	×
Jiang 2015[139]	✓	×	✓	×
Kobayashi 2016[140]	✓	×	✓	×
GEMINI 1[24]	✓	✓	✓	✓
VARSITY[141]	✓	✓	✓	✓

©Celgene Ltd, a BMS Company (2022). All rights reserved

Motoya 2019[143]	✓	✓	✓	✓
UNIFI[144]	✓	✓	✓	✓
Study A3921063[145]	✓	✓	×	×
OCTAVE 1[121]	×	×	×	×
OCTAVE 2[121]	×	×	×	×
OCTAVE 1 + 2[152]	✓	✓	✓	✓

Table 28: Trials included in maintenance period NMAs of clinical response and clinical remission

Trial name	Clinical Response		Clinical Remission	
	TNFi-naïve	TNFi-experienced	TNFi-naïve	TNFi-experienced
TRUE NORTH	✓	✓	✓	✓
ULTRA 2	✓	✓	✓	✓
Suzuki 2014	✓	×	✓	×
PURSUIT-M	✓	×	✓	×
PURSUIT-J	✓	×	✓	×
ACT 1	✓	×	✓	×
GEMINI 1	✓	✓	✓	✓
VISIBLE 1	×	×	✓	✓
VARSITY	×	×	✓	✓
Motoya 2019	✓	✓	✓	✓
UNIFI	✓	✓	✓	✓
OCTAVE SUSTAIN	✓	✓	✓	✓

Outcome definitions

In TRUENORTH, the two main efficacy outcomes (clinical response and clinical remission) were assessed using the 3-component Mayo score. To align with other trials that reported outcomes based on the 4-component Mayo score (which includes an additional PGA subscore) and hence reduce heterogeneity in the NMA, data from the analyses using the 4-compoment Mayo score from the TRUENORTH trial were used in the base case NMAs (See Section B.2.6.1.3). A separate sensitivity analysis was conducted to explore the influence of the 3-component TRUENORTH data on NMA findings (Section Section B.3.8).

The majority of trials defined endoscopy subscores on local readings, however TRUENORTH, TOUCHSTONE, VISIBLE 1, VARSITY and the three OCTAVE trials read the scores centrally. Differences in how endoscopic scores are read between trials (local vs central) is an additional source of heterogeneity when comparing results across trials. Centralised readings promote objectivity by being independent, having lower variability and reduce the placebo response rate and as such are recognised as being an important aspect in improving standards of clinical trials in IBD.[153] This has been highlighted for example when results of local endoscopic scoring in the OCTAVE trial were reported and led to higher clinical remission scores when compared to the centrally read scores.[152] The definitions of clinical remission and clinical response were generally

©Celgene Ltd, a BMS Company (2022). All rights reserved

consistent across trials and aligned with the 4-component definitions from TRUENORTH (Table 29 and Table 30). Any studies which based the definition using a score other than the Mayo clinic score were excluded from the analyses of the relevant outcome, in order to avoid introducing heterogeneity attributable to using an entirely different scale to assess an outcome. This approach is consistent with previous NMAs. As a result, Probert *et al.* 2003 and Sands *et al*. 2001 were excluded from the NMAs for using the UC Symptom and the modified Truelove and Witts scoring systems, respectively.[130,67]

Table 29: Summary of trial definitions of clinical remission

Clinical Remission Definition	Trials Using Definition in NMAs
Complete Mayo score of ≤2 points with no individual subscore >1 point	TRUENORTH, TOUCHSTONE, ULTRA 1, ULTRA 2, SERENE-UC, Suzuki 2014, PURSUIT-SC, PURSUIT-M, PURSUIT-J, ACT 1, ACT 2, Jiang 2015, Kobayashi 2016, GEMINI 1, VARSITY, VISIBLE 1, Motoya 2019, UNIFI, Study A3921063
Complete Mayo score of ≤2 points with no individual subscore >1 point and a rectal bleeding subscore of 0	OCTAVE 1, OCTAVE 2, OCTAVE SUSTAIN
Rectal bleeding subscore = 0, stool frequency subscore ≤1 (and a decrease of ≥1 point from the baseline stool frequency subscore), and endoscopy subscore ≤1	TRUENORTH (sensitivity analysis)
Ulcerative colitis symptom score ≤2	Probert 2003[a]

[a]Trial was excluded from NMAs.

Table 30: Summary of trial definitions of clinical response

Definition	Trials Using Definition
Decrease of ≥3 points and ≥30% in the complete Mayo score	UC-SUCCESS
Decrease of ≥3 points and ≥30% in the complete Mayo score and either an absolute rectal bleeding subscore of ≤1 point or a ≥1 point decrease in the rectal bleeding subscore	TRUENORTH, TOUCHSTONE, ULTRA 1, ULTRA 2, SERENE-UC, Suzuki 2014, PURSUIT-SC, PURSUIT-M, PURSUIT-J, ACT 1, ACT 2, Jiang 2015, Kobayashi 2016, GEMINI 1, VISIBLE 1, Motoya 2019, UNIFI, Study A3921063, OCTAVE 1, OCTAVE 2, OCTAVE SUSTAIN
Decrease of ≥2 point and ≥25% in the partial Mayo score (excludes endoscopy subscore) and either an absolute rectal bleeding subscore of ≤1 or a ≥1 point decrease in the rectal bleeding subscore	VARSITY
Decrease of ≥2 points and ≥35% in the 9-point Mayo and either an absolute rectal bleeding subscore of ≤1 point or a ≥1 point decrease in the rectal bleeding subscore	TRUENORTH
Modified Truelove and Witts score of <10 and a 5-point reduction compared with baseline	Sands 2001[a]

[a]Trial was excluded from NMAs.

Availability of common comparators

©Celgene Ltd, a BMS Company (2022). All rights reserved

All trials included a common comparator with another trial in the network, with the exception of SERENE-UC,[129] which compared an approved dose of adalimumab against an unapproved dose. SERENE-UC was therefore excluded from the NMA.

B.2.8.4 NMA methodology

The NMA was conducted using a Bayesian framework. Placebo was the chosen reference treatment for all analyses, given its presence as the anchor treatment across almost all included studies. Separate analyses were performed for the TNFi-naïve and TNFi-experienced populations, due to expected differences in clinical efficacy associated with prior treatment, in line with previous NMA publications in UC. Likewise, separate analyses were performed for studies reporting data at the induction (6–14 weeks) and maintenance (52–60 weeks) periods. Individual doses of the same active agent that had the same method of administration (e.g. infliximab 5 mg/kg and infliximab 10 mg/kg but not vedolizumab 300 mg and vedolizumab 108 mg SC) were pooled in the base case to increase the available data for each comparator and average the treatment effects observed between different regimens of the same treatment. A summary of the doses pooled in the base case NMA is provided in Table 31 below. A sensitivity analysis was performed using unpooled doses, the results of which are presented in Section B.2.8.5.3.

Table 31: Summary of dose pooling in the induction and maintenance period of the base case NMAs

Period	Treatment	Administration	Dose(s)
Induction	Ozanimod	Oral	• 1 mg OD
	Golimumab	SC	• 200/100 mg
	Adalimumab	SC	• 160/80/40 mg Q2W
	Infliximab	**IV (pooled)**	• **5 mg/kg** • **10 mg/kg**
	Tofacitinib	Oral	• 10 mg BID
	Vedolizumab	IV	• 300 mg
	Ustekinumab	**IV (pooled)**	• **130 mg** • **6 mg/kg**
Maintenance	Ozanimod	Oral	• 1 mg OD
	Infliximab pooled	**IV (pooled)**	• **5 mg/kg** • **10 mg/kg**
	Adalimumab	SC	• 40 mg Q2W
	Golimumab	**SC (pooled)**	• **100 mg Q4W** • **50 mg Q4W**
	Tofacitinib	**Oral (pooled)**	• **10 mg BID** • **5 mg BID**
	Vedolizumab	**IV (pooled)**	• **300 mg Q4W** • **300 mg Q8W**
		SC	• 108 mg Q2W
	Ustekinumab	**IV (pooled)**	• **90 mg Q12W** • **90 mg Q8W**

©Celgene Ltd, a BMS Company (2022). All rights reserved

Abbreviations: BID: twice daily; IV: intravenous; OD: once daily; Q4W: every 4 weeks; Q8W: every 8 weeks: Q12W: every 12 weeks; SC: subcutaneous.

Network diagrams were drawn to visualise the evidence base for each analysis (Section B.2.8.5). In these figures, lines that connect nodes signify the presence of one or more RCTs that directly compare treatments, with the thickness of each line reflecting the number of RCTs informing the comparison; thicker lines signify more RCTs comparing treatments. Treatments populated entirely with zero events were dropped from networks of evidence, although this was not required for any of the primary analyses.

An ordinal model with a probit link was used to assess clinical response and clinical remission as it allowed the correlation between the two outcomes to be taken into account, given they were both assessed using the Mayo score, with remission representing a stricter level of response (see Table 10). Use of a probit model is in line with the recommendation given in TSD2 for such ordered outcomes.[154] Precedence for ordinal probit modelling approach has also been established in previous UC NMAs, being leveraged by the recent ICER UC evidence report as well as the assessment group for TA329.[35, 155] In all cases, outcomes were transformed to odds ratios to facilitate clinical interpretation of findings, consistent with the standard outcome reporting method used by clinical trials in UC as well as previous NMAs conducted by evidence review groups.

All NMAs were performed using R (R Core Team, Vienna, Austria) and Just Another Gibbs Sampler (JAGS), based on the code outlined in the NICE Evidence Synthesis Decision Support Unit (DSU) Technical Support Document (TSD) Series. The JAGS code used in all NMA models explored can be found in Appendix B.4.2.

Outcome measures

Pairwise comparisons of interventions estimated from NMAs are presented through forest plots that report pairwise OR with 95% credible intervals (CrI). Statements regarding treatment differences are primarily informed by pairwise differences in effect estimates, with “statistically significant” conclusions derived from overlap of pairwise credible intervals with unity (i.e., no difference) (Section B.2.8.5).

Model effects

Fixed effect and random effects models were explored for each individual network. The preferred model was chosen based on a combination of statistical and clinical considerations. From a statistical standpoint, lower deviance information criteria (DIC) and residual deviance (ResDev) were favoured as outlined in NICE TSD 3.[156] However, due to the heterogeneity identified in Section B.2.8.3, random effects models were likely to have better clinical validity. As a result, random effects models were favoured by default, with a fixed effect model only being selected when estimates lacked face validity, to ensure that models were not generating conclusions contrary to the direct evidence observed in the clinical trials informing the network (such as in cases where the efficacy results of patients receiving the intervention were considered comparable to placebo in the random effects model).

This was also accompanied by an inspection of the networks of evidence available for each outcome; outcomes informed primarily by single-study connections can generate underpowered

©Celgene Ltd, a BMS Company (2022). All rights reserved

between-trial heterogeneity in the random effects models, potentially making fixed effect more suitable.[156.] The model selection rationale (fixed effects versus random effects) for each individual network is provided in Section B.2.8.5.

Model convergence

All analyses were performed using four unique sets of starting values and were based on burn-in and sampling durations of 20,000 iterations or more, with additional samples taken to achieve convergence when necessary. Convergence was monitored quantitatively using the latest implementation Gelman-Rubin diagnostic (Rhat) based on four chains.[157] This new implementation captured non-convergence from stationary but non-overlapping chains, over-lapping non-stationary chains, chains with heavy tails, and chains with different variance. Samples were considered to have converged if Rhat was equal to or less than 1.05. After convergence had been reached, concerns turned to whether there were sufficient independent samples for stable estimates. The newest version of effective sample size (ESS) and Monte Carlo standard error (MCSE) estimation were used to ensure sufficient post-convergence samples were taken to support inference.[157] If the rank-normalised effective sample size was greater than 400 (i.e., 100 per chain) then samples were taken to ensure that MCSE was small enough to allow for stable estimates to at least one decimal place.[157] All assessments of ESS and MCSE were made for each parameter that is reported.

Model priors

Vague prior distributions that assume no pre-existing information were assigned for the treatment effects, trial baselines, common regression terms (β), and between-study variance in all primary analyses (Table 32).

Table 32: Default model priors used across analyses

Parameter	Prior Distribution
Baselines, models (mu)	dnorm(0, 0.0001)
Basic parameters (d)	dnorm(0,0.0001)
Between-trial variation (sd)	dunif(0,2)
Meta-regression coefficient (B)	dnorm(0,0.0001)
Ordinal category cut-points (z)	dunif(0,3)

Model thinning

Across outcomes, models incorporated thinning such that 10,000 iterations of each parameter would be saved. For example, a model using 20,000 iterations given four independent chains would keep every eighth iteration. Thinned samples are still required to pass the same convergence diagnostics outlined above. This was done to accommodate the incorporation of NMA data into probabilistic sensitivity analyses in the cost-effectiveness analysis economic model, which requires a consistent amount of convergence diagnosis and output analysis (CODA) across outcomes.

Treat-through trial data re-calculations

As described in Section B.2.8.3, studies with a maintenance period were a mix of treat-through and re-randomised designs. To account for this heterogeneity in trial design, adjustments were considered for the maintenance treat-through data from the ACT 1, ULTRA 2, Suzuki 2014, and

©Celgene Ltd, a BMS Company (2022). All rights reserved

VARSITY trials to better align with what would be observed in a re-randomised trial of similar design. This approach was consistent with that taken in TA547 and the ERG 'maintenance only' scenario in TA633.[1, 88] As the majority of re-randomised trials only permitted enrolment of induction period responders in the maintenance period, for treat-through trials with available data for sustained clinical responders/remitters (i.e. response/remission amongst induction responders), the data were directly imputed into the treat-through to re-randomised analyses. In a case where these data were not available, the sustained response/remission rate was estimated by multiplying the total number of induction period responders in the trial by the proportion of sustained responders/remitters out of the total responders/remitters in the placebo arm of an alternative trial of the same active treatment included in the NMA. All the treat-through trials excluding VARSITY were adjusted. VARSTIY was not adjusted as sustained clinical response data were unavailable for sustained clinical responders or remitters among induction responders in the TNFi subgroup populations and imputation of these values was not feasible since the trial was head-to-head. As a result, VARSITY was not included in the base case analysis. Full details of the calculations performed for individual trials are presented in Appendix B.4.1.

Assessment of consistency

An assumption underlying NMAs is that the analysed network is consistent, meaning that there is no evidence of disagreement between the direct and indirect evidence being combined. An unrelated mean effects model (i.e., an inconsistency model) was used to test for inconsistency for key outcomes (clinical response and clinical remission). Similar posterior deviances between consistency and inconsistency models were observed across all outcomes (Appendix B.4.2). In addition, across all outcomes, there was significant overlap of the pairwise conclusions derived by the consistency and inconsistency models. Therefore, no evidence of significant inconsistency was observed.

Sensitivity analysis

In addition to the primary analytical framework described above, sensitivity analyses were conducted to explore the effect of various assumptions regarding analyses and the data contained within, and to evaluate the influence of potential sources of heterogeneity described in Section B.2.8.3.

B.2.8.5 NMA results

B.2.8.5.1 TNFi-naïve population

Induction period

The network for clinical response and clinical remission during the induction period for the TNFi naïve population is displayed in Figure 31. In total, 15 studies were included in the analysis. All interventions were assessed in one or more placebo-controlled studies, with some studies evaluating multiple doses of the same TNFi as well as a single head-to-head study (VARSITY).

Inspection of model fit statistics according to Table 33 suggested that the random effects model was associated with an improved fit relative to other models, given the residual deviance was lower than the fixed effect model and the posterior SD was reasonable. Although the deviance information criterion (DIC) was slightly (<5 points) larger in the random effects model, clinical heterogeneity, highlighted in Section B.2.8, favours the random effects model. Therefore, primary results for clinical

©Celgene Ltd, a BMS Company (2022). All rights reserved

response and remission during the induction period for the TNFi-naïve population were derived from the random effects model. An assessment of inconsistency determined there to be little evidence of inconsistency between direct and indirect estimates for either model (Appendix D.4.2).

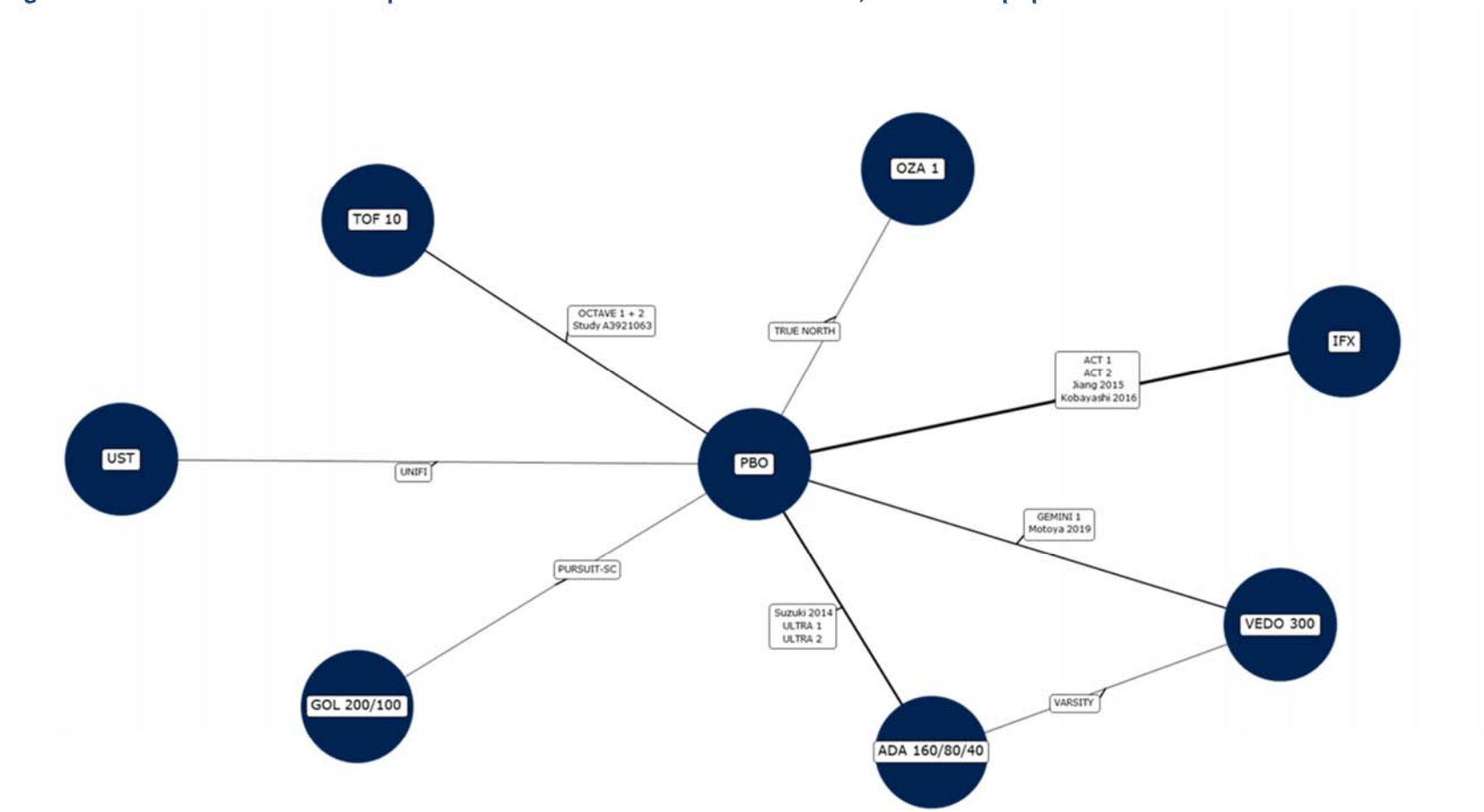
Table 33: Model fit statistics for clinical response and remission at induction (TNFi-naïve)

Diagnostic	Fixed effect model	Random effects model
DIC	122.98	124.39
ResDev (vs. 58 data points)	99.99	96.94
SD (95% CrI)	NA	0.123 (0.007 to 0.364)

Abbreviations: CrI: credible interval; DIC: deviance information criterion; NA: not applicable; ResDev: residual deviance; SD: between-trial heterogeneity.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 31: Network for clinical response and clinical remission at induction; TNFi-naïve population[a]



[a] NMA contains pooled infliximab and ustekinumab (Section B.2.8.4)
Footnote: Lines that connect nodes signify the presence of one or more RCTs that directly compare treatments, with the thickness of each line reflecting the number of RCTs informing the comparison; thicker lines signify more RCTs comparing treatments.
Abbreviations: ADA: adalimumab; GOL: golimumab; IFX: infliximab; OZA: ozanimod; PBO: placebo; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Clinical response and clinical remission

All active agents offered statistically significant improvement in clinical response over placebo (Figure 32). No statistically significant differences were found between ozanimod and other active therapies, however, point estimates were found to trend in favour of ozanimod versus adalimumab, golimumab, tofacitinib and ustekinumab.

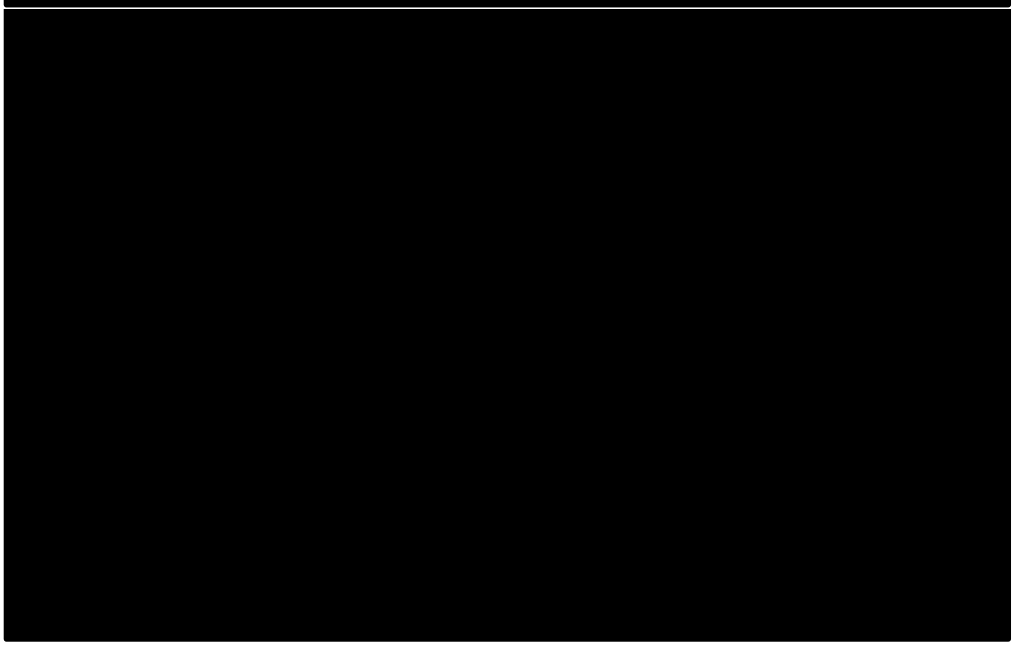
The network for clinical response and remission during the induction period for the TNFi-naïve population is displayed in Figure 31. All active agents offered statistically significant improvement in clinical remission over placebo (Figure 33). No statistically significant differences were found between ozanimod and other active therapies, however, point estimates trended in favour of ozanimod versus adalimumab, golimumab, tofacitinib, and ustekinumab.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 32: Forest plot for clinical response at induction; TNFi-naïve population (random effects – preferred model)

a)

b)



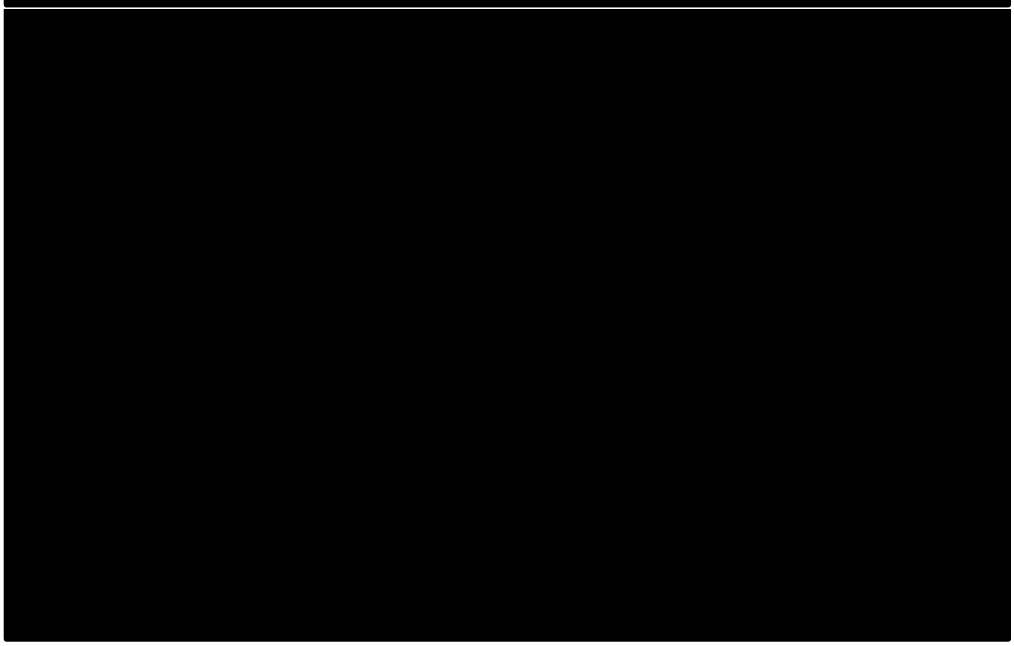
Forest plots displaying a) odds ratios of active treatments versus placebo and b) odds ratios of ozanimod versus active treatments. Treatments are sorted by descending median odds ratios. Bold values indicate statistical significance. **Abbreviations:** BID: twice a day; CrI: credible interval; IV: intravenous; OR: odds ratio; PBO: placebo; Q2W: every 2 weeks; QD: once a day; SC: subcutaneous.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 33: Forest plot for clinical remission at induction; TNFi-naïve population (random effects – preferred model)

a)

b)



Forest plots displaying a) odds ratios of active treatments versus placebo and b) odds ratios of ozanimod versus active treatments. Treatments are sorted by descending median odds ratios. Bold values indicate statistical significance.
Abbreviations: BID: twice a day; CrI: credible interval; IV: intravenous; OR: odds ratio; PBO: placebo; Q2W: every 2 weeks; QD: once a day; SC: subcutaneous.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Maintenance period

The network for clinical response and clinical remission at maintenance in the TNFi-naïve population is displayed in Figure 34. In total, 11 studies were included in the analysis. All interventions were assessed in one or more placebo-controlled studies, with some studies evaluating multiple doses of the same TNFi. Multiple doses of the same treatment were pooled, however vedolizumab 108 mg Q2W SC was considered separate to the vedolizumab IV owing to differing methods of administration (Section B.2.8.4).

Inspection of model fit statistics according to Table 34 suggested that the fixed effect model was associated with reasonable model fits in terms of DIC and residual deviance. The random effects model, however, did not converge. Therefore, primary results for clinical response and remission during the maintenance period for the TNFi-naïve population were derived from the fixed effect model. An assessment of inconsistency was not possible due to the inconsistency model failing to converge (Appendix D.4.2). However, given other analyses had no indications of inconsistency, it was considered unlikely that inconsistency would be present in the maintenance TNFi-naïve network.

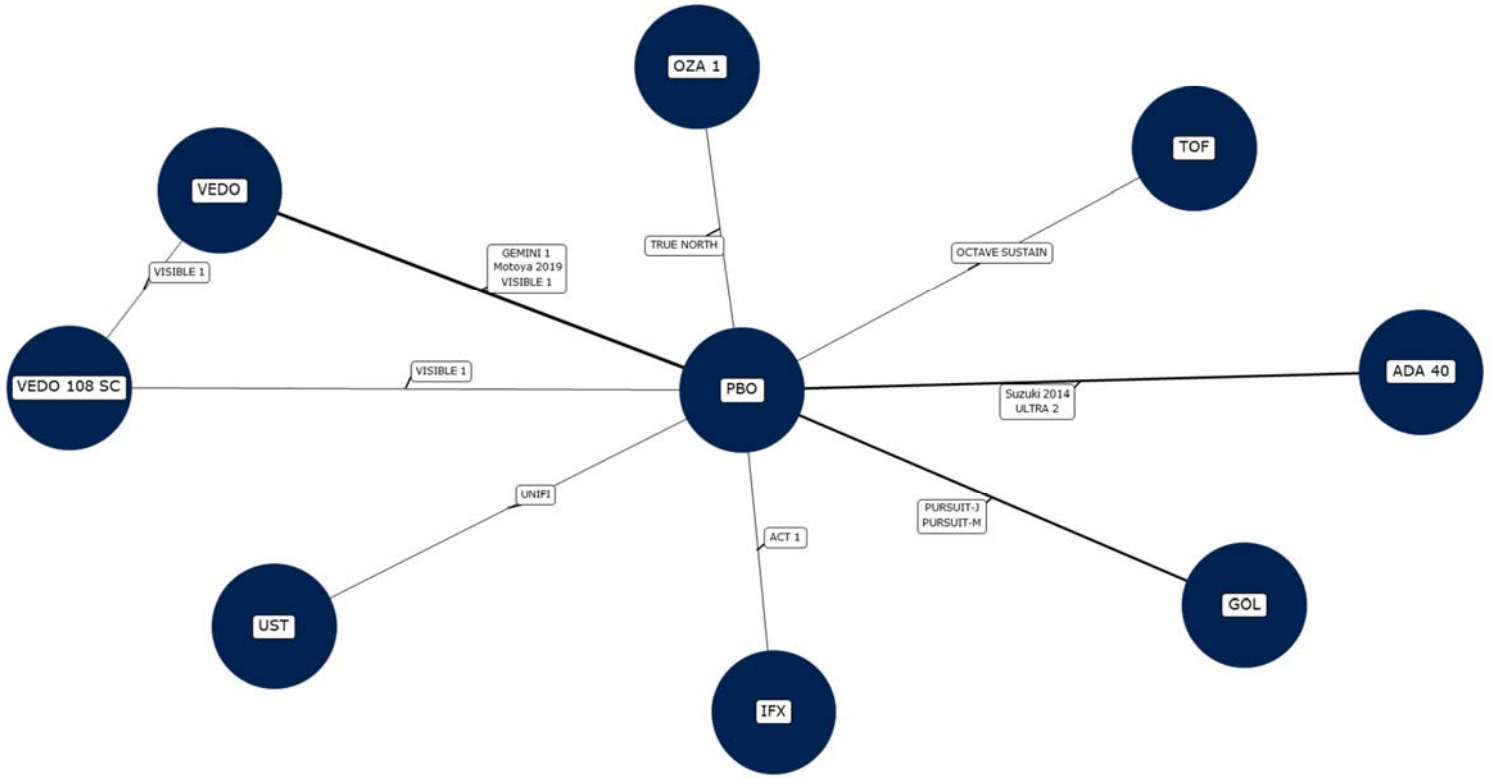
Table 34: Model fit statistics for clinical response and remission at maintenance (TNFi-naïve)

Diagnostic	Fixed effect model	Random effects model
DIC	113.15	DNC
ResDev (vs. 43 data points)	93.47	DNC
SD (95% CrI)	NA	DNC

Abbreviations: CrI: credible interval; DIC: deviance information criterion; DNC: did not converge; NA: not applicable; ResDev: residual deviance; SD: between-trial heterogeneity.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 34: Network for clinical response and clinical remission at maintenance; TNFi-naïve population[a]



[a]NMA contains pooled tofacitinib, vedolizumab, infliximab, golimumab and ustekinumab (Section B.2.8.4)
Footnote: Lines that connect nodes signify the presence of one or more RCTs that directly compare treatments, with the thickness of each line reflecting the number of RCTs informing the comparison; thicker lines signify more RCTs comparing treatments.
Abbreviations: ADA: adalimumab; GOL: golimumab; IFX: infliximab; OZA: ozanimod; PBO: placebo; Q4W: every 4 weeks; Q8W: every 8 weeks; Q12W: every 12 weeks; SC: subcutaneous; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Clinical response and clinical remission

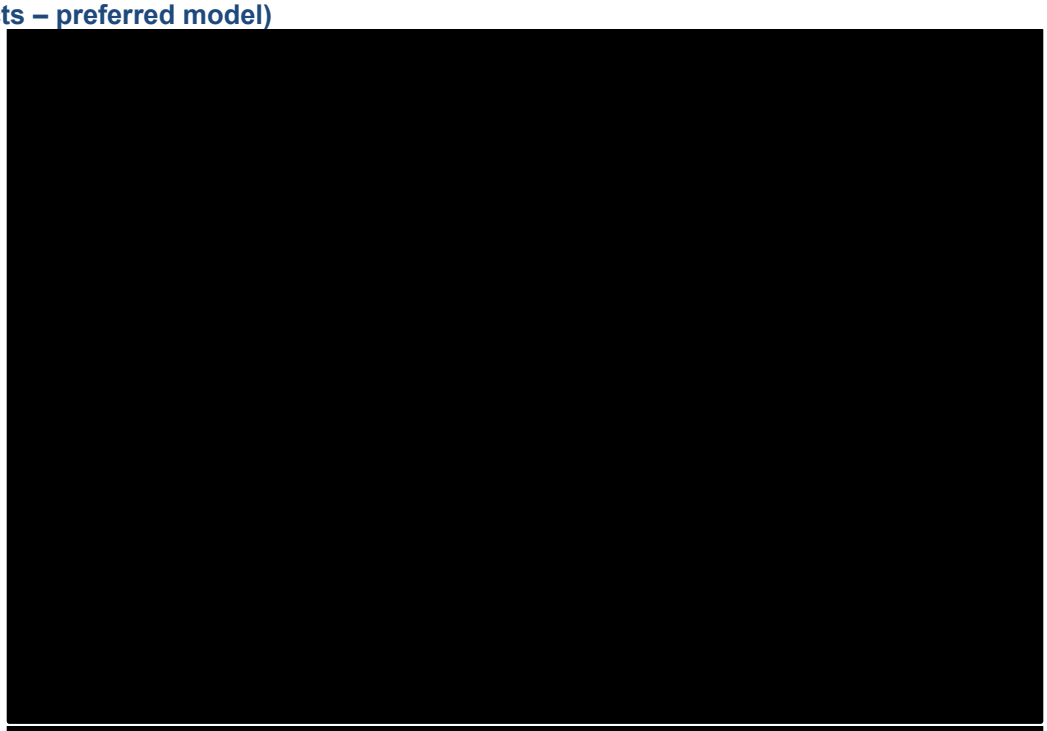
All active agents offered a statistically significant improvement in clinical response over placebo in the maintenance period (Figure 35). Point estimates trended in favour of ozanimod versus adalimumab. However, tofacitinib, and vedolizumab were found to be statistically superior to ozanimod. Ozanimod demonstrated comparable efficacy to all other active agents.

The network for clinical remission at maintenance in the TNFi-naïve population is displayed in Figure 34. All active agents offered statistically significant improvements in clinical remission compared to placebo in the maintenance period (Figure 36). Ozanimod demonstrated comparable efficacy to golimumab and ustekinumab and point estimates trended in favour of ozanimod versus infliximab and adalimumab. Tofacitinib and vedolizumab were found to be statistically superior to ozanimod.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 35: Forrest plot for clinical response at maintenance; TNFi-naïve population (fixed effects – preferred model)

a)



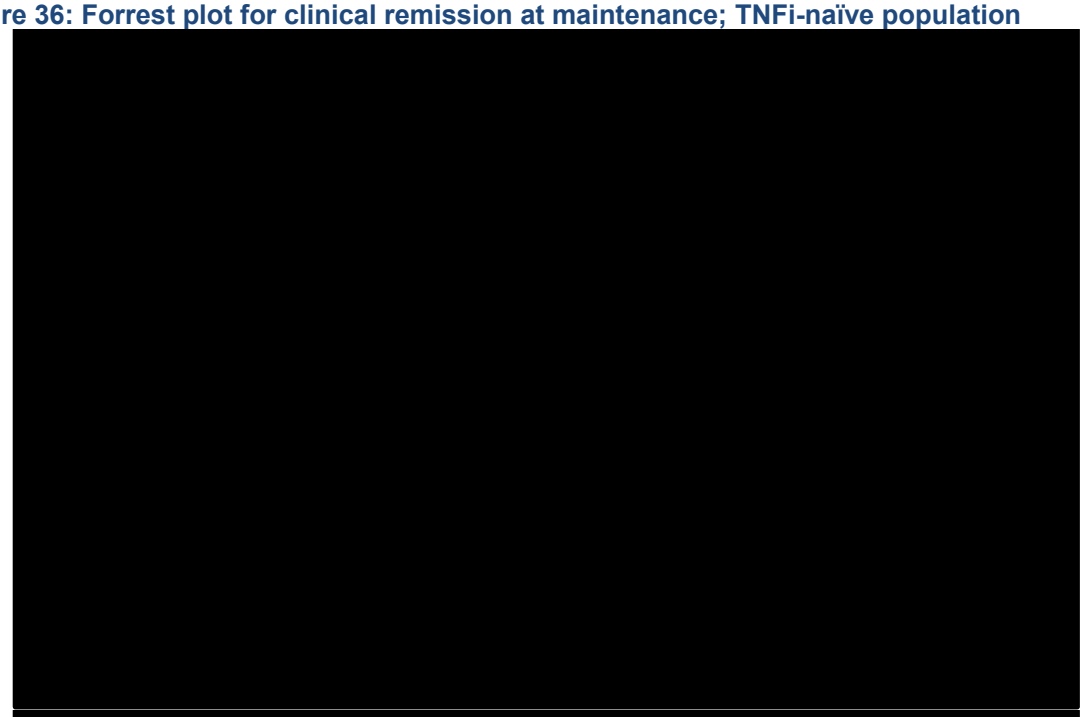
b)

Forest plots displaying a) odds ratios of active treatments versus placebo and b) odds ratios of ozanimod versus active treatments. Treatments are sorted by descending median odds ratios. Bold values indicate statistical significance. **Abbreviations:** BID: twice a day; CrI: credible interval; IV: intravenous; OR: odds ratio; PBO: placebo; Q2W: every 2 weeks; QD: once a day; SC: subcutaneous.

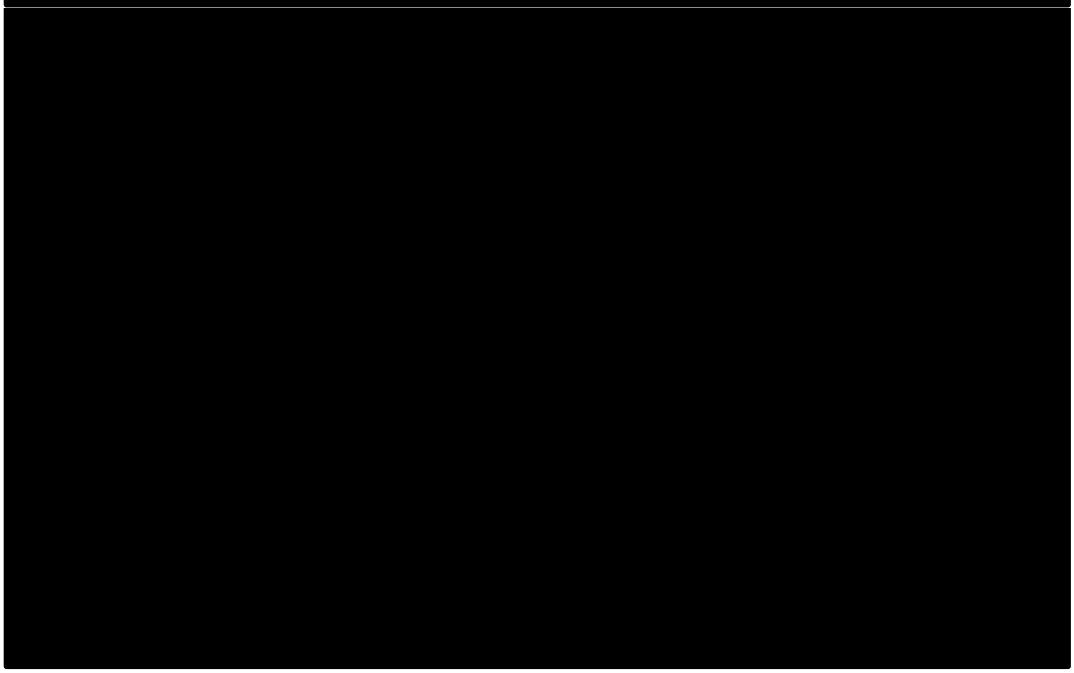
©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 36: Forrest plot for clinical remission at maintenance; TNFi-naïve population

a)



b)



Forest plots displaying a) odds ratios of active treatments versus placebo and b) odds ratios of ozanimod versus active treatments. Treatments are sorted by descending median odds ratios. Bold values indicate statistical significance. **Abbreviations:** BID: twice a day; CrI: credible interval; IV: intravenous; OR: odds ratio; PBO: placebo; Q2W: every 2 weeks; QD: once a day; SC: subcutaneous.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.8.5.2 TNFi-experienced population

Induction period

The network for clinical response and remission during the induction period for the TNFi-experienced population is displayed in Figure 37. In total, 8 studies were included in the analysis. Data for the TNFi-experienced analyses were sparser than those in the TNFi-naïve population since all data were retrieved from the TNFi-experienced subgroup in available mixed population trials. Of note, no data were available for golimumab and infliximab, since these treatments were studied in entirely TNFi-naïve populations. Of the remaining treatments, all were assessed in one or more placebo-controlled studies, with some studies evaluating multiple doses of the same TNFi as well as a single head-to-head study (VARSITY).[158]

Inspection of model fit statistics suggested that both models had similar fit (Table 35). Residual deviance was lower and the DIC was slightly (<5 points) lower in the random effects model, however, the posterior SD was highly uncertain in the random effects model. Although clinical heterogeneity highlighted in Section B.2.8, favours the random effects model the spare network structure meant the random effects model was highly uncertain leading to an underpowered SD, wherein all treatments were considered comparable to one-another and the upper bound on the pairwise 95% CrI for the odds ratios exceeded 100 in some comparisons. Therefore, primary results for clinical response and remission during the induction period for the TNFi-experienced population were derived from the fixed effect model. An assessment of inconsistency determined there to be little evidence of inconsistency between direct and indirect estimates for either model (Appendix D.4.2).

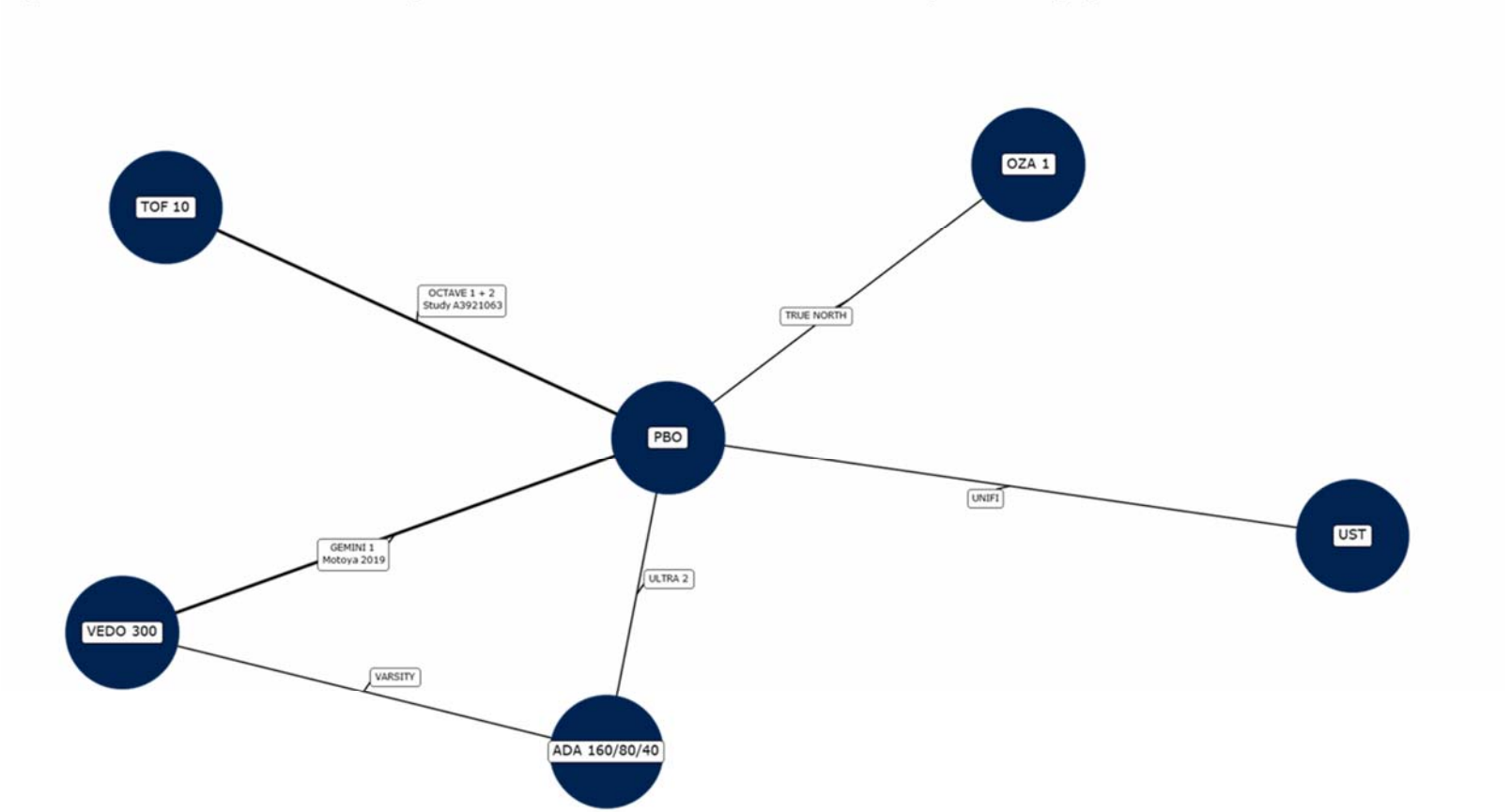
Table 35: Model fit statistics for clinical response and remission at induction (TNFi-experienced)

Diagnostic	Fixed effect model	Random effects model
DIC	67.33	67.10
ResDev (vs. 30 data points)	53.50	50.05
SD (95% CrI)	NA	0.389 (0.036 to 1.549)

Abbreviations: CrI: credible interval; DIC: deviance information criterion; NA: not applicable; ResDev: residual deviance; SD: between-trial heterogeneity.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 37: Network for clinical response and remission at induction; TNFi experienced population[a]



[a]NMA contains pooled ustekinumab (Section B.2.8.4)
Footnote: Lines that connect nodes signify the presence of one or more RCTs that directly compare treatments, with the thickness of each line reflecting the number of RCTs informing the comparison; thicker lines signify more RCTs comparing treatments.
Abbreviations: ADA: adalimumab; OZA: ozanimod; PBO: placebo; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Clinical response and clinical remission

All active agents except for adalimumab offered a statistically significant improvement in clinical response over placebo (Figure 38). No statistically significant differences were found between ozanimod and other active therapies apart from adalimumab, where ozanimod was statistically superior. Point estimates trended in favour of ozanimod versus adalimumab, vedolizumab, and ustekinumab.

The network for clinical response and remission during the induction period for the TNFi-experienced population is displayed in Figure 37. All active agents except for adalimumab offered statistically significant improvement in clinical remission over placebo (Figure 39). No statistically significant differences were found between ozanimod and other active therapies apart from adalimumab, where ozanimod was statistically superior. Point estimates trended in favour of ozanimod versus adalimumab, vedolizumab and ustekinumab.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 38: Forrest plots for clinical response at induction; TNFi experienced population (fixed effects – preferred model)

a)

b)

Forest plots displaying a) odds ratios of active treatments versus placebo and b) odds ratios of ozanimod versus active treatments. Treatments are sorted by descending median odds ratios. Bold values indicate statistical significance. **Abbreviations:** BID: twice a day; CrI: credible interval; IV: intravenous; OR: odds ratio; PBO: placebo; Q2W: every 2 weeks; QD: once a day; SC: subcutaneous.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 39: Forrest plots for clinical remission at induction; TNFi experienced population

a)

b)

Forest plots displaying a) odds ratios of active treatments versus placebo and b) odds ratios of ozanimod versus active treatments. Treatments are sorted by descending median odds ratios. Bold values indicate statistical significance. **Abbreviations:** BID: twice a day; CrI: credible interval; IV: intravenous; OR: odds ratio; PBO: placebo; Q2W: every 2 weeks; QD: once a day; SC: subcutaneous.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Maintenance period

The network for clinical response and remission at maintenance in the TNFi-experienced population is presented in Figure 40. As in the induction analyses, data for the TNFi-experienced maintenance analyses were sparser than those in the TNFi-naïve population since all data were retrieved from the TNFi-experienced subgroup in available mixed population trials. Similarly, no data were available for golimumab and infliximab as these treatments were studied in entirely TNFi-naïve populations. In total 7 studies were included in the analysis.

Inspection of model fit statistics according to suggested that the fixed effect model was associated with reasonable model fits in terms of DIC and residual deviance (Table 36). Random effects model, however, did not converge. Therefore, primary results for clinical response and remission during the maintenance period for the TNFi-experienced population were derived from the fixed effect model. An assessment of inconsistency determined there to be little evidence of inconsistency between direct and indirect estimates for the fixed effect model, while a random effects inconsistency model failed to converge (Appendix D.4.2).

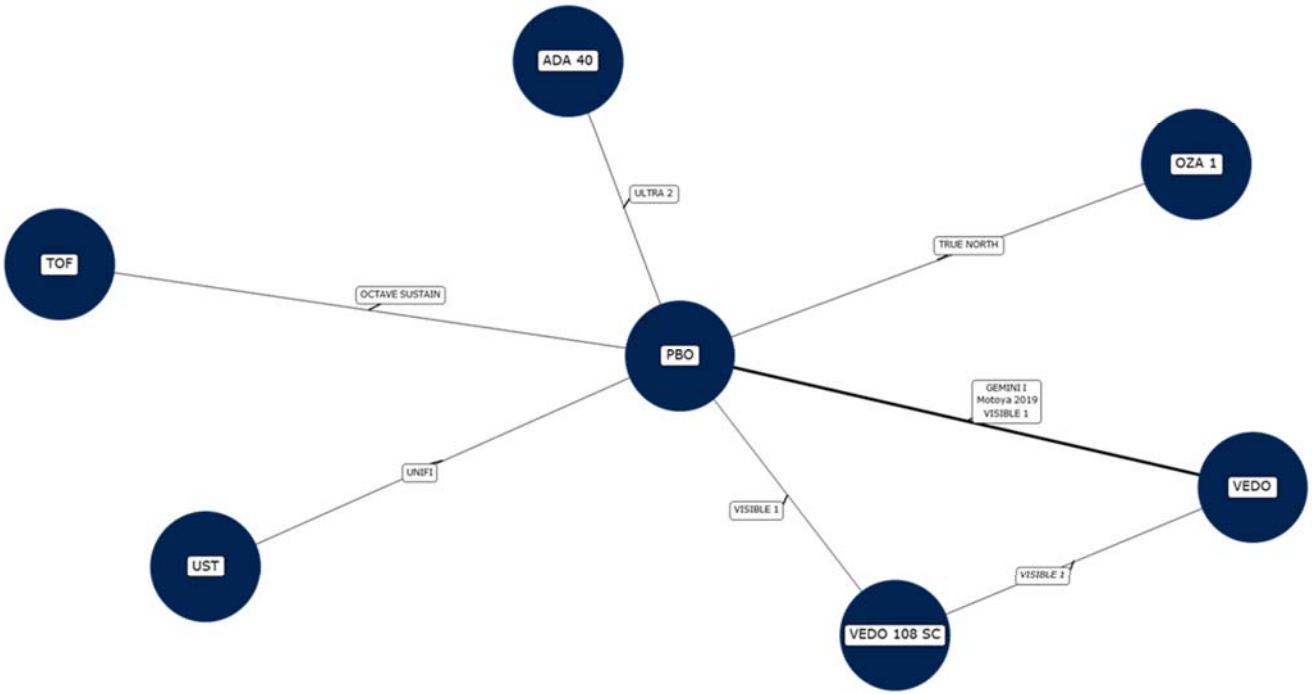
Table 36: Model fit statistics for clinical response and remission at maintenance (TNFi-experienced)

Diagnostic	Fixed effect model	Random effects model
DIC	57.25	DNC
ResDev (vs. 43 data points)	42.88	DNC
SD (95% CrI)	NA	DNC

Abbreviations: CrI: credible interval; DIC: deviance information criterion; DNC: did not converge; NA: not applicable; ResDev: residual deviance; SD: between-trial heterogeneity; vs: versus

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 40: Network for clinical response and remission at maintenance; TNFi-experienced population[a]



[a]NMA contains pooled tofacitinib, vedolizumab and ustekinumab (Section B.2.8.4)
Footnote: Lines that connect nodes signify the presence of one or more RCTs that directly compare treatments, with the thickness of each line reflecting the number of RCTs informing the comparison; thicker lines signify more RCTs comparing treatments.
Abbreviations: ADA: adalimumab; GOL: golimumab; IFX: infliximab; OZA: ozanimod; PBO: placebo; Q4W: every 4 weeks; Q8W: every 8 weeks; Q12W: every 12 weeks; SC: subcutaneous; TOF: tofacitinib; UST: ustekinumab; VDZ: vedolizumab.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Clinical response and clinical remission

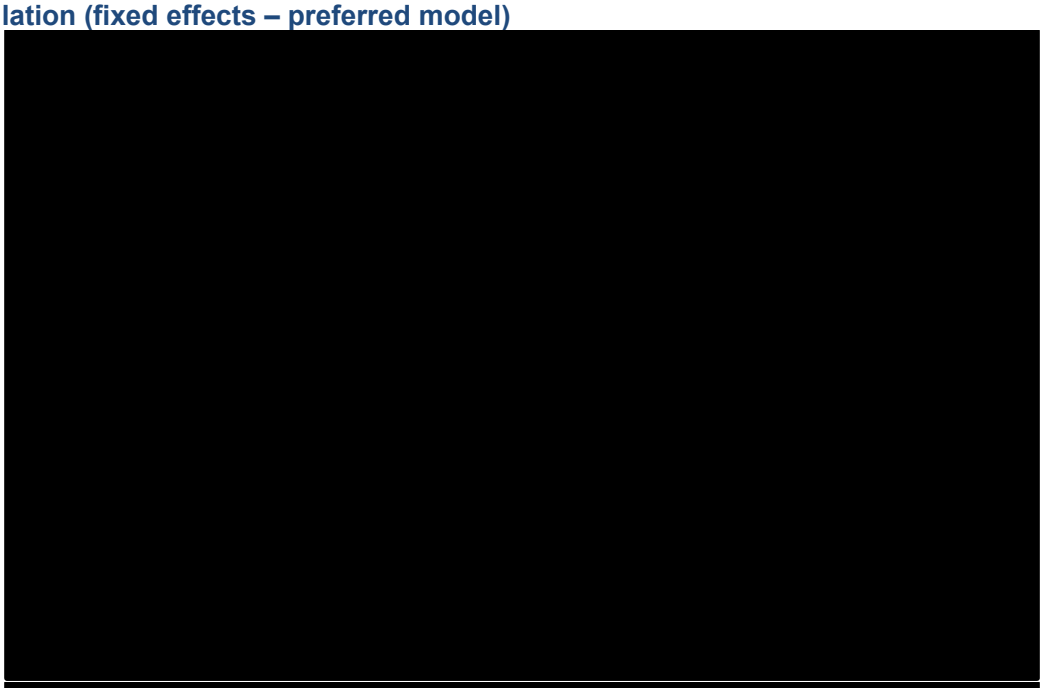
All active agents offered statistically significant improvement in clinical response over placebo in the TNFi-experienced maintenance period (Figure 41). Ozanimod was comparable to all active agents. Point estimates trended in favour of ozanimod versus ustekinumab and adalimumab.

The network for clinical response and remission at maintenance in the TNFi-experienced population is presented in Figure 40. All active agents offered statistically significant improvement in clinical remission over placebo in the TNFi-experienced population in the maintenance period (Figure 42). Ozanimod was comparable to all active agents, however, point estimates trended in favour of ozanimod versus ustekinumab and adalimumab.

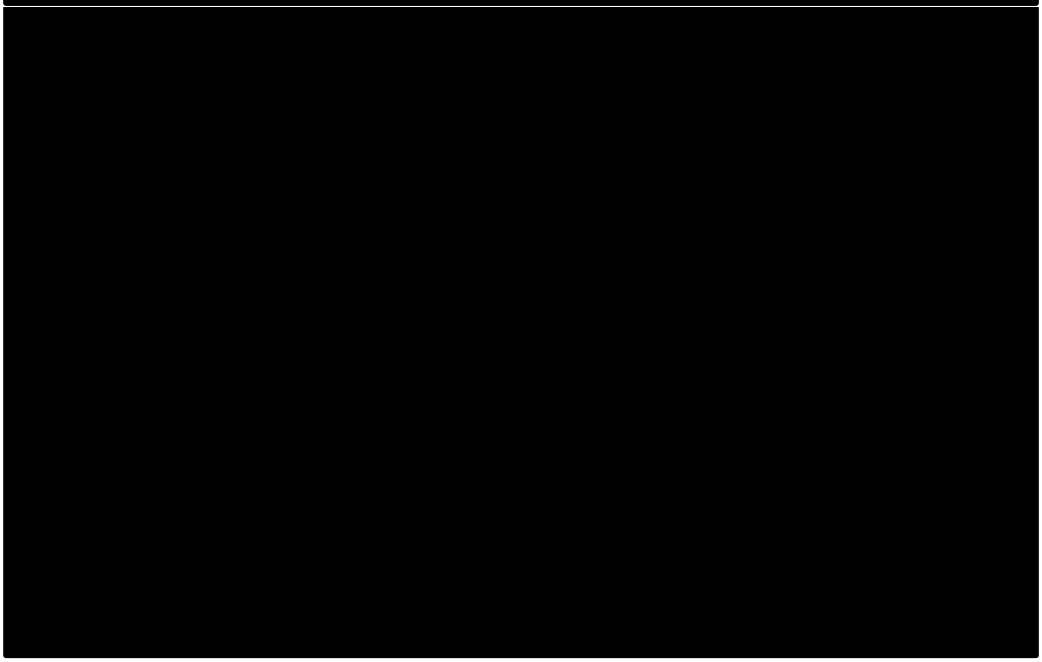
©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 41: Forrest plot for clinical response at maintenance; TNFi-experienced population (fixed effects – preferred model)

a)



b)

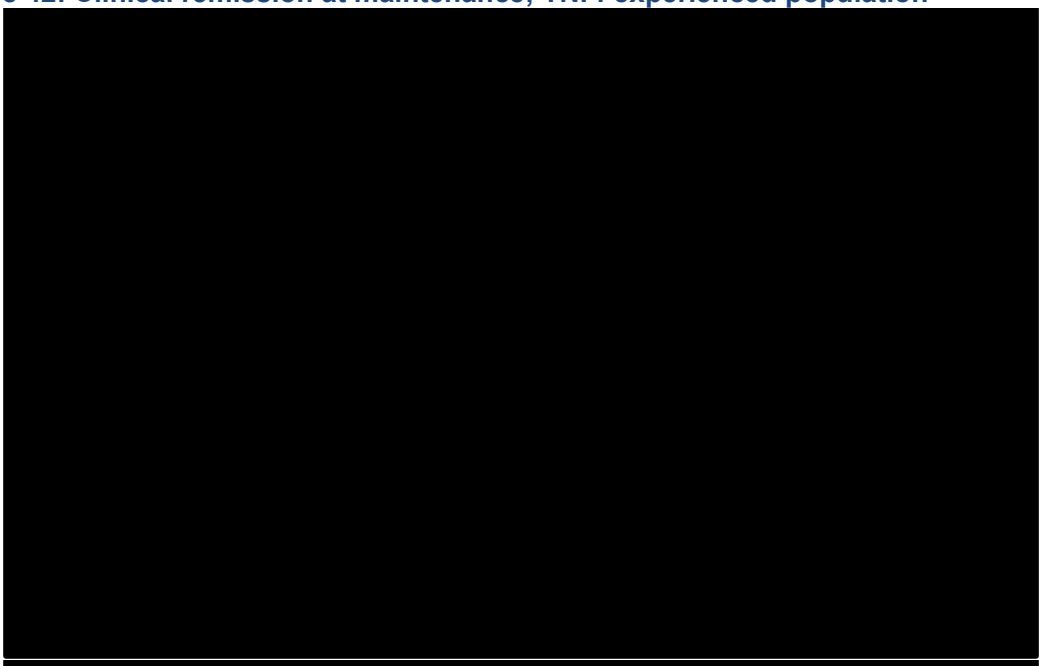


Forest plots displaying a) odds ratios of active treatments versus placebo and b) odds ratios of ozanimod versus active treatments. Treatments are sorted by descending median odds ratios. Bold values indicate statistical significance.
Abbreviations: CrI: credible interval; OR: odds ratio.

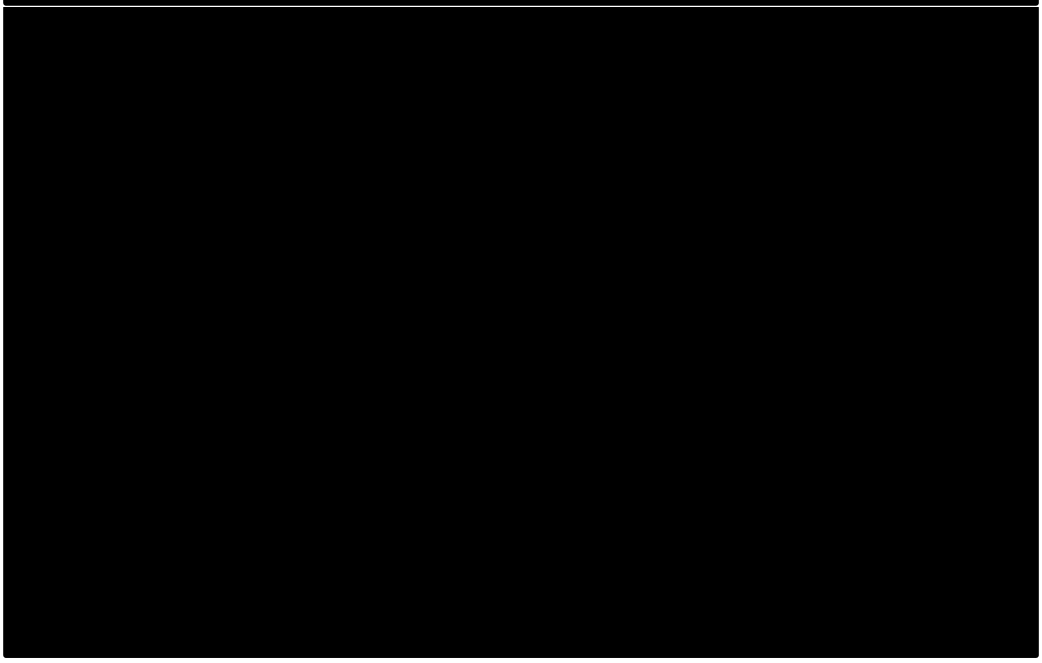
©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 42: Clinical remission at maintenance; TNFi-experienced population

a)



b)



Forest plots displaying a) odds ratios of active treatments versus placebo and b) odds ratios of ozanimod versus active treatments. Treatments are sorted by descending median odds ratios. Bold values indicate statistical significance.
Abbreviations: BID: twice a day; CrI: credible interval; IV: intravenous; OR: odds ratio; PBO: placebo; Q2W: every 2 weeks; QD: once a day; SC: subcutaneous.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.8.5.3 Sensitivity analyses

The conclusions of the sensitivity analyses to explore the effect of various assumptions used in the model as well as the influence of potential sources of heterogeneity (Section B.2.8.3), are presented below. The sensitivity analyses were performed on the primary outcomes of the NMA (clinical response and clinical remission). Overall, conclusions were unchanged relative to the base case, demonstrating the results of the NMA to be robust. The results of the sensitivity analysis are provided in Appendix B.4.5.

Unpooled doses

A sensitivity analyses was performed which did not assume dose pooling of the same active treatment if it had the same method of administration (Section B.2.8.4). In general, the results of the sensitivity analysis were aligned with the base case analyses (Appendix B.4.5).

Exclusion of treat-through trials

As discussed in Section B.2.8.2, owing to the heterogeneity in maintenance phase design (re-randomised design versus treat-through) a sensitivity analysis was performed where the re-calculated treat-through data were excluded from the maintenance results. The results of the sensitivity analysis can be found in Appendix B.4.5. While the analysis rendered comparisons versus adalimumab and infliximab in the maintenance period impossible, comparisons versus the remaining treatments were unaffected by the change with conclusions being identical to the base case. This confirmed that adjustments made did not affect the re-randomised treatments available, such as tofacitinib, vedolizumab, golimumab, and ustekinumab.

TRUENORTH 3-component Mayo scores

As described in Section B.2.8.2, to align with other trials, data from the TRUENORTH sensitivity analysis using the 4-component Mayo score was used in the base case NMAs. A separate sensitivity analysis was conducted to explore the influence of using the 3-component TRUENORTH data on NMA findings the results of which are presented Appendix B.4.5. The results of the sensitivity analyses were in agreement with the base case analyses.

Exclusion of Asian trials

As discussed in Section B.2.8.2, due to several of the included trials restricting recruitment to an entirely Asian population, a sensitivity analysis was performed in which Asian trials were excluded. Exclusion of trials with entirely Asian populations resulted in sparser networks and less data available on adalimumab, golimumab, and vedolizumab. Results were similar to the pooled dose base case NMAs with no changes to conclusions (Appendix B.4.5).

Inclusion of TOUCHSTONE in TNFi-naïve analysis

TOUCHSTONE reported no TNFi-naïve or TNFi-experienced subgroup data and therefore was not included in the base case NMA analysis (Section B.2.8.3). However, as to 82% of patients in TOUCHSTONE were TNFi-naïve, a sensitivity analysis was was conducted to explore the effect of including the trial in the TNFi-naïve induction analyses. Inclusion of the mixed TOUCHSTONE induction trial data introduced some heterogeneity to the network since 18% of patients in the trial were not TNFi-naïve. Despite this, conclusions made versus

©Celgene Ltd, a BMS Company (2022). All rights reserved

ozanimod in the induction period in the sensitivity analyses were unchanged relative to the base case, demonstrating consistency of ozanimod's treatment effect when considering Phase II data despite the potential heterogeneity introduced (Appendix B.4.5).

B.2.8.6 Uncertainties in the indirect and mixed treatment comparisons

As outlined in Section B.2.8.6, a thorough exploration of the various sources of heterogeneity within the trials included in the NMA was conducted to evaluate the potential influence on NMA results. Several key areas of heterogeneity were identified including trial design, definitions of the included subgroups, baseline characteristics and outcome availability and definitions.

Only reported study characteristics and baseline prognostic factors that were consistently reported across the relevant data sources could be considered in the feasibility assessment. Consequently, potential effect modifiers that were not consistently reported could not be accounted for in the analyses. The benefit of conducting an NMA based off randomised-stratified RCT data is that randomisation is retained and only treatment-effect modifying variables are of concern.

A number of steps were taken to reduce heterogeneity in the NMA, including restriction of eligible studies based on induction/maintenance length, statistical adjustment to treat-through trials to align with the data presented in re-randomised trials, and use of the 4-component Mayo data from TRUENORTH for alignment with other trials. Separate analyses were conducted in TNFi-naïve and TNFi-experienced patients, in line with all major previous NMAs in UC. In addition, several sensitivity analyses were conducted to control for the remaining key areas of heterogeneity. The sensitivity analyses showed that the results of the NMA to be robust to uncertainty.

Whilst there was some heterogeneity between baseline characteristics across trials, it was uncertain whether these characteristics represented treatment effect modifiers. As an adjusted NMA may introduce additional bias, a standard NMA was used in line with recommendations in NICE DSU TSD18 and the approaches taken in all previous NMAs in UC.[1, 88, 159] The decision to use a standard NMA in the base-case was further supported by the results of the sensitivity analyses which revealed the base-case results were robust to uncertainty and therefore supported the view that heterogeneities identified in the NMA were not likely to be treatment effect modifiers. Overall, a conservative approach to trial and data inclusion was taken to limit the influence of the heterogeneity described throughout.

Overall, the NMA found that ozanimod had similar efficacy in terms of clinical remission and clinical response in both the induction and maintenance period in TNFi-naïve and TNFi-experienced subgroups. These results were also consistent across a variety of sensitivity analyses exploring uncertainties in the NMA.

B.2.8.7 Conclusions of the NMA

In the TNFi-naïve subgroup, the NMA found no statistically significant differences between ozanimod and the relevant comparators (infliximab, adalimumab, golimumab and vedolizumab) for induction of clinical response and remission. At maintenance, ozanimod was found to have comparable results for clinical remission and response to all relevant comparators in the TNFi-naïve subgroup, with the exception of IV vedolizumab.

©Celgene Ltd, a BMS Company (2022). All rights reserved

In the TNFi-experienced subgroup, no statistically significant differences were found between ozanimod and the relevant comparators (vedolizumab and ustekinumab) for clinical response and remission in induction or maintenance.

Overall, ozanimod was found to have comparable treatment efficacy in both clinical response and clinical remission to the relevant comparators in both the induction and maintenance period in the TNFi-naïve and TNFi-experienced populations, respectively. Uncertainties in the results of the NMA were assessed in several sensitivity analyses and the results were found to be robust.

B.2.9 *Adverse reactions*

A summary of the safety data from TRUENORTH are presented below. Further safety data, such as details of dose interruptions as well as TEAEs by system organ class, can be found in Appendix K.

B.2.9.1 Safety summary

Treatment with ozanimod was well tolerated, with a similar overall incidence of TEAEs observed during the induction period in patients treated with ozanimod and placebo in Cohort 1 (40.1% versus 38.0%, respectively).[4] One death occurred in a patient in Cohort 2 during the induction period. This was considered to be unrelated to study drug by the Investigator or Sponsor.[4]

In the maintenance period, the overall incidence of TEAEs were slightly higher in patients randomised to maintenance with ozanimod compared to placebo (49.1% versus 36.6%, respectively).[4] However, the rates of serious TEAEs and TEAEs leading to study discontinuation were lower in patients randomised to ozanimod versus placebo. No deaths occurred in the maintenance period.

The safety profile of ozanimod was consistent with the known safety profile of ozanimod in MS and no unexpected AEs were observed.[160, 161] In line with clinical trials of ozanimod in MS infections were the most common TEAEs by system organ class (Appendix K). Overall ozanimod was found to have a tolerable safety profile with severe TEAEs occurring in [redacted] of patients in any treatment arm.

B.2.9.2 Treatment duration and exposure

In the induction period, the duration of exposure was comparable between the ozanimod arm ([redacted] SD:± [redacted] weeks) and the placebo arm ([redacted] SD:± [redacted] weeks). In the maintenance period, the duration of exposure was higher in patients re-randomised to ozanimod ([redacted] SD:± [redacted] weeks) compared to those re-randomised to placebo ([redacted] SD:± [redacted] weeks).

B.2.9.3 Adverse events

The safety and tolerability of ozanimod was evaluated at Week 10 (end of the induction period) and Week 52 (end of the maintenance period). The safety population consisted of all patients who received at least 1 dose of the investigational drug. For the induction period, a treatment-emergent adverse event (TEAE) was defined as any AE with date of first onset, or date of worsening in severity, on or after the first dose of the induction period. Within the maintenance period, a TEAE was defined as any AE with date of first onset, or date of

©Celgene Ltd, a BMS Company (2022). All rights reserved

worsening in severity, on or after the date of the first dose of the maintenance period. TEAEs with onset after the date of the 90-day safety follow-up visit were excluded.

Summary of TEAEs

Overall, for ozanimod compared to placebo in the induction period, there was a similar incidence of TEAEs (40.1% versus 38.0%), severe TEAEs ([redacted] versus [redacted]), serious TEAEs ([redacted] versus [redacted]), suspected related serious TEAEs ([redacted] versus [redacted]), TEAEs leading to discontinuation (3.3% versus 3.2%) and TEAEs leading to interruption ([redacted] versus [redacted]) in Cohort 1 (Table 37).[4] The most common TEAEs in the induction period, occurring in ≥3% in any arm were anaemia, nasopharyngitis, and headache. Serious TEAEs included UC exacerbation (occurring in ≥2 patients in either group), anaemia (occurring in ≥2 patients in either group) and infection.

In the maintenance period the overall incidence of TEAEs (49.1% versus 36.6%) and suspected related TEAEs ([redacted] versus [redacted]) were higher in patients re-randomised to ozanimod relative to those re-randomised to placebo (Table 37).[4] However, the incidence of serious TEAEs ([redacted] versus [redacted]) and TEAEs leading to discontinuation (1.3% versus 2.6%) were lower in patients re-randomised to ozanimod relative to those re-randomised to placebo.[4] The most common TEAEs >3% in any arm were ALT increased, headache, arthralgia, nasopharyngitis, and gamma-glutamyl transferase (GGT) increased. Serious TEAEs occurring in ≥2 patients in either group were UC exacerbation, anaemia, and appendicitis/complicated appendicitis (Table 37).

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 37: Overview of TEAEs in TRUENORTH (safety population)

Patients with at least 1, n (%)	Induction period			Maintenance period		
	Cohort 1		Cohort 2	Placebo (N = 69)	Re-randomised patients	
	Ozanimod (N =429)	Placebo (N = 216)	Ozanimod (N = 367)		Ozanimod – placebo (N = 227)	Ozanimod - Ozanimod (N = 230)
Any TEAE	172 (40.1)	82 (38.0)	146 (39.8)	[redacted]	83 (36.6)	113 (49.1)
Serious TEAE	17 (4.0)	7 (3.2)	23 (6.3)	[redacted]	18 (7.9)	12 (5.2)
Suspected related TEAE[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TEAE leading to discontinuation	14 (3.3)	7 (3.2)	14 (3.8)	[redacted]	6 (2.6)	3 (1.3)
Most common TEAEs (>3% in any arm in the induction or maintenance period)						
Anaemia	18 (4.2)	12 (5.6)	16 (4.4)	[redacted]	4 (1.8)	3 (1.3)
ALT increased	11 (2.6)	0	6 (1.6)	[redacted]	1 (0.4)	11 (4.8)
Headache	14 (3.3)	4 (1.9)	10 (2.7)	[redacted]	1 (0.4)	8 (3.5)
Arthralgia	10 (2.3)	3 (1.4)	5 (1.4)	[redacted]	6 (2.6)	7 (3.0)
Nasopharyngitis	15 (3.5)	3 (1.4)	10 (2.7)	[redacted]	4 (1.8)	7 (3.0)
GGT increased	5 (1.2)	0	6 (1.6)	[redacted]	1 (0.4)	7 (3.0)
Serious TEAS						
UC exacerbation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infection	4 (0.9)	1 (0.5)	6 (1.6)	[redacted]	4 (1.8)	2 (0.9)
Appendicitis/complicated appendicitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infections						
Nasopharyngitis	15 (3.5)	3 (1.4)	10 (2.7)	[redacted]	4 (1.8)	7 (3.0)
Upper respiratory tract infection	5 (1.2)	1 (0.5)	8 (2.2)	[redacted]	4 (1.8)	2 (0.9)
Herpes zoster infection	2 (0.5)	0	1 (0.3)	[redacted]	1 (0.4)	5 (2.2)

[a]Assessed as probably, possibly, or related to study drug by the investigator.
Note: Patients with multiple events reported for the same summary level are counted only once. Percentages are based upon the number of patients in the Safety population.
Abbreviations: AE: adverse event; GGT: gamma-glutamyl transferase increased; N: number of patients in trial arm; n: number of patients in category; TEAE: treatment-emergent adverse event.
Sources: TRUENORTH CSR: Table 42 and Table 14.3.1.1B.[5] Sandborn 2021.[4]

©Celgene Ltd, a BMS Company (2022). All rights reserved

AEs of special interest (induction period)

Based on the understanding of the physiological roles of S1P modulators, special attention was directed at assessing cardiac effects, hepatic effects, infections, lymphopenia, macular oedema, malignancies, and pulmonary effects during TRUENORTH. During the induction period the number of adverse events of special interest (AESI) were low. However, more AESIs that were investigator-coded or sponsor-identified (marked with an asterisk in Table 38) occurred in patients treated with ozanimod compared with placebo in Cohort 1 ([redacted] versus [redacted], respectively). The most frequently reported AESI in the randomised ozanimod group were hepatic effects (including [redacted] events of elevated liver enzymes, [redacted] events of increased ALT, [redacted] event of increased aspartate aminotransferase, and [redacted] event of increased liver function test) and infection (including 3 events of herpes zoster) (Table 38).[4] [redacted] of the [redacted] patients with AESIs in the ozanimod trial arm consisted of Sponsor-identified events including serious infections (appendicitis, nasopharyngitis, otitis externa, pyelonephritis, and vestibular neuronitis) and hepatic effects (transaminase increased). The [redacted] AESI in the placebo arm was a Sponsor-identified event of serious infection (bronchitis). As in the induction period the overall incidence of AESI was low, however AESIs occurred slightly more frequently in patients re-randomised to ozanimod than patients re-randomised to placebo ([redacted] versus [redacted], respectively). The most frequently reported AESIs were infection, hepatic effects and malignancy. Infections occurred more frequently in patients re-randomised to ozanimod. Investigator-identified infections in patients re-randomised to ozanimod included [redacted] events of herpes zoster and [redacted] event of *C. difficile* infection. There was [redacted] event investigator-identified complicated appendicitis in a patient re-randomised to placebo.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 38: Adverse events of special interest (safety population)

AESI category[a] Preferred term, n (%)	Induction period			Maintenance period		
	Cohort 1		Cohort 2	Placebo (N = 69)	Re-randomised patients	
	Ozanimod (N = 429)	Placebo (N = 216)	Ozanimod (N = 367)		Ozanimod – Placebo (N = 227)	Ozanimod – Ozanimod (N = 230)
Any AESI (Investigator coded or Sponsor- identified)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cardiac						
Bradycardia	2 (0.5)	0	3 (0.8)	0	0	0
Hepatic effects						
Alanine aminotransferase increased	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infection						
Herpes zoster*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Macular oedema	1 (0.2)	0	1 (0.3)	0	0	1 (0.4)
Pulmonary						
Dyspnoea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Asthma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Malignancy						
Basal cell carcinoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Rectal adenocarcinoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adenocarcinoma of the colon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Breast cancer	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vascular disorders						
Hypertension	6 (1.4)	0	7 (1.9)	0	3 (1.3)	4 (1.7)
Hypertensive crisis	1 (0.2)	0	0	0	1 (0.4)	1 (0.4)

*Asterisk denotes AESI identified by Sponsor review of TEAEs reports. [a] AESIs include bradycardia, heart conduction abnormalities (2nd degree and higher AV block), macular oedema, malignancy, serious or opportunistic infection, pulmonary effects, and hepatic effects and have been adjudicated by the safety review team per the safety management plan. Sponsor designated AESIs from AESI-Disposition CRF not categorised by the investigator will show up as 'Additional Event of Interest Defined by Sponsor'.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Note: TEAEs were coded using MedDRA version 22.1. Patients with multiple events reported for the same summary level will be counted only once. Percentages are based upon the number of patients in the Safety population.
Abbreviations: AE: adverse event; AESI: adverse events of special interest; AV: atrioventricular; CRF: case report form; MedDRA: Medical Dictionary for Regulatory Activities; N: number of patients in trial arm; n: number of patients in category; TEAE: treatment-emergent adverse event.
Source: TRUENORTH CSR: Table 54 –55.[5] Sandborn 2021.[4]

©Celgene Ltd, a BMS Company (2022). All rights reserved

Deaths

One death occurred in a patient in Cohort 2 during the induction period. The death was considered to be unrelated to study drug by the Investigator or Sponsor. No other deaths were recorded in any arm.[4]

B.2.10 *Ongoing studies*

The TRUENORTH and TOUCHSTONE trials are complete and no further trial readouts are anticipated.

A Phase III, multicentre, OLE trial (RPC01-202) evaluating the long-term safety and efficacy of ozanimod in patients with moderately to severely active UC is ongoing. However, as the study was single arm it is not anticipated to deliver any clinically relevant data to the decision problem. The study included patients who had previously participated in either TRUENORTH including both TNFi-naïve and experienced patients (Section B.2.3.1) or had completed ≥1 year of the open-label phase of TOUCHSTONE (Appendix J).

B.2.11 *Innovation*

Moderately to severely active UC is a debilitating disease associated with a high clinical burden. Patients suffer troublesome and distressing symptoms such as increased stool frequency requiring multiple bowel openings per day, bloody stools and faecal urgency. These symptoms have a profound impact on patients' HRQoL, psychological and emotional well-being. Whilst the availability of biologics has resulted in improvements in disease management for patients, there are still considerable limitations associated with the available treatments in both the TNFi-naïve and experienced populations.

Owing to the lifelong, chronic nature of the disease, patients often require long-term treatments. Patients may fail to respond to existing treatments or subsequently lose their response over time; approximately a third to a half of patients treated with a TNFi do not respond due to primary or secondary non-response.[105, 106] In clinical practice there remains a high rate of incomplete or non-response to UC medications, indicating a need for new therapeutic options.[108] As a result, the availability of new therapeutics options were identified as an area of unmet need by an expert group consensus in UC published in 2019.[108] When an existing treatment fails to control disease, fewer treatment options are available to patients before considering surgery. Due to surgery being associated with a multitude of short- and long-term risks surgery is viewed as an undesirable treatment option by most patients (Section B.1.3.2). A survey of 2,333 UC patients from Denmark, Italy, the Netherlands, Spain, Switzerland and the UK, found that 86.6% of patients would rather try a new UC drug than undergo surgery.[53] Ozanimod, with this novel mechanism of action, addresses this unmet need, providing patients a new therapeutic option to treat symptoms and induce remission. Importantly, ozanimod has been demonstrated to be efficacious in not only the TNF-naïve population but also the TNFi-experienced population (Section B.2.6.1), where there are fewer treatment options available and the disease often more difficult to treat.

In addition, ozanimod satisfies the particular unmet need for a therapy with a convenient method of administration and a tolerable safety profile. All available advanced therapies, excluding tofacitinib, which is associated with significant safety concerns (Section B.1.3.4), are administered either by IV or SC injection which can be viewed by patients as inconvenient and

©Celgene Ltd, a BMS Company (2022). All rights reserved

intrusive methods of administration.[2, 117, 118] The availability of treatments with convenient methods of administration are essential in a chronic disease such as UC where treatment is highly individualised, depending on both patient preference and clinician judgement. A study in 298 patients with IBD showed that patients preferred oral administration over IV or SC injection (91% versus 33% and 34%, respectively) and therefore the introduction of ozanimod would provide patients with another highly-desired oral treatment option, without the safety risks associated with tofacitinib.[119] Safety data from TRUENORH showed ozanimod has a tolerable safety profile, with a similar overall incidence of TEAEs reported during the induction period for both ozanimod and the placebo treatment arms (Section B.2.9). Similarly, withdrawals due to AEs occurred in 3.3% and 1.3% of patients in the induction and maintenance period of TRUENORTH, respectively, whereas withdrawals ranged from 0–9.7% in competitor therapies in both induction and maintenance.[4, 140, 152, 162]

The availability of oral treatment options is particularly important in the COVID-19 era as patients do not need to regularly attend hospital for treatment administration. This has the benefit of both reducing patient numbers in hospital (a priority goal during the COVID-19 era to reduce the risk of infection) and minimising the time patients spend travelling and attending hospital appointments, thus reducing the disease burden on patients.[120] A cross-sectional study of patients enrolled at the IBD unit of a French hospital revealed that patients receiving infliximab spent 6.5 hours away from home or work for each infliximab infusion.[109] These data show that ozanimod satisfies the unmet need for a highly desired convenient oral treatment option with a tolerable safety profile.

Finally, as ozanimod is a small molecule, unlike existing biologic therapies it is not associated with the production of anti-drug antibodies (immunogenicity).[163] Immunogenicity issues associated with biologics often leads to dose escalation, which is both expensive and leads to an increased risk of AEs.[110-112] Approximately 30% of patients taking TNFis receive dose escalation after 12 months due to loss of response, rising to 50% after 3 years.[113-116] Additionally, an SLR collecting RWE from 48 studies investigating interventions for the treatment of moderately to severely active UC found 35% of patients receiving vedolizumab received dose escalation.[74] Treatment with ozanimod is not anticipated to be associated with dose escalation beyond the initial dose titration week, thus reducing both the risks of AEs and increased costs associated with dose escalation.

Overall, ozanimod satisfies the clear unmet need amongst patients with moderately to severely active UC for a novel treatment option with a convenient oral method of administration and a tolerable safety profile.

B.2.12 *Interpretation of clinical effectiveness and safety evidence*

Principal findings from the clinical evidence base

In TRUENORTH, ozanimod 1 mg (the licensed dose) was effective at inducing clinical remission, response and endoscopic improvement in the overall population, demonstrating statistically significant improvements versus placebo.[4, 164] These results represent clinically meaningful improvements in UC, with clinical response/remission linked to a reduced symptom burden in patients and hence improved patient HRQoL and employment status due to reduced abseentism.[127] Similarly, mucosal healing, which had a stricter definition in the TRUENORTH study requiring improvements in both endoscopic and histologic healing, is associated with long-

©Celgene Ltd, a BMS Company (2022). All rights reserved

term remission of disease activity, decreased risk of surgery and improved HRQoL in UC patients.[23, 53, 80]

The results of TRUENORTH further demonstrate that ozanimod is able to maintain clinical remission in the longer term (one year), with a significantly higher proportion of patients achieving durable clinical remission and maintenance of remission at Week 52.[4, 165] Maintenance of remission is key to ensuring that patients do not relapse and therefore maintain their reduced symptom burden and improved HRQoL achieved during induction.[127] Disease remission is linked to better longer terms outcomes such as reduced risk of colectomy as well as potentially being linked to a reduced risk of dysplasia and colorectal cancer as carcinogenesis is thought to be promoted by the presence of chronic inflammation.[63] Further to this, endoscopic improvement (45.7% vs 26.4%, respectively; p<0.001), and mucosal healing (29.6% vs 14.1%, respectively; p<0.001) were found to be significantly improved with ozanimod versus placebo at 52 weeks, which as stated previously, are associated with long-term remission of disease activity. Long-term remission of disease activity is associated with reduced absenteeism, improved psychological and physical well-being as well as reduced resource utilitsation.[5, 23, 53, 80, 126]

Improvements in efficacy endpoints results for ozanimod compared to placebo were found to translate directly to improvements in patients HRQoL.[5] Improvements in patient HRQoL is particularly important in chronic diseases such as UC where patients may experience symptoms throughout the course of their life, and represents a key treatment aim for new treatments for UC.[16]

Furthermore, the TRUENORTH trial found ozanimod to have a tolerable safety profile which was consistent with the known longer-term safety profile for ozanimod in MS and the expected AE profile in patients with moderately to severely active UC (see Section B.2.9). No new safety signals were observed.

These results were reflected in the TNFi-naïve and experienced subgroups, however a slightly lower proportions of patients achieved the efficacy endpoints in the TNFi-experienced population versus the TNFi-naïve (Section B.2.6.1). This was in line with clinical expectations, as typically patients who have failed prior treatment have more severe and more difficult to treat disease.[5] The efficacy of ozanimod in the TNFi-experienced population is key as there is unmet need in these patients for an effective treatment option with a convenient oral method of administration and tolerable safety profile (Section B.1.3.5).

No head-to-head evidence was available for ozanimod versus relevant comparators in UK clinical practice. Therefore, an NMA based on clinical trials identified via an SLR was conducted. Overall, the NMA found that ozanimod had similar efficacy to relevant comparators in terms of clinical remission and clinical response in both the induction and maintenance period in TNFi-naïve and TNFi-experienced subgroups. These results were consistent across a variety of sensitivity analyses (Section B.2.8.5) investigating potential areas of heterogeneity. These results demonstrate that ozanimod, with its convenient oral method of administration provides patients comparable treatment efficacy to relevant comparators which are administered either IV or via SC, both of which can be viewed as inconvenient and intrusive methods of administration by patients.[117, 118] This is of particular importance in chronic diseases such as UC where patient preference plays a role in treatment choice.[99, 100]

Strengths and limitations of the clinical evidence base

©Celgene Ltd, a BMS Company (2022). All rights reserved

The evidence base informing this appraisal has been derived from a comprehensive clinical SLR investigating the efficacy and safety of a variety of treatment options, including ozanimod, in patients with moderately to severely active UC (see Section B.2.1). The SLR was conducted in accordance with the Cochrane Handbook for Systematic Reviews of Interventions and reported in alignment with the Preferred Reporting Items for Systematic Literature Reviews and Meta-Analyses (PRISMA) statement.[166, 167] The principal evidence for the safety and efficacy of ozanimod is provided by the TRUENORTH trial, which is a high quality, randomised, double-blind, placebo-controlled Phase III trial. Alongside TOUCHSTONE (a randomised, double-blind, placebo-controlled Phase II trial), these trials have been used as the basis of the submitted EMA marketing authorisation application.

The trial population of TRUENORTH is consistent with the licenced indication for ozanimod and the population specified in the NICE final scope (see Section B.1.1). The baseline characteristics for the patients in this trial are consistent with the target patient population in the UK; the average patient age in the trial was approximately 42 years which is consistent with the typical age of a UC patient.[40] Additionally, approximately twice as many patients had limited to left side colitis over extensive colitis which reflects the expected disease profile of UC.[40] Participants prior medications were also aligned with the typical treatment pathway patients would follow in UK clinical practice; TNFis (infliximab and adalimumab) were the most commonly used prior biologic medications reflecting their first-line use following treatment failure with CvT (Section B.1.3.4).

The primary endpoint of TRUENORTH was the proportion of patients in clinical remission at Week 10.[4] This is in alignment with prior appraisals (TA633) in this indication and therefore was considered a reasonable primacy efficacy endpoint.[88] This endpoint was defined according to the 3-component Mayo score and therefore accounted for multiple clinical markers, such as rectal bleeding, stool frequency and endoscopic improvement, representing an appropriate measure of the clinical benefit of ozanimod. TRUENORTH was 52 weeks in length and therefore enabled the generation of long-term evidence for ozanimod in terms of maintenance of clinical response and remission; there is an unmet need for novel therapeutics which are not only able to induce but maintain remission in the long term.

As with many clinical trials within UC a key limitation of the evidence base was the lack of a direct comparison versus relevant comparators to this appraisal (TNFis, vedolizumab, and ustekinumab). To address this limitation an NMA was conducted in order to obtain relative efficacy estimates to inform the economic analysis. The NMA was conducted in line with NICE Decision Support Unit (DSU) Technical Support Document (TSD) 2, utilising a Bayesian framework, with the best model (fixed-effects versus random-effects) chosen based on DIC and clinical considerations.[154] There are some limitations associated with these analyses, as presented in Section B.2.9.5. However, the results were consistent across a variety of sensitivity analyses exploring the impact of these limitations on the results.

A final limitation of the evidence base is that the placebo arm of re-randomised patients in the maintenance period of TRUENORTH may exhibit a carry-over effect. The arm included responders who were randomised to ozanimod during the induction period and therefore, it is possible that the trial design may facilitate a carry-over effect and hence bias results against ozanimod. The extended half-life of ozanimod metabolites may increase the likelihood of this carry-over effect.[12] This was an unavoidable limitation owing to the requirement of the study to investigate the impact of not continuing ozanimod treatment in the maintenance period and was aligned with the limitation discussed in a recent appraisal in the same indication with a similar study design (TA633).[1]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.2.12.1 Conclusion

Moderately to severely active UC is a debilitating disease, however there are considerable limitations associated with available treatments, namely resulting from the poor response durability, inconvenient methods of administration and concerns regarding safety profiles. As a result, up to 10–15% of patients have to undergo surgical resection in their lifetime despite the associated risks of severe and in some case irreversible complications, such as infertility, impaired sexual function or being left with a stoma.[168-171]

Ozanimod, with its novel mechanism of action offers patients a new therapeutic option, which effectively reduces symptoms and can control active disease by inducing and maintaining remission and may therefore reduce the risk of patients having to resort to surgery. A positive recommendation of ozanimod from NICE would provide patients with moderately to severely active UC a novel treatment option with a highly desired, convenient, oral method of administration, a tolerable safety profile and an innovative mechanism of action.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.3 Cost effectiveness

Summary of the *de novo* cost-effectiveness model

- A *de novo* cost-utility model was developed for the economic evaluation of ozanimod compared with relevant comparators in the UK, in accordance with the NICE reference case
- In line with the decision problem the analysis was conducted in two distinct trial populations: TNFi-naïve and TNFi-experienced patients
- The model was comprised of two parts: active treatment and post active treatment. The active treatment portion was further split into induction periods and maintenance periods which were modelled using a Markov cohort approach in line with prior submissions (TA633 and TA547)[1, 88]
- The model structure contained nine distinct health states: 'Remission', 'Response No Remission', 'Active UC', '1st Surgery', 'Post 1st Surgery Remission', 'Post 1st Surgery Complications', '2nd Surgery', 'Post 2nd Surgery Remission' and 'Death'
- The analysis was conducted from an NHS/PSS perspective, with a lifetime time horizon and costs and outcomes were discounted at 3.5% per annum
- Efficacy data for ozanimod and relevant comparators were derived from the base case NMA (Section B.2.8)
- EQ-5D-5L data were collected during the TRUENORTH trial, however the results were subject to several limitations (Section B.3.4.1). Health-state utility values derived from the literature were therefore used in the base-case analysis, in line with TA633[1]
- Resource use and costs included in the model were based on previous technology appraisals (TA633[1]) and appropriate published sources including the British National Formulary (BNF), electronic market information tool (eMIT) and National Schedule for NHS (2018/2019)
- Feedback from a UK clinician was sought to validate assumptions and inputs in the model

Base case cost-effectiveness results

- In the base case analysis in the TNFi-naïve population the incremental Net Health Benefit (NHB) for ozanimod versus adalimumab, infliximab, golimumab and vedolizumab was 0.003, 0.175, 0.100 and 0.205, respectively
- In the base case analysis in the TNFi-experienced population, incremental NHB for ozanimod versus ustekinumab and vedolizumab was 0.170 and 0.156, respectively
- In the fully incremental analysis ozanimod was the most cost-effective treatment option both in the TNFi-naïve and experienced populations, and were consistent with the pair-wise analyses
- Overall, the base case results for all comparisons demonstrated ozanimod to be cost-effective at a willingness-to-pay threshold £30,000 per QALY and thus ozanimod can be considered a cost-effective use of NHS resources in both the TNFi-naïve and TNFi-experienced populations

Sensitivity analyses

- Probabilistic sensitivity analyses (PSA) and deterministic sensitivity analyses (DSA) were conducted to assess uncertainty in the economic analysis and demonstrate that the base case cost-effectiveness results were robust to an extensive number of scenario analyses
- The DSA results identified a small number of key influential parameters such as the proportion of ozanimod patients achieving sustained response and remission at maintenance as well as the proportion of patients entering response no remission health states at induction, with the model being largely robust to uncertainty in the majority of parameters
- Scenario analyses conducted to address sources of uncertainty in the model such as an analysis which included extended induction demonstrated that whilst there was variation in the NHB, the cost-effectiveness conclusions remain the same and the majority of NHBs are considered cost-effective at a willingness-to-pay threshold of £30,000 per QALY

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.3.1 *Published cost-effectiveness studies*

An economic SLR was conducted on the 2nd September 2020 and was subsequently updated using the same review protocol on 14th October 2021 to identify all relevant literature published on the following topics:

- Economic evaluations of therapies for the treatment of moderately to severely active UC
- Healthcare resource use (HCRU) and cost studies on moderately to severely active UC

The SLR was conducted following current best practices, as recommended by the Cochrane Collaboration.[172] The reporting of the methods and results of the SLR were done in line with the guidance provided by the National Institute for Health and Care Excellence (NICE) and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[167, 173, 174] Full details of the economic SLR search strategy, study selection process and results are reported in Appendix E and Appendix G for economic evaluations HCRU, respectively.

In total, 13 unique UK economic evaluations in moderately to severely active UC were identified in the SLR, the details of which are presented in Table 39. No prior economic evaluations were identified for ozanimod in the population of relevance to this submission.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 39: Summary list of published cost-effectiveness studies

Study	Year	Summary of model	Patient population	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Punekar Y (2010)[175]	2010	CUA; Hybrid decision tree (first year) and Markov model (year 1 and beyond); 10 years; 3.5% discount rate; 4 health states	Patients hospitalised with an acute exacerbation of UC	• IFX: QALY= 0.80 • SoC: QALY= 0.68 • Ciclosporin: QALY= 0.70 • Surgery: QALY= 0.58	**British pounds, 2006-2007** • IFX, total: cost= £19,847 • SoC, total: cost= £18,524 • Ciclosporin, total: cost= £18,122 • Surgery, total: cost= £17,067	• IFX vs. SoC: £397 • IFX vs. Ciclosporin: £5,731
Wilson MR (2017)[66]	2017	CUA; Hybrid decision tree (induction) and Markov model (maintenance); NHS; lifetime horizon; 3.5% discount rate	Patients with moderately-to-severely active UC who have had an inadequate response with, lost response to, or are intolerant to either a CvT or an TNFi, TNFi naïve	**Mixed (ITT) population:** • VDZ: QALY= 10.516 • CvT: QALY= 10.181 **TNFi-naïve population:** • VDZ: QALY= 10.549 • CvT: QALY= 10.168 **TNFi-failure population:** • VDZ: QALY= 10.416 • CvT: QALY=	**British pounds, 2013-2014** **Mixed (ITT) population:** • VDZ, total: cost= £205,362 • CvT, total: QALY= cost= £203,991 **TNFi-naïve population:** • VDZ, total: cost= £205,521 • CvT, total: cost= £203,917 **TNFi-failure population:** • VDZ, total: cost= £206,133	**VDZ + CvT vs. CvT alone:** • Mixed (ITT) population: £4,095 • TNFi-naïve population: £4,423 • TNFi-failure population: £5,972

©Celgene Ltd, a BMS Company (2022). All rights reserved

				10.150	• CvT, total: cost= £204,547	
Lohan C (2019)[176]	2019	CUA; Markov model; NHS; lifetime; 3.5% discount rate; 3 health states for TOFA and 6 for CvT and surgery	Adult patients with moderately to severely active UC, with or without prior exposure to TNFi therapies	**TNFi-naïve population:** • ADA: QALY= 9.191 • GLM: QALY= 9.286 • IFX: QALY= 9.346 • VDZ: QALY= 9.462 • TOF: QALY= 9.536 • CvT: QALY= 8.991 **TNFi-experienced population:** • ADA: QALY= 9.051 • GLM: QALY= 9.051 • IFX: QALY= 9.051 • VDZ: QALY= 9.146 • TOF: QALY= 9.240 • CvT: QALY= 8.903	**British pounds, 2016-2017** **TNFi-naïve population:** • ADA, total: cost= £138,534 • GLM, total: cost= £141,360 • IFX, total: cost= £145,660 • VDZ, total: cost= £152,694 • TOF, total: cost= £143,963 • CvT, total: cost= £132,349 **TNFi-experienced population:** • ADA, total: cost= £137,035 • GLM, total: cost= £138,088 • IFX, total: cost= £140,661 • VDZ, total: cost= £145,380 • TOF, total: cost= £140,399 • CvT, total: cost= £132,712	**TNFi-naïve population:** • ADA vs. conventional: £30,982 (fully incremental: extendedly dominated) • GLM vs. CvT: £30,602 (fully incremental: extendedly dominated) • IFX vs. CvT: £37,495 (fully incremental: dominated) • VDZ vs. CvT: £43,205 (fully incremental: dominated) • TOF vs. CvT: £21,388 (fully incremental: £21,388) **TNFi-experienced population:** • ADA vs. CvT: £29,284 (fully incremental: extendedly 4dominated) • Golimumab vs. con7ventional: £53,831 (fully incremental: dominated) • IFX vs. CvT: £36,403 (fully incremental: extendedly dominated) • VDZ vs. CvT: £52,275 (fully incremental: dominated) • TOF vs. CvT: £22,816 (fully incremental: £22,816) **Scenario analysis, overall ITT population:**

©Celgene Ltd, a BMS Company (2022). All rights reserved

				Scenario analysis, overall ITT population: • VDZ: QALY= 9.301 • TOF: QALY= 9.397 • CvT: QALY= 8.948	**Scenario analysis, overall ITT population:** • VDZ, total: cost= £147,822 • TOF, total: cost= £141,500 • CvT, total: cost= £132,508	• VDZ vs. CvT: £43,485 (fully incremental: dominated) • TOF vs. CvT: £20,038 (fully incremental: £20,038)
Tappenden P (2016)[177]	2016	CUA; Markov model; 5 health states; NHS perspective; lifetime horizon (60 years); 3.5% discount rate	Patients with moderate to severe UC for whom at least one prior therapy has failed.	**Colectomy not an option:** • ADA: QALY= 10.82 • IFX: QALY= 10.81 • GLM: QALY= 10.63 • CvT: QALY= 10.47 **Colectomy is an option:** • Colectomy: QALY= 14.71 • ADA: QALY= 10.82 • IFX: QALY= 10.81 • GLM: QALY= 10.63 • CvT: QALY= 10.47	**British pounds, 2013-2014** **Colectomy not an option:** • ADA, total: cost= £91,222 • IFX, total: cost= £96,595 • GLM, total: cost= £90,087 • CvT, total: cost= £73,620 **Colectomy is an option:** • Colectomy, total: cost= £56,268 • ADA, total: cost= £91,222 • IFX, total: cost= £96,595 • GLM, total: cost= £90,087	• For the population in whom colectomy is not an acceptable option, IFX was dominated by ADA (although the difference in expected QALYs between these options is very small), whereas GLM is dominated by ADA and conventional non-biologic therapy • ADA versus CvT: £50,278 • In the population in whom surgery is an option colectomy is expected to dominate IFX, ADA, GOL and CvT

©Celgene Ltd, a BMS Company (2022). All rights reserved

					• CvT, total: cost= £73,620	
Williams JG (2016)[178]	2016	CUA; Markov model; NHS; 3.5% discount rate; 5 health states	Patients admitted to hospitals with acute severe UC that is refractory to IV hydrocortisone (2-5 days of treatment)	• IFX, mean (SD): QALY = 1.900 (0.16) • Ciclosporin, mean (SD): QALY = 1.921 (0.18)	**British pounds, 2012-2013:** • IFX, mean (SD): cost = £20,241 (£695) • Ciclosporin, mean (SD): cost = £14,609 (£593)	NR
Wilson MR (2018)[66]	2018	CUA; Hybrid decision tree (induction) and Markov model (maintenance); NHS; lifetime horizon; 3.5% discount rate; 6 health states	Patients with moderately to severely active UC who have had an inadequate response with, lost response to, or are intolerant to a CvT such that they have switched to a treatment with a biologic, TNFi naive	• VDZ: QALY= 14.077 • ADA: QALY= 13.872 • IFX: QALY= 13.788 • GLM: QALY= 13.809	**British pounds, 2012-2013** • VDZ, total: cost= £199,431 • ADA, total: cost= £194,765 • IFX, total: cost= £206,066 • GLM, total: cost= £200,018	• VDZ versus ADA: £22,775 • VDZ was dominant compared with IFX • VDZ versus GLM: £10,618 • VDZ versus ADAL: £59,466
NICE TA329[35] (ERG Model)	2015	Hybrid decision tree (induction) and Markov model (maintenance) NHS; Lifetime; 3.5% discount rate	Moderate-to-severe with at least one prior failed therapy	Colectomy: QALY= 14.72	British pounds, NR Colectomy: £41,921	**Colectomy is not an option:** • GOL versus CvT: £97,149 • ADA versus CvT: £50,624 **Colectomy is an option:** • ADA, GOL, IFX, and CvT were dominated by colectomy; that is, they provided fewer QALYs at a higher cost than colectomy.

©Celgene Ltd, a BMS Company (2022). All rights reserved

NICE TA329[35] (Company Model)	2015	Markov model (maintenance); NHS;10 years; 3.5% discount rate; 8 states	Moderate-to-severe with at least one prior failed therapy *TNFi naïve* *TNFi exposed*	NR	**British pounds, 2012-2013, NR** NR	• GOL versus Colectomy: £27,994 • GOL versus IFX: £80,318 • GOL versus ADA: NR
NICE TA329[35] (Company Model)	2015	Hybrid decision tree (induction) and Markov model (maintenance); NHS; 10 years; 3.5% discount rate; 8 health states	Moderate-to-severe with at least one prior failed therapy	NR	**British pounds, 2012-2013, NR** NR	• IFX versus Colectomy: £38,307 • IFX versus ADA: £54,564 • IFX versus GOL: £75,998
NICE TA342[85] 2015	2015	Hybrid decision tree (induction) and Markov model (maintenance); NHS; 10 years; 3.5% discount rate; 6 health states for both induction and maintenance	Patients with moderate-to-severe UC with inadequate response or intolerant to CvT *Mixed* *TNFi naïve* *TNFi exposed*	**Mixed (ITT) population:** • VDZ: QALY= 5.551 • CvT: QALY= 5.397 • Surgery: QALY= 4.281 **TNFi-naïve population:** • VDZ: QALY= 5.898 • CvT: QALY= 5.555 • IFX: QALY= 5.818 • ADA: QALY= 5.760 • GOL: QALY= 5.790 • Surgery: QALY= 4.281	**British pounds, NR** **Mixed (ITT) population:** • VDZ, total: cost= £77,056 • CvT, total: cost= £71,925 • Surgery, total: QALY= cost= £107,831 **TNFi-naïve population:** • VDZ, total: cost= £69,075 • CvT, total: cost= £67,406 • IFX, total: QALY= cost= £73,952 • ADA, total: QALY= cost= £68,157 • GOL, total:	**VDZ vs. CvT:** • Mixed population: £33,297 • TNF-naïve population: £4,862 • TNF-failure population: £64,999 **VDZ vs. surgery:** • Mixed population: VDZ dominates • TNF-naïve population: VDZ dominates • TNF-failure population: VDZ dominates **VDZ vs. IFX:** • TNF-naïve population: VDZ dominates **VDZ vs. ADA:** • TNF-naïve population: £6,634

©Celgene Ltd, a BMS Company (2022). All rights reserved

				TNFi-failure population: • VDZ: QALY= 5.463 • CvT: QALY= 5.373 • Surgery: QALY= 4.281	QALY= cost= £70,387 • Surgery, total: QALY= cost= £107,831 **TNFi-failure population:** • VDZ, total: cost= £78,409 • CvT, total: cost= £72,570 • Surgery, total: cost= £107,831	**VDZ vs. GOL:** • TNF-naïve population: VDZ dominates
NICE TA547[88]	2018	Markov model; NHS; Lifetime; 3.5% discount rate	Patients with moderate-to-severe UC *TNFi naïve* *TNFi exposed*	NR	**British pounds, NR** NR	**TNFi-naïve:** • CT: £8,554 • ADA: Dominated • GOL: Dominated • IFX: Dominated • VDZ: £615,057 **TNF-experienced:** • CT: £10,302 • VDZ: £7,838,238
NICE TA633[1] 2020	2020	Hybrid decision tree (induction) and Markov model (maintenance); NHS; Lifetime; 3.5% discount rate; 9 health states	Patients with moderate-to-severe UC *Biologic/JAK naïve* *Biologic/JAK exposed*	NR	**British pounds, NR** NR	**Biologic non-failure population (vs. CvT):** • UST: £33,192 • VDZ: Dominated • IFX: Dominated • TOFA: Extended Dominated • GOL: Dominated • IFX-bio: Dominated

©Celgene Ltd, a BMS Company (2022). All rights reserved

						• ADA: Dominated • ADA-bio: Extended Dominated **Biologic failure population (vs. CvT):** • UST: £37,023 • VDZ: Dominated • TOFA: Dominated • ADA: Dominated • ADA-bio: Extended Dominated
Diamantopoulos, 2019[179]	2019	Markov model; NHS; Lifetime; NR; NR; health states determined by treatment and level of disease control but not specified	Patients with moderate-to-severe UC *TNFi-naïve* *TNFi-experienced*	NR	**British pounds, NR** NR	**TNFi-naïve (vs CvT):** • TOFA: £21,338 **TNF-experienced (vs CvT):** • TOFA : £22,816

Abbreviations: ADA: adalimumab; CvT: conventional therapy; CUA: cost utility analysis; EQ-5D: EuroQoL five-dimensions; GOL: golimumab; ICER: incremental cost-effectiveness ratio; IFX: infliximab; ITT: intention to treat; IV: intravenous; QALYs: quality-adjusted life years; SoC: standard of care; TNFi: tumour necrosis factor alpha inhibitor; TOF: tofacitinib; UC: ulcerative colitis; VDZ: vedolizumab.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.3.2 *Economic analysis*

The economic SLR found no economic evaluations investigating ozanimod for the treatment of moderately to severely active UC and therefore a *de novo* cost-effectiveness analysis has been conducted for the purpose of this appraisal and is described below. The cost-effectiveness model employed for this economic analysis was built in Microsoft Excel®.

The objective of this economic analysis was to assess the cost-effectiveness of ozanimod within its marketing authorisation for the treatment of adult patients with moderately to severely active UC who have had an inadequate response with, lost response to, or were intolerant to CvT compared with current clinical management in the UK. As discussed in Section B.1.1, the base case cost-effectiveness analyses were conducted in sub-populations based on prior TNFi exposure.

In line with the NICE reference case the analysis was conducted from the perspective of the NHS and Personal Social Services (PSS) in the United Kingdom (UK) and included direct medical costs over a lifetime horizon.

B.3.2.1 Patient population

In line with the decision problem addressed in this submission (Section B.1.1) and the licensed indication for ozanimod, the patient population considered in the economic analyses was split into two distinct populations:

- **TNFi-naïve**: This subgroup includes adult patients with moderately to severely active UC who have not previously received a TNFi. Patients in this group had recently progressed on CvT
- **TNFi-experienced:** This subgroup of includes adult patients with moderately to severely active UC that have experienced treatment failure whilst receiving TNFis. Treatment failure includes intolerance to treatment (for example, discontinuation due to adverse events [AEs]), lack of treatment efficacy (patients fail to experience response to treatment) and loss of response

These populations were in line with the subgroups of patients from TRUENORTH by exposure to prior TNFi therapy, as presented in Section B.2.6.1.

B.3.2.2 Model structure

A *de novo* hybrid decision-analytic model was developed in Microsoft Excel® to evaluate the cost-effectiveness of ozanimod versus relevant comparators in the UK as a treatment for patients with moderately to severely active UC. The model was comprised of two parts: active treatment and post active treatment. The active treatment portion was further split into induction periods and maintenance periods.

Justification of model structure

Prior economic models developed for UC used a hybrid decision tree for the initial induction period and a Markov cohort model for the maintenance period. Similarly, the induction period in the *de novo* model was modelled via tunnel states within the Markov cohort model trace which were able to capture the effective decision tree at the end of the induction period that determines the initial health state distribution in the maintenance period. However, these tunnel

©Celgene Ltd, a BMS Company (2022). All rights reserved

states have the added benefit of allowing patients to enter the maintenance period at any model cycle, therefore enabling the variable length of induction periods (Section B.3.3.3) between treatments to be captured. This cohort model structure was considered to be sufficient to capture the treatment pathway for these patients, in line with the models submitted in prior appraisals (TA633, TA547).[1, 88]

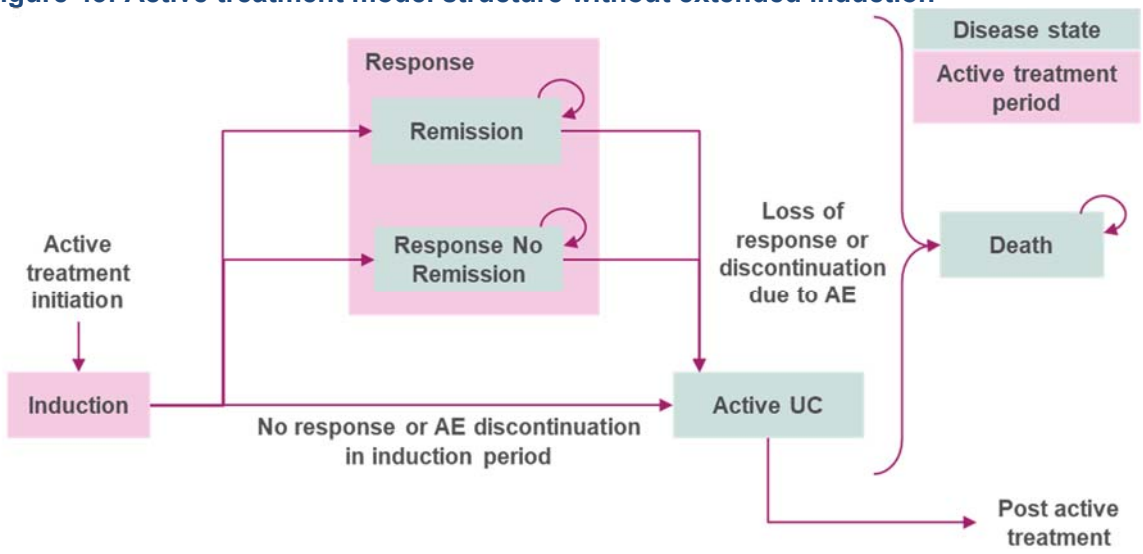
Fundamentally the model structure was divided into two parts. The first part modelled active treatment; Figure 43 shows this model structure (without extended induction, as per the base case analysis). The second part modelled the movement of patients after failure on active treatment, where surgery remains the last available treatment option. This part of the model structure is shown in Figure 45. Overall, the model structure contained nine distinct health states: 'Remission', 'Response No Remission', 'Active UC', '1st Surgery', 'Post 1st Surgery Remission', 'Post 1st Surgery Complications', '2nd Surgery', 'Post 2nd Surgery Remission' and 'Death'. Patients experiencing a 'response' in the model included patients with remission, i.e. response included both patients in the 'Remission' and 'Response No Remission' health states. This model structure was used for both TNFi-naïve and TNFi-experienced patients, with comparators differing between the two populations, as discussed in B.3.2.1.

Active treatment – induction period

The induction periods for active treatment were modelled through a series of tunnel states. According to the SmPCs for ozanimod and the relevant comparators for this submission, patients are assessed for response level after 6 to 10 weeks of treatment. In the model, patients initiating active treatment progressed through a series of 2-week tunnel states representative of the induction period of that treatment, thereby accounting for the variable length of initial induction periods between treatments. For example, ozanimod has a 10-week induction period and so patients could progress through 5 tunnel states. Following the final induction tunnel state patients were distributed into 'Remission', 'Response No Remission' or 'Active UC' health states (see Section B.3.3.4). To ensure the face validity of the induction period, patients could move to the 'death' health state during these tunnel states. Patients could also experience adverse events leading to treatment discontinuation during the induction period. This was captured in the model as an additional proportion of patients transitioning to the 'Active UC' state. It was assumed that patients who discontinued due to adverse events during the induction period still incurred the full cost associated with a course of induction treatment given they 'purchased' the full course at the beginning of the induction period.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 43: Active treatment model structure without extended induction



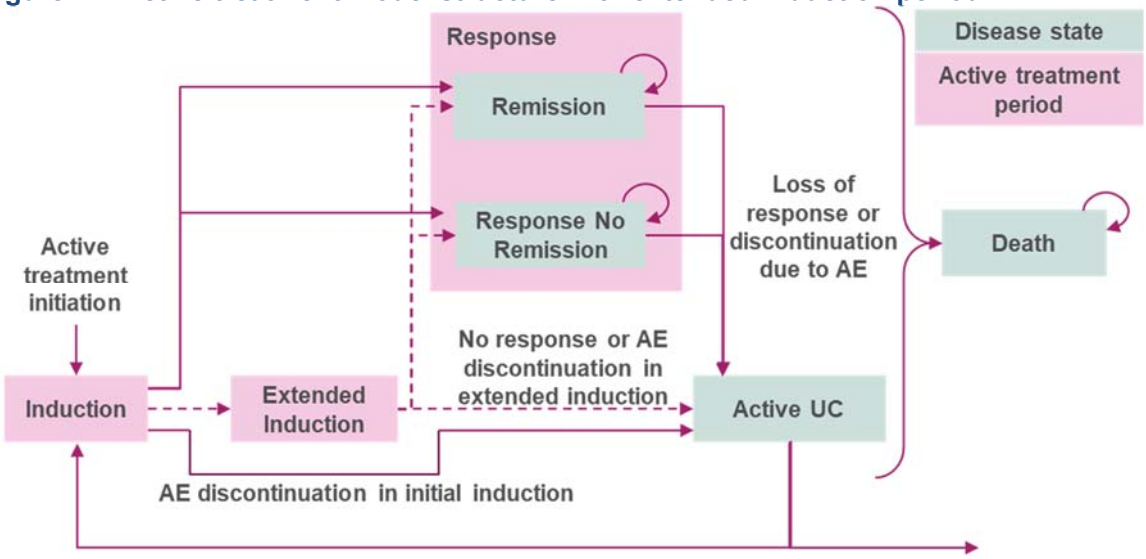
Abbreviations: AE, Adverse event; UC: ulcerative colitis.

TA633 included functionality to model extended induction periods where patients who initially did not respond in the standard induction remained in an extended induction state while they received additional treatment before reassessment for response at a later timepoint.[1] However, the Committee in TA633 noted that no conclusive clinical evidence was provided that demonstrated extended induction always occurs in clinical practice. In addition, clinical consultation conducted as part of this appraisal noted that while some patients may experience an extended induction it is not considered standard clinical practice in the UK.

However, as extended induction does occur in some cases and is included in the SmPC for ustekinumab, it was explored in a scenario analysis (see Section B.3.8.3). The structure of the active treatment part of the model when an extended induction period is included is depicted in Figure 44. Following the standard induction period, a proportion of patients could transition to the 'Response No Remission' and 'Remission' health states. A proportion of patients could experience discontinuation due to an adverse event during the standard induction period, and transition directly to the 'Active UC' health state. Patients who do not experience a response but do not discontinue after the initial induction period enter an extended induction period after which they could transition to 'Remission', 'Response No Remission' or 'Active UC' health states based on their response status (see Figure 44).

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 44: Active treatment model structure with extended induction period



Abbreviations: AE, Adverse event; UC: ulcerative colitis.

Active treatment – maintenance period

As mentioned above, upon completion of induction a proportion of patients enter the 'Remission' or 'Response No Remission' health states. Patients remain in the 'Remission' and 'Response No Remission' health states until they lose their initial response or discontinue treatment due to AEs.

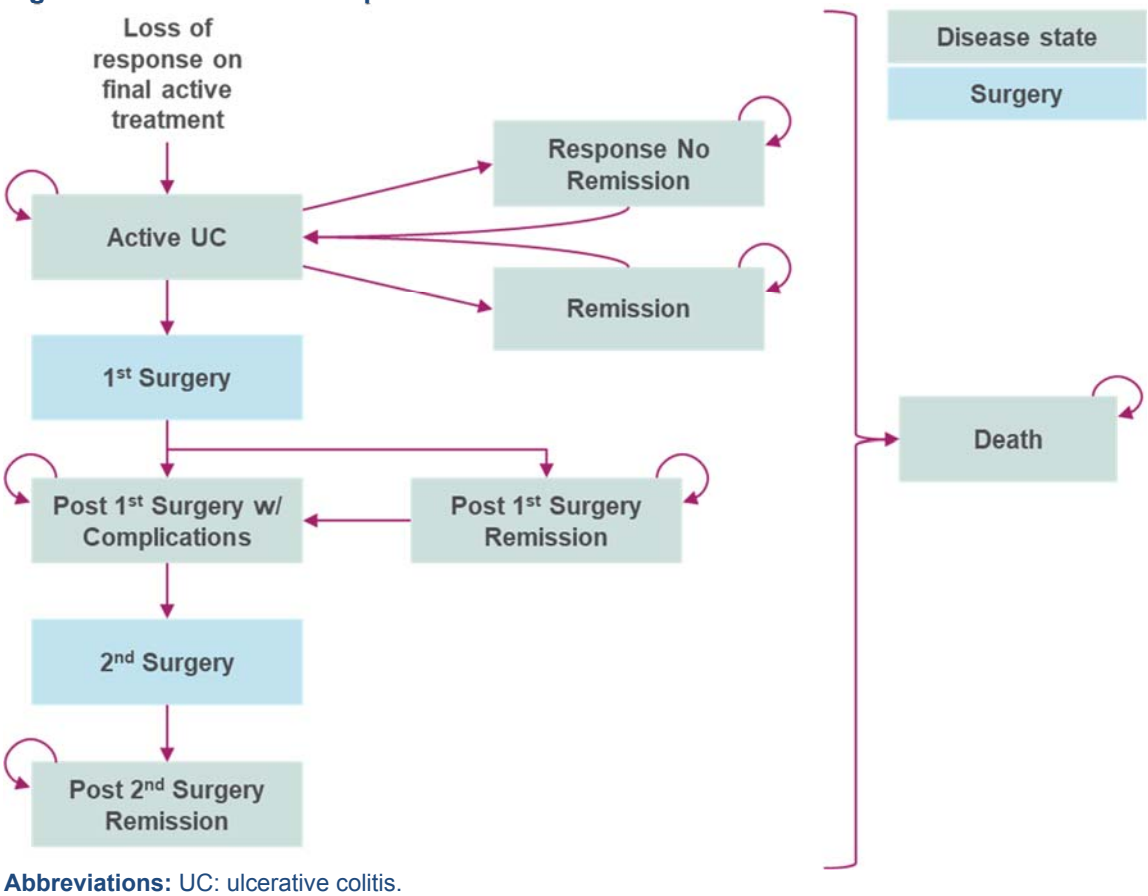
Given the availability of multiple treatment options in clinical practice, feedback from a clinical expert consulted as part of this appraisal indicated that patients in both the TNFi-naïve and TNFi-experienced subgroups may receive subsequent biologic therapies following treatment failure on initial therapy. However, due to the lack of robust efficacy data available and the uncertainty surrounding the subsequent treatments patients typically receive, subsequent treatments were not considered in the base case analysis. The exclusion of subsequent therapies in the base-case analysis is in line with prior appraisals in UC (TA633 and TA547).[1, 88] A scenario analysis was explored where patients who transitioned to 'Active UC' following treatment failure in the TNFi-naïve population could initiate a subsequent treatment in the following model cycle. Patients then transition through the induction period of this subsequent treatment in the same manner as the initial treatment. Throughout this second induction patients remained in an 'Active UC' health state. Given the lack of data to inform efficacy of biologic therapies in 3rd-line or later and feedback from clinical consultation indicating that treatment decisions after failure of two biologics are highly variable, modelling of subsequent treatments was limited to 2nd-line for the TNFi-naïve subgroup (Section B.3.3.5).

Post active treatment period

Following progression from active treatment (due to failure to achieve response, loss of response or discontinuation due to AEs) patients were assumed to transition initially to an 'Active UC' health state (shown in Figure 45) in which they received no further active treatments but could continue on best supportive care (BSC), comprising components of CvT. There was no induction period associated with this transition as patients receive components of CvT concomitantly whilst on active treatments, and therefore BSC does not represent initiation of a new treatment.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 45: Model structure post active treatment



Abbreviations: UC: ulcerative colitis.

As per the preferred assumptions of the ERG in TA633, it was possible for patients who discontinued active therapies (receiving BSC) to move from 'Active UC' to 'Remission' and 'Response No Remission' and back (relapse) as shown in Figure 45.[1] A lack of such transitions would imply that patients follow a chronic active or progressive form of disease, which would lack face validity. A proportion of the patients in the post active treatment 'Active UC' health state could move to a '1st Surgery' health state during each cycle. In line with prior submissions (TA633)[1] surgery was considered to take place over a fixed amount of time (13 model cycles or ~6 months) and therefore this '1st Surgery' health state was a series of tunnel states following which patients could move into either a 'Post 1st Surgery Remission' or 'Post 1st Surgery Complications' health state. Patients could also progress from 'Post 1st Surgery Remission' to 'Post 1st Surgery Complications'. A proportion of the patients in the 'Post 1st Surgery Complications' health state could move on to receive a '2nd Surgery' which was modelled as another tunnel health state following which patients remained in a 'Post 2nd Surgery Remission' health state until death. The progression of surgery health states was in line with previous NICE appraisals in UC (TA633 and TA342).[1, 85]

Mortality

Patients could transition to the 'Death' health state from any of the health states, including tunnel states which represent treatment induction periods or surgery.

Features of the *de novo* analysis (base case)

©Celgene Ltd, a BMS Company (2022). All rights reserved

Costs and health-related utilities were allocated to each health state and multiplied by state occupancy to calculate the weighted costs and QALYs per cycle. The following costs were considered in the model: treatment costs (acquisition and administration), costs associated with the management of AEs, surgery costs, health state related resource use costs and monitoring costs for interventions and comparators. Effectiveness measures included life years (LYs) and quality-adjusted life years (QALYs).

The analysis was conducted from the perspective of the NHS in England, including direct medical costs and PSS costs over a lifetime horizon of the patient cohort from the initiation of treatment, in line with prior economic models developed for UC submissions to NICE (TA633 and TA547) and the NICE NHS reference case.[1, 88, 180] In order to capture the variety of different treatment regimens accurately, a 2-week cycle length was deemed most appropriate. Given the short cycle length half-cycle correction was not considered necessary. An annual discount rate of 3.5% was applied to both costs and benefits, in line with the NICE reference case.[180] A summary of the main characteristics of the model are provided in Table 40.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 40: Features of the economic analysis

Factor	Previous appraisals				Current appraisal	
	TA633[1]	TA547[88]	TA342[85]	TA329[35]	Chosen values	Justification
Model structure	• Hybrid decision tree/ Markov cohort model • The Markov structure contained nine distinct health states: 'Remission', 'Response No Remission', 'Active UC', '1st Surgery', 'Post 1st Surgery Remission', 'Post 1st Surgery Complications', '2nd Surgery', 'Post 2nd Surgery Remission' and 'Death'	• Markov cohort model • The Markov structure consisted of 9 health states: 'Active UC', 'Response No Remission', and 'Remission', two post-surgery health states with and without long-term complications and an absorbing state ('Death')	• Hybrid decision tree/ Markov cohort model • The Markov structure contained 10 health states: 'Remission', 'Mild disease', 'Moderate-to-severe disease', 'Surgery', 'Post-surgery remission', 'Post-surgery complications', 'Discontinue' and 'Death'	• Hybrid decision tree/ Markov cohort model • The Markov structure contained 8 health states, 'Remission', 'Mild disease' and 'Moderate-to-severe disease', 'Surgery' and 4-post surgery health states, 'Post-surgery remission', 'Short-term transient complications', 'Long-term chronic complications' and 'Death related to surgery'	• Cohort level Markov model • The Markov structure contained nine distinct health states: 'Remission', 'Response No Remission', 'Active UC', '1st Surgery', 'Post 1st Surgery Remission', 'Post 1st Surgery Complications', '2nd Surgery', 'Post 2nd Surgery Remission' and 'Death'	• Modelling the induction period via tunnel states within the Markov cohort model accounts for the variable length of initial induction periods between treatments • Markov cohort approach for the maintenance period in line with all prior submissions and is sufficient to capture the treatment pathway for patients with UC • Health states consistent with TA633[1] and TA547[88]
Time horizon	Lifetime	Lifetime	10-year	10-year	Lifetime	• Sufficient for capturing the full treatment health effects and outcomes given the chronic nature of the disease • In alignment with the time horizons used in TA633[1]

©Celgene Ltd, a BMS Company (2022). All rights reserved

						and TA547[88] and preferred by the ERG in TA342[85] and TA329[35]
Cycle length	2-week	8-week	8-week	Adalimumab model: • 2-week Golimumab model: • 2-month	2 weeks	Sufficient to capture the variety of different treatment regimens accurately and in alignment with TA633[1] and the adalimumab model in TA329[35]
Discount rate	3.5% for both costs and benefits	3.5% for both costs and benefits	3.5% for both costs and benefits	3.5% for both costs and benefits	3.5% for both costs and benefits	In line with the NICE reference case[181]
Perspective	NHS/PSS in England	NHS/PSS in England	NHS/PSS in England	NHS/PSS in England	NHS/PSS in England	In line with the NICE reference case[181]
Treatment waning effect?	No	No	No	No	No treatment waning was applied in the base case. The model includes a scenario where a 25% treatment waning effect after 2 years is applied, in line with a scenario explored in TA547[88]	No data are available to inform estimates of a loss of response in the second and subsequent years, so treatment waning was not considered in the base case analysis. This is in line with previous appraisal.
Source of utilities	• Woehl *et al.* 2008[182] • Arseneau *et al.* 2006[183]	Woehl *et al.* 2008[182]	• GEMINI 1 (vedolizumab) • Punekar *et al.* 2010[175] • Utility decrements for AE were taken from clinical trials	• ULTRA 2 (adalimumab) • Swinburn *et al.* 2012[184] • Tsai *et al.* 2008[185]	• Woehl *et al.* 2008[182] • Arseneau *et al.* 2006[183]	Utility values collected from TRUENORTH trials were associated with several limitations including: • There are key differences between patients who did not achieve response

©Celgene Ltd, a BMS Company (2022). All rights reserved

						or remission in TRUENORTH and the modelled 'Active UC' health state. Patients in the trial continued to receive ozanimod, whereas the model assumes no further treatment would be received • The utility values were not consistent with published literature • There were no utility data available to inform the surgery model health states Therefore, literature values were used in the base case analysis. The impact of the TRUENORTH utility values on the results was explored in a scenario analysis. This approach is in line with TA633[1]
Source of costs	• 2017/2018 NHS reference cost • British National	• 2016/201 7 NHS reference costs • Electronic	• NHS list price • BNF December 2013	Published literature	• BNF • eMIT • National Schedule	Established sources of costs within the NHS. In line with the NICE reference

©Celgene Ltd, a BMS Company (2022). All rights reserved

	Formulary (BNF) • Monthly Index of Medical Specialities (MIMS) • Previous submissions • Published literature	Market Information Tool (eMIT) • MIMS • Personal Social Services Research Unit (PSSRU)			for NHS (2019/2020) • Personal Social Services Research Unit (PSSRU)	case[181] and previous appraisals in UC
Source of resource use	• Tsai *et al.* 2008[185] • Buchanan *et al.* 2011[186]	Tsai *et al.* 2008[185]	• Tsai *et al.* 2008[185] • Buchanan *et al.* 2011[186]	Tsai *et al.* 2008[185]	• Tsai *et al.* 2008[185] • Buchanan *et al.* 2011[186]	In line with previous appraisals
Measure of health effects	QALYs	QALYs	QALYs	QALYs	QALYs	In line with the NICE reference case[181]

Abbreviations: AE: adverse event; BNF: British National Formulary; eMIT: electronic Market Information Tool; ERG: evidence review group; MIMS: monthly index of medical specialties; PSSRU: Personal Social Services Research Unit; TA: technology appraisal.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.3.2.1 Intervention technology and comparators

The intervention considered in the cost-effectiveness analysis was ozanimod, as described in Section B.1.2. In alignment with the NICE final scope, the phase III TRUENORTH trial supporting this submission and the draft SmPC for ozanimod, the model included a 1-week dose escalation period (0.25 mg OD administered orally on Days 1–4, increased to 0.50 mg OD administered orally on Days 5–7), followed by 1 mg orally OD in continuous cycles of 14 days from Day 8.

The comparators varied depending on the whether the TNFi-naïve or TNFi-experienced population was selected in the model; further details of which are presented below. Treatment dosing schedules for comparators were based on the relevant SmPCs, including additional doses in the extended induction period (explored in scenario analyses) and dose escalation in the maintenance period (employed in the base-case analysis) (Table 41).

TNFi-naïve population

As mentioned in Section B.1.1, following failure with CvT patients are typically initially treated with TNFis (infliximab, adalimumab or golimumab). Whilst NICE recommendations for vedolizumab and tofacitinib do not restrict use to patients who have failed, or are intolerant to TNFis, based on clinical opinion received in TA633 and as part of this appraisal, non-TNFi biologics are commonly only considered in TNFi-naïve patients who are contraindicated or have specific safety concerns surrounding the use of TNFis.[1] It was noted that in these instances patients would receive vedolizumab, owing to safety concerns with tofacitinib limiting its use in early lines in clinical practice. The NICE recommendation for ustekinumab is restricted to patients who have failed CvT or a biologic AND who have failed a TNFi or for whom a TNFi cannot be tolerated, or is unsuitable, and therefore ustekinumab was not considered a relevant comparator in the TNFi-naïve population. [1]

Therefore, the key comparators considered in the TNFi-naïve population in the economic model are:

- Infliximab/biosimilar
- Adalimumab/biosimilar
- Golimumab
- Vedolizumab

In line with the accepted assumption in TA633, it was assumed that available TNFi biosimilars had similar efficacy to their originator counterparts.[1] Originator TNFis and biosimilars were therefore combined into single comparators.[1] It was expected that biosimilars would be significantly cheaper than their equivalent branded originator and therefore as a conservative assumption the cost of the cheapest biosimilar was used in the base case as opposed to the originator.

TNFi-experienced population

Relevant comparators for patients in the TNFi-experienced population include ustekinumab and vedolizumab. Tofacitinib was not considered a relevant comparator in the TNFi-experienced population since it is not routinely used in UK clinical practice and when used is typically reserved for later treatment lines (Section B.1.1).[1] TNFis were also not considered relevant comparators in the TNFi-experienced population as TNFi switching is no longer routine clinical

©Celgene Ltd, a BMS Company (2022). All rights reserved

practice (Section B.1.3.4). The exclusion of TNFis is in line with the accepted assumption in TA633.[1]

As a result, the key comparators for the TNFi-experienced population analysed in the economic model were:

- Ustekinumab
- Vedolizumab

Recently a subcutaneous (SC) formulation of vedolizumab was licensed for use in the maintenance period of treatment and therefore a weighted average of the two different formulations based on their use in clinical practice in England was used for vedolizumab in the maintenance period of the model.

Table 41: Recommended dose regimens for therapies

Drug	Induction		Maintenance	
	Standard dose (duration)	Extended dose (duration)	Standard dose	Escalated dose
OZA	0.25 mg daily on days 1–4; 0.5 mg daily on days 5–7; 1 mg once daily thereafter[a]	No extended induction	1 mg once daily	No escalated dose
UST	6 mg/kg IV at week 0 (8 weeks)	90 mg SC week 8 consider stopping if no evidence of benefit by week 16 (+8 weeks)	SC 90 mg every 12 weeks	May reduce to every 8 weeks, if response is lost.
VDZ	300 mg IV at weeks 0, 2 & 6 (6 weeks)	300 mg IV at week 6 discontinue if no response by week 10 (+ 4 weeks)	300 mg IV every 8 weeks	Consider 4-weekly if decrease in response
			108 mg SC every 2 weeks	No escalated dose
GOL	200 mg SC at week 0; 100 mg at week 2 (6 weeks)	Reassess if no response after 12–14 weeks (+8 weeks)	50 mg SC every 4 weeks	100 mg every 4 weeks if ≥80 kg or inadequate response
IFX[b]	5 mg/kg IV at weeks 0, 2 & 6 (8 weeks)	Discontinue if no response after 3 doses (+6 weeks)	5 mg/kg IV every 8 weeks	10 mg/kg IV every 8 weeks (Not recommended in SmPC, but common in practice)
ADA[b]	160 mg SC at week 0; 80 mg at week 2; 40 mg at weeks 4 & 6 (8 weeks)	No extended induction	40 mg SC every 2 weeks	40 mg once per week if necessary

[a] Dose referred to here, represents the dose of ozanimod hydrochloride (HCl); a 0.25 mg, 0.5 mg, and 1 mg of ozanimod HCl equivalents to 0.23 mg, 0,46, and 0.92 mg of ozanimod respectively, as described in the SmPC.
[b]Available biosimilars will be assumed to have the same regimens, effects and safety parameters
Abbreviations: ADA; adalimumab; GOL; golimumab; IFX infliximab; IV: intravenous administration; SC: subcutaneous injection; TOF: tofacitinib; UST: ustekinumab; VDZ: vedolizumab.
Source: TA633 ACD Committee Papers, page 103.[1]

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.3.3 *Clinical parameters and variables*

B.3.3.1 Summary of clinical trial data

The clinical inputs presented in the base case analysis for ozanimod for TNFi-naïve and TNFi-experienced populations were derived from the relevant populations of the TRUENORTH trial. In the absence of head-to-head evidence between ozanimod and relevant comparator therapies in these two populations, the NMA described in Section B.2.8 was used to inform the base case economic analysis.

B.3.3.2 Baseline characteristics

The baseline characteristics for the two populations in terms of age, gender distributions and weight were derived from data collected during the TRUENORTH trial and can be found in Table 42.

Table 42: Baseline characteristics for the two populations used in the economic model

Characteristic	Population	
	TNFi-naive	TNFi-experienced
Mean patient weight, n (SD), kg	[redacted]	[redacted]
Proportion of female patients	[redacted]	[redacted]
Patient age, n (SD)	[redacted]	[redacted]

Abbreviations: SD: standard deviation; TNFi: tumour necrosis factor alpha inhibitor.
Source: TRUENORTH CSR: Tables 14.1.5.1.2.1A –2A

B.3.3.3 Induction period transitions

A discussed in Section B.3.2.2, patients transition to the 'Remission', 'Response No Remission' or 'Active UC' states upon completion of the induction period. Induction periods varied in length for ozanimod and comparators, based on recommendations in the respective SmPCs

The proportion of patients achieving clinical 'Remission' and 'Response No Remission' at the end of the initial induction period was informed by the NMA (Section B.2.8). Mean absolute probabilities were derived from the following NMA outputs: baseline anchor, response effect, remission effect, and standardised mean difference (SMD) versus baseline for a given treatment in the induction period. The mean absolute probabilities used in the model are presented in Table 43.

Table 43: Clinical efficacy at the end of the induction period

Drug	Induction length (weeks)	Remission	Response no remission	No response (Active UC)
TNFi-naïve				
Ozanimod	10	[redacted]	[redacted]	[redacted]
Golimumab	6	[redacted]	[redacted]	[redacted]
Infliximab/biosimilar	8	[redacted]	[redacted]	[redacted]
Adalimumab/biosimilar	8	[redacted]	[redacted]	[redacted]
Vedolizumab[a]	6	[redacted]	[redacted]	[redacted]
TNFi-experienced				

©Celgene Ltd, a BMS Company (2022). All rights reserved

Ozanimod	10	[redacted]	[redacted]	[redacted]
Ustekinumab	8	[redacted]	[redacted]	[redacted]
Vedolizumab[a]	6	[redacted]	[redacted]	[redacted]
BSC[b]	10	[redacted]	[redacted]	[redacted]

[a]Vedolizumab SC is only licensed for maintenance treatment.
[b]Patients will not experience an induction period related to CvT within the model however the clinical efficacy after induction is used to calculate the transition probabilities in the maintenance period for patients receiving BSC alone. Given patients on BSC in the model have failed multiple prior treatments, the data for TNFi-experienced patients receiving placebo in the NMA was considered the best available data to inform these transitions.
Abbreviations: BSC: best supportive care; CvT: conventional therapy; IV, intravenous; TNFi: tumour necrosis factor alpha inhibitor; UC: ulcerative colitis.

It was not possible to perform an NMA (including ozanimod) which assessed response after extended induction, as TRUENORTH did not include an extended induction period, in line with the SmPC for ozanimod. As such, direct clinical trial data were used to inform patient distribution into the 'Remission', 'Response No Remission' and 'Active UC' health states for the scenario analysis where extended induction was selected. This approach is in line with TA633, where probabilities of response and remission at the end of extended induction were derived directly from trial data, using results for individual treatment arms. However, use of within-trial data results in 'breaking of trial randomisation'. Given the limitations of the available data to inform response and remission rates after extended induction, and that clinical consultation received as part of this appraisal noted that extended induction is not considered standard in UK clinical practice (Section B.3.2.2), extended induction was not included in the base case analysis. The efficacy data used for extended induction for the other active treatments are presented in Table 44.

Table 44: Clinical efficacy after extended induction period

Drug	Extended induction length (weeks)	Remission	Response No Remission	No response (Active UC)
TNFi-naïve				
Ozanimod	NA			
Golimumab	8	[redacted]	[redacted]	[redacted]
Infliximab/biosimilar	6	[redacted]	[redacted]	[redacted]
Adalimumab/biosimilar	NA			
Vedolizumab[a]	4	[redacted]	[redacted]	[redacted]
TNFi-experienced				
Ozanimod	NA			
Ustekinumab	8	[redacted]	[redacted]	[redacted]
Vedolizumab	4	[redacted]	[redacted]	[redacted]

[a]Vedolizumab must be administered IV for the first 2 doses and therefore can only be IV at induction.
Abbreviations: CvT: conventional therapy; IV, intravenous; TNFi: tumour necrosis factor alpha inhibitor; UC: ulcerative colitis.

B.3.3.4 Maintenance period transitions

Transition probabilities for the 'Remission', and 'Response No Remission' health states in the maintenance period were informed by the maintenance NMA. The mean absolute probabilities indicating the likelihood of sustained response or sustained remission (i.e., maintaining response

©Celgene Ltd, a BMS Company (2022). All rights reserved

or remission at the end of the maintenance period of the trial) are shown in Table 45. These probabilities were only applicable to patients who entered the maintenance period in the 'Remission' or 'Response No Remission' health states, i.e. patients who showed a response following induction. These mean absolute probabilities for both health states were derived from the following NMA outputs: baseline anchor, response effect, remission effect, and standardised mean difference (SMD) versus baseline for a given treatment in the maintenance period.

Table 45: Clinical efficacy in the maintenance period

Drug	Sustained Remission	Sustained Response
TNFi-naïve		
Ozanimod	[redacted]	[redacted]
Golimumab	[redacted]	[redacted]
Infliximab/biosimilar	[redacted]	[redacted]
Adalimumab/biosimilar	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]
Vedolizumab (IV)	[redacted]	[redacted]
Vedolizumab (SC)	[redacted]	[redacted]
TNFi-experienced		
Ozanimod	[redacted]	[redacted]
Ustekinumab	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]
Vedolizumab (IV)	[redacted]	[redacted]
Vedolizumab (SC)	[redacted]	[redacted]
BSC[a]	[redacted]	[redacted]

[a] Patients will not experience an induction period related to CvT within the model however the clinical efficacy after induction is used to calculate the transition probabilities in the maintenance period for patients receiving BSC alone. Given patients on BSC in the model have failed multiple prior treatments, the data for TNFi-experienced patients receiving placebo in the NMA was considered the best available data to inform these transitions.

Abbreviations: CvT: conventional therapy; IV, intravenous; TNFi: tumour necrosis factor alpha inhibitor; UC: ulcerative colitis.

The absolute mean probabilities derived from the NMA outcomes (presented in Table 45) were translated to per cycle transition probabilities for 'loss of response' and 'loss of response no remission', presented in Table 46. The calculations used for these transition probabilities are presented in Appendix L. These transition probabilities did not include the probability of discontinuation due AEs and represented discontinuation due to loss of response only. Before application in the engine of the model the transition probabilities in Table 45 were adjusted to account for a proportion of the patients discontinuing due to AEs (see Section B.3.3.6). The transition probabilities for 'Retained Response' and 'Retained Response No Remission' were simply calculated as the complement of the below transition probabilities for 'Loss of Response' and 'Loss of Response No Remission'.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 46: Transition probabilities for loss of response in the maintenance period

Drug	Duration of maintenance period	Loss of Response	Loss of Response No Remission
TNFi-naïve			
Ozanimod	42	[redacted]	[redacted]
Golimumab	54	[redacted]	[redacted]
Infliximab/biosimilar	46	[redacted]	[redacted]
Adalimumab/biosimilar	44	[redacted]	[redacted]
Vedolizumab	46	[redacted]	[redacted]
TNFi-experienced			
Ozanimod	42	[redacted]	[redacted]
Ustekinumab	44	[redacted]	[redacted]
Vedolizumab	46	[redacted]	[redacted]
BSC	42	[redacted]	[redacted]

Abbreviations: CvT, conventional therapy; IV, intravenous; TNFi: tumour necrosis factor alpha inhibitor; UC: ulcerative colitis.

In each model cycle, the number of patients losing response was calculated by applying the transition probability for loss of response (inc. AE discontinuation) to patients in the 'Remission' and 'Response No Remission' health states in the previous cycle. These patients joined the 'Active UC' health state. The number of overall responders was then reduced by the number of patients losing response or discontinuing due to AEs and transitioning to 'Active UC' each cycle.

The number of patients continuing in the 'Response No Remission' health state from the previous cycle was calculated by multiplying [the patients in the 'Response No Remission' health state in the previous cycle] by [1 - the transition probability of loss of 'Response No Remission' (including AE discontinuation)]. The number of patients continuing in the 'Remission' health state each cycle was given by the [remaining overall responders] - [the remaining patients in the 'Response No Remission' health state]. This allowed for patients to implicitly transfer between the 'Remission' and 'Response No Remission' health states in the maintenance period.

The approach assumed a constant loss of response within and beyond the trial duration of one-year, an assumption which is in line with previous submissions.[1] There was no publicly available data to validate the estimates of a loss of response in the second and subsequent years.

In line with a scenario explored in TA547, the model included a scenario to introduce a 25% treatment waning effect after 2 years to the active treatment (Section B.3.8.3).[88] This factor was applied to the maintenance transition probabilities for retained response and retained response no remission as a percentage reduction, e.g. an expected probability of 90% for retained response would be reduced to 67.5% after 2 years.

B.3.3.5 Subsequent treatments

As discussed in Section B.3.2.2, given the availability of multiple treatment options in clinical practice, scenarios were explored where patients who transitioned to 'Active UC' following treatment failure could initiate a subsequent treatment. The subsequent treatments explored for

©Celgene Ltd, a BMS Company (2022). All rights reserved

ozanimod and relevant comparators were informed by clinical expert opinion and represent the most common subsequent treatments received by patients in clinical practice.

Scenarios were explored in the TNFi-naïve population where both vedolizumab and ustekinumab were modelled as subsequent treatments following receipt of either ozanimod or TNFis first line; this approach is similar to that taken in TA633.[1] Ustekinumab was explored as a subsequent treatment for the small subset of TNFi-naïve patients who receive vedolizumab first line.

No efficacy data were available to inform efficacy of biologic therapies specifically in the third-line or later, and clinical consultation indicated that treatment decisions after failure on multiple biologics are patient-dependent and highly variable. Therefore, no subsequent treatments could be modelled for the TNFi-experienced population and subsequent treatments could only be modelled for second-line treatments in the TNFi-naïve population. Efficacy of subsequent treatments for the TNFi-naïve population was informed by data for the TNFi-experienced subgroup of the NMA without adjustment.

B.3.3.6 Treatment discontinuation due adverse events

Discontinuation from AEs was captured in this model to account for anticipated discontinuation in clinical practice, which other submissions were criticised for omitting.[1, 88] Treatment discontinuation due to AEs (Table 47) was applied in addition to patients discontinuing treatment due to loss of response. Patients who discontinued treatment due to AEs were considered to transfer to the 'Active UC' health state at the end of the induction period. The proportion of patients who discontinued each model cycle of the maintenance period were derived from the trials included in the NMA and are presented in Table 47. Reported rates were converted to probabilities over the respective trials' maintenance length.

Table 47: Treatment discontinuation after induction period

Drug	Probability of discontinuation due to AEs over the course of the induction period[a]	Per cycle probability of discontinuation due to AEs in the maintenance period[a]
TNFi-naïve		
Ozanimod	[redacted]	[redacted]
Golimumab	0.30%	0.27%
Infliximab/biosimilar	2.78%	0.33%
Adalimumab/biosimilar	4.93%	0.26%
Vedolizumab[b]	2.96%	0.14%
Vedolizumab (IV)	2.96%	0.13%
Vedolizumab (SC)	N/A	0.14%
TNFi-experienced		
Ozanimod	[redacted]	[redacted]
Ustekinumab	0.00%	0.16%
Vedolizumab[b]	2.96%	0.14%
Vedolizumab (IV)	2.96%	0.13%
Vedolizumab (SC)	N/A	0.14%
BSC	NA	0.00%

Probabilities were derived from the ITT population (i.e. both TNFi-naïve and TNFi-experienced patients) due to lack of subgroup-specific data. [b]Vedolizumab SC is only licensed for maintenance treatment.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Abbreviations: AE: adverse event: CvT: conventional therapy; IV, intravenous; TNFi: tumour necrosis factor alpha inhibitor; UC: ulcerative colitis.

B.3.3.7 Best supportive care (BSC)

As discussed in Section B.3.2.2, following progression from active treatment (due to failure to achieve response, loss of response or discontinuation due to AEs) patients transitioned to an 'Active UC' health state in which they received no further active treatment. Some patients may continue to receive components of CvT (which were being prescribed concomitantly alongside active treatment) as BSC. As BSC does not represent initiation of a new therapy, no explicit induction period was modelled for patients entering this state. As per the preferred assumptions in TA633, patients who discontinued all treatment were able to experience 'Remission' or 'Response No Remission', since UC (a relapse-remitting disease) can worsen or improve with the underlying disease course.[1]

It was assumed that all placebo arms from the RCTs included in NMA were representative of BSC and therefore these arms were combined to provide an estimate of patients sustaining response or sustaining remission (Table 45). These data were used to inform transition probabilities for 'Loss of Response' and 'Lose of Response No Remission' as described previously.

As noted, patients could experience spontaneous 'Remission' or 'Response No Remission' from the 'Active UC' health state post active treatment. No trial data was available to inform this transition. A scenario presented by the ERG in TA633 assumed that 1% of patients would enter spontaneous 'Remission' from the 'Active UC' health state and 1% of patients would enter spontaneous 'Response No Remission' from the 'Active UC' health state at each model cycle. The committee considered the ERG's assumption to be an overestimate but agreed that there may be a small number of people who would improve without treatment.[1] Therefore, in the base case, 0.5% of patients were modelled to enter spontaneous 'Remission' from the 'Active UC' or 'Response No Remission' from the 'Active UC' health state at each model cycle (0.25% entering each health state, respectively). Scenario analyses were conducted which assumed 0% and 1% of patients experienced spontaneous remission/response from the 'Active UC' health state (Section B.3.8.3).

B.3.3.8 Surgeries

A proportion of patients in the post active treatment 'Active UC' health state could receive surgery in each model cycle. This was derived from Misra *et al.* (2016) which reported annual probability of patients receiving surgery, which were then converted to a probability per model cycle. As Misra *et al.* 2016 did not present data for patients receiving 2nd surgery it was assumed the transition probability of 2nd surgery from 'Post 1st Surgery Complications' health state was equal to the probability of 1st surgery, in line with the accepted assumption made in TA633.[1] Post-surgery transitions were aligned with recent NICE submissions (TA633 and TA547) and complemented by targeted review to address any data gaps.[1, 88] These probabilities were converted to the 2-week cycle probabilities as presented in Table 49. As a simplifying assumption, surgery-related model inputs were assumed to be the same in both the TNFi-naïve and TNFi-experienced populations.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 48: Surgery related probabilities

Input	Value
Annual probability of first surgery[a]	0.0047[187]
Proportion of patients with complications after surgery[b]	33.5%[188]
Annual probability of complications from post-surgery remission	0.0325[189]
Annual probability of second surgery[c]	0.0047[187]

[a] Misra *et al. 2016* report 5,044 colectomies over a 15-year period among a total of 71,966 patients. It is assumed that these surgeries happen at a constant rate and therefore one would expect 336 colectomies in a single year. Assuming a constant population of 71,966 patients this equates to an annual probability surgery = 336/71,966 = 0.0047.
[b] Weighted average of 32% of patients with elective surgery and 35% of patients with non-elective surgery.
[c] Assumed equal to first surgery.

Table 49: Surgery related transition probabilities

From health state	To health state	Per cycle probability
Active UC (post active treatments)	1st Surgery	0.00018
1st Surgery	1st Surgery Complications	0.33500
1st Surgery	1st Surgery Remission	0.66500
1st Surgery Remission	1st Surgery Remission	0.99876
1st Surgery Remission	1st Surgery Complications	0.00124
1st Surgery Complications	2nd Surgery	0.00018

Abbreviations: UC: ulcerative colitis

B.3.3.9 Adverse events

AEs associated with UC treatments during the induction and the maintenance period for each population of interest were derived from the pivotal RCTs for biologics and the TRUENORTH trial for ozanimod. In alignment with previous submissions, serious infections were the only AE included in the model due to their high costs.[1, 88] The AEs incidence reported across the relevant clinical trials were converted to 2-week probabilities (Table 50). In line with TA633, patients were assumed to be at constant risk of experiencing AEs over the model time horizon.[1]

Table 50: Per cycle probability of serious infections

Treatment	TNFi-naïve		TNFi-experienced	
	Induction period	Maintenance period[a]	Induction period	Maintenance period[a]
Ozanimod	[redacted]	[redacted]	[redacted]	[redacted]
Ustekinumab	N/A	N/A	0.08%	0.10%
Vedolizumab[a]	0.09%	0.10%	1.07%	0.10%
Vedolizumab IV	N/A	0.05%	N/A	0.15%
Vedolizumab SC	N/A	0.06%	N/A	0.06%[a]
Golimumab	0.20%	0.12%	N/A	0.12%
Infliximab/biosimilar	0.28%	0.17%	0.28%	0.17%
Adalimumab/biosimilar	0.24%	0.06%	0.31%	0.06%
BSC	N/A	0.12%	NA	0.12%

©Celgene Ltd, a BMS Company (2022). All rights reserved

[a]Vedolizumab at maintenance calculated as the weighted average of vedolizumab IV and SC based on relative use in UK clinical practice which is assumed to be 50:50 split based on feedback received from clinical consultation.
Abbreviations: BSC: best supportive care; IV: intravenous; SC: subcutaneous; N/A: not applicable; UC: ulcerative colitis

B.3.3.10 Mortality

The model allowed patients to transition to the 'Death' health state from any other health state at any time, based on age- and gender-adjusted background (i.e., all-cause) mortality using the life tables published by the UK Office for National Statistics.[190] The model included the flexibility to apply health state-specific mortality risk.

The standard mortality ratios (SMRs) used in the base case are presented in Table 51. These represent the increased risk of mortality associated with a particular health state. In the base case, patients in the 1st Surgery health state were assumed to have an increased risk of mortality of 30%, based off values reported in Jess *et al*. 2007 and in line with TA633 and TA547.[1, 88, 191] Clinical feedback received as part of this appraisals noted that UC does not significantly impact mortality and therefore no increase in mortality was assumed for the remaining health states (SMR equal to 1).

Table 51: Standardise mortality ratios for each health state

From health state	Standardised mortality ratio
Remission	1.0
Response No Remission	1.0
Active UC	1.0
1st Surgery	1.3[191]
1st Surgery Remission	1.0
1st Surgery Complications	1.0
2nd Surgery	1.3[191]
2nd Surgery Remission	1.0

Abbreviations: UC: ulcerative colitis

B.3.4 *Measurement and valuation of health effects*

As mentioned in Section B.3.2.2 the model includes nine mutually exclusive health states; 'Remission', 'Response No Remission', 'Active UC', '1st Surgery', 'Post 1st Surgery Remission', 'Post 1st Surgery Complications', '2nd Surgery', 'Post 2nd Surgery Remission', and 'Death'. Each distinct health state had a specific utility value, which was considered independent of the treatment received.

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

The TRUENORTH trial assessed HRQoL via the EQ-5D health utilities instrument.[192] Health state utility data were collected at Baseline and at the end of the induction and maintenance periods (Week 10 and Week 52, respectively). Utility data are presented in Table 52 for all patients at Baseline, as well as for patients who achieved 'Remission' and 'Response no remission' by the end of the induction and maintenance period based on the 4-component Mayo score, in line with the clinical data informing the NMA (Table 53 and Table 53 for induction and maintenance, respectively).

©Celgene Ltd, a BMS Company (2022). All rights reserved

The EQ-5D-5L data collected during the TRUENORTH trial were cross-walked to EQ-5D-3L index scores using the algorithm presented in van Hout *et al*. 2012 (based on the UK value set by Dolan *et al*. 1997)[193] and the weighted average across treatment arms was used to inform each health state (Table 52).[194] Utility values for each health state were assumed to be independent of treatment arm; i.e. only dependent on health state. This appears reasonable given the similarities in utility values observed across the ozanimod and placebo arms. Surgeries were not included in the TRUENORTH trial and therefore no HRQoL data were available to inform the surgery health states. Utility values for the surgery health states in the model were obtained from the literature (Section B.3.4.4).

Table 52: Utility data from the TRUENORTH trial (Induction)

Health state		Cohort 1		Cohort 2	Weighted average
		Ozanimod 1 mg	Placebo	Ozanimod 1 mg	
Baseline (Active UC)	n	[redacted]	[redacted]	[redacted]	[redacted]
	EQ-5D summary index score (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Remission at Week 10	n	[redacted]	[redacted]	[redacted]	[redacted]
	EQ-5D summary index score (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Response no remission at Week 10	n	[redacted]	[redacted]	[redacted]	[redacted]
	EQ-5D summary index score (SD)	[redacted]	[redacted]	[redacted]	[redacted]
No response or remission at Week 10 (Active UC)	n	[redacted]	[redacted]	[redacted]	[redacted]
	EQ-5D summary index score (SD)	[redacted]	[redacted]	[redacted]	[redacted]

Response and remission based off 4-component Mayo definition.
Abbreviations: UC: ulcerative colitis.

Table 53: Utility data from the TRUENORTH trial (Maintenance)

Health state		Placebo (N=55)	Re-randomised Patients		Weighted average
			Ozanimod – Placebo	Ozanimod– Ozanimod	
Remission at Week 52	n	[redacted]	[redacted]	[redacted]	[redacted]
	EQ-5D summary index score	[redacted]	[redacted]	[redacted]	[redacted]
Response no remission at Week 52	n	[redacted]	[redacted]	[redacted]	[redacted]
	EQ-5D summary index score	[redacted]	[redacted]	[redacted]	[redacted]
No response or remission at Week 52 (Active UC)	n	[redacted]	[redacted]	[redacted]	[redacted]
	EQ-5D summary index score	[redacted]	[redacted]	[redacted]	[redacted]

Response and remission based off 4-component Mayo definition.
Abbreviations: UC: ulcerative colitis.

The weighted average utility values for patients with remission and response no remission were similar between Week 10 and Week 52. The Week 52 values were marginally higher, indicating a

©Celgene Ltd, a BMS Company (2022). All rights reserved

potential improvement in utility for patients sustaining remission and response no remission. These utility values were largely in agreement with the published literature (Section B.3.4.2).

The weighted average utility values at baseline were similar to the values for patients with no response or remission at Week 10, but were typically higher than the results from the published literature. The utility data collected at Baseline may represent a reasonable proxy for 'Active UC', as to be eligible for the trial patients were required to have active UC, defined as a Mayo score of between 6–12 points inclusive as well as evidence of UC extending ≥15 cm from the anal verge as determined by Baseline endoscopy. However, this trial health state differs from the modelled health state, where patients have failed active treatment. Similarly, patients with no response or remission at Week 10 during TRUENORTH were receiving ozanimod whilst in this active UC state, whereas the modelled health state assumes no further treatment would be received.

The utility values for patients with no response or remission at Week 52 were found to be considerably higher than the values at Baseline, for patients with no response or remission at Week 10, and the published literature. The TRUENORTH trial only permitted patients who had achieved at least a response in the induction period to continue into the maintenance period for assessment, and therefore this utility value is derived from a select subgroup of patients who had previously achieved remission or response but subsequently lost this response by Week 52. In addition, these patients continued to receive ozanimod whilst in this active UC state, whereas the modelled health state assumes no further treatment would be received. These patients are not reflective of the modelled Active UC state as a whole and may have less severe disease following loss of response compared with those patients who were not able to achieve a response by Week 10. Further to this, this utility value is based on smaller patient numbers and hence is subject to greater uncertainty.

Additionally, the length of the trial follow-up was not considered to be long enough to assess the change in utility over time and no trial data could be used to inform the surgical health states. Finally, assumptions were required to classify the health states that each EQ-5D value corresponded to for patients with missing response and remission data for EQ-5D time points and required the use of partial Mayo scores.

Given these limitations, the utility values from the TRUENORTH trial were only included as a scenario in the model and published data were used to inform the base-case in line with prior submissions (TA633, TA547 and TA342).[1, 85, 88]

B.3.4.2 Health-related quality-of-life studies

An SLR was conducted to identify relevant HRQoL data in patients with moderately to severely active UC. Searches were conducted on 2nd September 2021 and updated on 14th October 2021. The SLR was conducted following current best practices, as recommended by the Cochrane Collaboration.[172] The reporting of the methods and results of the SLR was done in line with the guidance provided by NICE[173] and following the PRISMA guidelines.[167, 174] Full details of the SLR search strategy, study selection progress and results are reported in Appendix F.

In total 27 unique studies were identified that reported on HRQoL data in patients with moderately to severely active UC and 9 HTA submissions. Of these, the utility values used in the NICE TA633 for Ustekinumab (2020) were considered most relevant to inform HRQoL inputs in the model as they provided utility values that matched the health states in the model and they

©Celgene Ltd, a BMS Company (2022). All rights reserved

had been used in prior appraisals in UC.[1, 88] The utility values from TA633 are provided in Table 54 below.

Full results for all the identified studies are presented in Appendix F.

Table 54: Utility inputs from NICE TA633 (2020) used to inform the base-case analysis

State	Utility value: mean	Source
Remission	0.87	Woehl *et al.* 2008
Response no remission	0.76	Woehl *et al.* 2008
Active UC	0.41	Woehl *et al.* 2008
1st Surgery	0.61	Arseneau *et al.* 2006
Post 1st surgery remission	0.72	Woehl *et al.* 2008
Post 1st surgery complications	0.34	Arseneau *et al.* 2006

Abbreviations: UC: ulcerative colitis; SE: standard error.
Source: NICE TA633 (2020).[1]

B.3.4.3 Adverse reactions

Decrements in utility for AEs associated with treatment with ozanimod or relevant comparators were captured in the model via the application of disutility values and estimated AE duration, where necessary. As mentioned in Section B.3.4.3, in line with prior appraisals it was assumed that adverse events occurred uniformly throughout the induction period and uniformly throughout the maintenance period. As noted in Table 50 the likelihood of serious infections was different in the induction period compared to the maintenance period. In alignment with the most recent NICE submissions (TA547 and TA633) the adverse event disutility for serious infections (0.156/year) was calculated based on data from Stevenson *et al.* 2016 and was applied for a 4-week period to patients experiencing serious infeciton.[1, 88, 195]

B.3.4.4 Health-related quality-of-life data used in the cost-effectiveness analysis

As discussed in Section B.3.4.1, there are several limitations in the utility values derived from the TRUENORTH trial, and as such these values may not accurately reflect the 'Active UC' state in the model. The TRUENORTH trial only permitted patients who had achieved at least a response in the induction period to continue into the maintenance period for assessment, and therefore the utility values for response and remission were derived from a select subgroup of patients who had previously achieved remission or response but subsequently lost this response by Week 52 (Section B.3.3.6). Accordingly, the utility values for the Active UC state obtained from TRUENORTH were not consistent with (higher than) those in the published literature. Finally, no surgeries were recorded in the trial and therefore utility values for the surgery-related health states could not be obtained from TRUENORTH.

Due to the above limitations in the utility values derived from TRUENORTH, they were not used to inform the base-case model analysis. Instead, literature-derived utility values from Woehl *et al.* 2008[182] and Arseneau *et al*. 2006[183] were used in alignment with values used in the most recent NICE submission in the same indication (TA633).[1] A scenario analyses was conducted exploring the use of the TRUENORTH-derived utility values (Section B.3.8.3).

©Celgene Ltd, a BMS Company (2022). All rights reserved

A summary of the various utility values used in the base-case analysis are provided in Table 55. The utility data collected at Week 52 were used to inform the 'Remission' and 'Response no remission' states. The utility data collected at Baseline was used to inform the Active UC state, since this may represent a reasonable proxy for 'Active UC' (see Section B.3.4.1).

Table 55: Summary of utility values for base case cost-effectiveness analysis

State	Utility value: mean	Reference in submission (section and page number)	Justification
Remission	0.87	Section B.3.4.2	Aligned with values used in TA633[1]
Response no remission	0.76		
Active UC	0.41		
1st Surgery	0.61		
Post 1st surgery remission	0.72		
Post 1st surgery complications	0.34		
2nd surgery	0.61	Assumed equal to first surgery	Aligned with TA633 assumption[1]
Post 2nd surgery remission	0.72		
Serious infection	-0.156	B.3.4.3	Aligned with value used in TA633[1]

Abbreviations: HS: health state; AR: adverse reaction
Sources: Woehl *et al.* 2008.[182] Arseneau *et al*. 2006.[183]

Table 56: Summary of utility values for TRUENORTH scenario analysis

State	Utility value: mean	Reference in submission (section and page number)	Source/ justification
Remission	0.90	Section B.3.4.1	TRUENORTH[5]
Response no remission	0.84		
Active UC	0.68		
1st Surgery	0.61	Section B.3.4.2	Aligned with values used in TA633[1]
Post 1st surgery remission	0.72		
Post 1st surgery complications	0.34		
2nd surgery	0.61	Assumed equal to first surgery	Aligned with TA633 assumption[1]
Post 2nd surgery remission	0.72		
Serious infection	-0.156	B.3.4.3	Aligned with value used in TA633[1]

Abbreviations: HS: health state; AR: adverse reaction
Sources: TRUENOTH CSR.[5]

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

An SLR was conducted to identify relevant cost or resource use studies for incorporation in the model. The searches were run on the 2nd September 2020 and updated on 14th October 2021.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Full details of the SLR search strategy, study selection process and results are presented in Appendix G .

In total, 133 unique studies reporting on cost or healthcare resource use data and 9 HTA submissions in moderately to severely active UC were identified. Eight of the 133 included studies and 4 of the HTA submissions were conducted in the UK and presented data specifically for a UK patient population.

The following cost categories were included in the model:

- Drug acquisition costs
- Administration costs
- Treatment-specific monitoring
- Health state related resource use
- AE management costs
- Surgery costs

The economic analysis was conducted from an NHS and PSS perspective and therefore included only costs that would be incurred by the NHS and PSS. Cost inputs were based on British National Formulary (BNF), electronic market information tool (eMIT), National Schedule for NHS (2019/2020), and Personal Social Services Research Unit (PSSRU).

B.3.5.1 Intervention and comparators' costs and resource use

Acquisition costs

Drug acquisition costs for treatment regimens were calculated based on cost per pack and dosing regimen reported in the respective prescribing information and/or clinical trials. Dose escalation was accounted for in the base-case model. In line with TA633 and RWE collected on dose escalation it was assumed that 30% of patients would be receiving an escalated dose in the maintenance period at any one time (Section B.1.3.5). This proportion was validated by clinical consultation conducted as part of this appraisal. A corresponding higher drug acquisition cost was applied to all patients receiving dose escalation.[1] In line with feedback received from clinical consultation and the results of an SLR collecting RWE on dose escalation, infliximab was also modelled to experience dose escalation, despite being off-label.[74] This was in line with the accepted assumption in TA633.[1] Scenario analyses were run in which dose escalation was assumed to be 0% and 50% in the maintenance period to assess uncertainties surrounding dose escalation in the model.

Unit cost per pack were derived from publicly available databases (BNF/eMIT) in the UK. If more than one formulation of treatment with similar strength was available, the conservative assumption of selecting the lowest priced treatment was used to provide the cost. For treatments with multiple formulations with different strengths, the least expensive formulation that aligns with the prescribed dosing was selected. The drug pack details and acquisition costs are presented in Table 57 for ozanimod (with PAS).

©Celgene Ltd, a BMS Company (2022). All rights reserved

Vedolizumab can be administered either IV or SC during the maintenance period.[196] In line with feedback received from clinical experts it was assumed that the proportion of patients receiving vedolizumab SC and IV during maintenance was evenly distributed.

Table 57: Drug pack acquisition costs

Drug	Description	Pack unit size (quantity)	Pack unit strength (mg)	Pack cost
Active Treatment				
Ozanimod (with PAS)	7-day oral tablet starter packer	4	0.25	[redacted]
		3	0.5	
	Oral tablet pack	28	1.0	[redacted]
Ustekinumab	Vial for IV	1	130	£2,147.00
	Prefilled pen for SC	1	90	£2,147.00
Vedolizumab	Vial for IV	1	300	£2,050.00
	Prefilled pen for SC	1	108	£512.50
Golimumab	Prefilled pen for SC	1	50	£762.97
Infliximab/biosimilar[a]	Vial for IV	1	100	£377.00
Adalimumab/biosimilar[a]	Prefilled pen for SC	1	40	£633.60
Concomitant therapies				
Balsalazide	Oral tablet pack	130	750	£30.42
Mesalazine	Oral tablet pack	120	400	£22.10
Olsalazine	Oral tablet pack	60	500	£161.00
Sulfasalazine	Oral tablet pack	112	500	£6.74
Prednisolone	Oral tablet pack	28	20	£2.93
Hydrocortisone	Oral tablet pack	30	20	£3.29
Azathioprine	Oral tablet pack	56	50	£1.46
6 mercaptopurine	Oral tablet pack	25	50	£34.39
Methotrexate	Oral tablet pack	100	2.5	£4.32
Budesonide	Oral tablet pack	50	3	£37.53

[a]Only cheapest shown, Flixabi for infliximab and Amegevita for adalimumab
Abbreviations: IV, intravenous; SC: subcutaneous.

The recommended dosing of drugs was based on EMA-approved prescribing information (e.g. fixed-dose, based on mean patient-weight, or based on average body surface area) as given by the relevant SmPCs. Drug wastage was not considered in the base case analysis, but a scenario was conducted where vial sharing was included. It was assumed that patients receiving oral therapies used up each pack before acquiring a new one. The dosing schedules for the induction period, extended induction period (explored in a scenario analysis in Section B.3.8.3) and maintenance period are presented in Table 58, Table 59 and Table 60 respectively.

Patients were modelled to incur the full cost the of induction dosing as a one-off cost at the initiation of treatment. Likewise, patients entering extended induction in the scenario analysis also incurred the one-off cost of extended induction upon initiation of treatment. Where treatment dosing was based on weight the average number of vials required per patient was based on a

©Celgene Ltd, a BMS Company (2022). All rights reserved

distribution of patient weights; patient weight was assumed to have a normal distribution based on the average baseline patient weight and standard deviation.

Table 58: Induction period dosing schedule

Drug	Description	Administration	Dose (mg)		Doses during induction
			TNFi-naive	TNFi-experienced	
Ozanimod	0.25 mg[a] OD from days 1–4; 0.5 mg[a] OD on days 5–7; 1 mg[a] OD on day 8 till Week 10	Oral	0.25[a]	0.25[a]	3
			0.5[a]	0.5[a]	4
			1[a]	1[a]	63
Ustekinumab	6 mg/kg at Week 0	IV	403	410	1
Vedolizumab	300 mg at Weeks 0 and 2	IV	300	300	2
Golimumab	200 mg at Week 0	SC	100	100	1
	100 mg at Week 2		200	200	1
Infliximab/ biosimilar	5 mg/kg at weeks 0, 2 and 6	IV	417	424	3
Adalimumab/ biosimilar	160 mg at Week 0	SC	160	160	1
	80 mg at Week 2		80	80	1
	40 mg at Weeks 4 and 6		40	40	2

[a]Dosing based on dose of ozanimod hydrochloride (HCl). 0.25 mg, 0.50 mg and 1.0 mg of ozanimod HCl equivalates to 0.23 mg, 0.46 mg and 0.92 mg of ozanimod, respectively.
Abbreviations: IV, intravenous; mg, milligrams; OD, once daily; SC: subcutaneous.

Table 59: Extended induction period dosing schedule

Drug	Description	Administration	Dose (mg)	Doses during extended induction
Ozanimod	NA			
Ustekinumab	90 mg SC at week 8	SC	90	1
Vedolizumab	300 mg at week 6	IV	300	1
Golimumab	100 mg at week 6	SC	100	1
Infliximab/biosimilar	No further treatment administered in extended induction			
Adalimumab/biosimilar	NA			

Abbreviations: IV, intravenous; mg, milligrams; OD, once daily; SC: subcutaneous.

Table 60: Maintenance dosing schedule

Drug	Description	Administration	Dose (mg)		Doses per cycle
			TNFi-naive	TNFi-naive	
Ozanimod	1 mg once daily	Oral	1.0[a]	1.0[a]	14
Ustekinumab std.	90 mg once per 12 weeks	IV	90	90	0.16
Ustekinumab esc.	90 mg once per 8 weeks		90	90	0.26

©Celgene Ltd, a BMS Company (2022). All rights reserved

Drug	Description	Administration	Dose (mg)		Doses per cycle
			TNFi-naive	TNFi-naive	
Vedolizumab (IV) std.	300 mg once per 8 weeks	IV	300	300	0.26
Vedolizumab (IV) esc.	300 mg once per 4 weeks		300	300	0.50
Golimumab std.	50 mg once per 4 weeks	SC	50	50	0.50
Golimumab esc.	100 mg once per 4 weeks		100	100	0.50
Infliximab/biosimilar std.	5 mg/kg once per 8 weeks	IV	417	424	0.26
Infliximab/biosimilar esc.	10 mg/kg once per 8 weeks		783	798	0.26
Adalimumab/biosimilar std.	40 mg once per 2 weeks	SC	40	40	1.00
Adalimumab/biosimilar esc.	40 mg once per week		40	40	2.00

[a]Dosing based on dose of ozanimod hydrochloride (HCl). 0.25 mg, 0.50 mg and 1.0 mg of ozanimod HCl equivalates to 0.23 mg, 0.46 mg and 0.92 mg of ozanimod, respectively.
Abbreviations: esc: escalated; IV: intravenous; mg: milligrams; std: standard.

Concomitant therapy (CcT)

For all treatments, CcT costs were included in the acquisition costs alongside the cost of the active treatment as per the committee preference in TA633.[1] The pack acquisition costs and details of CcT are provided in Table 57. The model considered CcT usage to remain constant throughout the model at the concomitant therapy usage level, including for the treatment of flares for patients on BSC. The patient usage of CcT for all active treatments was based off the assumptions in prior technology appraisals in UC (TA342 and TA633 [Table 61]) and was validated by clinical consultation. The CcT distributions for patients receiving ozanimod did not include certain CcT owing to contraindications specified in the SmPC. The weighted cost of CcT for each active treatment per cycle is provided in Table 61.

The cost of CcT applied during active treatment is subsequently applied as the BSC cost after patients have progressed from active treatments, assuming no contraindications.

Table 61:CcT dosage and anticipated concomitant usage

Drug	Dose description	Patient usage for ozanimod	Patient usage for all other active treatments (TA342/TA633)[85]
Balsalazide	1.5 g twice daily	0.0%	0.0%
Mesalazine	1.2 g/day (divided doses)	13.0%	13.0%
Olsalazine	500 mg twice daily	0.0%	0.0%
Sulfasalazine	500 mg 4 times daily	0.0%	0.0%
Prednisolone	20.0 mg/day for two weeks	36.0%	36.0%
Hydrocortisone	20 mg/day	0.0%	0.0%
Azathioprine[a,b]	2.5 mg/kg/day	0.0%	39.0%

©Celgene Ltd, a BMS Company (2022). All rights reserved

Drug	Dose description	Patient usage for ozanimod	Patient usage for all other active treatments (TA342/TA633)[85]
6-mercaptopurine[a]	1.5 mg/kg/day	0.0%	15.0%
Methotrexate[a]	17.5 mg/week	0.0%	9.0%
Budesonide	3.0 mg/3xday for eight weeks	1.0%	1.0%
Cost per average patient per cycle	N/A	£1.85	£11.14

[a]Patients receiving ozanimod are contraindicated to azathioprine, 6 mercaptopurine and methotrexate and would therefore not receive these concomitantly. Patients receiving tofacitinib are contraindicated to azathioprine and would therefore not receive it concomitantly.
Abbreviations: IV, intravenous; SC: subcutaneous.

Administration costs

The unit costs of administration were derived from the National schedule of NHS costs and are presented in Table 62.[197] Based on the prescribing information and billing guidelines published by manufacturers of each treatment, any resources or monitoring associated with each administration were identified and included under treatment-specific monitoring costs.

Table 62: Treatment administration costs

Administration method	Cost per dose or per pack	Sources
Intravenous injection	£186.37	Average of Consultant led and non-consultant led non-admitted face-to-face attendance, follow-up, WF01A 300 (general Medicine NHS National schedule of NHS costs 2019/2020)[197]
Subcutaneous injection	£0.00	Assumed that for subcutaneous injections most patients self-inject their medication and as such there is no associated administration cost. It has also been assumed that the one-off nurse training cost to teach patients how to self-administer the injection is covered by the manufacturer. This assumption is in line with clinical feedback provided to the ERG TA633[1]
Oral tablet	£0.00	Assumed zero administration costs

Abbreviations: NHS: National health service.

Treatment-specific monitoring costs

Clinical consultation conducted as part of this appraisal confirmed that the monitoring requirements were similar for ozanimod and existing treatments for moderately to severely active UC in UK clinical practice. Therefore, monitoring costs were not considered in the economic model, in line with previous appraisals (TA633, TA547 and TA342)[1, 85, 88] with the exception of a single electrocardiogram (ECG) for ozanimod during induction; this is a specific requirement for ozanimod as specified in the SmPC. The ECG cost was obtained from the National Schedule of NHS costs 2019/2020 and is provided in Table 63.

Table 63: Monitoring test cost per test

Test	Unit cost	Source
ECG	£61.80	National Schedule of NHS Costs 2019/20, EY51Z, Directly Accessed Diagnostic Services, Electrocardiogram Monitoring or Stress Testing

©Celgene Ltd, a BMS Company (2022). All rights reserved

Abbreviations: ECG, electrocardiogram.

B.3.5.2 Health-state unit costs and resource use

Disease management costs

Previously conducted economic analyses, NICE submission reports and published papers identified in the economic SLR were reviewed to identify the healthcare resources recommended for disease management in patients with UC in the UK. The unit costs of the resources identified were sourced from the National Schedule of NHS Costs 2019/2020, PSS Research Unit (PSSRU) and the literature, where relevant. The healthcare resource use per health state and the total health state cost per cycle can be found in Table 64 and Table 65, respectively.

Disease management costs were specific to each model health state, independent of treatment and were assumed to be similar across all populations of interest. Disease management costs incurred during induction (or extended induction in the relevant scenario analysis) were incurred as a one-off cost upon treatment initiation (or the initiation of extended induction).

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 64: Healthcare resource use per year by health state

Drug	Unit costs	Remission	Response No Remission	Active UC	1st/2nd surgery	Post-1st/2nd surgery remission	Post-1st surgery complications
Outpatient							
Consultant visit[185]	£183.43[a]	2.00	4.50	6.50	6.50	1.50	1.75
Blood test[185]	£1.81[b]	3.25	3.90	6.50	6.50	1.50	3.25
Inpatient							
Emergency endoscopy[185]	£814.46[c]	0.00	0.25	0.75	0.75	0.50	0.13
Elective endoscopy[185]	£330.51	0.20	0.50	2.00	2.00	1.25	0.65
Care without colectomy[1]	£2,301.47[e]	0.00	0.00	0.15	0.15	0.00	3.25
Stoma care (post-colectomy)[1]	£541.75[f]	NA	NA	NA	1.00	NA	NA

[a] Consultant led non-admitted face-to-face attendance, first, WF01B-300 general medicine, National Schedule of NHS Costs 2018/2019

[b] Directly assessed pathology services, Integrated blood services, DAPS03, National Schedule of NHS Costs 2018/2019

[c] Non elective short stay wireless capsule endoscopy, 19 years and over FE50A, National Schedule of NHS Costs 2018/2019

[d] Outpatient Procedures, Gastroenterology, service code 301 - wireless capsule endoscopy, 19 years and over - FE50A, National Schedule of NHS Costs 2018/2019

[e] 2018-2019 NHS reference costs for IBD without interventions, CC score 0-5+ (average considered), National Schedule of NHS Costs 2018/2019

[f] Stoma care costs included as per TA547 ERG review assuming 40% of patients have a stoma. Stoma consumable: Buchanan 2011, £66.75 per month which was inflated to 2020 value. Nursing costs: Specialist Nursing, Stoma Care Services, Adult, Face to face, N24AF, National Schedule of NHS costs 2018/19, assuming 6 visits per year in line with TA547 ERG

Abbreviations: HBV: hepatitis B virus.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 65: Health-state cost per cycle

Health state	Cost per cycle, mean	Reference to section in submission
Remission	£16.82	Unit cost and resource use presented in Table 64
Response no remission	£46.05	
Active UC	£108.13	
1st Surgery	£128.90	
Post 1st surgery remission	£42.09	
Post 1st surgery complications	£311.52	
2nd Surgery	£128.90	
Post 2nd surgery remission	£42.09	

Abbreviations: UC: ulcerative colitis.

Surgery-related costs

Based on feedback received from clinical consultation the cost of the 1st surgery was based on the cost of a subtotal colectomy, formation of an ileo-anal pouch and stoma closure. It was assumed that the surgery related to a very complex large intestine procedure (elective long stay only) and costs were sourced from the NHS Schedule Costs 2019/2020.[197] The weighted average was taken which yielded a 1st surgery cost of £14,309.51. Further to this, in line with expert clinical feedback received as part of the appraisal the cost of the 2nd surgery was based on the cost of pouch failure or adhesive small bowel obstruction. It was assumed this related to a complex large intestine procedure (non-elective long stay only). The weighted average of costs sourced from NHS Schedule costs 2019/20 yielded a value of £10,438.22.[197]

B.3.5.3 Adverse reaction unit costs and resource use

As mentioned in Section B.3.3.9 serious infection was the only adverse events included in the model owing to its high associated cost, an approach which is in line with TA633.[1] The cost associated with a serious infection adverse event was £2,992.17. This was calculated as a weighted average of six different infections included in the National Schedule of NHS Costs: sepsis, tuberculosis, pneumonia, soft tissue infections, bone and joint infections and urinary tract infections.[197]

B.3.5.4 Miscellaneous unit costs and resource use

There were no further unit costs or resource use included in the model.

B.3.6 *Summary of base-case analysis inputs and assumptions*

B.3.6.1 Summary of base-case analysis inputs

A summary of the variables applied in the base case economic analysis is presented in Table 66.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 66: Summary of variables applied in the cost effectiveness analysis

Variable	Inputs		Cross-reference
	TNFi-naive	TNFi-experienced	
Model settings			
Discount rate costs, %	3.5		Section B.3.2.2
Discount rate benefits, %	3.5		
Time horizon	Lifetime		
Perspective	NHS and PSS		
Patient characteristics			
Baseline patient age, years (SD)	[redacted]	[redacted]	Section B.3.3.2
Proportion female	[redacted]	[redacted]	
Weight, kg (SD)	[redacted]	[redacted]	
Clinical inputs			
Induction period transitions	Various – the proportion of patients achieving clinical 'Remission' and 'Response No Remission' at the end of the induction period was informed by the NMA		Section B.3.3
Maintenance period transitions	Various – mean absolute probabilities of sustained response or sustain remission derived from the NMA were translated to per cycle transition probabilities for 'loss of response' and 'loss of response no remission'		Section B.3.3
Discontinuation	Various – the proportion of patients discontinuing treatment due to AE were derived from the pivotal RCTs for biologics and the TRUENORTH trial for ozanimod		Section B.3.3.6
AEs	Serious infections were the only AEs included in the model due to their high costs. Serious infection associated with UC treatments during the induction and the maintenance period for each population of interest were derived from the pivotal RCTs for biologics and the TRUENORTH trial for ozanimod		Section B.3.3.9
Utility inputs			
Remission	0.87 (sourced from Woehl *et al.* 2008[182])		Section B.3.4
Response no remission	0.76 (sourced from Woehl *et al.* 2008[182])		
Active UC	0.41 (sourced from Woehl *et al.* 2008[182])		
1st Surgery	0.61 (sourced from Arseneau *et al*. 2006[183])		
Post 1st Surgery Remission	0.72 (sourced from Woehl *et al.* 2008[182])		
Post 1st Surgery Complications	0.34 (sourced from Arseneau *et al*. 2006[183])		
2nd Surgery	0.61 (assumed equal to 1st surgery)		

©Celgene Ltd, a BMS Company (2022). All rights reserved

2nd Surgery Remission	0.72 (assumed equal to 1st surgery)				
AE decrement	Serious infection: 0.156 (sourced from Stevenson *et al.* 2016)				
Cost inputs					
Intervention and comparator costs per cycle	**Acquisition**		**Administration**		
	Induction	**Maintenance**	**Induction**	**Maintenance**	
Ozanimod	[redacted]	[redacted]	£0.00		Section B.3.5.1
Infliximab	£4,710.67	£496.18	£599.10	£46.59	Section B.3.5.1
Adalimumab	£2,534.40	£411.84	£0.00		
Golimumab	£4,577.82	£495.53	£0.00		
Vedolizumab	£4,100.00	£589.38	£372.73	£30.28	
Ustekinumab	£6,655.97	£411.51	£186.37	£0.00	Section B.3.5.1
Health state costs per cycle, mean					
Remission	£16.82				Section B.3.5.2
Response no remission	£46.05				
Active UC	£108.13				
1st Surgery	£128.90				
Post 1st surgery remission	£42.09				
Post 1st surgery complications	£311.52				
2nd Surgery	£128.90				
Post 2nd surgery remission	£42.09				
Adverse events	Serious infection: £2,992.17 (sourced from the National Schedule of NHS Costs 2019/2020)[197]				Section B.3.5.3
Monitoring costs	The only monitoring test included in the model was the ECG test for ozanimod at induction: £61.80 (Sourced from the National Schedule of NHS Costs 2019/20, EY51Z, Directly Accessed Diagnostic Services, Electrocardiogram Monitoring or Stress Testing)[197]				Section B.3.5.2
Surgery-related costs	**1st surgery**		**2nd surgery**		Section B.3.5.2
	£14,310 (sourced from NHS Schedule costs 2019/20)[197]		£10,438 (sourced from NHS Schedule costs 2019/20) [197]		

Abbreviations: AE: adverse events; IV: intravenous; NMA; network meta-analysis; SC: subcutaneous; SD: standard deviation; TNFi: tumour necrosis factor-alpha inhibitor; UC: ulcerative colitis.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.3.6.2 Assumptions

The assumptions used in the base case analysis are described in Table 67.

Table 67: List of assumptions for the base case analysis model

Assumption	Justification	Reference to section in submission	Addressed in scenario analysis
The patient population in TRUENORTH trial is generalisable to UK clinical practice	The baseline characteristics for the patients in this trial are consistent with the target patient population in the UK; the average patient age in the trial was approximately 42 years which is consistent with the typical age of a UC patient.[40] Additionally, approximately twice as many patients had limited to left side colitis over extensive colitis which reflects the expected disease profile of UC.[40] Participants prior medications were also aligned with the typical treatment pathway patients would follow in UK clinical practice; TNFis (infliximab and adalimumab) were the most commonly used prior biologic medications reflecting their first-line use following treatment failure with CvT (Section B.1.3.4).	Section B.3.2.1 and Section B.2.12	-
An extended induction period was not modelled in the base case analysis	Clinical consultation conducted as part of this appraisal noted while extended induction may be offered to patients in some cases it is not standard clinical practice in the UK. This view was supported by the Committee in TA633 which noted that there was no conclusive clinical evidence that extended induction always occurs in clinical practice. In addition, use of data from the TRUENORTH trial for extended induction was associated with several limitations such as breaking of trial randomisation.[1]	Section B.3.2.2	Extended induction was explored in a scenario analysis
Patients who discontinued active therapies (receive BSC) were able to move from 'Active UC' to both the 'Remission' and 'Response No Remission' health states and back (relapse)	The natural course of the disease follows a relapsing-remitting pattern, meaning some patients with active UC may go into response, as the disease can worsen or improve regardless of active therapy. This assumption is in line with preferred assumption of the ERG in TA633.[1] No trial data were available to inform this transition. A scenario presented by the ERG in TA633 assumed that 1% of patients would enter spontaneous 'Remission' and 'Response No Remission' from the 'Active	Section B.3.2.2	Scenario analyses were conducted which assumed 0% and 1% of patients experienced spontaneous remission/response from the 'Active UC' health state

©Celgene Ltd, a BMS Company (2022). All rights reserved

	UC' health state at each model cycle. The committee considered the ERG's assumption to be an overestimate but agreed that there may be a small number of people who would improve without treatment.[1] Therefore, in the base case, 0.5% of patients were modelled to enter spontaneous 'Remission' or 'Response No Remission' from the 'Active UC' health state at each model cycle.		
Following progression from active treatment (due to failure to achieve response, loss of response or discontinuation due to AEs) patients transition initially to an 'Active UC' health state in which they received no further active treatments but could continue on BSC alone	Reflective of UK clinical practice where once patients have failed multiple available treatment options, they have a choice of either remaining on BSC or progressing to surgery.	Section B.3.2.2	-
Patients who discontinued due to AE during the induction period still incurred the full cost associated with a course of induction treatment	In clinical reality a small proportion of patients do discontinue treatments due to AEs. This view is supported by prior appraisals (TA633)[1] which were criticised for not explicitly modelling discontinuation due to AEs. Patients who discontinue due to AE are assumed to still have 'purchased' the full course of induction before they start so the full cost is incurred regardless of whether they discontinue.	Section B.3.2.2	-
There was a constant loss of response amongst patients within and beyond the trial duration of one-year	Clinical consultation conducted as part of this appraisal noted that there are no publicly available data to assess the potential loss of response in the second and subsequent years. Therefore a constant loss of response was assumed in line with the accepted assumption in previous submissions (TA633).[1]	Section B.3.3.4	A scenario was conducted where a 25% treatment waning effect was applied after 2 years to the active treatment, in line with TA633 scenario[1]
The transition probability of 2nd surgery from 'Post 1st Surgery Complications' health state was equal to the probability of '1st surgery'	There are no published data available for patients receiving 2nd surgery and therefore in line with the assumption in TA633 it was assumed that the transition probability was equal to the probability of 1st surgery	Section B.3.3.8	-
Patients were at constant risk of experiencing AEs over the model time horizon	It was assumed that in clinical practice the risk of AEs was not dependent on time on treatment. This assumption was in line with TA633[1]	Section B.3.3.9	-

©Celgene Ltd, a BMS Company (2022). All rights reserved

All UC treatment included the model were assumed to have no impact on general mortality	Clinical consultations conducted as part of this appraisal noted that UC does not significantly increase mortality and therefore treatments have little effect on mortality rates	Section B.3.3.10	-
A proportion of patients (30% of patients on TNFis in the base case) were assumed to receive an escalated dose in the maintenance period at any one time, and a corresponding higher drug acquisition cost was applied to these patients	Clinical consultation conducted as part of this appraisal noted that dose escalation is commonplace in UK clinical practice. This is in line with the accepted assumption in TA633[1]	Section B.3.5.1	Scenario analyses were run in which dose escalation was assumed to be 0% and 50% in the maintenance period to assess uncertainties surrounding dose escalation in the model
Whilst being off label patients receiving infliximab were also modelled to experience dose escalation	In line with clinicians in TA633[1], clinical feedback conducted as part of this appraisal noted that patients receiving infliximab would receive dose escalation in clinical practice		
ECG test is the only relevant monitoring cost in the economic analysis	Clinical consultation conducted as part of this appraisal confirmed that the monitoring requirements are similar for ozanimod and existing treatments for moderately to severely active UC in UK clinical practice. Therefore, monitoring costs were not considered in the economic model in line with previous appraisals (TA633,[1] TA547[88] and TA342[85]), with the exception of a single ECG for ozanimod at induction; this is a specific requirement for ozanimod as specified in the SmPC.[12]	Section B.3.5.1	-

Abbreviations: AE: adverse event; ECG: electrocardiogram; TA: technology appraisal; TNFi: tumour necrosis factor inhibitor; UC: ulcerative colitis; UK: United Kingdom.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.3.7 *Base-case results*

Results of the economic analysis in both the TNFi-naïve and TNFi-experienced population are presented in Section B.3.7.1 below.

B.3.7.1 Base-case cost-effectiveness analysis results

Table 68 and Table 69 present the base case pairwise results of the economic evaluation for the TNFi-naïve and TNFi-experienced population, respectively. Fully incremental analyses are presented in and Table 71. In both cases the PAS price of ozanimod has been used. As some of the comparisons resulted in a south-west (SW) quadrant, to improve the readability of the results, the net-health benefit (NHB) of ozanimod at a willingness-to-pay (WTP) threshold of £30,000 compared to the comparators has been included in each of the results tables. NHB was selected in line with the preference expressed in the consultation on the NICE methods for health technology evaluations 2020.[198]

Ozanimod was found to be cost-effective compared to all relevant comparators in the TNFi-naïve population yielding a NHB of 0.175, 0.003, 0.100 and 0.205 when compared to infliximab, adalimumab, golimumab and vedolizumab, respectively. Similarly, ozanimod was cost-effective compared to both vedolizumab and ustekinumab in the TNFi-experienced population with a NHB of 0.170 and 0.156, respectively. Ozanimod dominated ustekinumab in the TNFi-experienced population.

Clinical outcomes from the cost-effectiveness model, the proportion of the cohort in each health state over time (Markov trace), and the disaggregated results of the base case incremental cost-effectiveness analysis are reported in Appendix H.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 68: Base-case pair-wise cost-effectiveness results – TNFi-naïve population

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER for ozanimod versus comparator (£/QALY)	NHB (at a WTP threshold of £30,000)
Ozanimod	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£28,686	0.003
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£167,024[a]	0.175
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£71,023[a]	0.100
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£52,736[a]	0.205

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Table 69: Base-case pair-wise cost-effectiveness results – TNFi-experienced population

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER for ozanimod versus comparator (£/QALY)	NHB (at a WTP threshold of £30,000)
Ozanimod	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£199,551[a]	0.170
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Ozanimod dominant	0.156

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

©Celgene Ltd, a BMS Company (2022). All rights reserved

The base case fully incremental analyses were carried out in both the TNFi-naïve and the TNFi-experienced subpopulations. The base case fully incremental analysis in the TNFi-naïve population showed ozanimod to be most cost-effective treatment option, with the lowest fully incremental ICER of £28,686 as compared to adalimumab. In the TNFi-experienced population, the results of the fully incremental analysis were consistent with the pair-wise analysis, with a fully incremental ICER of £199,551 saved per QALY forgone as compared to vedolizumab.

Table 70: Base-case fully incremental cost-effectiveness results – TNFi-naïve population

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	-
Ozanimod	[redacted]	[redacted]	[redacted]	[redacted]	£28,686
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	Extendedly Dominated
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£52,736

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Table 71: Base-case fully incremental cost-effectiveness results – TNFi-experienced population

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Ozanimod	[redacted]	[redacted]	-	-	-
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£199,551

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.3.8 *Sensitivity analyses*

B.3.8.1 Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was conducted in order assess the simultaneous effect of uncertainty in the different model parameters and to demonstrate whether the model results are robust to those variations. A Monte-Carlo simulation with 1,000 iterations was performed where model inputs were randomly sampled from the specified probability distributions. The following key parameters were varied:

- Absolute mean probabilities for response (no remission), remission, and no response, no remission using CODA output generated by NMA
- The rate of each AE in the induction phase, and the rate of each AE in the maintenance phase, using a log-normal distribution
- The proportions of males and patients discontinuing treatment due to AEs in the induction and maintenance phases, health-state utilities, and AE disutility's, using a beta distribution
- The AE management costs, and disease management costs, using a gamma distribution

Where a standard error or CI was not available for a selected parameter, variation of 20% of the mean was applied. A table containing a list of the inputs used in PSA is presented in Appendix H.2.

The results of the PSA with 1,000 iterations are presented in Table 72 and Table 73 for the TNFi-naïve and experienced population, respectively. Overall, the results of the PSA were similar to the base case, with ozanimod remaining cost-effective versus all comparators, indicating the results to be robust to uncertainty. However, the cost-effectiveness of ozanimod in the TNFi-experienced population was improved in the PSA, with ICERs (SW quadrant) increasing from £199,551 to £1,324,054 when comparing ozanimod to vedolizumab, and with ozanimod continuing to dominate ustekinumab. The large increase in ICER versus vedolizumab in the TNFi-experienced population occurred because the base case incremental QALYs were small ([redacted]) and therefore marginal variation in this value resulted in significant variations in the ICER. Based on this analysis, the probability that ozanimod is cost-effective in the TNFi-naïve and TNFi-experienced populations is estimated to be [redacted]% and [redacted]%, respectively, at a willingness-to-pay threshold of £30,000 per QALY.

Table 72: Probabilistic results (TNFi-naïve population)

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Ozanimod	-	-	-	[redacted]
Adalimumab	[redacted]	[redacted]	£28,934	[redacted]
Infliximab	[redacted]	[redacted]	£155,144[b]	[redacted]
Golimumab	[redacted]	[redacted]	£71,945[b]	[redacted]
Vedolizumab	[redacted]	[redacted]	£63,862[b]	[redacted]

[a]The probability of ozanimod being cost-effective versus the comparator technology at a cost-effectiveness threshold of £30,000/QALY.
[b]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year gained; TNFi: tumour necrosis factor-alpha inhibitor.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 73: Probabilistic results (TNFi-experienced population)

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Ozanimod	-	-	-	[redacted]
Vedolizumab	[redacted]	[redacted]	£1,324,054[b]	[redacted]
Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]

[a]The probability of ozanimod being cost-effective versus the comparator technology at a cost-effectiveness threshold of £30,000/QALY.
[b]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year gained.

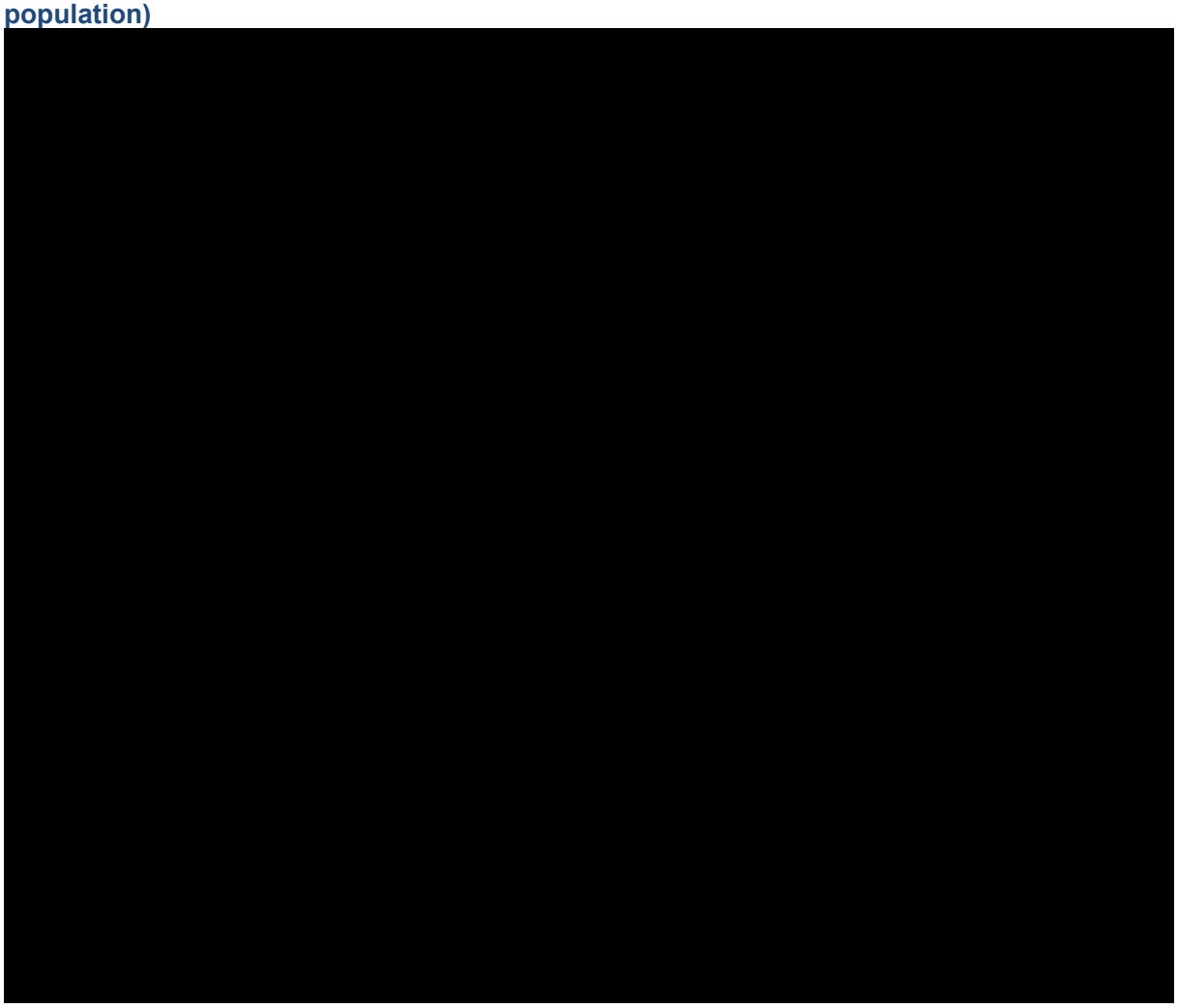
TNFi-naïve

Scatter plots showing the incremental costs and QALYs for ozanimod versus infliximab, adalimumab, golimumab and vedolizumab in the TNFi-naïve population across all iterations in the PSA are presented in Figure 46, Figure 47, Figure 48 and Figure 49, respectively. The scatterplots suggest considerable correlation in incremental QALYs and costs, which may be expected given treatment costs are only applied to responders (e.g., a higher response rate means more patients achieve higher utility, resulting in higher QALYs gained, but also higher costs given more patients remain on treatment).

Cost-effectiveness acceptability curves for ozanimod, infliximab, adalimumab, golimumab and vedolizumab are presented in Figure 49. As is to be expected, comparators that were more costly but more effective than ozanimod for the majority of iterations (infliximab, golimumab and vedolizumab) were found to have a greater probability of being cost-effective at higher WTP thresholds. Conversely, adalimumab, which was less costly but less effective than ozanimod for the majority of iterations was more likely to be cost-effective at a lower WTP threshold.

©Celgene Ltd, a BMS Company (2022). All rights reserved

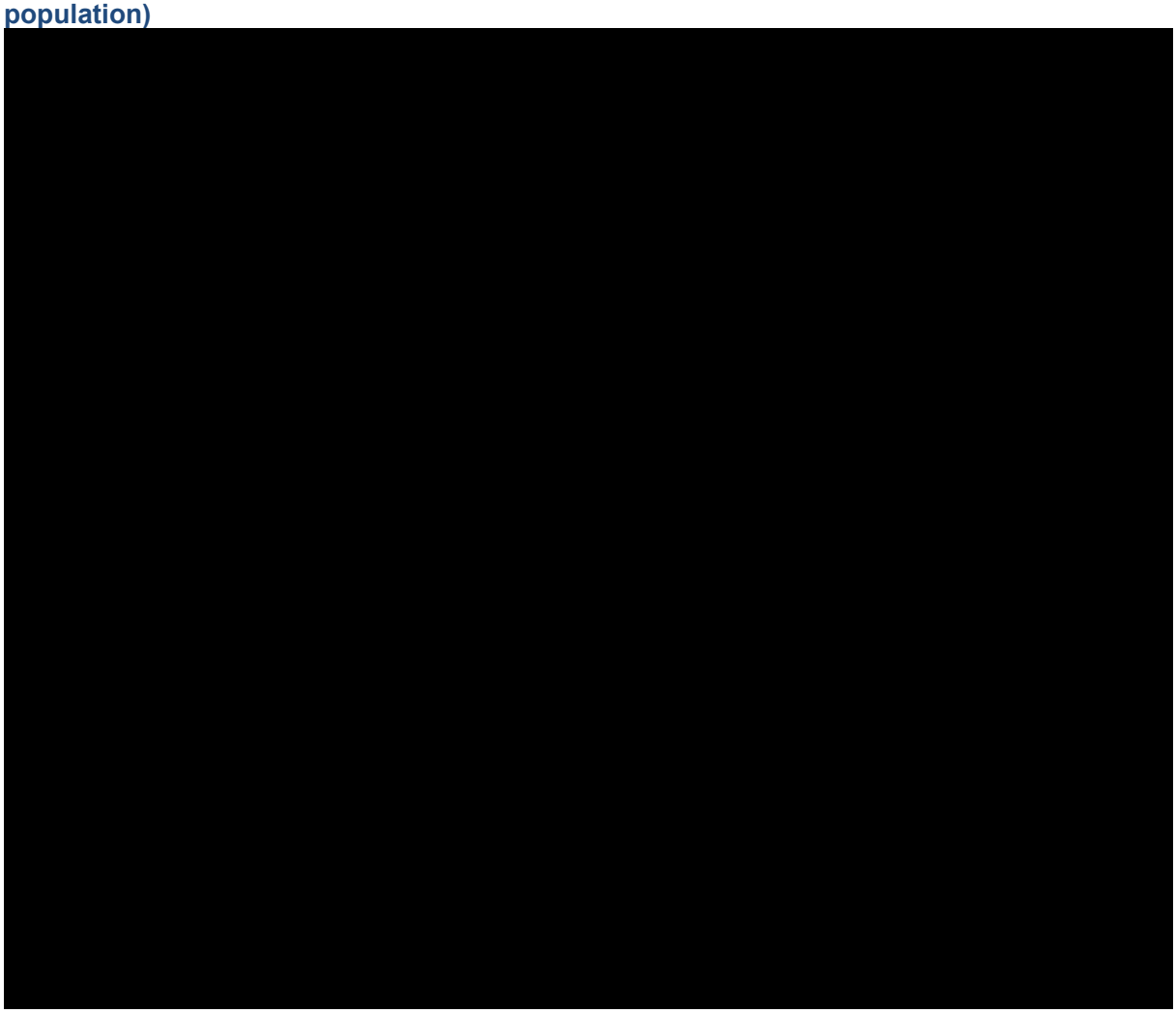
Figure 46: Cost-effectiveness plane for ozanimod versus infliximab (TNFi-naïve population)



Abbreviations: QALYs: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; WTP: willingness-to-pay threshold.

©Celgene Ltd, a BMS Company (2022). All rights reserved

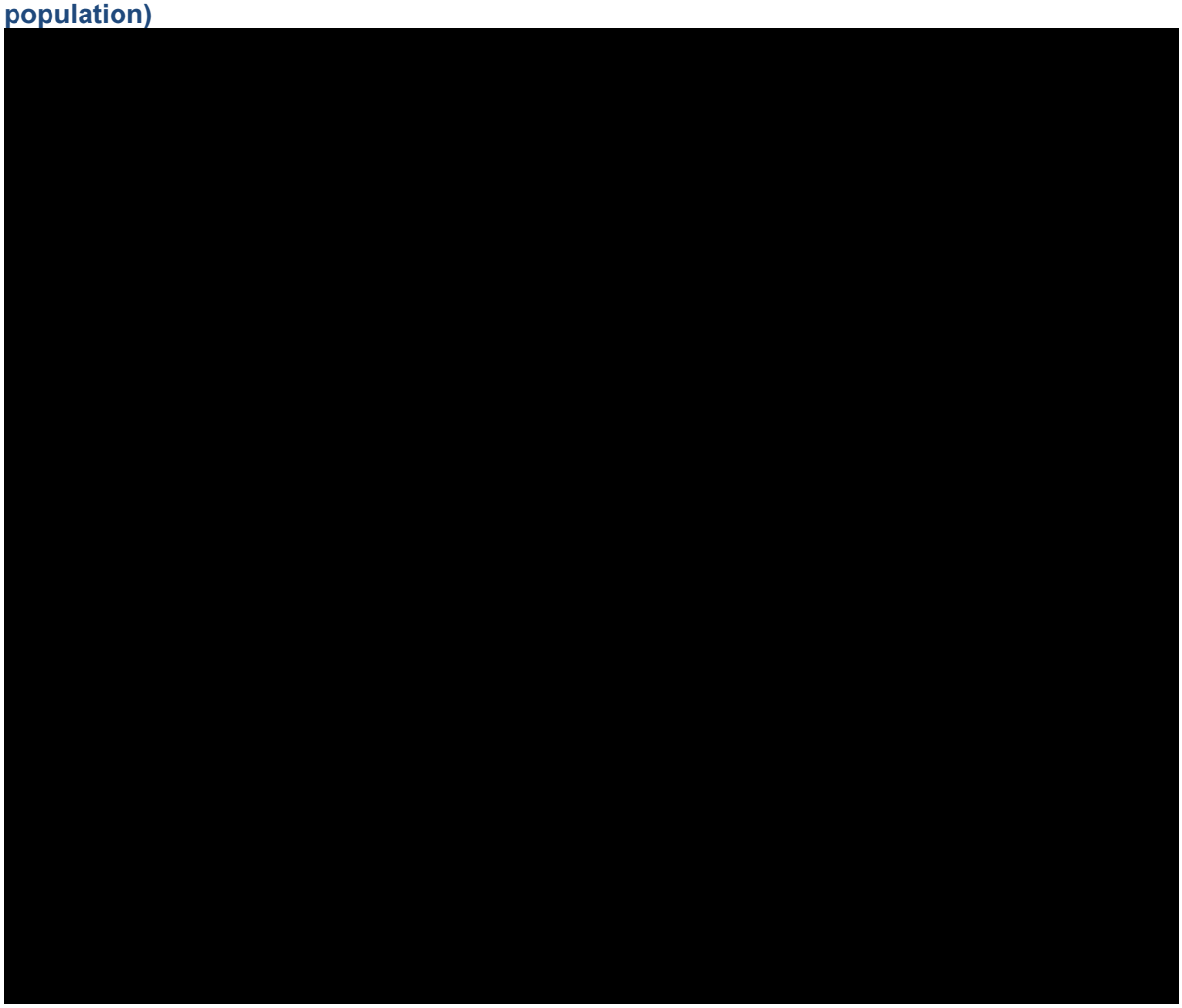
Figure 47: Cost-effectiveness plane for ozanimod versus adalimumab (TNFi-naïve population)



Abbreviations: QALYs: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness-to-pay threshold.

©Celgene Ltd, a BMS Company (2022). All rights reserved

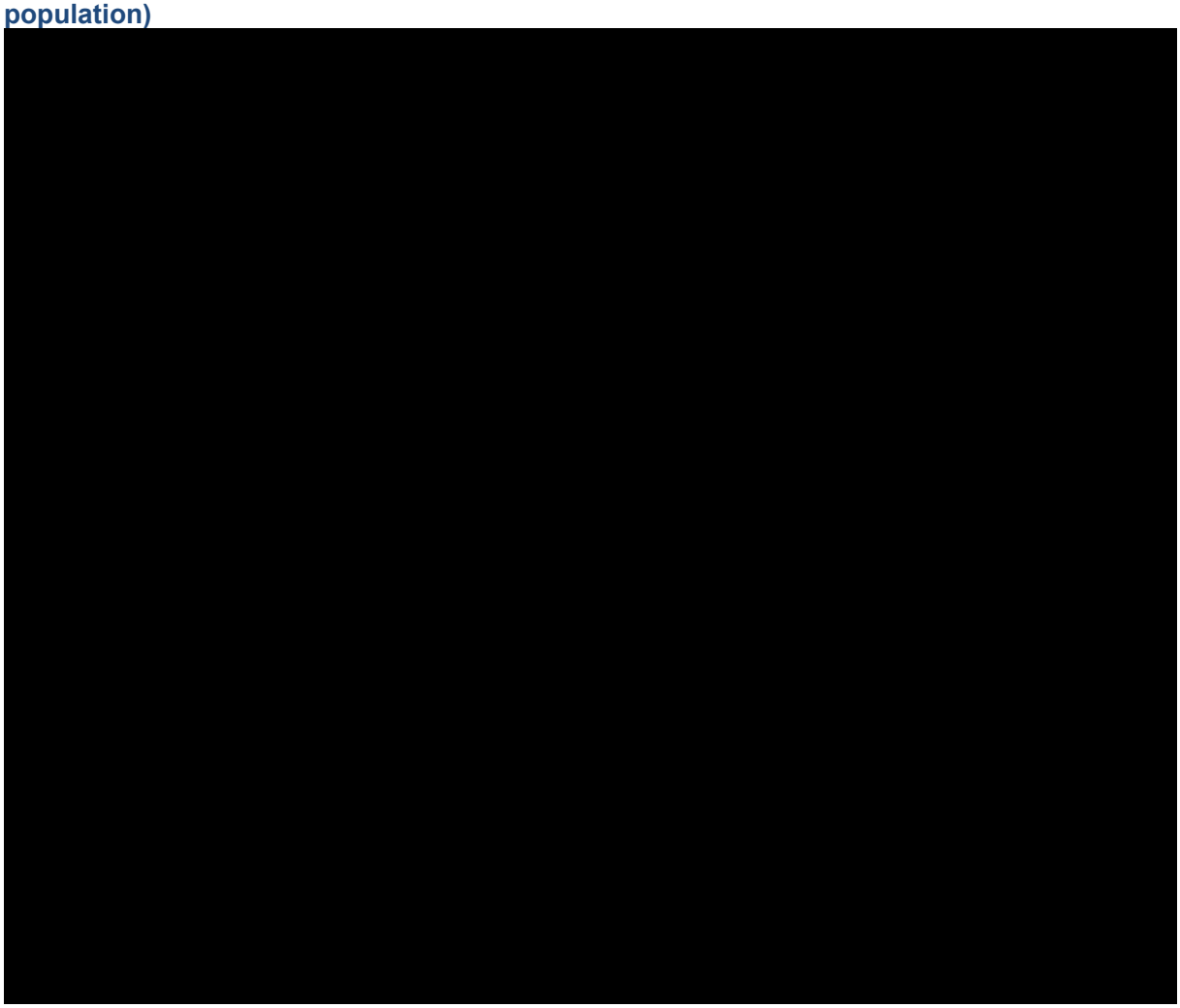
Figure 48: Cost-effectiveness plane for ozanimod versus golimumab (TNFi-naïve population)



Abbreviations: QALYs: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; WTP: willingness-to-pay threshold.

©Celgene Ltd, a BMS Company (2022). All rights reserved

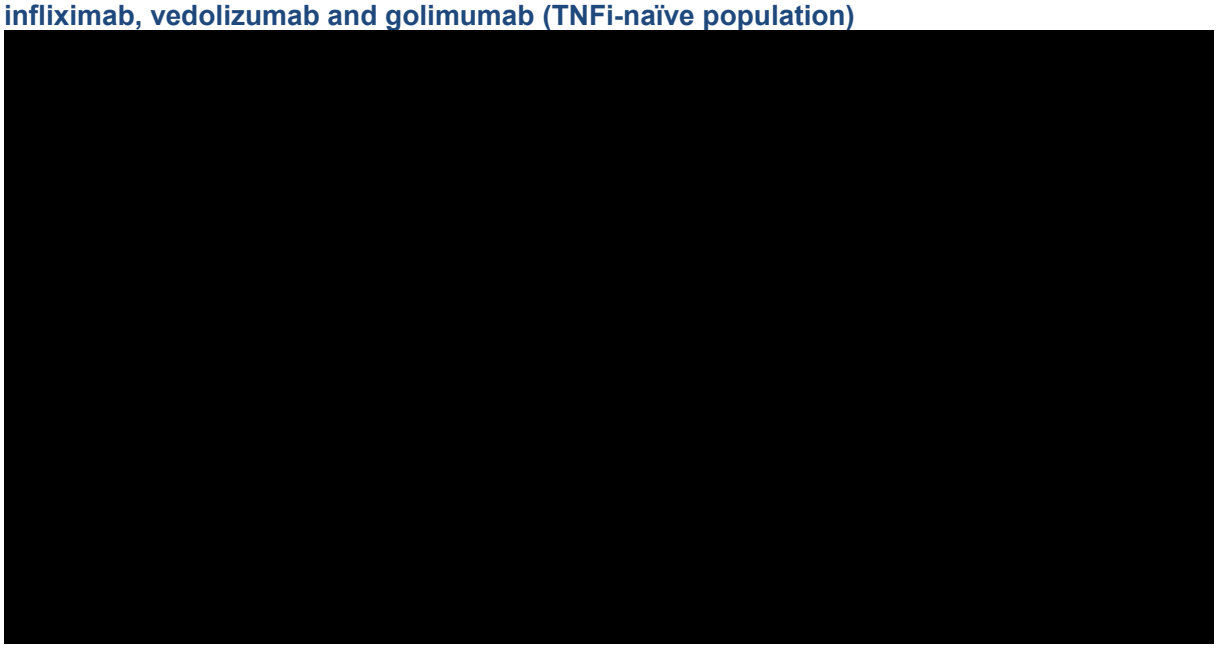
Figure 49: Cost-effectiveness plane for ozanimod versus vedolizumab (TNFi-naïve population)



Abbreviations: QALYs: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness-to-pay threshold.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 50: Cost-effectiveness acceptability curve for ozanimod versus adalimumab, infliximab, vedolizumab and golimumab (TNFi-naïve population)



Abbreviations: TNFi: tumour necrosis factor-alpha inhibitor.

TNFi-experienced

Scatter plots showing the incremental costs and QALYs for ozanimod versus ustekinumab, and vedolizumab in the TNFi-experienced population across all iterations in the PSA are presented in Figure 51 and Figure 52, respectively. Cost-effectiveness acceptability curves for ozanimod, ustekinumab and vedolizumab are presented in Figure 53. Ozanimod was found to have the highest probability of being cost-effective at all WTP thresholds.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 51: Cost-effectiveness plane for ozanimod versus vedolizumab (TNFi-experienced population)

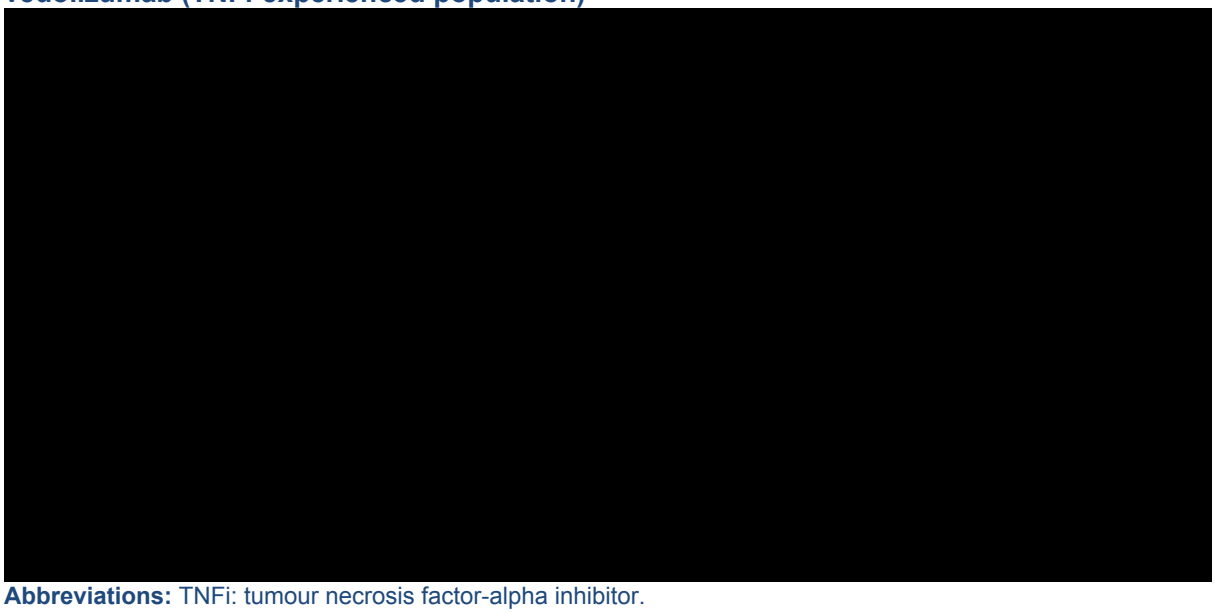
Abbreviations: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness to pay.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 52: Cost-effectiveness plane for ozanimod versus ustekinumab (TNFi-experienced population)

Abbreviations: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness to pay.

Figure 53: Cost-effectiveness acceptability curve for ozanimod versus ustekinumab and vedolizumab (TNFi-experienced population)



Abbreviations: TNFi: tumour necrosis factor-alpha inhibitor.

©Celgene Ltd, a BMS Company (2022). All rights reserved

B.3.8.2 Deterministic sensitivity analysis

In order to assess the robustness of the base case cost-effectiveness results, deterministic sensitivity analyses (DSA) were conducted by varying the input for each parameter in the model by ±20% of their mean value, whilst keeping all other inputs the same. For certain parameters where standard errors of the mean were available, the lower and upper limits were defined by the 95% CI around the mean. The inputs used in the DSA are presented in Appendix H.3.

As some of the comparisons resulted in a south-west (SW) quadrant ICER, to improve the readability of the results, the NHB of ozanimod at a WTP threshold of £30,000 compared to the comparators has been presented. Net Health Benefit (NHB) calculated at the upper and lower bounds for the 10 most influential parameters are shown graphically in tornado plots in Figure 54, Figure 55, Figure 56 and Figure 57 for the TNFi-naïve population and Figure 58 and Figure 59 for the TNFi-experienced population.

As shown in Figure 54, Figure 55 and Figure 56 the parameters with the greatest impact on the NHB in the TNFi-naïve population were those related to the proportion of patients achieving both sustained clinical response and remission at maintenance. These parameters also had the highest impact on the NHB in the TNFi-experienced population for ozanimod versus ustekinumab (Figure 59). The values used in the base case economic analysis were derived from the base case NMA (Section B.2.8) and are considered to represent the most suitable inputs available.

Further to this, the proportion of patients receiving vedolizumab SC also impacted the NHB results for ozanimod versus vedolizumab in both populations (Figure 57 and Figure 58). The proportion of patients receiving vedolizumab SC/IV at maintenance were validated by clinical consultation conducted as part of this appraisal and are hence considered to represent the most suitable inputs available.

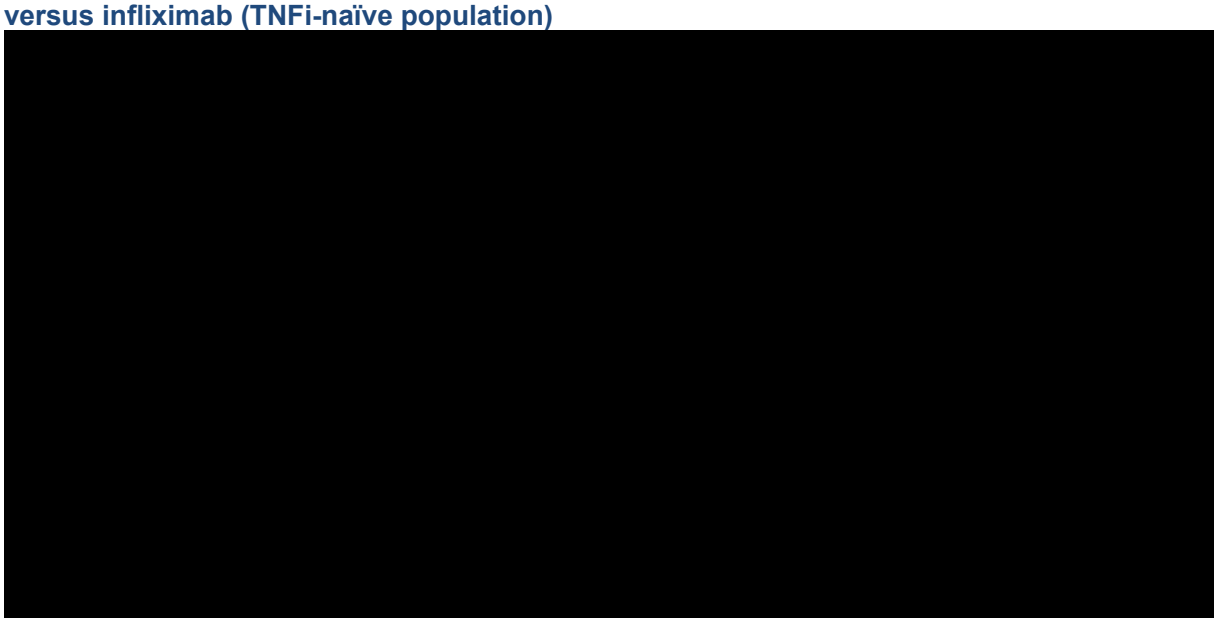
TNFi-naive

Figure 54: Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus adalimumab (TNFi-naïve population)

Abbreviations: DSA: deterministic sensitivity analysis; IFX: infliximab; OZA: ozanimod; RNR: response no remission; TNFi: tumour necrosis factor alpha inhibitor.

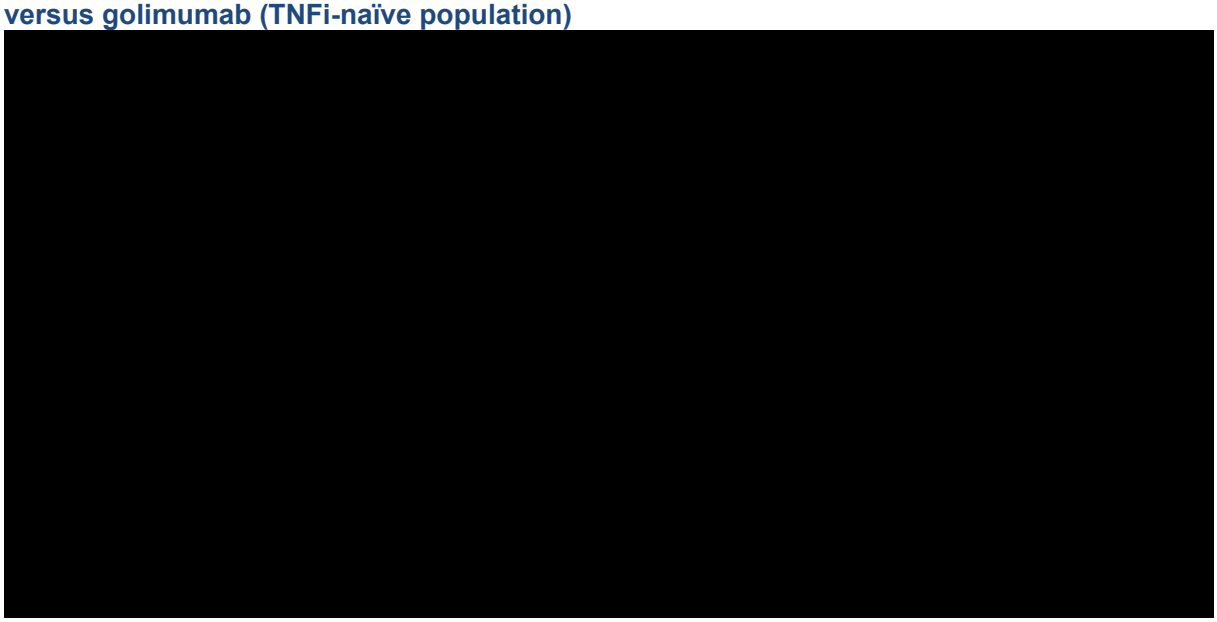
©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 55: Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus infliximab (TNFi-naïve population)



Abbreviations: ADA: adalimumab; BSC: best supportive care; DSA: deterministic sensitivity analysis; OZA: ozanimod; RNR: response no remission; TNFi: tumour necrosis factor alpha inhibitor.

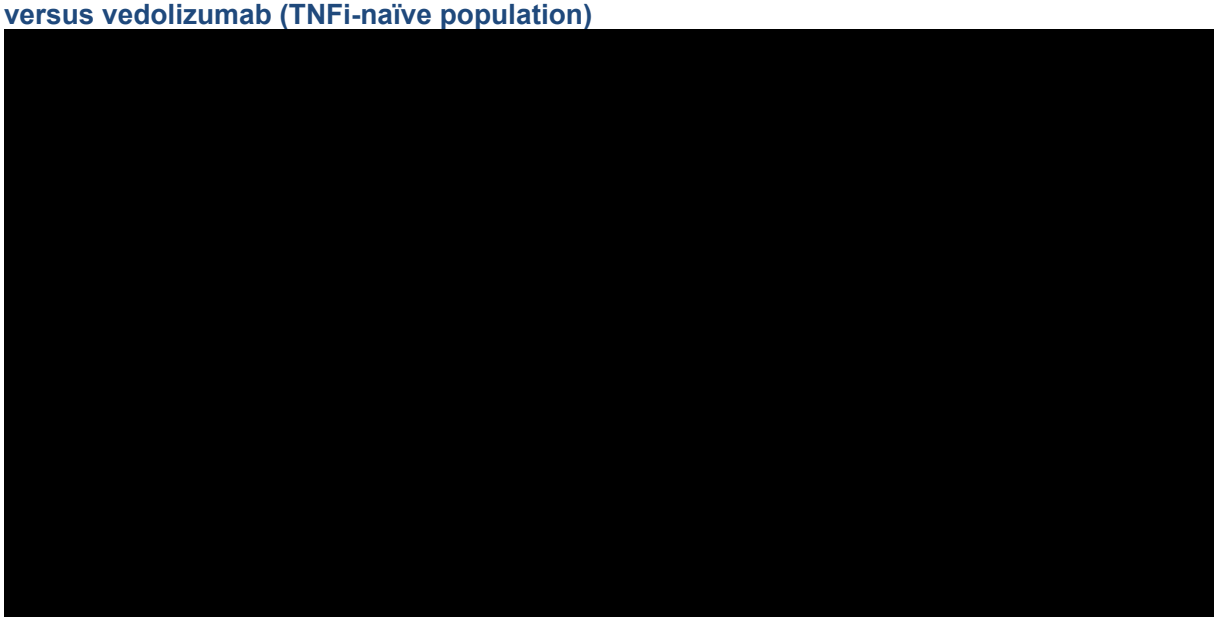
Figure 56: Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus golimumab (TNFi-naïve population)



Abbreviations: ADA: adalimumab; BSC: best supportive care; DSA: deterministic sensitivity analysis; OZA: ozanimod; RNR: response no remission; TNFi: tumour necrosis factor alpha inhibitor.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Figure 57: Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus vedolizumab (TNFi-naïve population)



Abbreviations: DSA: deterministic sensitivity analysis; OZA: ozanimod; RNR: response no remission; TNFi: tumour necrosis factor alpha inhibitor; VDZ: vedolizumab.

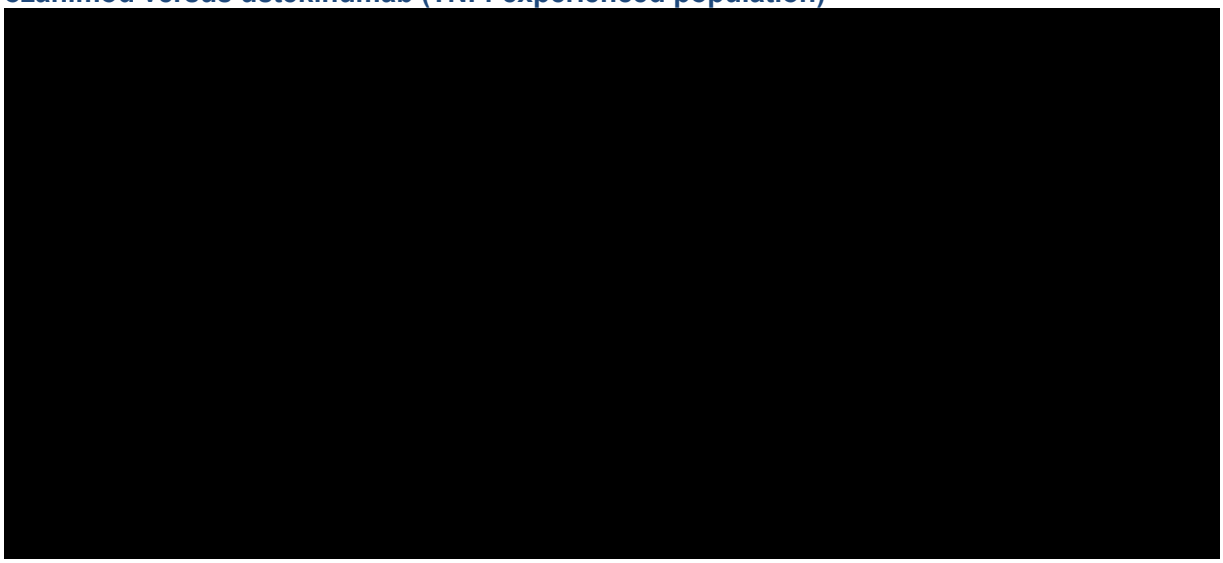
©Celgene Ltd, a BMS Company (2022). All rights reserved

TNFi-experienced

Figure 58: Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus vedolizumab (TNFi-experienced population)

Abbreviations: CI: confidence interval; DSA: deterministic sensitivity analysis; NHB: net health benefit; QALY: quality-adjusted life year; TNFi: tumour necrosis factor alpha inhibitor.

Figure 59: Tornado diagram for the driver of NHB results – top ten parameters for ozanimod versus ustekinumab (TNFi-experienced population)



Abbreviations: CI: confidence interval; DSA: deterministic sensitivity analysis; IV: intravenous; QALY: quality-adjusted life year; NHB: net health benefit; OZA: ozanimod; SC: subcutaneous; TNFi: tumour necrosis factor alpha inhibitor; VDZ: vedolizumab.

B.3.8.3 Scenario analysis

As described in Sections B.3.2.2 –B.3.5.2, scenario analyses were conducted to explore the impact of structural assumptions and alternative inputs on the results of the cost-effectiveness model. The results of the scenario analyses are presented below.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Overall, the results of the scenario analyses were not too dissimilar from the results of the base case analysis, demonstrating the results to be robust to uncertainties in the model inputs and assumptions.

In general, the results from the spontaneous remission scenario analyses were consistent with the base case results. Overall, in the TNFi-naïve population the NHB marginally decreased when spontaneous remission was not included in the model and marginally increased when it was assumed to be 1%, except for adalimumab which had negligible NHB in both cases (Table 74). In the TNFi-experienced population the NHB was marginally increased when spontaneous remission was not included and marginally decreased for ustekinumab when it was assumed to be 1% (Table 74). As it is clinically possible for patients to experience spontaneous response/remission owing to the relapsing remitting nature of the disease and a spontaneous response/remission rate of 1% has been considered too high (Section B.3.2.2), the proportion of 0.5% used in the base case analysis is considered the most clinically relevant value.[1]

The results in the TNFi-naïve population when extended induction was included as a scenario were similar to the base case results, with only marginal increases to NHB (Table 75). In the TNFi-experienced population the NHB of ozanimod compared to relevant comparators slightly increased upon inclusion of extended induction. Clinical consultation conducted as part of this appraisal noted that extended induction may be offered to patients in some cases, but that it is not standard clinical practice. As such, the decision to exclude extended induction in the base case analysis is considered appropriate. However, it may fail to capture to some of the potential benefits of ozanimod.

Overall, the results of the scenario analyses exploring dose escalations were similar to the base case results (Table 76). However, removal of dose escalation from the model generally reduced the NHB of ozanimod compared to relevant comparators while increasing dose escalation to 50% generally increased the NHB of ozanimod compared to relevant comparators in both the TNFi-naïve and experienced populations. The dose escalation of 30% used in the base case analysis was validated by clinical consultation conducted as part of this appraisal and is therefore considered the most appropriate assumption.

The results of the scenario analyses which included treatment waning (Table 77), vial sharing (Table 78) and alternative distributions of CcT/BsC (Table 81) were similar to the base analysis in both populations. Similarly, the results for the scenario analyses which included subsequent treatments in the TNFi-naïve population were similar to the base case results (Table 79).

The results of the scenario analyses which explored alternative utility inputs were varied but were generally aligned with the base case analysis (Table 80). With the exception of adalimumab and ustekinumab the NHB for ozanimod compared to relevant comparators in both the TNFi-naïve and TNFi-experienced population was marginally increased upon the use of utility values derived from the TRUENORTH trial, TA342 and TA547. As discussed in Section B.3.4.1 the TRUENORTH utility values were associated with several limitations and therefore the utility values used in the base case analysis (aligned with TA633) were considered the most appropriate.

Finally, the results from the scenario analyses assessing the effect of the proportion of patients receiving SC vedolizumab found in both scenarios (0% and 30%) ozanimod was associated with reduced incremental costs and hence increased NHB for ozanimod compared to vedolizumab.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Table 74: Result from spontaneous response/remission scenario analyses

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024[a]	0.175
	Golimumab	[redacted]	[redacted]	£71,023[a]	0.101
	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
0% spontaneous remission	Adalimumab	[redacted]	[redacted]	£26,437	0.008
	Infliximab	[redacted]	[redacted]	£161,415[a]	0.174
	Golimumab	[redacted]	[redacted]	£68,735[a]	0.098
	Vedolizumab	[redacted]	[redacted]	£50,196[a]	0.190
1% spontaneous remission	Adalimumab	[redacted]	[redacted]	£30,986	-0.002
	Infliximab	[redacted]	[redacted]	£172,375[a]	0.177
	Golimumab	[redacted]	[redacted]	£73,195[a]	0.103
	Vedolizumab	[redacted]	[redacted]	£55,215[a]	0.218
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551[a]	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.156
0% spontaneous remission	Vedolizumab	[redacted]	[redacted]	£202,670[a]	0.171
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.162
1% spontaneous remission	Vedolizumab	[redacted]	[redacted]	£196,832[a]	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.151

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: TNFi: tumour necrosis factor alpha inhibitor.

Table 75: Result from extended induction scenario analysis

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024[a]	0.175
	Golimumab	[redacted]	[redacted]	£71,023[a]	0.101
	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
Extended induction included	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£95,490[a]	0.178
	Golimumab	[redacted]	[redacted]	£53,607[a]	0.116
	Vedolizumab	[redacted]	[redacted]	£49,151[a]	0.250
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.156
	Vedolizumab	[redacted]	[redacted]	£81,131	0.234

©Celgene Ltd, a BMS Company (2022). All rights reserved

Extended induction included	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.184

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: TNFi: tumour necrosis factor alpha inhibitor.

Table 76: Result from dose escalation scenario analyses

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024[a]	0.175
	Golimumab	[redacted]	[redacted]	£71,023[a]	0.101
	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
0% dose escalation	Adalimumab	[redacted]	[redacted]	£52,734	-0.047
	Infliximab	[redacted]	[redacted]	£105,530[a]	0.097
	Golimumab	[redacted]	[redacted]	£32,908[a]	0.007
	Vedolizumab	[redacted]	[redacted]	£41,492[a]	0.104
50% dose escalation	Adalimumab	[redacted]	[redacted]	£12,655	0.036
	Infliximab	[redacted]	[redacted]	£208,020[a]	0.228
	Golimumab	[redacted]	[redacted]	£96,434[a]	0.163
	Vedolizumab	[redacted]	[redacted]	£60,233[a]	0.272
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551[a]	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.156
0% dose escalation	Vedolizumab	[redacted]	[redacted]	£147,551[a]	0.118
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.134
50% dose escalation	Vedolizumab	[redacted]	[redacted]	£234,217[a]	0.205
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.171

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: TNFi: tumour necrosis factor alpha inhibitor.

Table 77: Result from treatment waning scenario analyses

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024[a]	0.175
	Golimumab	[redacted]	[redacted]	£71,023[a]	0.101
	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
25% treatment waning after 2 years	Adalimumab	[redacted]	[redacted]	£28,729	0.002
	Infliximab	[redacted]	[redacted]	£169,131[a]	0.172
	Golimumab	[redacted]	[redacted]	£76,121[a]	0.100
	Vedolizumab	[redacted]	[redacted]	£54,939[a]	0.198

©Celgene Ltd, a BMS Company (2022). All rights reserved

TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551[a]	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.156
25% treatment waning after 2 years	Vedolizumab	[redacted]	[redacted]	£244,071[a]	0.166
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.154

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: TNFi: tumour necrosis factor alpha inhibitor.

Table 78: Result from vial sharing scenario analysis

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024[a]	0.175
	Golimumab	[redacted]	[redacted]	£71,023[a]	0.101
	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
Vial sharing	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£124,892[a]	0.121
	Golimumab	[redacted]	[redacted]	£71,023[a]	0.100
	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551[a]	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.156
Vial sharing	Vedolizumab	[redacted]	[redacted]	£199,551[a]	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.149

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: TNFi: tumour necrosis factor alpha inhibitor.

Table 79: Result from subsequent treatment scenario analyses

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024[a]	0.175
	Golimumab	[redacted]	[redacted]	£71,023[a]	0.101
	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
Subsequent treatment with vedolizumab	Adalimumab	[redacted]	[redacted]	£28,033	0.004
	Infliximab	[redacted]	[redacted]	£168,704[a]	0.175
	Golimumab	[redacted]	[redacted]	£71,143[a]	0.099
Subsequent treatment with ustekinumab	Adalimumab	[redacted]	[redacted]	£28,153	0.004
	Infliximab	[redacted]	[redacted]	£167,957[a]	0.175
	Golimumab	[redacted]	[redacted]	£70,899[a]	0.099

©Celgene Ltd, a BMS Company (2022). All rights reserved

	Vedolizumab	[redacted]	[redacted]	£52,503[a]	0.200

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: TNFi: tumour necrosis factor alpha inhibitor.

Table 80: Result from utilities scenario analyses

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024[a]	0.175
	Golimumab	[redacted]	[redacted]	£71,023[a]	0.101
	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
TRUENORTH	Adalimumab	[redacted]	[redacted]	£54,046	-0.026
	Infliximab	[redacted]	[redacted]	£337,782[a]	0.195
	Golimumab	[redacted]	[redacted]	£143,381[a]	0.138
	Vedolizumab	[redacted]	[redacted]	£103,454[a]	0.337
TA342	Adalimumab	[redacted]	[redacted]	£29,933	0.000
	Infliximab	[redacted]	[redacted]	£170,401[a]	0.176
	Golimumab	[redacted]	[redacted]	£72,272[a]	0.102
	Vedolizumab	[redacted]	[redacted]	£54,142[a]	0.212
TA547	Adalimumab	[redacted]	[redacted]	£64,906	-0.032
	Infliximab	[redacted]	[redacted]	£418,880[a]	0.198
	Golimumab	[redacted]	[redacted]	£175,903[a]	0.144
	Vedolizumab	[redacted]	[redacted]	£123,157[a]	0.359
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551[a]	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.156
TRUENORTH	Vedolizumab	[redacted]	[redacted]	£440,991[a]	0.187
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.121
TA342	Vedolizumab	[redacted]	[redacted]	£197,216[a]	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.153
TA547	Vedolizumab	[redacted]	[redacted]	£517,373[a]	0.189
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.115

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: TA: technology appraisal; TNFi: tumour necrosis factor alpha inhibitor.

Table 81: Result from CcT/BsC scenario analysis

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024[a]	0.175
	Golimumab	[redacted]	[redacted]	£71,023[a]	0.101

©Celgene Ltd, a BMS Company (2022). All rights reserved

	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
CcT/BsC TA547 treatment distribution	Adalimumab	[redacted]	[redacted]	£32,191	-0.005
	Infliximab	[redacted]	[redacted]	£161,376[a]	0.168
	Golimumab	[redacted]	[redacted]	£68,073[a]	0.093
	Vedolizumab	[redacted]	[redacted]	£51,939[a]	0.198
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551[a]	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.156
CcT/BsC TA547 treatment distribution	Vedolizumab	[redacted]	[redacted]	£193,100[a]	0.164
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.150

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: BsC: best supportive care; CcT: concomitant therapy; TNFi: tumour necrosis factor alpha inhibitor.

Table 82: Result from vedolizumab SC scenario analyses

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Vedolizumab	[redacted]	[redacted]	£52,736[a]	0.205
0% patients receive SC		[redacted]	[redacted]	£68,803[a]	0.330
30% patients receive SC		[redacted]	[redacted]	£59,039[a]	0.256
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551[a]	0.170
0% patients receive SC		[redacted]	[redacted]	£1,982,556[a]	0.231
30% patients receive SC		[redacted]	[redacted]	£338,194[a]	0.196

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: TNFi: tumour necrosis factor alpha inhibitor.
Abbreviations: SC: subcutaneous; TNFi: tumour necrosis factor alpha inhibitor.

B.3.8.4 Summary of sensitivity analyses results

The impact of uncertainty and alternative inputs/assumptions in the model were explored as part of sensitivity analyses. The results of the cost-effectiveness analysis were seen to be sensitive to changes in parameters related to the proportion patients achieving sustained response or remission at maintenance and the proportion of patients receiving vedolizumab SC (Section B.3.8.2). The values used in the base case economic analysis were derived from the base case NMA (Section B.2.8) and are considered to represent the most suitable inputs available.

The results of the cost-effectiveness analysis were also seen to be sensitive to changes in the utility inputs used in the model, in particular the inclusion of utility values collected during TRUENORTH trial or the use of utility values used in TA547 (Section B.3.8.3). The utility values collected during TRUENORTH were not deemed appropriate for use in the model owing to the

©Celgene Ltd, a BMS Company (2022). All rights reserved

differences between the ‘Active UC’ health state of TRUENORTH and the ‘Active UC’ health state of the model. Further to this, the utility values collected from TRUENORTH were not aligned with literature values. The utility values used in the base-case analysis are considered to represent the most suitable model inputs.

Overall, the base-case cost-effectiveness results were found to be robust to uncertainties surrounding model inputs and assumptions.

B.3.9 *Subgroup analysis*

No further subgroup analyses were performed beyond the subpopulations informing the base case analysis: TNFi-naïve patients and TNFi-experienced patients.

B.3.10 *Validation*

B.3.10.1 Validation of cost-effectiveness analysis

The model methodology was designed to align with NICE’s preferred methods. The model was built to align with the NICE reference case, and used an NHS and PSS perspective and discount rates for cost and benefits of 3.5%.[181] The model used a lifetime time horizon in order to capture all costs and QALY gains associated with the interventions.

Economic model verification

Quality-control procedures were undertaken to ensure the programming and physical implementation of the conceptual model was completed correctly. An independent modelling team undertook a cell-by-cell verification process facilitating a check of all input calculations, formulae and Visual Basic code. Any discrepancies were identified, discussed and corrected as required.

Validation of economic model outputs against clinical expert opinion

Clinician opinion was used to conceptualise the economic model wherever possible, in order to ensure face validity of model structure, inputs and assumptions. Feedback from a Consultant in IBD in the UK was sought. Specifically, clinical feedback was used to validate the cost and resource use inputs described in the model, including inputs relating to the cost of 1st and 2nd surgery and also to validate the assumptions regarding current clinical management of UC (i.e. use of TNFis and tofacitinib in UK clinical practice). Where possible, UK source were used for model inputs and similar inputs and approaches to those used in prior appraisal were adopted.

B.3.11 *Interpretation and conclusions of economic evidence*

Summary of cost-effectiveness evidence

In order to assess the cost-effectiveness of ozanimod versus relevant comparators for the treatment of moderately to severely active UC in adult patients in the UK, a *de novo* cost-effectiveness analysis was conducted from the perspective of the NHS and PSS in England. In line with the anticipated licensed indication for ozanimod, the model considered the use of ozanimod in patients who were both naïve and experienced to prior TNFi therapy.

©Celgene Ltd, a BMS Company (2022). All rights reserved

In the base case analysis ozanimod was found to be cost-effective compared to all relevant comparators in both the TNFi-naïve and TNFi-experienced populations. In the TNFi-naïve population the NHB for ozanimod versus adalimumab, infliximab, golimumab and vedolizumab was 0.003, 0.175, 0.100 and 0.205, respectively. In the base case analysis in the TNFi-experienced population the NHB for ozanimod versus ustekinumab and vedolizumab was 0.170 and 0.156, respectively.

The PSA analyses demonstrated that the probability that ozanimod is cost-effective in the TNFi-naïve and TNFi-experienced populations is estimated to be [redacted]% and [redacted]%, respectively, at a willingness-to-pay threshold of £30,000 per QALY.

The DSA results identified a small number of key influential parameters (the proportion of ozanimod patients achieving sustained response and remission at maintenance, and the proportion of patients receiving vedolizumab SC) with the model being largely robust to uncertainty in the majority of parameters. Scenario analyses conducted to address sources of uncertainty in the model such as the inclusion of extended induction, demonstrated that whilst there was variation in the NHB, the cost-effectiveness conclusions remain largely the same, with ozanimod remaining cost-effective at a willingness-to-pay threshold of £30,000 per QALY across the vast majority of scenarios.

Overall, the base case ICERs for all comparisons demonstrated ozanimod to be cost-effective at a willingness-to-pay threshold £30,000 per QALY and thus ozanimod can be considered a cost-effective use of NHS resources in both the TNFi-naïve and TNFi-experienced populations.

Strengths

The clinical effectiveness evidence presented in this submission has been derived from an SLR of clinical trials investigating the efficacy and safety of a variety of treatment options, including ozanimod, for the treatment of moderately to severely active UC. Results from the TRUENORTH trial have demonstrated that ozanimod was associated with improved response and remission compared to placebo. Efficacy for ozanimod in the model was based on the primary and secondary outcomes of the TRUENORTH trial, which represents the primary source of evidence for ozanimod in this indication. The baseline characteristics of patients in the TRUENORTH trial were considered to be reflective of patients with moderately to severely active UC in the UK and therefore the outcomes of the TRUENORTH trial were considered generalisable to UK clinical practice. Further to this, TRUENORTH was 52-weeks in length and therefore enabled the generation of long-term evidence for ozanimod in terms of maintenance of clinical response and remission to inform the model. Maintenance of response in novel therapeutics is key in order to prevent the need for extended corticosteroid use.

An NMA was conducted to compare ozanimod to relevant comparators in clinical practice, which found ozanimod to have similar efficacy in terms of achieving and maintaining remission and response (Section B.2.9). A cost-utility analysis was selected to assess the cost-effectiveness of ozanimod in this indication. The model was built to align with the NICE reference case, adopting an NHS and PSS perspective, a lifetime time horizon to capture fully all costs and QALY gains associated with the interventions, and discount rates for costs and benefits of 3.5%.[181] As mentioned in Section B.3.10.1, the model underwent extensive internal and external validation. Clinical expert opinion was sought to validate the model structure, inputs and assumptions.

Limitations

©Celgene Ltd, a BMS Company (2022). All rights reserved

A key limitation of the clinical evidence base was the lack of head-to-head evidence for ozanimod versus relevant comparators to this appraisal (TNFis, vedolizumab and ustekinumab). An NMA was conducted in order to obtain relative efficacy estimates to inform the economic analysis, however these were subject to uncertainty due to heterogeneity between the trial populations and design incorporated in the networks, as discussed in Section B.2.8.2.

Further to this, discrepancies between the definition of active UC between the TRUENORTH trial and model and lack of available HRQoL data for the surgery health states meant literature derived utility values were used in the base case analysis (Section B.3.4.1). This limitation in utility data is in line with prior economic evaluations in UC.[1]

Finally, due to the restriction on the use of ustekinumab to patients who have failed conventional therapy or a biologic and who have failed a TNFi or for whom a TNFi cannot be tolerated or is unsuitable, and the variation in use of relevant comparators in clinical practice, the decision problem necessitated comparisons restricted by TNFi-experience.[1]

Conclusion

There remains a considerably high unmet need amongst adult patients with moderately to severely active UC in both the TNFi-naïve and TNFi-experienced populations for novel, safe, convenient therapies in order to offer patients a further treatment option before resorting to surgery. Ozanimod has demonstrated comparable efficacy to relevant comparators in UK clinical practice at achieving and maintaining remission and response (Section B.2.8) which, as demonstrated in the TRUENORTH trial, is associated with improved patient HRQoL. Ozanimod, with its novel mechanism of action, oral method of administration and tolerable safety profile could therefore offer a much-needed additional treatment option for patients with moderately to severely active UC. Overall, the base case ICERs for all comparisons demonstrated ozanimod to be cost-effective at a willingness-to-pay threshold £30,000 per QALY and thus ozanimod can be considered a cost-effective use of NHS resources in both the TNFi-naïve and TNFi-experienced populations.

©Celgene Ltd, a BMS Company (2022). All rights reserved

Appendices

©Celgene Ltd, a BMS Company (2022). All rights reserved

References

1. National Institute for Health and Care Excellence (NICE). TA633: Ustekinumab for treating moderately to severely active ulcerative colitis [ID1511] available at: https://www.nice.org.uk/guidance/ta633/documents/committee-papers-3 Last accessed: [05/11/20].
2. EMA confirms Xeljanz to be used with caution in patients at high risk of blood clots. European Medicines Agency (EMA). Available at: https://www.ema.europa.eu/en/news/ema-confirms-xeljanz-be-used-caution-patients-high-risk-blood-clots. [Last accessed: 30/06/21].
3. Ozanimod NICE Final Scope. Available at: https://www.nice.org.uk/guidance/gid-ta10732/documents/final-scope. [Last accessed: 19/11/2021].
4. Sandborn WJ, Feagan BG, D'Haens G, et al. Ozanimod as Induction and Maintenance Therapy for Ulcerative Colitis. New England Journal of Medicine 2021;385:1280-1291.
5. Celgene (a BMS Company). Data on file. TRUENORTH CSR.
6. Scott F, Clemons B, Brooks J, et al. Ozanimod (RPC1063) is a potent sphingosine-1-phosphate receptor-1 (S1P1) and receptor-5 (S1P5) agonist with autoimmune disease-modifying activity. British journal of pharmacology 2016;173:1778-1792.
7. Masayo A, Hiroaki A, Rajesh R. Hait Nitai C., Takabe Kazuaki. Sphingosine-1-Phosphate Signaling in Immune Cells and Inflammation: Roles and Therapeutic Potential. Mediators of Inflammation 2016;2016:1-11.
8. Danese S, Furfaro F, Vetrano S. Targeting S1P in inflammatory bowel disease: new avenues for modulating intestinal leukocyte migration. Journal of Crohn's and Colitis 2018;12:S678-S686.
9. Guan Q. A comprehensive review and update on the pathogenesis of inflammatory bowel disease. Journal of immunology research 2019;2019.
10. Montrose DC, Scherl EJ, Bosworth BP, et al. S1P1 localizes to the colonic vasculature in ulcerative colitis and maintains blood vessel integrity. Journal of lipid research 2013;54:843-851.
11. Zeposia Marketing Authorisation. EMA. Available at: https://ec.europa.eu/health/documents/community-register/html/h1442.htm. [Last accessed: 02/12/2021].
12. Celgene (a BMS Company). Data on file. Draft SmPC for Zeposia®.
13. da Silva BC, Lyra AC, Rocha R, et al. Epidemiology, demographic characteristics and prognostic predictors of ulcerative colitis. World J Gastroenterol 2014;20:9458-67.
14. Magro F, Rodrigues A, Vieira AI, et al. Review of the disease course among adult ulcerative colitis population-based longitudinal cohorts. Inflamm Bowel Dis 2012;18:573-83.
15. National Clinical Guideline Centre (UK). Ulcerative Colitis: Management in Adults, Children and Young People. London: Royal College of Physicians (UK). 2013. (NICE Clinical Guidelines, No. 166.) 1, Introduction. Available at: https://www.ncbi.nlm.nih.gov/books/NBK258842/. [Last accessed: 18/01/21].
16. National Institute for Health and Care Excellence (NICE). NG130. Ulcerative colitis: management. Available at: https://www.nice.org.uk/guidance/ng130/resources/ulcerative-colitis-management-pdf-66141712632517. [Last accessed: 10/11/20].
17. King D, Reulen RC, Thomas T, et al. Changing patterns in the epidemiology and outcomes of inflammatory bowel disease in the United Kingdom: 2000-2018. Alimentary pharmacology & therapeutics 2020;51:922-934.
18. Truelove SC, Witts LJ. Cortisone in Ulcerative Colitis: final report on a therapeutic trial. British Medical Journal 1955;2:1041-1048.
19. Canadian Agency for Drugs and Technologies in Health. Clinical Review Report: adalimumab (Humira) 2016. Appendix 5, Validity of Outcome Measures. Available at: https://www.ncbi.nlm.nih.gov/books/NBK539018/?report=classic. [Last accessed:19/01/21].

©Celgene Ltd, a BMS Company (2022). All rights reserved

20. Schroeder KW, Tremaine WJ, Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. N Engl J Med 1987;317:1625-9.
21. Stange EF, Travis SP, Vermeire S, et al. European evidence-based Consensus on the diagnosis and management of ulcerative colitis: Definitions and diagnosis. J Crohns Colitis 2008;2:1-23.
22. Mayo Clinic. Ulcerative Colitis Symptoms. Available at: https://www.mayoclinic.org/diseases-conditions/ulcerative-colitis/symptoms-causes/syc-20353326. [Last accessed: 10/11/20].
23. 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;11:649-670.
24. Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med 2013;369:699-710.
25. Macaluso FS, Maida M, Ventimiglia M, et al. Factors Affecting Clinical and Endoscopic Outcomes of Placebo Arm in Trials of Biologics and Small Molecule Drugs in Ulcerative Colitis: A Meta-Analysis. Inflammatory Bowel Diseases 2018;25:987-997.
26. Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab maintains clinical response in patients with moderate-to-severe ulcerative colitis. Gastroenterology 2014;146:96-109.e1.
27. Magro F, Portela F, Lago P, et al. Inflammatory bowel disease: a patient's and caregiver's perspective. Digestive diseases and sciences 2009;54:2671-2679.
28. Mohsenizadeh S, Manzari Z, Vosoghinia H, et al. Family caregivers burden in inflammatory bowel diseases: An integrative review. Journal of Education and Health Promotion 2020;9:289-289.
29. Parekh NK, Shah S, McMaster K, et al. Effects of caregiver burden on quality of life and coping strategies utilized by caregivers of adult patients with inflammatory bowel disease. Annals of gastroenterology 2017;30:89-95.
30. Harbord M, Eliakim R, Bettenworth D, 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;11:769-784.
31. Waljee A, Waljee J, Morris AM, et al. Threefold increased risk of infertility: a meta-analysis of infertility after ileal pouch anal anastomosis in ulcerative colitis. Gut 2006;55:1575-80.
32. Brown C, Gibson PR, Hart A, et al. Long-term outcomes of colectomy surgery among patients with ulcerative colitis. SpringerPlus 2015;4:573.
33. 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;142:257-265.e251-253.
34. Ulcerative Colitis. IBD Clinic. Available at: http://www.ibdclinic.ca/what-is-ibd/ulcerative-colitis/. Last accessed: [09/11/20].
35. National Institute for Health and Care Excellence (NICE). TA329: Infliximab, adalimumab and golimumanb for treating moderately to severely active ulcerative colitis after the failure of conventional therapy. available at: https://www.nice.org.uk/guidance/ta329/resources/infliximab-adalimumab-and-golimumab-for-treating-moderately-to-severely-active-ulcerative-colitis-after-the-failure-of-conventional-therapy-pdf-82602495307717. Last accessed [13/11/20].
36. Liverani E, Scaioli E, Digby RJ, et al. How to predict clinical relapse in inflammatory bowel disease patients. World J Gastroenterol 2016;22:1017-33.
37. Lamb CA, Kennedy NA, Raine T, et al. British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. Gut 2019;68:s1.
38. Rubin D, Panés J, Lindsay JO, et al. P558 The SF-36® Health Survey distinguishes disease burden on functioning and well-being between patients with active vs inactive ulcerative colitis. Journal of Crohn's and Colitis 2017;11:S364-S365.

©Celgene Ltd, a BMS Company (2022). All rights reserved

39. Bamias G, Kaltsa G, Ladas S. Cytokines in the Pathogenesis of Ulcerative Colitis. Discovery medicine 2011;11:459-67.
40. Ungaro R, Mehandru S, Allen PB, et al. Ulcerative colitis. Lancet 2017;389:1756-1770.
41. Strober W, Fuss I, Mannon P. The fundamental basis of inflammatory bowel disease. J Clin Invest 2007;117:514-21.
42. National Institute for Health and Care Excellence (NICE). NG130 Full Guideline. Ulcerative Colitis: Management in adults, children and young people. Available at: https://www.nice.org.uk/guidance/ng130/evidence/full-guideline-pdf-6777262190. [Last accessed: 10/11/20].
43. The World Bank. Population, total - United Kingdom. Available at: https://data.worldbank.org/indicator/SP.POP.TOTL?name_desc=true&locations=GB. [Last accessed 17/11/20].
44. Ulcerative Colitis and Proctitis. BonheurMD. Available at: https://www.bonheurmd.com/ulcerative-colitis-and-proctitis/#:~:text=The%20usual%20age%20of%20diagnosis%20with%20ulcerative%20colitis,in%20non-Jews.%20IBD%20tends%20to%20run%20in%20families.. [Last acessed: 18/01/21].
45. The fact about inflammatory bowel disease. Crohn's and Colitis Foundation of America. Available at: ps://www.crohnscolitisfoundation.org/sites/default/files/legacy/assets/pdfs/updatedibdfactbook.pdf. [Last accessed: 18/01/21].
46. Overview of ulcerative colitis. Crohn's and Colitis Foundation. Available at: https://www.crohnscolitisfoundation.org/what-is-ulcerative-colitis/overview. [Last accessed: 18/01/21].
47. Alexakis C, Davies G, Stephens J, et al. Perspectives and attitudes of young patients with inflammatory bowel disease: symptoms, burden of disease and communication with their healthcare professionals. Frontline gastroenterology 2014;5:197-202.
48. Opheim R, Fagermoen MS, Bernklev T, et al. Fatigue interference with daily living among patients with inflammatory bowel disease. Qual Life Res 2014;23:707-17.
49. Vaizey CJ, Gibson PR, Black CM, et al. Disease status, patient quality of life and healthcare resource use for ulcerative colitis in the UK: an observational study. Frontline Gastroenterol 2014;5:183-189.
50. Pallis AG, Mouzas IA, Vlachonikolis IG. The Inflammatory Bowel Disease Questionnaire: A Review of Its National Validation Studies. Inflammatory Bowel Diseases 2004;10:261-269.
51. McMullan C, Pinkney TD, Jones LL, et al. Adapting to ulcerative colitis to try to live a 'normal' life: a qualitative study of patients' experiences in the Midlands region of England. BMJ Open 2017;7:e017544.
52. Research Partnership. Living with ulcerative colitis US 2017. A syndicated patient report. Available at: https://www.researchpartnership.com/solutions/syndicated/living-with/living-with-ulcerative-colitis-2020-report [Last accessed: December 2021].
53. 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. Journal of Crohn's and Colitis 2007;1:10-20.
54. Treatment satisfaction, preferences and perception gaps between patients and physicians in the ulcerative colitis CARES study: A real world-based study. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2016b; 48(6): 601-60.
55. Sajadinejad MS, Asgari K, Molavi H, et al. Psychological Issues in Inflammatory Bowel Disease: An Overview. Gastroenterology Research and Practice 2012;2012:106502.
56. Willian MK, D'Haens G, Yarlas A, et al. Changes in health-related quality of life and work-related outcomes for patients with mild-to-moderate ulcerative colitis receiving short-term and long-term treatment with multimatrix mesalamine: a prospective, open-label study. J Patient Rep Outcomes 2018;2:22.
57. Larsson K, Lööf L, Nordin K. Stress, coping and support needs of patients with ulcerative colitis or Crohn's disease: a qualitative descriptive study. J Clin Nurs 2017;26:648-657.

©Celgene Ltd, a BMS Company (2022). All rights reserved

58. Lönnfors S, Vermeire S, Avedano L. IBD and health-related quality of life — Discovering the true impact ☆. Journal of Crohn's and Colitis 2014;8:1281-1286.
59. Zand A, Kim BJ, van Deen WK, et al. The effects of inflammatory bowel disease on caregivers: significant burden and loss of productivity. BMC health services research 2020;20:1-10.
60. Vavricka SR, Brun L, Ballabeni P, et al. Frequency and Risk Factors for Extraintestinal Manifestations in the Swiss Inflammatory Bowel Disease Cohort. Official journal of the American College of Gastroenterology | ACG 2011;106.
61. Vavricka SR, Schoepfer A, Scharl M, et al. Extraintestinal Manifestations of Inflammatory Bowel Disease. Inflamm Bowel Dis 2015;21:1982-92.
62. Jess T, Rungoe C, Peyrin–Biroulet L. Risk of Colorectal Cancer in Patients With Ulcerative Colitis: A Meta-analysis of Population-Based Cohort Studies. Clinical Gastroenterology and Hepatology 2012;10:639-645.
63. Rutter M, Saunders B, Wilkinson K, et al. Severity of inflammation is a risk factor for colorectal neoplasia in ulcerative colitis. Gastroenterology 2004;126:451-9.
64. Trikudanathan G, Venkatesh PGK, Navaneethan U. Diagnosis and Therapeutic Management of Extra-Intestinal Manifestations of Inflammatory Bowel Disease. Drugs 2012;72:2333-2349.
65. Peyrin-Biroulet L, Germain A, Patel AS, et al. Systematic review: outcomes and post-operative complications following colectomy for ulcerative colitis. Alimentary Pharmacology & Therapeutics 2016;44:807-816.
66. 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 Outcomes Res 2017;9:641-652.
67. Farmer RG, Easley KA, Rankin GB. Clinical patterns, natural history, and progression of ulcerative colitis. A long-term follow-up of 1116 patients. Dig Dis Sci 1993;38:1137-46.
68. Torres J, Billioud V, Sachar DB, et al. Ulcerative colitis as a progressive disease: the forgotten evidence. Inflamm Bowel Dis 2012;18:1356-63.
69. Van Assche G, Peyrin-Biroulet L, Sturm A, 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;48:592-600.
70. Sammut J, Scerri J, Xuereb RB. The lived experience of adults with ulcerative colitis. J Clin Nurs 2015;24:2659-67.
71. Ding Z, Obando C, Izanec J, et al. Relationship between partial mayo score and patient-reported outcomes for ulcerative colitis patients in the United States. Am J Gastroenterol 2019;114:S417-S418.
72. COHEN RD, YU AP, WU EQ, et al. Systematic review: the costs of ulcerative colitis in Western countries. Alimentary Pharmacology & Therapeutics 2010;31:693-707.
73. Black CM, Yu E, McCann E, et al. 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:e0149692.
74. Singh H, Wilson APL, Tencer T, et al. 107 REAL-WORLD DOSE ESCALATION IN PATIENTS WITH MODERATE-TO-SEVERE ULCERATIVE COLITIS: A SYSTEMATIC REVIEW OF LITERATURE. Gastroenterology 2021;160:S-26.
75. Feuerstein JD, Cheifetz AS. Ulcerative Colitis: Epidemiology, Diagnosis, and Management. Mayo Clinic Proceedings 2014;89:1553-1563.
76. 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.
77. Satsangi J, Silverberg MS, Vermeire S, et al. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. Gut 2006;55:749-53.
78. Kornbluth A, Sachar DB. Ulcerative colitis practice guidelines in adults: American College Of Gastroenterology, Practice Parameters Committee. Am J Gastroenterol 2010;105:501-23; quiz 524.

©Celgene Ltd, a BMS Company (2022). All rights reserved

79. Hopkins Medicine. Ulcerative colits. Available from: https://www.hopkinsmedicine.org/gastroenterology_hepatology/_pdfs/small_large_intestine/ulcerative_colitis.pdf. [Last accessed: 09/11/20].
80. Neurath MF, Travis SPL. Mucosal healing in inflammatory bowel diseases: a systematic review. Gut 2012;61:1619.
81. Feakins RM. Inflammatory bowel disease biopsies: updated British Society of Gastroenterology reporting guidelines. Journal of clinical pathology 2013;66:1005-1026.
82. Gajendran M, Loganathan P, Jimenez G, et al. A comprehensive review and update on ulcerative colitis. Disease-a-Month 2019;65:100851.
83. Danese S, Fiocchi C. Ulcerative colitis. N Engl J Med 2011;365:1713-25.
84. 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). Scand J Gastroenterol 2009;44:431-40.
85. National Institute for Health and Care Excellence (NICE). TA342: Vedolizumab for treating moderately to severely active ulcerative colitis. Available at: https://www.nice.org.uk/guidance/ta342. Last accessed: [14/12/20].
86. Taylor Meadows KR, Steinberg MW, Clemons B, et al. Ozanimod (RPC1063), a selective S1PR1 and S1PR5 modulator, reduces chronic inflammation and alleviates kidney pathology in murine systemic lupus erythematosus. PLoS One 2018;13:e0193236.
87. Lewis JD, Chuai S, Nessel L, et al. Use of the noninvasive components of the mayo score to assess clinical response in Ulcerative Colitis. Inflammatory Bowel Diseases 2008;14:1660-1666.
88. National Institute for Health and Care Excellence (NICE). TA547: Tofacitinib for treating moderately to severely active ulcerative colitis. Available at https://www.nice.org.uk/guidance/ta547. [Last accessed: 24/11/20].
89. D'Haens G, Sandborn WJ, Feagan BG, et al. A Review of Activity Indices and Efficacy End Points for Clinical Trials of Medical Therapy in Adults With Ulcerative Colitis. Gastroenterology 2007;132:763-786.
90. Pai RK, Jairath V, Casteele NV, et al. The emerging role of histologic disease activity assessment in ulcerative colitis. Gastrointestinal endoscopy 2018;88:887-898.
91. Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: an update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) initiative of the International Organization for the Study of IBD (IOIBD): determining therapeutic goals for treat-to-target strategies in IBD. Gastroenterology 2021;160:1570-1583.
92. Rubin DT, Ananthakrishnan AN, Siegel CA, et al. ACG clinical guideline: ulcerative colitis in adults. American Journal of Gastroenterology 2019;114:384-413.
93. Kim DB, Lee K-M, Lee JM, et al. Correlation between histological activity and endoscopic, clinical, and serologic activities in patients with ulcerative colitis. Gastroenterology research and practice 2016;2016.
94. Geboes K, Riddell R, Ost A, et al. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. Gut 2000;47:404-9.
95. Torres J, Bonovas S, Doherty G, et al. ECCO Guidelines on Therapeutics in Crohn's Disease: Medical Treatment. Journal of Crohn's and Colitis 2019;14:4-22.
96. Heyman MB, Kirschner BS, Gold BD, et al. Children with early-onset inflammatory bowel disease (IBD): analysis of a pediatric IBD consortium registry. The Journal of pediatrics 2005;146:35-40.
97. Sandhu BK, Fell JM, Beattie RM, et al. Guidelines for the management of inflammatory bowel disease in children in the United Kingdom. Journal of Pediatric Gastroenterology and Nutrition 2010;50:S1-S13.
98. Feuerstein JD, Moss AC, Farraye FA. Ulcerative Colitis. Mayo Clinic Proceedings 2019;94:1357-1373.
99. Early Intervention with Anti-Tumour Necrosis Factor in Ulcerative Colitis: the Missing Piece of the Puzzle? Available from: https://www.emjreviews.com/gastroenterology/symposium/early-intervention-with-anti-tumour-necrosis-factor-in-ulcerative-colitis-the-missing-piece-of-the-puzzle/. [Last accessed: 14/04/21].

©Celgene Ltd, a BMS Company (2022). All rights reserved

100. Ulcerative Colitis Treatment. John Hopkin's Medicine. Available from: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/ulcerative-colitis-treatment. [Last accessed: 14/04/21].
101. Adult Drugs Treatment Pathway for Moderately to Severey Active Ulcerative Colitis (UC)-based on NICE TAs. NHS Herts Valleys. Available at: https://www.enhertsccg.nhs.uk/sites/default/files/content_files/Prescribing/Local_Decisions/Gastro_Intestinal_System/Ulcerative%20colitis%20in%20adults%20%E2%80%93%20Drug%20treatment%20pathway%20for%20moderately%20to%20severely%20active%20ulcerative%20colitis.pdf. [Last accessed: 16/11/21].
102. Hendy P, Hart A, Irving P. Anti-TNF drug and antidrug antibody level monitoring in IBD: a practical guide. Frontline Gastroenterology 2016;7:122.
103. Sandborn WJ. Tailoring Treatment to the Individual Patient: Drug Monitoring. Digestive Diseases 2014;32(suppl 1):88-91.
104. FDA Issues Drug Safety Communication Related to Current XELJANZ Label. Pfizer. Available at: https://cdn.pfizer.com/pfizercom/2021-09/Xeljanz-Label-Press-Statement-Final.pdf?YdYcCWSVWP4qILXeMQ1vI.bxpTYQOub4. [Last accessed: 09/12/21].
105. Nielsen OH, Ainsworth MA. Tumor Necrosis Factor Inhibitors for Inflammatory Bowel Disease. New England Journal of Medicine 2013;369:754-762.
106. Olesen CM, Coskun M, Peyrin-Biroulet L, et al. Mechanisms behind efficacy of tumor necrosis factor inhibitors in inflammatory bowel diseases. Pharmacol Ther 2016;159:110-9.
107. Papamichael K, Rivals-Lerebours O, Billiet T, et al. Long-term outcome of patients with ulcerative colitis and primary non-response to infliximab. Journal of Crohn's and Colitis 2016;10:1015-1023.
108. Danese S, Allez M, van Bodegraven AA, et al. Unmet Medical Needs in Ulcerative Colitis: An Expert Group Consensus. Dig Dis 2019;37:266-283.
109. Danese S, Allez M, van Bodegraven AA, et al. Unmet Medical Needs in Ulcerative Colitis: An Expert Group Consensus. Digestive Diseases 2019;37:266-283.
110. Boland B, Sandborn W, Chang J. Update on Janus Kinase Antagonists in Inflammatory Bowel Disease. Gastroenterology Clinics of North America 2014;43.
111. Yanai H, Hanauer S. Assessing Response and Loss of Response to Biological Therapies in IBD. The American journal of gastroenterology 2011;106:685-98.
112. Vermeire S, Gils A, Accossato P, et al. Immunogenicity of biologics in inflammatory bowel disease. Therap Adv Gastroenterol 2018;11:1756283x17750355.
113. Lindsay JO, Armuzzi A, Gisbert JP, et al. Indicators of suboptimal tumor necrosis factor antagonist therapy in inflammatory bowel disease. Digestive and Liver Disease 2017;49:1086-1091.
114. Fausel R, Afzali A. Biologics in the management of ulcerative colitis - comparative safety and efficacy of TNF-α antagonists. Therapeutics and clinical risk management. Volume 11, 2015:63-73.
115. Patel H, Lissoos T, Rubin DT. Indicators of suboptimal biologic therapy over time in patients with ulcerative colitis and Crohn's disease in the United States. PLoS ONE 2017;12:e0175099.
116. Rostholder E, Ahmed A, Cheifetz AS, et al. Outcomes after escalation of infliximab therapy in ambulatory patients with moderately active ulcerative colitis. Aliment Pharmacol Ther 2012;35:562-7.
117. Testa A, Castiglione F, Nardone OM, et al. Adherence in ulcerative colitis: an overview. Patient Prefer Adherence 2017;11:297-303.
118. Kane SV. Systematic review: adherence issues in the treatment of ulcerative colitis. Aliment Pharmacol Ther 2006;23:577-85.
119. Denesh D, Carbonell J, Kane JS, et al. Patients with inflammatory bowel disease (IBD) prefer oral tablets over other modes of medicine administration. Expert review of gastroenterology & hepatology 2021:1-6.
120. Kennedy NA, Jones G-R, Lamb CA, et al. British Society of Gastroenterology guidance for management of inflammatory bowel disease during the COVID-19 pandemic. Gut 2020;69:984-990.

©Celgene Ltd, a BMS Company (2022). All rights reserved

121. Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. New England Journal of Medicine 2017;376:1723-1736.
122. Reinisch W, Sandborn WJ, Hommes DW, et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. Gut 2011;60:780.
123. Rutgeerts P, Sandborn WJ, Feagan BG, et al. Infliximab for induction and maintenance therapy for ulcerative colitis. New England Journal of Medicine 2005;353:2462-2476.
124. Papp KA, Fonjallaz P, Casset-Semanaz F, et al. Analytical approaches to reporting long-term clinical trial data. Current medical research and opinion 2008;24:2001-2008.
125. Theede K, Kiszka-Kanowitz M, Nordgaard-Lassen I, et al. The Impact of Endoscopic Inflammation and Mucosal Healing on Health-related Quality of Life in Ulcerative Colitis Patients. J Crohns Colitis 2015;9:625-32.
126. Sandborn, W., D'Haens G, Wolf DC, et al. Ozanimod as maintenance therapy in patients with moderate-to-severe ulcerative colitis: Results from the Phase III, randomized, double-blind, placebo-controlled TRUENORTH study [Oral presentation: LB10]. Presented at: UEG Week Virtual 2020; October 11-13, 2020. 2020b. Available at: https://journals.sagepub.com/doi/full/10.1177/2050640620968709. [Last accessed:12/07/21].
127. Reinisch W, Sandborn WJ, Bala M, 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 2007;13:1135-1140.
128. CRD's guidance for undertaking reviews in health care: University of York, 2009.
129. ClinicalTrials.gov. Study to Evaluate the Safety and Efficacy of Two Adalimumab Dosing Regimens in Subjects With Moderate to Severe Ulcerative Colitis.
130. Probert CS, Hearing SD, Schreiber S, et al. Infliximab in moderately severe glucocorticoid resistant ulcerative colitis: a randomised controlled trial. Gut 2003;52:998-1002.
131. Sands BE, Tremaine WJ, Sandborn WJ, et al. Infliximab in the treatment of severe, steroid-refractory ulcerative colitis: a pilot study. Inflammatory Bowel Diseases 2001;7:83-88.
132. Sarl CII. Phase 3, multicenter, randomized, double-blind, placebo-controlled trial of oral RPC1063 as induction and maintenance therapy for moderate to severe ulcerative colitis - TRUE NORTH, 2020.
133. Sandborn WJ, Feagan BG, Wolf DC, et al. Ozanimod Induction and Maintenance Treatment for Ulcerative Colitis. New England Journal of Medicine 2016;374:1754-1762.
134. Reinisch W, Sandborn WJ, Hommes DW, et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. Gut 2011;60:780-787.
135. 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;49:283-294.
136. Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab induces clinical response and remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology 2014;146:85-95; quiz e14-85.
137. 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;52:1101-1111.
138. 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;146:392-400.e393.
139. 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;49:582-588.
140. 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

©Celgene Ltd, a BMS Company (2022). All rights reserved

prospective randomized controlled trial and its post hoc analysis. J Gastroenterol 2016;51:241-51.
141. Sands BE, Peyrin-Biroulet L, Loftus EV, et al. Vedolizumab versus Adalimumab for Moderate-to-Severe Ulcerative Colitis. New England Journal of Medicine 2019;381:1215-1226.
142. Sandborn WJ, Baert F, Danese S, et al. Efficacy and Safety of Vedolizumab Subcutaneous Formulation in a Randomized Trial of Patients With Ulcerative Colitis. Gastroenterology 2020;158:562-572.e512.
143. Motoya S, Watanabe K, Ogata H, et al. Vedolizumab in Japanese patients with ulcerative colitis: A Phase 3, randomized, double-blind, placebo-controlled study. PLoS ONE [Electronic Resource] 2019;14:e0212989.
144. Sands BE, Sandborn WJ, Panaccione R, et al. Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis. New England Journal of Medicine 2019;381:1201-1214.
145. Sandborn WJ, Ghosh S, Panes J, et al. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. New England Journal of Medicine 2012;367:616-624.
146. Gordon JP, McEwan PC, Maguire A, et al. Characterizing unmet medical need and the potential role of new biologic treatment options in patients with ulcerative colitis and Crohn's disease: a systematic review and clinician surveys. European journal of gastroenterology & hepatology 2015;27:804.
147. Hoentjen F, Haarhuis BJ, Drenth JP, et al. Elective switching from infliximab to adalimumab in stable Crohn's disease. Inflammatory bowel diseases 2013;19:761-766.
148. Patil SA, Rustgi A, Langenberg P, et al. Comparative effectiveness of anti-TNF agents for Crohn's disease in a tertiary referral IBD practice. Digestive diseases and sciences 2013;58:209-215.
149. Van Assche G, Vermeire S, Ballet V, et al. Switch to adalimumab in patients with Crohn's disease controlled by maintenance infliximab: prospective randomised SWITCH trial. Gut 2012;61:229-234.
150. NICE. Ustekinumab for treating moderately to severely active ulcerative colitis [ID1511]; Single Technology Appraisal, 2020.
151. NICE. Tofacitinib for previously treated active ulcerative colitis [ID1218]; Single Technology Appraisal, 2018.
152. Dubinsky MC, Peyrin-Biroulet L, Melmed GY, et al. Efficacy of Tofacitinib in Patients With Ulcerative Colitis by Prior Tumor Necrosis Factor Inhibitor Treatment Status: Results From OCTAVE Induction and Maintenance Studies: 640. Official journal of the American College of Gastroenterology | ACG 2017;112.
153. Gottlieb K, Daperno M, Usiskin K, et al. Endoscopy and central reading in inflammatory bowel disease clinical trials: achievements, challenges and future developments. Gut 2021;70:418-426.
154. 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. 2011.
155. ICER. Targeted Immune Modulators for Ulcerative Colitis: Effectiveness and Value; Final Evidence Report and Meeting Summary, 2020.
156. Dias S, Sutton AJ, Welton NJ, et al. NICE DSU Technical Support Document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment. 2011.
157. Vehtari A, Gelman A, Simpson D, et al. Rank-normalization, folding, and localization: An improved $\widehat {R} $ for assessing convergence of MCMC. 2019.
158. Schreiber S, Peyrin-Biroulet L, Loftus EV, et al. VARSITY: A double-blind, double-dummy, randomised, controlled trial of vedolizumab versus adalimumab in patients with active ulcerative colitis. Journal of Crohn's and Colitis 2019;13 (Supplement 1):S612-S613.
159. NICE DSU TECHNICAL SUPPORT DOCUMENT 18: METHODS FOR POPULATION-ADJUSTED INDIRECT COMPARISONS IN SUBMISSIONS TO NICE. NICE DSU. Available at: http://nicedsu.org.uk/wp-content/uploads/2018/08/Population-adjustment-TSD-FINAL-ref-rerun.pdf [Last accessed: 24/07/21].

©Celgene Ltd, a BMS Company (2022). All rights reserved

160. Cohen JA, Comi G, Arnold DL, et al. Efficacy and safety of ozanimod in multiple sclerosis: Dose-blinded extension of a randomized phase II study. Mult Scler 2019;25:1255-1262.
161. Cohen JA, Comi G, Selmaj KW, et al. Safety and efficacy of ozanimod versus interferon beta-1a in relapsing multiple sclerosis (RADIANCE): a multicentre, randomised, 24-month, phase 3 trial. Lancet Neurol 2019;18:1021-1033.
162. Alcala M, Sulleiro S, Hernando T, et al. Efficacy of Ustekinumab at the end of maintenance in Ulcerative Colitis patients receiving 6 mh/kg induction posology. Data from UNIFI trial. United European Gastroenterology Journal 2020;8:414.
163. 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;146:392-400.e3.
164. Ramos GP, Papadakis KA. Mechanisms of disease: inflammatory bowel diseases, In Mayo Clinic Proceedings, Elsevier, 2019.
165. Pasvol TJ, Horsfall L, Bloom S, et al. Incidence and prevalence of inflammatory bowel disease in UK primary care: a population-based cohort study. BMJ open 2020;10:e036584.
166. Higgins JPT GS. Cochrane handbook for systematic reviews of interventions version 5.1.0. Chichester (UK): John Wiley & Sons, 2011.
167. Moher D, Liberati A. A., Tetzlaff, J., & Altman, DG (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ;339:b2535.
168. Cosnes J, Gower-Rousseau C, Seksik P, et al. Epidemiology and Natural History of Inflammatory Bowel Diseases. Gastroenterology 2011;140:1785-94.
169. Cottone M, Scimeca D, Mocciaro F, et al. Clinical course of ulcerative colitis. Digestive and Liver Disease 2008;40:S247-S252.
170. Mehta SJ, Silver AR, Lindsay JO. Review article: strategies for the management of chronic unremitting ulcerative colitis. Alimentary Pharmacology & Therapeutics 2013;38:77-97.
171. Surgery for ulcerative colitis. Crohn's & Colitis UK> Available from: https://www.crohnsandcolitis.org.uk/about-crohns-and-colitis/publications/surgery-for-ulcerative-colitis. [Last accessed: 30/06/21].
172. Cochrane Handbook for Systematic Reviews of Interventions version 6.1. In: Julian Higgins JT, Jacqueline Chandler, Miranda Cumpston, Tianjing Li, Matthew Page, Vivian Welch, ed. Volume 2020.
173. National Institute for Health and Care Excellence (NICE). Single technology appraisal: User guide for company evidence submission template. Available from: https://www.nice.org.uk/process/pmg24/chapter/instructions-for-companies. [Last accessed: 02/03/21].
174. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Annals of internal medicine 2015;162:777-784.
175. Punekar YS, Hawkins N. Cost-effectiveness of infliximab for the treatment of acute exacerbations of ulcerative colitis. Eur J Health Econ 2010;11:67-76.
176. Lohan C, Diamantopoulos A, LeReun C, et al. Tofacitinib for the treatment of moderately to severely active ulcerative colitis: a systematic review, network meta-analysis and economic evaluation. BMJ Open Gastroenterol 2019;6:e000302.
177. 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;34:1023-38.
178. Williams JG, Alam MF, Alrubaiy L, et al. Comparison Of iNfliximab and ciclosporin in STeroid Resistant Ulcerative Colitis: pragmatic randomised Trial and economic evaluation (CONSTRUCT). Health Technol Assess 2016;20:1-320.
179. Diamantopoulos A, Lohan C, LeReun C, et al. An Economic Evaluation of Tofacitinib in the Treatment of Moderately to Severely Active Ulcerative Colitis in the United Kingdom. Value Health 2019;22:S621.

©Celgene Ltd, a BMS Company (2022). All rights reserved

180. National Institute of Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013: The reference case. Available at: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case [Last accessed: 9th October 2020].
181. National Institute for Health and Care Excellence (NICE). NICE reference case, Guide to the methods of technology appraisal 2013. 2013.
182. Woehl A ABH, and P. McEwan,. The Relation Between Disease Activity, Quality of Life and Health Utility in Patients With Ulcerative Colitis. Gut 2008;57:153-154.
183. Arseneau KO, Sultan S, Provenzale DT, et al. Do Patient Preferences Influence Decisions on Treatment for Patients With Steroid-Refractory Ulcerative Colitis? Clinical Gastroenterology and Hepatology 2006;4:1135-1142.
184. Swinburn P, Elwick H, Bean K, et al. The Impact of Surgery on Health Related Quality of Life in Ulcerative Colitis, In Gut Conference: Digestive Disorders Federation Meeting, Liverpool, UK, 2012.
185. Tsai HH, Punekar YS, Morris J, et al. A model of the long-term cost effectiveness of scheduled maintenance treatment with infliximab for moderate-to-severe ulcerative colitis. Aliment Pharmacol Ther 2008;28:1230-9.
186. 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;5:301-316.
187. 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. Can J Gastroenterol Hepatol 2016;2016:8723949.
188. Royal College of Physicians. National clinical audit of inpatient care for adults with ulcerative colitis, UK inflammatory bowel disease (IBD) audit, National Report, 2014.
189. Segal JP, McLaughlin SD, Faiz OD, et al. Incidence and Long-term Implications of Prepouch Ileitis: An Observational Study. Diseases of the Colon & Rectum 2018;61:472-475.
190. Office for National Statistics. Population Estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2019. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2019. [Last accessed: 14/12/2020].
191. Jess T, Gamborg M, Munkholm P, et al. Overall and cause-specific mortality in ulcerative colitis: meta-analysis of population-based inception cohort studies. Am J Gastroenterol 2007;102:609-17.
192. BMS. Data on file. TRUENORTH trial protocol. Version 5. , May 2018.
193. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35:1095-108.
194. Van Hout B, Janssen M, Feng Y-S, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in health 2012;15:708-715.
195. Stevenson M, Archer R, Tosh J, et al. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying antirheumatic drugs and after the failure of conventional disease-modifying antirheumatic drugs only: systematic review and economic evaluation. Health Technol Assess 2016;20:1-610.
196. Entyvio (2014). Vedolizumab summary of product characteristics. Denmark, Takeda Pharma A/S. 2020.
197. NHS Schedule costs 2019/2020. NHS England. Available at: https://www.england.nhs.uk/national-cost-collection/. [Last accessed: 27/07/21].
198. The NICE methods of HTA evalusation: the case for change 2020. Available at: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.nice.org.uk%2FMedia%2FDefault%2FAbout%2Fwhat-we-do%2Four-programmes%2Fnice-guidance%2Fchte-methods-consultation%2FNICE-methods-of-health-technology-evaluation-case-for-change.docx&wdOrigin=BROWSELINK. [Last accessed: 10/12/21].

©Celgene Ltd, a BMS Company (2022). All rights reserved

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Ozanimod for treating moderately to severely active ulcerative colitis [ID3841]

Clarification questions

2nd February 2022

File name	Version	Contains confidential information	Date
ID3841 Ozanimod_Response to ERG Clarification Questions_ACIC	1.0	ACIC	23rd February 2022

Section A: Clarification on effectiveness data

Literature searches

A1. Appendix D.1 states “an additional SLR was conducted to identify non-randomised controlled trials of ozanimod for patients with moderately to severely active UC which yielded no results”. Please provide further details of these searches.

A systematic literature review (SLR) was conducted to identify evidence from non-randomised or non-controlled studies investigating the use of ozanimod for the treatment of patients with ulcerative colitis (UC). Full details of the searches conducted are provided in Appendix 1, and an overview is provided below.

The SLR was conducted following current best practices, as recommended by the Cochrane Collaboration[1] and in accordance with the methodological principles of conduct for systematic reviews as detailed in the University of York's Centre for Reviews and Dissemination Handbook.[2]

The electronic database searches were undertaken on 29th October 2020 and databases were searched from inception. The conference proceedings of major gastroenterology congresses from the last two years (i.e. January 2018 to October 2020) were manually hand-searched in October 2020. The exclusion of abstracts from conferences prior to 2018 was justified under the assumption that high-quality research would since have been published in a peer-reviewed journal.

Additional supplementary searches included querying the ClinicalTrials.gov website on 3rd November 2020 and hand-searching the bibliographies of any relevant SLRs identified during the course of the SLR. Full details of search strategies including search terms used can be found in Appendix 1.

Studies were selected based on the inclusion criteria provided in Table 1. Studies were selected for inclusion in two stages: first, the titles and abstracts of the search results were reviewed for relevance. Second, the full texts of potentially relevant articles were screened in order to obtain the final list of included studies.

Table 1: Eligibility criteria for the SLR

Domain	Inclusion criteria	Exclusion criteria
Patient population	Patients with UC	• Patients without UC or with multiple conditions • Studies with mixed populations that do not present results for patients with UC separately
Intervention	Ozanimod (Zeposia®)	Studies not investigating ozanimod
Comparator	Any or none	NA
Outcomes	Clinical, safety, HRQoL and other patient burden outcomes, including economic: • Clinical remission	Data not related to clinical, safety or quality of life outcomes

	• Clinical response • Endoscopic improvement / mucosal healing • Histologic remission • Steroid-free remission • Durable remission • Adverse events • Serious adverse events • AEs leading to discontinuation • Specific AE categories • Health-related quality of life outcomes • Patient-reported outcomes • Resource use • Productivity • Mortality	
Study design	• Non-randomised, prospective clinical studies • Uncontrolled clinical studies • Database/registry studies • Case series • Cross-sectional studies • Cohort studies (prospective or retrospective)	• RCTs • Narrative reviews • Comments • Editorials • Case reports (n-of-1) • Pharmacokinetic studies
	Systematic reviews and NMAs were considered relevant at the title/abstract review stage and hand searched for relevant primary studies, but were excluded during the full-text review stage unless they presented primary research	
Language	Abstract or full-text articles in English language	Non-English language
Other	Studies on human subjects	Animal studies

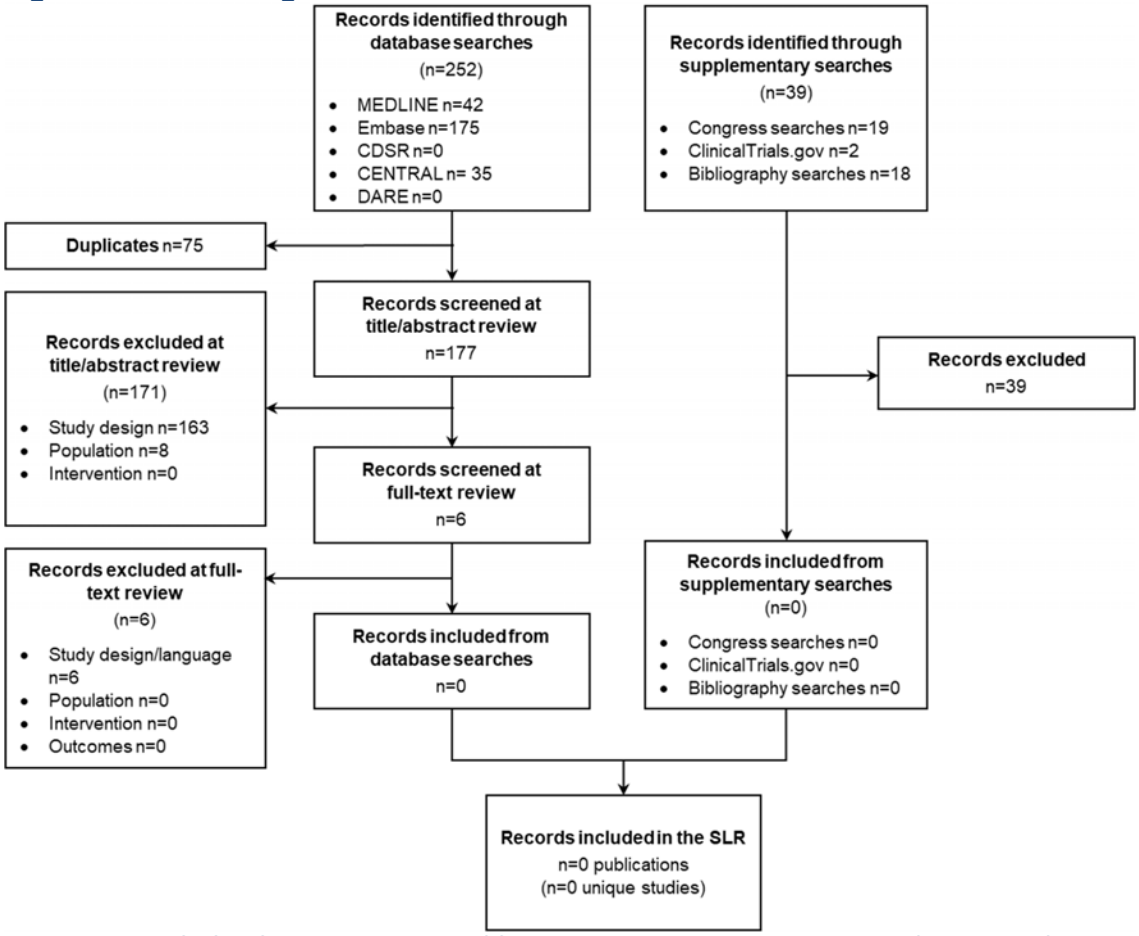
Abbreviations: AE: adverse event; NMA: network meta-analysis; RCT: randomised controlled trial; UC: ulcerative colitis

A total of 252 records were retrieved by the electronic database searches. After de-duplication of results, 177 unique records were suitable for review. After title and abstract review, 6 records were selected to be reviewed at the full-text stage. Of these, 0 records were found to fulfil the eligibility criteria for inclusion in the SLR.

Supplementary searches of conferences, SLR bibliographies and clinical trials registries yielded 39 records. Of these, 0 records fulfilling the eligibility criteria were identified. A PRISMA diagram showing the flow of records through each stage of the review process is presented in Figure 1.

As clinical data on all relevant comparators in the submission were obtained from the clinical SLR conducted in randomised controlled trials (RCT) (Appendix D of the company submission), the gold standards for clinical evidence, it was not considered necessary to collect additional data from non-randomised studies and therefore no updates to the non-randomised SLR were performed.

Figure 1: PRISMA diagram of included and excluded studies for the SLR



Abbreviations: CDSR: Cochrane Database of Systematic Reviews; DARE: Database of Abstracts of Review of Effects; SLR: systematic literature review.

A2. Appendix D.2.2, Table 7. The eligibility criteria for the clinical systematic literature review outlined in Table 7 include certolizumab as a relevant intervention, but this is not listed in the final scope and search strategies do not include terms for certolizumab. Please clarify whether certolizumab should be included as an intervention in this table.

Certolizumab is approved by the European Medicines Agency (EMA) for active rheumatoid arthritis, axial spondylarthritis, psoriatic arthritis and plaque arthritis, but not for ulcerative colitis.[3] Certolizumab was incorrectly included in Table 7 in the submission, and no searches for certolizumab were conducted as part of the clinical SLR.

A3. Please clarify whether any specific searches were conducted for adverse effects of ozanimod.

No specific searches were conducted for the adverse effects of ozanimod. However, in line with outcomes reported in Table 7 in Appendix D.2.2 of the Appendices of the company submission, adverse events, serious adverse events, withdrawals, withdrawals due to adverse events and

several specific AE categories (serious infections, injection-site reactions, and malignancies) were extracted from the studies included in the clinical SLR. This meant that data were collected on the adverse effects of ozanimod reported in the randomised controlled trials (RCTs) included in the clinical SLR (TRUENORTH and TOUCHSTONE).[4, 5]

Network meta-analysis (NMA)

A4. Priority question. Company submission (CS), Document B, section B.2.8.5. For each arm of each trial included in the NMAs, please provide the trial outcome results. Specifically, please provide the numbers achieving clinical response and clinical remission (according to the outcome definitions used as described in Tables 29 and 30) for each response in each arm of each trial, together with the total number of patients, for each arm of each trial.

The trial outcome results utilised in the base case analysis, in the form of number of patients achieving clinical response and clinical remission at both induction and maintenance in each arm of the included trials, are presented in Appendix 2.1. The trial outcome results utilised in the sensitivity analyses are presented in Appendix 2.2.

A5. Appendix D.4.3. Please explicitly describe the meaning of the following variables defined within the JAGS code: 'T', 'SUCRA', 'prob', 'PASI', e.g. SUCRA is 'surface under cumulative ranking curve'. Please provide summary results (e.g. posterior means, medians and 95% credible intervals) for each of these (T, SUCRA, prob, PASI), or help locate these results in the submission.

Definitions for each of the parameters included within the JAGS code for the company NMAs are provided in Table 2 below. The code used to run the NMAs was implemented in line with the code presented in NICE TSD 2, Appendix Example 6 (multinomial; probit link; Program 6(a) for random effects, and Program 6(b) for fixed effect).[6] The JAGS code files for both fixed and random effects models used for the base case and sensitivity NMAs presented in the company submission are provided in Appendix 3.1.

Summary results including the posterior means, medians, and 95% credible intervals for T and PASI are provided in Appendix 3.3. SUCRA, prob, and all other parameters to do with ranks and rank probabilities were included as additional parameters in the code but were not summarised or used to inform analyses. Summary results for rank probabilities are therefore not presented in these tables.

Table 2: Summary of JAGS code parameters used in company NMA

Parameter	Meaning	Reference
Cmax	Maximum number of response categories across trials in the given NMA. Since every submitted NMA has at least one trial with both response data and remission data, Cmax is always equal to 3.	NICE DSU TSD 2, pg. 87
d[k]	Treatment effect for treatment k relative to treatment 1 (placebo) expressed on the standard normal scale	NICE DSU TSD 2, pg. 87
z[1] z[2]	Category effect for clinical response (z[1]) and clinical remission (z[2]) relative to the previous category expressed on the standard normal scale	NICE DSU TSD 2, pg. 84,87
A	The absolute probability of no response, no remission for treatment 1 (placebo) (i.e., "baseline anchor"), estimated as the unweighted average, on the standard normal distribution scale, across placebo arms in trials from the NMA	NICE DSU TSD 2, pg. 87, 88
T[j,k]	Absolute probability of achieving at least clinical response (j = 1) or clinical remission (j = 2) on treatment k.	NICE DSU TSD 2, pg. 87-88
PASI[1,k] PASI[2,k] PASI[3,k]	Absolute probability of no clinical response, no clinical remission (PASI[1,k]); clinical response, no clinical remission (PASI[2,k]); clinical remission (PASI[3,k]), for each treatment k. Please note that the name "PASI" was used to align with the psoriasis example from NICE TSD 2, where the psoriasis area activity score (PASI) was the underlying continuous outcome of interest. As such, the name "PASI" has no particular meaning.	This is a convenient transformation of T[j,k] to absolute probabilities for each category or clinical response and clinical remission.
NP	Number of placebo arms across trials within each network. Used to facilitate calculation of A.	None.
rk[k][a]	Rank of treatment k	NICE TSD 2, pg. 87
best[k] [a]	Calculates the probability of treatment k being ranked first/best (i.e., rk[k] = 1).	NICE TSD 2, pg. 60
prob[k, h][a]	Calculates the probability of treatment k being ranked h'th (rk[k] = h).	NICE TSD 2, pg. 60
cumeffectiveness[a]	Cumulative sum of rank probabilities to facilitate SUCRA calculations	None. This is a simple calculation to support SUCRA values.
SUCRA[a]	Surface under the cumulative ranking curve	None.
pbetter_interest [a]	A constant value representing the treatment number for which it was of interest to estimate the probability of that treatment being better than other treatments. Set to 2 for submitted analyses, which is the treatment number associated with the company's product, ozanimod.	None.

[a]These parameters were included as additional parameters in the code but were not sumarised or used to inform analyses results. These parameters have been defined here for transparency purposes only.

Abbreviations: DSU: decision support unit: NICE: National Institute of Health and Care Excellence; TSD: technical support document.
Source: NICE DSU TSD2.[7]

A6. Priority question. Appendix D.4.3. Please provide any additional JAGS code to that already in Appendix D.4.3 required to execute the NMAs in JAGS, including for any sensitivity analyses. Please supply versions of code for both fixed and random effects models. Please include code for data set up (including data corresponding to clarification question A4), values set for any constants (including 'NP', 'pbetter_interest', 'Cmax'), and any specifications of initial values and seed settings.

The JAGS code for both fixed and random effects models used for the base case and sensitivity NMAs are presented in Appendix 3.1. A list of data inputs, value sets for any constants, and initial values and seeds used for base case analyses are provided in Appendix 3.2.

A7. Appendix D.4.3. The JAGS code provided in Appendix D.4.3. specifies an uninformative prior for parameter A. However, this prior specification is not included in the list in Table 32 of Document B. Please explain the meaning of parameter A in the context of this decision problem. Please provide the rationale for not using an informative prior and justify the use of an uninformative prior.

The definition for parameter A is provided in Table 2. When transformed to the absolute probability scale (PASI[1,1]), this represents the probability of a patient who receives placebo not achieving response nor remission. Parameter A was not listed in Table 32 of Document B of the company submission as it was not given a prior distribution in the JAGS code. Parameter A was calculated by averaging across placebo arms in trials from the NMA (per response to Question A15), which were, themselves, given uninformative priors.

A8. Priority question. Appendix D.4. Please comment on any differences in baseline patient characteristics between the placebo arms of the trials in the NMAs in both induction and maintenance phases, as shown in Tables 13 and 14. For example, the CRP (mg/L) for the placebo arm ranges from 3.2 (ULTRA1) to 35.1 (Jiang 2015) in the induction phase. For 'extensive disease' in the induction phase, the placebo arm ranges from 7.1% (VISIBLE1) to 80.8% (Kobayashi 2016).

The baseline characteristics of patients enrolled the placebo arms in trials included in the NMA were generally similar, with comparable mean age, years since disease diagnosis and Mayo scores in the induction period. Differences in certain baseline patient characteristics do exist between placebo arms across different trials, such as C-reactive protein (CRP) levels and extent of disease. However, these differences were generally minor, and sensitivity analyses indicate that they have limited

impact on the results of the NMA (as discussed below).[8] It should be noted that clinical experts in TA633 highlighted that CRP is a non-specific inflammatory marker and may vary among patients with similar levels inflammation.[8] A more clinically relevant and specific assessment for UC disease activity would be the Mayo score, which is well balanced between the groups.

Baseline characteristics were much more sparsely reported for maintenance trials, making the interpretation of similarities or differences between these placebo arms more difficult. It should be noted that these trials were included in previous NMAs in the same indication submitted to NICE and therefore equivalent heterogeneity in baseline characteristics were present in prior NMAs.[8]

In the induction period NMA the difference in baseline patient CRP levels between placebo arms of included trials is greatly reduced when data from Jiang 2015 are excluded (35.1 mg/L). Disregarding Jiang 2015, the remaining trials have similar baseline CRP levels across placebo arms, with CRP levels ranging from 3.2 (ULTRA1) to 17 (ACT) mg/L. As the Jiang 2015 trial exclusively recruited Asian patients any heterogeneity introduced into the NMA resulting from the higher CRP levels in these trials were tested as part of a sensitivity analysis conducted which excluded trials containing only Asian participants (see Section B.2.8.5.3 of the company submission).

Similarly, Kobayashi 2016 reports a much higher level of extensive disease than other trial placebo arms at induction (80.8%), with the next highest value reported as 62.2% (Motoya 2019 [Probert 2003 was excluded following feasibility assessment]). As with Jiang 2015, Kobayashi 2016 only recruited Asian participants. Indeed, the three highest percentages of extensive disease reported in placebo arms were in trials recruiting exclusively Asian patients (Kobayashi 2016, Motoya 2015 and Suzuki 2014), giving a range of extensive disease of 7.1% (VISIBLE1) to 56.2% (ULTRA1) when excluding trials which only recruited Asian participants. As such, the sensitivity analysis where these trials were excluded implicitly tested the impact of these differences in extent of disease on the results. The results of the sensitivity analysis were similar to the base case, indicating that heterogeneity in baseline characteristics across the placebo arms of trials included in the NMA had limited impact on the results of the NMA (Document B, Appendix 4.5.5).

A9. Document B, Table 32 refers to a prior distribution for a ‘meta-regression coefficient’. However, there is no further mention of meta-regression in the company submission. Please clarify.

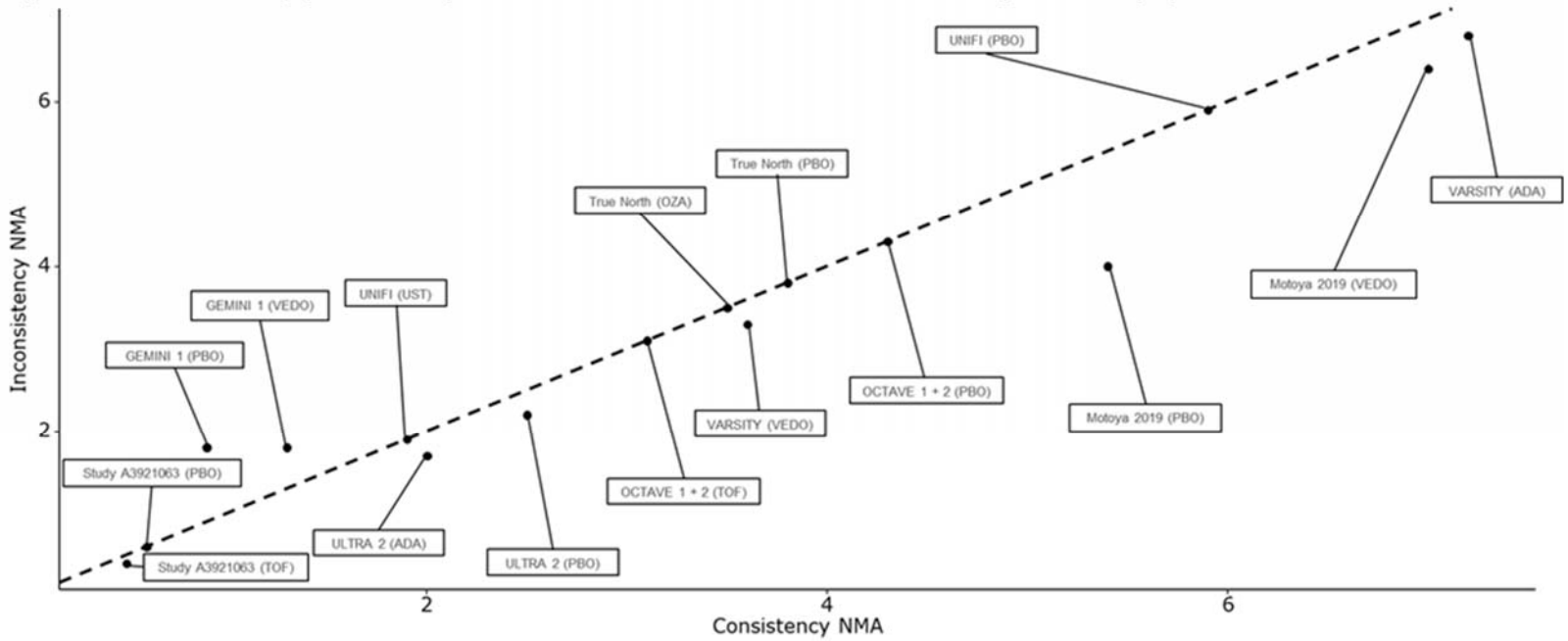
No meta-regressions were run on the base case NMA and therefore the meta-regression coefficient is was included in error and can be removed from Table 32 of Document B – this was a parameter accompanying other code from NICE TSD 2, but was not relevant for the company’s NMAs.[6] A standard NMA was used in line with recommendations in National Institute of Health and Care Excellence (NICE) DSU TSD18 and the approaches taken in all previous NMAs in UC.[6, 9, 10]

A10. Appendix D 4.2. In the ‘Assessment of consistency’, the company states that ‘*some trials in the TNFi-experienced induction analyses had slightly lower deviance in*

the inconsistency model'. Please reproduce Figure 3, fixed effect model (left hand plot), labelling each point with its trial name.

An updated version of Figure 3 from Appendix D.4.2 of the company submission, with each point labelled with the corresponding trial is provided in Figure 2 below.

Figure 2: Inconsistency plot for response-remission at Induction in the TNFi-experienced population – Fixed effects



Abbreviations: ADA: adalimumab; OZA: ozanimod; PBO: placebo; TNFi: tumour necrosis factor alpha inhibitor; TOF: tofacitinib; VEDO: vedolizumab.

A11. Please clarify why:

(i) The VARSITY trial is included in the TNFI-experienced induction network (Document B, Figure 37), when Appendix D.4, Table 11 lists '*Naïve to TNFi therapy* or discontinuation of TNFi therapy (except ADA) for reasons other than safety' among the inclusion criteria for this study [italics ERG emphasis].

The inclusion criteria for the VARSITY trial states:[11] *"Previous exposure to a tumour necrosis factor inhibitor other than adalimumab was allowed in up to 25% of patients."* As a result, the VARSITY trial included both TNFi-naïve and TNFi-experienced patients and therefore clinical response and clinical remission outcomes were able to be collected in both populations in the induction period (see Table 27, Section B.2.8.3 of the company submission) and were subsequently included in the both the TNFi-naïve and TNFi-experienced induction NMAs. The information reported in Table 11 of Appendix D.4 is therefore correct, as both TNFi-naïve and TNFi-experienced patients were included in the VARSITY trial, where the TNFi-experienced patients were those who had discontinued treatment with a TNFi which was not adalimumab and therefore had prior exposure to infliximab or golimumab.

(ii) The ULTRA2 trial is included in the TNFI-experienced induction and maintenance networks (Document B, Figures 37 and 40) when Appendix D.4, Table 11 lists 'Treatment with IFX, ADA or other TNFi agent or biologic agent' among the exclusion criteria for this study.

The exclusion criteria for the ULTRA2 trial states:[12] *"Prior treatment with infliximab was permitted if it had been discontinued because of loss of response or drug intolerance for more than 8 weeks."* Patients were therefore permitted previous use of TNFi agents other than if treatment had been discontinued due to a loss of response or intolerance for longer than 8 weeks. As a result, the ULTRA2 trial included both TNFi-naïve and TNFi-experienced patients and therefore clinical response and clinical remission outcomes were able to be collected in both populations in the induction and maintenance periods (see Table 27 and Table 28, Section B.2.8.3 of the company submission) and were subsequently included in the both the TNFi-naïve and TNFi-experienced induction NMAs. The information reported in Table 11 of Appendix D.4 is therefore incomplete when reporting exclusion criteria regarding prior use of infliximab and TNFi agents, and should make clear that these were allowed if discontinued due to loss of response for longer than 8 weeks.

A12. Document B, section B.2.8.4. The company submission states that 'Individual doses of the same active agent that had the same method of administration … were pooled in the base case'. In Table 31, these pooled doses are shown in rows shaded

blue. Please explain why the entries for golimumab SC 200/100 mg and adalimumab SC 160/80/40 mg Q2W in the induction period are not shown as pooled.

Golimumab 200/100 mg and adalimumab 160/80/40 mg once every two weeks (Q2W) each represent a single treatment regimen as opposed to different treatment doses, so should not be shown as pooled in Table 31: patients in the golimumab 200/100 mg group received golimumab 200 mg and then 100 mg, 2 weeks apart; patients in the adalimumab 160/80/40 mg Q2W group received adalimumab 160 mg at week 0, 80 mg at week 2, and then 40 mg at weeks 4 and 6. As there were no other golimumab or adalimumab induction treatment regimens across the trials included in the NMAs, no pooling of trial results was necessary.

A13. The company provided results matrices for sensitivity analyses on the key outcomes of the NMA, clinical response and remission (Appendix D.4.5, Figures 5 to 42). However, in its submission, the company only provides the forest plots for the primary analyses of the NMA (Document B, Figures 32, 33, 35, 36, 38, 39, 41 and 42). Please provide the results matrices that correspond to these forest plots for the primary analyses to aid interpretation of the robustness of results, as the plots do not reflect comparisons between treatments other than ozanimod or placebo.

League tables (results matrices) for the base case analyses in the TNFi-naïve population are provided in Figure 3 –Figure 6 below. League tables for the base case analyses in the TNFi-experienced populations are presented Figure 7–Figure 10 below.

Induction period – TNFi-naïve

Figure 3: League table for clinical response at induction in the TNFi-naïve population (random effects)[a]

[a]Treatments are ordered from top left to bottom right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a purple background to indicate significance.
Abbreviations: ADA: adalimumab; GOL: golimumab; IFX: infliximab; OZA: ozanimod; PBO: placebo; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

Figure 4: League table for clinical remission at induction in the TNFi-naïve population (random effects)[a]

[a]Treatments are ordered from top left to bottom right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a purple background to indicate significance.
Abbreviations: ADA: adalimumab; GOL: golimumab; IFX: infliximab; OZA: ozanimod; PBO: placebo; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

Maintenance period – TNFi-naïve

Figure 5: League table for clinical response at maintenance in the TNFi-naïve population (fixed effects)[a]

[a]Treatments are ordered from top left to bottom right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a purple background to indicate significance.
Abbreviations: ADA: adalimumab; GOL: golimumab; IFX: infliximab; OZA: ozanimod; PBO: placebo; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

Figure 6: League table for clinical remission at maintenance in the TNFi-naïve population (fixed effects)[a]

[a]Treatments are ordered from top left to bottom right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a purple background to indicate significance.
Abbreviations: ADA: adalimumab; GOL: golimumab; IFX: infliximab; OZA: ozanimod; PBO: placebo; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

Induction period – TNFi-experienced

Clinical response

Figure 7: League table for clinical response at induction in the TNFi-experienced population (fixed effects)[a]

[a]Treatments are ordered from top left to bottom right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a purple background to indicate significance.
Abbreviations: ADA: adalimumab; GOL: golimumab; IFX: infliximab; OZA: ozanimod; PBO: placebo; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

Clinical remission

Figure 8: League table for clinical remission at induction in the TNFi-experienced population (fixed effects)[a]

[a]Treatments are ordered from top left to bottom right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a purple background to indicate significance.
Abbreviations: ADA: adalimumab; GOL: golimumab; IFX: infliximab; OZA: ozanimod; PBO: placebo; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

Maintenance period –TNFi-experienced

Clinical response

Figure 9: League table for clinical response at maintenance in the TNFi-experienced population (fixed effects)[a]

[a]Treatments are ordered from top left to bottom right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a purple background to indicate significance.
Abbreviations: ADA: adalimumab; OZA: ozanimod; PBO: placebo; SC: subcutaneous; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

Clinical remission

Figure 10: League table for clinical remission at maintenance in the TNFi-experienced population (fixed effects)[a]

[a]Treatments are ordered from top left to bottom right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a purple background to indicate significance.
Abbreviations: ADA: adalimumab; OZA: ozanimod; PBO: placebo; SC: subcutaneous; TOF: tofacitinib; UST: ustekinumab; VEDO: vedolizumab.

A14. Please confirm whether NMAs were conducted on any adverse event or treatment discontinuation outcomes. If such analyses have been conducted, please provide the results along with the corresponding data and JAGS code.

No NMAs were conducted for either adverse event or treatment discontinuation outcomes and therefore no results are available.

A15. Priority question. Please provide exact details on how the baseline risk in the anchor treatment (placebo) was calculated in the NMA.

Parameter A represents the absolute baseline probability of no response, no remission for treatment 1 (placebo), which was calculated as the unweighted average, on the standard normal distribution scale, across placebo arms in trials from the NMA. This calculation was performed at each iteration of the Markov Chain Monte Carlo (MCMC) process and captures joint variability around prior distributions for treatment effect, category effect, and between-trial heterogeneity parameters.

Imputation/adjustment of trial data

A16. Document B, section B.2.3.3. In Table 12, the company indicates that 'non-responder imputation (NRI)' was used for the TRUENORTH trial. Patients with insufficient data to determine endpoint status at Week 10 and/or Week 52 were classified as non-responders, as well as patients meeting the criteria for treatment failure. Please clarify how many patients required NRI in each arm. Please clarify whether the same or a similar imputation process was applied to the results of comparator trials.

NRI is a conservative analysis method in which participants with missing data are assumed to be non-responders, regardless of actual response status, at the time of dropout and therefore may underestimate treatment effect. The number of patients in each arm of the TRUENORTH trial in the ITT population imputed using non-responder imputation (NRI) at induction and maintenance are reported in Table 3 for clinical response and clinical remission. As both definitions required values to be obtained at baseline and latest follow-up visit, the NRI counts for the 3-component Mayo score for clinical remission and clinical response were the same. The definitions for both clinical remission and clinical response according to the 3-component Mayo scale used in the TRUENORTH trial are provided in Table 10 of Section B.2.3.1 of the company submission.

Table 3: Number of patients in the ITT population imputed using NRI for clinical response and remission at both induction (Week 10) and maintenance (Week 52) (3-component Mayo score)

Induction period ITT population			Maintenance period ITT population		
Cohort 1		Cohort 2	Placebo	Re-randomised patients	
Ozanimod (N = 429) n (%)	Placebo (N = 216) n (%)	Ozanimod (N = 367) n (%)	Placebo (N = 69) n (%)	Ozanimod – Placebo (N = 227) n (%)	Ozanimod – Ozanimod (N = 230) n (%)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ITT: intention-to-treat; NRI: non-responder input.

All but two of the comparator trials included in the NMA used NRI to account for patients with missing data, with the exception of the Kobayashi 2016 trial, which used last observation carried forward imputation (LOCF) (Table 4). Imputation methods were not reported for the PURSUIT-J trial.

Table 4: Imputation approach for non-responders utilised by comparator trials included in the NMAs

Trial	VERBATIM: Method used to impute dichotomous outcome (response/remission)	Imputation method
VISIBLE 1[13]	"For dichotomous (i.e., proportion-based) end points, any patient with missing information for determination of endpoint status was considered as a non-responder in the analysis."	NRI
OCTAVE 1[14]	"Patients with missing data were considered as not having had a response."	NRI
OCTAVE 2[14]	"Patients with missing data were considered as not having had a response."	NRI
OCTAVE SUSTAIN[14]	"Patients with missing data were considered as not having had a response."	NRI
VARSITY[11]	"Missing values for binary outcomes were imputed as nonresponses"	NRI
UNIFI[15]	"For dichotomous end points, including all end points that were controlled for multiple comparisons, patients with missing data were considered not to have reached the end points."	NRI
Motoya 2019[16]	"Clinical response, clinical remission, and mucosal healing were considered as noresponse, no-remission, or no-mucosal healing, when adjudication for these endpoints were missing at the time of evaluation."	NRI
PURSUIT-J[17]	Not reported (for continuous, large observation carried forward)	Not reported
Jiang 2015[18]	"In addition, patients with insufficient data for the assessment of response were not considered to have had a clinical response, to be in clinical remission, or to have had mucosal healing at that visit"	NRI
Suzuki 2014[19]	"Nonresponder imputation, whereby the patient was assumed to not have efficacy, was used for patients with missing data or those who moved to the rescue arm for all efficacy endpoints."	NRI
PURSUIT-M[20]	"Patients with missing data for a dichotomous end point were considered as not having achieved the end point"	NRI
PURSUIT-SC[21]	"Patients with missing data for an end point were considered not to have achieved the respective end point for dichotomous end points"	NRI
Study A3921063[22a]	"Patients with missing data were considered to be non-responders"	NRI
ULTRA 2[23]	"Missing or incomplete data as well as values at or after switch to open-label treatment of adalimumab were handled using the nonresponder imputation methods."	NRI

ULTRA 1[24]	"Results for the ADA 160/80 and ADA 80/40 groups were compared with results for the placebo group using the chi-2 test for dichotomous endpoints, with missing or incomplete data handled using nonresponder imputation"	NRI
ACT 1[25]	"In addition, patients with insufficient data for the assessment of a response were not considered to have had a clinical response, to be in clinical remission, or to have had mucosal healing at that visit."	NRI
ACT 2[25]	“In addition, patients with insufficient data for the assessment of a response were not considered to have had a clinical response, to be in clinical remission, or to have had mucosal healing at that visit.”	NRI
TRUE NORTH[4]	NRI	NRI
Kobayashi 2016[26]	"Efficacy was assessed in the full analysis set. Patients who took prohibited medication because of worsening UC (lack of efficacy or loss of response to the study medication), who discontinued the study medication because of worsening UC, including 8-week non-responders, or who underwent colectomy or colostomy were not considered to have had a clinical response, clinical remission, or MH, and their post-procedure CAI score was used as the baseline value from the time of the procedure onward. For other patients who withdrew prematurely, the last observation was carried forward."	LOCF
GEMINI 1[27]	"Treatment was considered to have failed in patients who withdrew prematurely"	NRI

[a]The prespecified method in the trial protocol was to limit analysis to the full analysis set and therefore not include any data for non-responders. However, the data reported in the trial’s primary publication, and therefore incorporated in the company’s NMA, were based on NRI.
Abbreviations: CMH: Cochran–Mantel–Haenszel; LOCF: last observation carried forward; NRI: non-response imputation; UC: ulcerative colitis.

A17. Document B, section B.2.8.3. The company submission states that '*Heterogeneity generated by differences in trial design in UC may influence the maintenance NMA findings if maintenance results for induction responders were directly compared to those of non-induction responders'* (p87). Please clarify the following:

(i) Please expand on the term 'induction responder'. Does this term include all patients whose condition responded during induction (including those randomised to control at first randomisation), or only to patients initially randomised to active treatment whose condition responded during induction?

The use of the phrase 'induction responders' on p.87 in Section B.2.8.3 of the company submission refers to patients exhibiting response at the end of the induction period in re-randomised trials only. As specified in Section B.2.8.3 of the company submission, re-randomised trials involved an additional randomisation step at the end of the induction period, on top of the initial randomisation that occurs at baseline in treat-through trial designs. With the exception of OCTAVE SUSTAIN and UNIFI, trials with a re-randomised trial design only re-randomised responders receiving active treatment in the induction period. OCTAVE SUSTAIN also re-randomised placebo responders in the induction phase and UNIFI re-randomised delayed responders, defined as patients who did not respond to placebo during the 8-week induction phase but who then received ustekinumab at Week 8 and were responders at Week 16. All of these responders in the re-randomised trials were classed as 'induction responders' in the company submission.

(ii) Please explain the term 'non-induction responder'.

'Non-induction' responders refer to all patients in trials with a 'treat-through' maintenance design who exhibited clinical response at induction and continued on into the maintenance period of the trial.

(iii) Please explain the rationale for comparing 'non-induction responders' with 'induction responders'.

As discussed in Section B.2.8.3 of the company submission, there was heterogeneity in the maintenance trial designs of the trials included in the NMA, which were a mix of treat-through and re-randomised maintenance trial designs. Comparisons across re-randomised and treat-through trials are difficult, since they include heterogeneous groups with respect to study drug exposure. In re-randomised trials, 'induction responders' (see definition in the response to question A.17 i) re-randomised to placebo in maintenance may "carry over" the effect of the induction therapy into the maintenance period, resulting in a heightened level of response at maintenance. In comparison, 'non-induction responders' (see definition in the response to question A.17 ii) from treat-through trials receiving placebo may have a comparatively lower level of response at maintenance, since none of these patients had received prior active treatment. The levels of response at maintenance in the placebo arms of these trials can therefore not be directly compared.

As specified in Section B.2.8.3 of the company submission, to account for these differences in trial design, statistical adjustments were made to treat-through trials to align with the data presented in re-randomised trials during the maintenance period. This approach is in alignment with that used in TA547 and the ERG 'maintenance only' scenario in TA633.[9, 10] A sensitivity analysis was also performed where the re-calculated treat-through data were excluded from the maintenance results to assess potential bias introduced from the adjustments to treat-through trials. The results of the sensitivity analysis can be found in Appendix D.4.5 of the company submission and demonstrate the results of the NMA to be robust to heterogeneity in maintenance period trial design.

A18. Document B, section B.2.8.3. The company submission states that *'statistical adjustments were made to treat-through trials to align with the data presented in re-randomised trials during the maintenance period*' (p87), and refers to Section B.2.8.4 for full details. Section B.2.8.4. states that '*for treat-through trials with available data for sustained clinical responders/remitters (i.e. response/remission amongst induction responders), the data were directly imputed into the treat-through to re-randomised analyses'*. Please explain this imputation process in full, including the role and identity of any predictor variables. Please clarify whether and how the re-randomisation process within re-randomised trials is emulated by the imputation process on the treat-through data.

The re-randomised trials randomised patients who had a clinical response at the end of the induction period, prior to entering the trial maintenance period. In contrast, treat-through trials may allow both induction responders and non-responders (or delayed responders) to continue into the maintenance period. To emulate a re-randomised trial design, the company leveraged sustained clinical response and sustained clinical remission data in the maintenance period (i.e., patients with clinical response/remission at the end of the maintenance period, amongst patients who had clinical response/remission at the end of induction) so that outcomes from treat-through trials reflected only the induction responders (Table 5). Imputation approaches leveraging predictor variables, which depend on individual patient-level data, were not used. The exact calculation methods differed by trial due to variations in data availability. Detailed explanations of the imputation calculations performed for each trial are provided in Appendix D.4.1 of the company submission and summarised in Table 5 below.

Table 5: Re-calculated treat-through trial data used in maintenance NMAs

Study	Population	Treatment arm	Sustained clinical responders	Clinical remitters among induction responders	Explanation for inputs
ACT 1[25]	Naïve	Placebo	17[a]	*20 (reported)* 10 = 17*61.7%	Applied average maintenance responders-to-remitters ratio from biologic-naïve placebo arms from re-randomized trials

					(61.7%) to 17 sustained clinical responders
		Infliximab 5 mg/kg	47[a]	32 = 84*38%	Applied ratio of induction responders to maintenance remitters among responders from adalimumab in biologic-naïve arm of ULTRA 2 (38%) to induction responders
		Infliximab 10 mg/kg	45[a]	28 = 75*38%	
ULTRA 2[23]	Overall	Placebo	30	17	-
		Adalimumab 40 mg	59	44	-
	Naïve	Placebo	24	15	-
		Adalimumab 40 mg	44	34	-
	Experienced	Placebo	6	2	-
		Adalimumab 40 mg	15	10	-
Suzuki 2014[19]	Naïve	Placebo	12 = 17*69%	8 = 12*62.5%	Applied ratio of biologic-naïve responders to sustained responders available in the biologic-naïve placebo arm of ULTRA 2 (69%) to the 17 clinical responders at maintenance, then applied ratio of remitters-to-responders among induction responders from same arm in ULTRA 2 (62.5%)
		Adalimumab 40 mg	50	38	-
VARSITY[11]	Overall	Adalimumab 40 mg	NR	NR	-
		Vedolizumab 300 mg Q8W	NR	NR	-
	Naïve	Adalimumab 40 mg	NR	NR	-
		Vedolizumab 300 mg Q8W	NR	NR	-
	Experienced	Adalimumab 40 mg	NR	NR	-
		Vedolizumab 300 mg Q8W	NR	NR	-

[a]Defined as those in clinical response at weeks 8, 30, and 54.
Abbreviations: NMA: network meta-analysis; Q4W: every 4 weeks; Q8W: every 8 weeks.

Section B: Clarification on cost-effectiveness data

Literature searches

B1. Appendices G and H. Please provide the search strategies for MEDLINE, EconLit, DARE and NHS EED for the cost-effectiveness systematic literature review and health-related quality of life systematic literature review. The Embase strategy is the only search strategy provided.

The search strategies for the MEDLINE, NHS EED, DARE, and EconLit database searches for the economic SLR are provided in Table 6–Table 9 below. MEDLINE was the only database searched for the health-related quality of life (HRQoL) SLR. The search strategy for the HRQoL SLR for the MEDLINE database is provided in Table 10 below.

Table 6: Economic - Medline (via OvidSP) Search Strategy

Search #	Search Algorithm
1	Inflammatory Bowel Diseases/ or exp Colitis, Ulcerative/
2	(inflammatory bowel disease or ulcerative Colitis).ti,ab.
3	1 or 2
4	exp Health Care Costs/ or exp Drug Costs/ or exp Cost of Illness/ or exp Hospital Costs/ or exp Economics, Pharmaceutical/ or (treatment cost$ or direct cost$ or direct medical cost$ or nonmedical cost$ or non-medical cost$ or total cost or total costs or cost per patient treated or budget impact or cost burden or societal cost$ or administrative cost$ or travel cost$ or travel time or disease cost or cost of drugs).ti,ab.
5	(indirect cost$ or disability or functional status or physical function or impairment or disabilities or productivity or employment or retirement or work disability or absenteeism or presenteeism or sick leave or sick day or worktime loss or opportunity loss or job performance or (work adj2 loss)).ti,ab.
6	exp Hospitalization/ or (healthcare resource or healthcare resources or medical resource or medical resources or healthcare resource or health resource consumption or health care consumption or medical resource consumption or hospitalization or hospital admission or hospital admissions or icu admission or icu admissions or emergency department visit or emergency department visits or emergency room visit or emergency room visits or er visit or er visits or ed visit or ed visits or inpatient visit or inpatient visits or outpatient visit or outpatient visits or specialist visit or specialist visits or unscheduled doctor visit or unscheduled doctor visits or unscheduled physician visit or unscheduled physician visits or general practitioner visit or general practitioner visits).ti,ab.
7	exp Quality-Adjusted life Years/ or (quality adjusted life year$ or qaly or qalys or life year or life years).ti,ab. or (daly or disability adjusted life years).mp.
8	Markov chains/ or (cost minimi?ation or cost-utilit$ or health utility$ or economic evaluation$ or economic review$ or cost outcome or cost analys?s or economic analys?s or cost-benefit analys?s or cost-effectiveness analys?s).ti,ab. or ((cost or economic$) and (costs or cost-effectiveness or markov)).ab,ti. or ((cost effectiveness and ratio) or icer or incremental cost effectiveness ratio or incremental cost-effectiveness ratio).ti,ab. or (cost-consequence or cost-minimisation or cost minimisation or cost-minimization or cost minimization).ti,ab.
9	or/4-8
10	3 and 9

Search #	Search Algorithm
11	exp Longitudinal Studies/ or (longitudinal study or retrospective study or prospective study or cohort$ or follow up or cross-sectional study or cross sectional study or followup study or observational study or registry or registries or real world or cross sectional).ti,ab. or exp Retrospective studies/ or exp Prospective studies/ or exp Cohort Studies/ or exp Cross-Sectional Study/ or exp Cohort Studies/ or exp Observational Study/
12	exp Cost-Benefit Analysis/ or exp "cost effectiveness analysis"/ or exp “Costs and Cost Analysis”/
13	10 and (11 or 12)
14	(Ephemera or "Introductory Journal Article" or News or "Newspaper Article" or Editorial or Comment or Overall or Letter or Short Survey or Tombstone or Books).pt. or in vitro Techniques/ or in vitro study/ or (commentary or editorial or comment or letter or mice or rat or mouse or animal or murine).ti.
15	(case report or case study or case series or woman or man or child or adolescent or female or male or boy or girl or infant).ti.
16	case reports/ or case study/ or case report$.jw.
17	((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) not ((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) and adults)).ti.
18	review.pt. not (((systematic or meta) and analy$) or ((indirect or mixed) and treatment comparison)).ti,ab.
19	or/14-18
20	13 not 19
21	(animals/ not humans/)
22	20 not 21
23	limit 22 to yr=”2010-Current”
24	limit 23 to english language

Table 7: Economic - NHS EED (via OvidSP) Search Strategy

Search #	Search Algorithm
1	Inflammatory Bowel Diseases/ or exp Colitis, Ulcerative/
2	(inflammatory bowel disease or ulcerative Colitis).af.
3	1 or 2
4	exp Health Care Costs/ or exp Drug Costs/ or exp Cost of Illness/ or exp Hospital Costs/ or exp Economics, Pharmaceutical/ or (treatment cost$ or direct cost$ or direct medical cost$ or nonmedical cost$ or non-medical cost$ or total cost or total costs or cost per patient treated or budget impact or cost burden or societal cost$ or administrative cost$ or travel cost$ or travel time or disease cost or cost of drugs).af.
5	(indirect cost$ or disability or functional status or physical function or impairment or disabilities or productivity or employment or retirement or work disability or absenteeism or presenteeism or sick leave or sick day or worktime loss or opportunity loss or job performance or (work adj2 loss)).af.
6	exp Hospitalization/ or (healthcare resource or healthcare resources or medical resource or medical resources or healthcare resource or health resource consumption or health care consumption or medical resource consumption or hospitalization or hospital admission or hospital admissions or icu admission or icu admissions or emergency department visit or emergency department visits or emergency room visit or emergency room visits or er visit or er visits or ed visit or ed visits or inpatient visit or inpatient visits

Search #	Search Algorithm
	or outpatient visit or outpatient visits or specialist visit or specialist visits or unscheduled doctor visit or unscheduled doctor visits or unscheduled physician visit or unscheduled physician visits or general practitioner visit or general practitioner visits).af.
7	exp Quality-Adjusted life Years/ or (quality adjusted life year$ or qaly or qalys or life year or life years or daly or disability adjusted life years).af.
8	exp Cost-Benefit Analysis/ or exp "cost effectiveness analysis"/ or exp "Costs and Cost Analysis"/ or Markov chains/ or (cost minimi?ation or cost-utilit$ or health utility$ or economic evaluation$ or economic review$ or cost outcome or cost analys?s or economic analys?s or cost-benefit analys?s or cost-effectiveness analys?s).af. or ((cost or economic$) and (costs or cost-effectiveness or markov)).af. or ((cost effectiveness and ratio) or icer or incremental cost effectiveness ratio or incremental cost-effectiveness ratio).af. or (cost-consequence or cost-minimisation or cost minimisation or cost-minimization or cost minimization).af.
9	or/4-8
10	3 and 9
11	limit 10 to yr="2010 -Current"

Table 8: Economic - DARE (via OvidSP) Search Strategy

Search #	Search Algorithm
1	Inflammatory Bowel Diseases/ or exp Colitis, Ulcerative/
2	(inflammatory bowel disease or ulcerative Colitis).af.
3	1 or 2
4	exp Health Care Costs/ or exp Drug Costs/ or exp Cost of Illness/ or exp Hospital Costs/ or exp Economics, Pharmaceutical/ or (treatment cost$ or direct cost$ or direct medical cost$ or nonmedical cost$ or non-medical cost$ or total cost or total costs or cost per patient treated or budget impact or cost burden or societal cost$ or administrative cost$ or travel cost$ or travel time or disease cost or cost of drugs).af.
5	(indirect cost$ or disability or functional status or physical function or impairment or disabilities or productivity or employment or retirement or work disability or absenteeism or presenteeism or sick leave or sick day or worktime loss or opportunity loss or job performance or (work adj2 loss)).af.
6	exp Hospitalization/ or (healthcare resource or healthcare resources or medical resource or medical resources or healthcare resource or health resource consumption or health care consumption or medical resource consumption or hospitalization or hospital admission or hospital admissions or icu admission or icu admissions or emergency department visit or emergency department visits or emergency room visit or emergency room visits or er visit or er visits or ed visit or ed visits or inpatient visit or inpatient visits or outpatient visit or outpatient visits or specialist visit or specialist visits or unscheduled doctor visit or unscheduled doctor visits or unscheduled physician visit or unscheduled physician visits or general practitioner visit or general practitioner visits).af.
7	exp Quality-Adjusted life Years/ or (quality adjusted life year$ or qaly or qalys or life year or life years or daly or disability adjusted life years).af.
8	exp Cost-Benefit Analysis/ or exp "cost effectiveness analysis"/ or exp "Costs and Cost Analysis"/ or Markov chains/ or (cost minimi?ation or cost-utilit$ or health utility$ or economic evaluation$ or economic review$ or cost outcome or cost analys?s or economic analys?s or cost-benefit analys?s or cost-effectiveness analys?s).af. or ((cost or economic$) and (costs or cost-effectiveness or markov)).af. or ((cost effectiveness and ratio) or icer or incremental cost effectiveness ratio or incremental cost-effectiveness ratio).af. or (cost-consequence or cost-minimisation or cost minimisation or cost-minimization or cost minimization).af.

Search #	Search Algorithm
9	or/4-8
10	3 and 9
11	limit 10 to yr="2010 -Current"

Table 9: Economic - EconLit (via OvidSP) Search Strategy

Search #	Search Algorithm
1	(Inflammatory bowel disease or Ulcerative Colitis).ti,ab.
2	(treatment cost$ or direct cost$ or direct medical cost$ or nonmedical cost$ or non-medical cost$ or total cost or total costs or cost per patient treated or budget impact or cost burden or societal cost$ or administrative cost$ or travel cost$ or travel time or disease cost or cost of drugs).ti,ab.
3	(indirect cost$ or disability or functional status or physical function or impairment or disabilities or productivity or employment or retirement or work disability or absenteeism or presenteeism or sick leave or sick day or worktime loss or opportunity loss or job performance or (work adj2 loss)).ti,ab.
4	(healthcare resource or healthcare resources or medical resource or medical resources or healthcare resource or health resource consumption or health care consumption or medical resource consumption or hospitalization or hospital admission or hospital admissions or icu admission or icu admissions or emergency department visit or emergency department visits or emergency room visit or emergency room visits or er visit or er visits or ed visit or ed visits or inpatient visit or inpatient visits or outpatient visit or outpatient visits or specialist visit or specialist visits or unscheduled doctor visit or unscheduled doctor visits or unscheduled physician visit or unscheduled physician visits or general practitioner visit or general practitioner visits).ti,ab.
5	(quality adjusted life year$ or qaly or qalys or life year or life years).ti,ab. or (daly or disability adjusted life years).mp.
6	(cost minimi?ation or cost-utilit$ or health utility$ or economic evaluation$ or economic review$ or cost outcome or cost analys?s or economic analys?s or budget$ impact analys?sor cost-benefit analys?s or cost-effectiveness analys?s).ti,ab. or ((cost or economic$) and (costs or cost-effectiveness or markov)).ti,ab. or ((cost effectiveness and ratio) or icer or incremental cost effectiveness ratio or incremental cost-effectiveness ratio).ti,ab.
7	or/2-6
8	1 and 7
9	limit 8 to yr="2010 -Current"

Table 10: HRQoL - Medline (via OvidSP) Search Strategy

Search #	Search Algorithm
1	Inflammatory Bowel Diseases/ or exp Colitis, Ulcerative/
2	(inflammatory bowel disease or ulcerative Colitis).ti,ab.
3	1 or 2
4	(hrql or hrqol or patient reported outcome$ or satisfaction or preference or disability adjusted life or daly$ or activities of daily living or adl).ab,ti.
5	((health adj3 (utility$ or status)) or (utilit$ adj3 (valu$ or measur$ or health or life or estimate$ or elicit$ or disease or score$ or weight)) or (disutility$ and health) or (disutility$ and scor$) or (disutility$ and valu$) or standard gamble or time trade off or

Search #	Search Algorithm
	time tradeoff or tto or rosser or willingness to pay or visual analog scale or visual analogue scale or discrete choice experiment or qwb or 15d or health utilities index or hui or hui1 or hui2 or hui3).ab,ti.
6	(sf36 or sf 36 or sf6 or sf 6 or short form 6 or sf6d or sf 6d or short form 6d or eq 5d or eq5d or euroqol or euro qol or health status or hye or hyes or rosser index or quality of wellbeing or qwb or CUCQ or (Crohn$ adj1 Ulcerative Colitis Questionnaire) or RFIPC or Rating Form of Inflammatory Bowel Disease Patient Concerns or IBDQ or IBDQ-32 or Inflammatory Bowel Disease Questionnaire or SIBDQ or Short Inflammatory Bowel Disease Questionnaire or (health$ adj year$ adj equivalent$)).ti,ab.
7	3 and (4 or 5 or 6)
8	exp Longitudinal Studies/ or (longitudinal study or retrospective study or prospective study or cohort$ or follow up or cross-sectional study or cross sectional study or followup study or observational study or registry or registries or real world or cross sectional).ti,ab. or exp Retrospective studies/ or exp Prospective studies/ or exp Cohort Studies/ or exp Cross-Sectional Study/ or exp Cohort Studies/ or exp Observational Study/
9	7 and 8
10	(Ephemera or "Introductory Journal Article" or News or "Newspaper Article" or Editorial or Comment or Overall or Letter or Short Survey or Tombstone or Books).pt. or in vitro Techniques/ or in vitro study/ or (commentary or editorial or comment or letter or mice or rat or mouse or animal or murine).ti.
11	(case report or case study or case series or woman or man or child or adolescent or female or male or boy or girl or infant).ti.
12	case reports/ or case study/ or case report$.jw.
13	((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) not ((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) and adults)).ti.
14	review.pt. not (((systematic or meta) and analy$) or ((indirect or mixed) and treatment comparison)).ti,ab.
15	(exp animal/ or nonhuman/) not exp human/
16	or/10-15
17	9 not 16
18	Limit 17 to yr=”2010-Current”
19	Limit 18 to English language

B2. Appendices G and H. The Embase searches (Table 29, line #23 and Table 35, line #19) have been limited to ‘Article’ or ‘Article in Press’. This approach appears to exclude conference abstracts from database search results. As a result, any conference abstracts published in proceedings that were not hand-searched may have been missed. Table 30 and Table 36 both list conference abstracts in the inclusion criteria for these systematic literature reviews. Please provide the rationale for limiting the Embase searches to ‘Article’ or ‘Article in Press’.

The limit to ‘Article’ or ‘Article in Press’ was intended to avoid duplication with separate searches that were conducted in Embase to identify relevant conference abstracts from the past 4 four years (2018-2021). These stand-alone searches were designed to limit scope and avoid unnecessary screening of abstracts unlikely to report data of interest for the most recent years. It

also allowed for screeners to review abstracts separately from the peer-reviewed publications to ensure that they were screening those references based on the level of detail reported in the abstract knowing that no further information would be available. The search terms used for the economic and HRQoL SLR are provided in Table 11 and Table 12, respectively.

Table 11: Conference Abstracts – Economic SLR: Embase (via OvidSP) Search Strategy

Search #	Search Algorithm
1	inflammatory bowel disease/ or exp ulcerative colitis/
2	(inflammatory bowel disease or Ulcerative Colitis).ti,ab.
3	1 or 2
4	exp "health care cost"/ or exp "drug cost"/ or exp "cost of illness"/ or exp "hospital cost"/ or exp pharmacoeconomics/ or (treatment cost$ or direct cost$ or direct medical cost$ or nonmedical cost$ or non-medical cost$ or total cost or total costs or cost per patient treated or budget impact or cost burden or societal cost$ or administrative cost$ or travel cost$ or travel time or disease cost or cost of drugs).ti,ab.
5	(indirect cost$ or disability or functional status or physical function or impairment or disabilities or productivity or employment or retirement or work disability or absenteeism or presenteeism or sick leave or sick day or worktime loss or opportunity loss or job performance or (work adj2 loss)).ti,ab.
6	exp Hospitalization/ or (healthcare resource or healthcare resources or medical resource or medical resources or healthcare resource or health resource consumption or health care consumption or medical resource consumption or hospitalization or hospital admission or hospital admissions or icu admission or icu admissions or emergency department visit or emergency department visits or emergency room visit or emergency room visits or er visit or er visits or ed visit or ed visits or inpatient visit or inpatient visits or outpatient visit or outpatient visits or specialist visit or specialist visits or unscheduled doctor visit or unscheduled doctor visits or unscheduled physician visit or unscheduled physician visits or general practitioner visit or general practitioner visits).ti,ab.
7	exp quality adjusted life year/ or (quality adjusted life year$ or qaly or qalys or life year or life years).ti,ab. or (daly or disability adjusted life years).mp.
8	Markov chain/ or (cost minimi?ation or cost-utilit$ or health utility$ or economic evaluation$ or economic review$ or cost outcome or cost analys?s or economic analys?s or budget$ impact analys?s or cost-benefit analys?s or cost-effectiveness analys?s).ti,ab. or ((cost or economic$) and (costs or cost-effectiveness or markov)).ab,ti. or ((cost effectiveness and ratio) or icer or incremental cost effectiveness ratio or incremental cost-effectiveness ratio).ti,ab.
9	or/4-8
10	3 and 9
11	exp longitudinal study/ or (longitudinal study or retrospective study or prospective study or cohort$ or follow up or cross-sectional study or cross sectional study or followup study or observational study or registry or registries or real world or cross sectional).ti,ab. or exp retrospective study/ or exp prospective study/ or exp cohort analysis/ or exp cross-sectional study/ or exp cohort analysis/ or exp observational study/
12	exp "cost benefit analysis"/ or exp "cost utility analysis"/ or exp economic evaluation/ or exp "cost effectiveness analysis"/ or exp "cost minimization analysis"/
13	10 and (11 or 12)
14	(Ephemera or "Introductory Journal Article" or News or "Newspaper Article" or Editorial or Comment or Overall or Letter or Short Survey or Tombstone or Books).pt. or in vitro Techniques/ or in vitro study/ or (commentary or editorial or comment or letter or mice or rat or mouse or animal or murine).ti.

Search #	Search Algorithm
15	(case report or case study or case series or woman or man or child or adolescent or female or male or boy or girl or infant).ti.
16	case reports/ or case study/ or case report$.jw.
17	((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) not ((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) and adults)).ti.
18	review.pt. not (((systematic or meta) and analy$) or ((indirect or mixed) and treatment comparison)).ti,ab.
19	or/14-18
20	13 not 19
21	(exp animal/ or nonhuman/) not exp human/
22	20 not 21
23	Limit 22 to yr=”2018-Current”
24	23 and United European Gastroenterology Week.cf,cg
25	23 and Digestive Disease Week.cf,cg.
26	23 and ecco.cf,cg.
27	23 and American College of Gastroenterology.cf,cg.
28	23 and Advances in Inflammatory Bowel Diseases.cf,cg.
29	23 and colitis congress.cf,cg.
30	23 and ISPOR.cf,cg.
31	23 and AMCP.cf,cg.
32	or/24-31

Table 12: Conference Abstracts – HRQoL SLR: Embase (via OvidSP) Search Strategy

Search #	Search Algorithm
1	inflammatory bowel disease/ or exp ulcerative colitis/
2	(inflammatory bowel disease or Ulcerative Colitis).ti,ab.
3	1 or 2
4	(hrql or hrqol or patient reported outcome$ or satisfaction or preference or disability adjusted life or daly$ or activities of daily living or adl).ab,ti.
5	((health adj3 (utility$ or status)) or (utilit$ adj3 (valu$ or measur$ or health or life or estimate$ or elicit$ or disease or score$ or weight)) or (disutility$ and health) or (disutility$ and scor$) or (disutility$ and valu$) or standard gamble or time trade off or time tradeoff or tto or rosser or willingness to pay or visual analog scale or visual analogue scale or discrete choice experiment or qwb or 15d or health utilities index or hui or hui1 or hui2 or hui3).ab,ti.
6	(sf36 or sf 36 or sf6 or sf 6 or short form 6 or sf6d or sf 6d or short form 6d or eq 5d or eq5d or euroqol or euro qol or health status or hye or hyes or rosser index or quality of wellbeing or qwb or CUCQ or (Crohn$ adj1 Ulcerative Colitis Questionnaire) or RFIPC or Rating Form of Inflammatory Bowel Disease Patient Concerns or IBDQ or IBDQ-32 or Inflammatory Bowel Disease Questionnaire or SIBDQ or Short Inflammatory Bowel Disease Questionnaire or (health$ adj year$ adj equivalent$)).ti,ab. or exp psychological well-being/
7	3 and (4 or 5 or 6)
8	exp longitudinal study/ or (longitudinal study or retrospective study or prospective study or cohort$ or follow up or cross-sectional study or cross sectional study or followup study

Search #	Search Algorithm
	or observational study or registry or registries or real world or cross sectional).ti,ab. or exp retrospective study/ or exp prospective study/ or exp cohort analysis/ or exp cross-sectional study/ or exp cohort analysis/ or exp observational study/
9	7 and 8
10	(Ephemera or "Introductory Journal Article" or News or "Newspaper Article" or Editorial or Comment or Overall or Letter or Short Survey or Tombstone or Books).pt. or in vitro Techniques/ or in vitro study/ or (commentary or editorial or comment or letter or mice or rat or mouse or animal or murine).ti.
11	(case report or case study or case series or woman or man or child or adolescent or female or male or boy or girl or infant).ti.
12	case reports/ or case study/ or case report$.jw.
13	((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) not ((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) and adults)).ti.
14	review.pt. not (((systematic or meta) and analy$) or ((indirect or mixed) and treatment comparison)).ti,ab.
15	or/10-14
16	9 not 15
17	(exp animal/ or nonhuman/) not exp human/
18	16 not 17
19	Limit 18 to yr=”2018-Current”
20	19 and United European Gastroenterology Week.cf,cg.
21	19 and Digestive Disease Week.cf,cg.
22	19 and ecco.cf,cg.
23	19 and American College of Gastroenterology.cf,cg.
24	19 and Advances in Inflammatory Bowel Diseases.cf,cg.
25	19 and colitis congress.cf,cg.
26	19 and ISPOR.cf,cg.
27	19 and AMCP.cf,cg.
28	or/20-27

B3. Appendix G. Please provide details of the search filter used to limit searches for cost-effectiveness studies. The combination of terms appears to restrict search results to only those records containing the EMTREE terms in line #12 (Table 29), and this approach may have excluded relevant studies that have not been adequately indexed by the database.

To ensure all relevant studies were captured, line #8 of the economic search in Table 29 of Appendix G of the company submission provides title and abstract terms for economic evaluations. This search was intended to capture relevant literature not otherwise identified by the index terms in line #12.

B4. Appendix G. The systematic literature review reported in Appendix G identified 13 unique UK economic evaluations from 40 published economic evaluations.

Please provide details of the non-UK relevant economic evaluations that were excluded at this stage.

As reported in Appendix G of the company submission in total, 40 publications reporting on economic evaluations were incorporated in the economic SLR. An additional 9 Health Technology Assessments reporting on 11 unique economic evaluations were also included in the economic SLR. Of these economic evaluations, 13 were UK based (including 7 UK publications and 6 NICE Health technology assessment (HTA) models, across 4 NICE HTA submissions), as reported in Table 31 in Appendix G.

Of the remaining 33 publications, 6 economic evaluations were US-based and 27 were conducted in countries from outside of the US. A further 5 HTA evaluations from outside the UK (4 Canadian Agency for Drugs and Technologies in Health [CADTH] and 1 Institute of Clinical and Economic Review [ICER]) were included following hand-searches, resulting in a total of 38 non-UK economic evaluations. The majority of studies were conducted in Canada (n=5) and Spain (n=5) followed by Poland (n=4), Japan (n=3), and Germany (n=2). One study each was conducted in the following countries: Brazil, Chile, China and the US, Greece, Iran, Netherlands, and Thailand. The details of these economic evaluations are presented in below.

Table 13 below.

Table 13: Non-UK economic evaluations included in the SLR

Author, Year	Reference
Milev, 2019[28]	Milev S, DiBonaventura MD, Quon P, Wern Goh J, Bourret J, Peeples-Lamirande K, Soonasra A, Cappelleri JC, Quirk D. An economic evaluation of tofacitinib for the treatment of moderately-to-severely active ulcerative colitis: modeling the cost of treatment strategies in the United States. J Med Econ. 2019 Sep;22(9):859-868.
Scott, 2020[29]	Scott FI, Luo M, Shah Y, Lasch K, Vajravelu RK, Mamtani R, Fennimore B, Gerich ME, Lewis JD. Identification of the Most Cost-effective Position of Vedolizumab Among the Available Biologic Drugs for the Treatment of Ulcerative Colitis. J Crohns Colitis. 2020 Jun 19;14(5):575-587.
Yokomizo, 2016[30]	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 Gastroenterol. 2016 May 3;3(1):e000093.
Nguyen, 2020[31a]	Nguyen, V, Carlson, J. J, Bloudek, L. PGI26 The Cost-Effectiveness of Pharmacologic Vs Surgical Intervention for Patients Hospitalized with Acute Severe Ulcerative Colitis in the United States. Value in Health. 2020 May 1; Volume 23, S147.
Shaffer, 2021[32]	Shaffer SR, Huang E, Patel S, Rubin DT. Cost-Effectiveness of 5-Aminosalicylate Therapy in Combination With Biologics or Tofacitinib in the Treatment of Ulcerative Colitis. Am J Gastroenterol. 2021 Jan 1;116(1):125-133.
ICER 2020[33]	Ollendorf DA, Bloudek L, Carlson JJ, Pandey R, Fazioli K, Chapman R, Bradt P, Pearson SD. Targeted Immune Modulators for Ulcerative Colitis: Effectiveness and Value; Evidence Report. Institute for Clinical and Economic Review. Available at https://icer.org/assessment/ulcerative-colitis-2020/. [Accessed 15 February 2022]
Bloudek, 2021[34]	Bloudek LM, Pandey R, Fazioli K, Ollendorf DA, Carlson JJ. Optimal treatment sequence for targeted immune modulators for the treatment of moderate to severe ulcerative colitis. J Manag Care Spec Pharm. 2021 Aug;27(8):1046-1055.
Wu, 2018[35]	Wu B, Wang Z, Zhang Q. Cost-Effectiveness of Different Strategies for the Treatment of Moderate-to-Severe Ulcerative Colitis. Inflamm Bowel Dis. 2018 Oct 12;24(11):2291-2302.
Ung, 2014[36]	Ung V, Thanh NX, Wong K, Kroeker KI, Lee T, Wang H, Ohinmaa A, Jacobs P, Fedorak RN. Real-life treatment paradigms show infliximab is cost-effective for management of ulcerative colitis. Clin Gastroenterol Hepatol. 2014 Nov;12(11):1871-8.e8.
Chaudhary, 2013[37]	Chaudhary MA, Fan T. Cost-Effectiveness of Infliximab for the Treatment of Acute Exacerbations of Ulcerative Colitis in the Netherlands. Biol Ther. 2013;3(1):45-60. doi: 10.1007/s13554-012-0007-0. Epub 2012 Dec 21.
Trigo-Vicente, 2019[38]	Trigo-Vicente C, Gimeno-Ballester V, López-Del Val A. Cost-effectiveness analysis of infliximab, adalimumab, golimumab, vedolizumab and tofacitinib for moderate to severe ulcerative colitis in Spain. Eur J Hosp Pharm. 2020 Nov;27(6):355-360.
Trigo-Vicente, 2018[39]	Trigo-Vicente C, Gimeno-Ballester V, Montoiro-Allué R, López-Del Val A. Cost-effectiveness analysis of infliximab, adalimumab, golimumab and vedolizumab for moderate to severe ulcerative colitis in Spain. Expert Rev Pharmacoecon Outcomes Res. 2018 Jun;18(3):321-329.
Stawowczyk, 2016[40]	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. Eur J Clin Pharmacol. 2016 Nov;72(11):1319-1325.

Author, Year	Reference
Stawowczyk, 2016[41]	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 Aug 5;11(8):e0160444.
Stawowczyk, 2016[42]	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 May;36(5):472-81.
Petryszyn, 2020[43]	Petryszyn P, Ekk-Cierniakowski P, Zurakowski G. Infliximab, adalimumab, golimumab, vedolizumab and tofacitinib in moderate to severe ulcerative colitis: comparative cost-effectiveness study in Poland. Therap Adv Gastroenterol. 2020 Aug 25;13:1756284820941179.
Hernandez, 2020[44]	Hernandez L, Kuwabara H, Shah A, Yamabe K, Burnett H, Fahrbach K, Koufopoulou M, Iwakiri R. Cost-Effectiveness Analysis of Vedolizumab Compared with Other Biologics in Anti-TNF-Naïve Patients with Moderate-to-Severe Ulcerative Colitis in Japan. Pharmacoeconomics. 2020 Jan;38(1):69-84.
Moradi, 2016[45]	Moradi, N., et al. (2015). "Economic Evaluation of Infliximab for Treatment of Refractory Ulcerative Colitis In Iran: Cost-Effectiveness Analysis." Value in Health 18(7): A628.
Beilman, 2016[46]	Beilman CL, Thanh NX, Ung V, Ma C, Wong K, Kroeker KI, Lee T, Wang H, Ohinmaa A, Jacobs P, Halloran BP, Fedorak RN. Real-Life Treatment Paradigms Show Adalimumab Is Cost-Effective for the Management of Ulcerative Colitis. Can J Gastroenterol Hepatol. 2016;2016:5315798.
Xie, 2009[47]	Xie F, Blackhouse G, Assasi N, Gaebel K, Robertson D, Goeree R. Cost-utility analysis of infliximab and adalimumab for refractory ulcerative colitis. Cost Eff Resour Alloc. 2009 Dec 11;7:20.
Toor, 2015[48]	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. J Med Econ. 2015 Jun;18(6):437-46.
Sruamsiri, 2018[49a]	Sruamsiri R. Cost-Effectiveness of Vedolizumab for the Management of Moderate-To-Severe Ulcerative Colitis in Thailand. Value in Health 2018;21:S2.
Menchen, 2019[50a]	Menchén L, de Andrés-Nogales F, García S, et al. PGI16 Cost-Effectiveness of Tofacitinib for the Treatment of Moderate-to-Severe Ulcerative Colitis after Biologic Failure or Intolerance in Spain. Value in Health 2019;22:S619.
Taxonera, 2019[51a]	Taxonera C, de Andrés-Nogales F, García S, et al. PGI17 Cost-Effectiveness of Tofactinib for the Treatment of Moderate-to-Severe Ulcerative Colitis after Conventional Therapy in Spain. Value in Health 2019;22:S619.
Hernandez, 2019[44a]	Hernandez L, Kuwabara H, Shah A, Yamabe K, Burnett H, Fahrbach K, Koufopoulou M, Iwakiri R. Cost-Effectiveness Analysis of Vedolizumab Compared with Other Biologics in Anti-TNF-Naïve Patients with Moderate-to-Severe Ulcerative Colitis in Japan. Pharmacoeconomics. 2020 Jan;38(1):69-84.
Quon, 2019[52a]	Quon P, Sardesai A, Milev S, et al. PGI9 Cost-Effectiveness of Tofacitinib Compared with Infliximab, Adalimumab, Golimumab and Vedolizumab for the Treatment of Moderate to Severe Ulcerative Colitis in Germany. Value in Health 2019;22:S617.

Author, Year	Reference
De La Puente, 2018[53a]	De La Puente C, Hurtado V. PSY100 - Cost-Effectiveness Analysis of Second Line Treatment for Patients with Ulcerative Colitis. Value in Health 2018;21:S453.
Brito, 2020[54a]	Brito N, Mulinari E. PGI18 Cost-Effectiveness of Biologics for Moderate to Severe Ulcerative Colitis in Brazil. Value in Health 2020;23:S146.
Stern, 2018[55]	Stern S, Ward AJ, Saint-Laurent Thibault C, 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. J Med Econ 2018;21:27-37.
Vellopoulou, 2021[56]	Vellopoulou K, Stefanou G, Tzanetakos C, et al. Cost-effectiveness of tofacitinib for the treatment of moderate to severe active ulcerative colitis in Greece. Eur J Gastroenterol Hepatol 2021;33:325-333.
Taxonera, 2021[57]	Taxonera C, de Andrés-Nogales F, García-López S, et al. Cost-effectiveness analysis of using innovative therapies for the management of moderate-to-severe ulcerative colitis in Spain. Expert Rev Pharmacoecon Outcomes Res 2022;22:73-83.
Sardesai, 2021[58]	Sardesai A, Dignass A, Quon P, et al. Cost-effectiveness of tofacitinib compared with infliximab, adalimumab, golimumab, vedolizumab and ustekinumab for the treatment of moderate to severe ulcerative colitis in Germany. J Med Econ 2021;24:279-290.
Uda, 2020[59]	Uda A, Eto Y, Li Y, et al. Evaluation of the Optimal Position for Vedolizumab in the Japanese Treatment Paradigm for Ulcerative Colitis Using Markov Modeling. Crohn's & Colitis 360 2020;2.
Sheng, 2020[60]	Sheng Y, Ma L, Chen J, et al. PGI9 Cost-Effectiveness Comparison of Vedolizumab, Infliximab and Conventional Therapy for the Treatment of Biologic-Naïve, Moderately to Severely ACTIVE Ulcerative Colitis Patients in China. Value in Health 2020;23:S534.
CADTH 2016[61]	Canadian Agency for Drugs and Technologies in Health. Common Drug Review. Pharmacoeconomic Review Report. Adalimumab (Humira). Available at https://cadth.ca/sites/default/files/cdr/pharmacoeconomic/SR0450_Humira_UC_PE_Report.pdf. [Accessed 16 February 2022].
CADTH 2019[62]	Canadian Agency for Drugs and Technologies in Health. Common Drug Review. Pharmacoeconomic Review Report. Tofactininb (Xeljanz). Available at https://www.cadth.ca/sites/default/files/cdr/pharmacoeconomic/sr0572-xeljanz-pharmacoeconomic-report.pdf. [Accessed 16 January 2022].
CADTH 2020[63]	Common Drug Review. Pharmacoeconomic Review Report. Ustekinumab (Stelara/Stelara IV). Available at https://www.cadth.ca/sites/default/files/cdr/pharmacoeconomic/sr0627-stelara-pharmacoeconomic-review-report.pdf. [Accessed 16 January 2022]
CADTH 2020[64]	Canadian Agency for Drugs and Technologies in Health. Common Drug Review. Pharmacoeconomic Review Report. Vedolizumab (Entyvio SC). Available at https://www.cadth.ca/sites/default/files/cdr/pharmacoeconomic/sr0635-entyvio-pharmacoeconomic-review-report.pdf. [Accessed 16 January 2022].

[a]Abstract

Abbreviations: CADTH: Canadian Agency for Drugs and Technologies in Health; ICER: Institute for Clinical and Economic Review.

Model structure

B5. Document B, section B.3.2.2, p141. The company submission highlights that previous models developed for ulcerative colitis used a hybrid decision tree for the initial induction period. Please provide further rationale for opting to use tunnel states for the induction phase of the model.

Tunnel states were chosen in the induction period as these allowed for greater flexibility when modelling different components:

- The use of tunnel states to model the initial induction period decision tree allowed patients to enter the Markov maintenance trace at a variety of endpoints within a single model engine design.
- Having flexibility in when patients enter the Markov maintenance trace enables the variety of induction period lengths seen across treatment options to be modelled. In addition to this, the tunnel state approach to the initial decision tree enables extended induction periods (again of a flexible length) to be captured.
- Using a single model engine design for all treatments allows greater transparency and clarity of the engine design which provides an engine which is easier to quality control. The tunnel state approach to the initial induction period decision tree also allows improved granularity of patient tracking during the induction period. In particular, it enables death due to background mortality to be captured on a per cycle basis within the induction period. A decision tree would be unable to provide this granularity in results.

Transition probabilities and clinical data

B6. Priority question. Document B, section B.3.3.3 and B.3.3.4. With respect to the estimation of induction and maintenance period transitions, the company submission states that '*Mean absolute probabilities were derived from the following NMA outputs: baseline anchor, response effect, remission effect, and standardised mean difference (SMD) versus baseline for a given treatment' in the induction/maintenance period.* Please provide the calculations used to estimate the probabilities at the end of the induction period (Table 43), after extended induction period (Table 44) and in the maintenance period (Table 45).

In response to Question A5, the summary tables of posterior means, medians, and 95% credible intervals have been provided in Appendix 3.3 (Table 50–Table 53), along with definitions for each parameter (Table 2). The values presented in Document B, section B.3.3.3 Table 43 and Document B, section B.3.3.4 Table 45 are those named PASI[1,k], PASI[2,k], and PASI[3,k], and are provided in these summary tables. Specifically, for treatment k, these are defined as the mean of the posterior distribution of absolute probabilities of mutually exclusive categories of "no response, no remission" (PASI[1,k] = P(No response, No Remission) for treatment k), "response,

no remission" (PASI[2,k] = P(Response, No Remission) for treatment k), and "response and remission" (PASI[3,k] = P(Remission) for treatment k). At each MCMC iteration, of the Bayesian estimation process, these absolute probabilities were estimated according to the model specifications detailed in the NMA code (following NICE TSD 2, Appendix Example 6; now included for each analysis in response to Question A6). Their estimates are a direct transformation from T[j,k] and calculated based on parameters A, d[k], z[1], and z[2], each defined in the table of definitions provided in response to Question A5

Specifically, for each posterior sample b:

$$P(Remission)_b = 1 - \Phi(A_b + d[k]_b + z[2]_b)$$

$$P(Response, No\ Remission)_b = 1 - \Phi(A_b + d[k]_b + z[1]_b) - P(Remission)_b$$

$$P(No\ Response, No\ Remission)_b = (1 - P(Remission)_b - P(Response, No\ Remission)_b$$

Here, Φ represents the standard normal distribution, A represents the effects of treatment 1 (placebo) on the standard normal scale (the "baseline anchor"), d[k] represents the treatment effect for treatment k relative to treatment 1 (placebo) on the standard normal scale (the "standardised mean difference versus baseline"), z[1] represents the category effect for clinical response not clinical remission (the "response effect") on the standard normal scale, and z[2] represents the category effect for clinical remission (the "remission effect") on the standard normal scale. The posterior samples were later used for the PSA iterations for these probabilities (see response to QB.13). Since these calculations were performed at each iteration of the MCMC, correlation between all parameters were inherently captured through estimation of their values.

B7. Document B, section B.3.3.3, p154. The company submission states that direct trial data were used to inform patient distribution into the 'Remission', 'Response No Remission' and 'Active UC' health states for the scenario analysis where extended induction was selected. Please clearly state the data sources used to estimate the clinical efficacy (after extended induction) for each treatment in Table 44.

The data used to inform patient distributions following extended induction in Table 44 of the company submission are in line with the values presented in TA633 which were obtained from the relevant clinical trials.[8]

B8. Excel model, Document B, sections B.3.3.3 and B.3.3.4. In the Excel model, for the TNFi naïve population, please clarify from where the probability of 'Remission' and 'No Remission' for conventional therapy (following the induction period) were derived (See 'Efficacy active treatment' sheet, cells J21 and K21). Please clarify from where the probability of 'Sustained Remission' and 'Sustained Response' for BSC (best supportive care) were derived (See 'Efficacy active treatment' sheet, cells J47

and K47). These values do not appear to be in Tables 43 to 46 in the company submission.

Tables 43, 45 and 46 have been updated to include BSC health state distributions in the TNFi-naïve population after the induction and maintenance periods, and loss of response probabilities in the maintenance period. These are shown below in Table 14, Table 15, and Table 16, respectively. It should be noted that these values do not directly inform the model, as there is no induction period for BSC. Rather, these data are used to inform rate of loss of response or remission following spontaneous response or remission in the maintenance period for patients in the Active UC health state. These probabilities were derived from the absolute probabilities for placebo from the NMA. The data for patients receiving placebo in the NMA were considered the best available data to inform these transitions. Please note that, as BSC was not modelled as having an extended induction period, Table 44 and the corresponding table in the economic model do not include health state probabilities for BSC and therefore no correction is required.

Table 14: Clinical efficacy at the end of the induction period

Drug	Induction length (weeks)	Remission	Response no remission	No response (Active UC)
TNFi-naïve				
Ozanimod	10	[redacted]	[redacted]	[redacted]
Golimumab	6	[redacted]	[redacted]	[redacted]
Infliximab/biosimilar	8	[redacted]	[redacted]	[redacted]
Adalimumab/biosimilar	8	[redacted]	[redacted]	[redacted]
Vedolizumab[a]	6	[redacted]	[redacted]	[redacted]
BSC[b]	10	[redacted]	[redacted]	[redacted]
TNFi-experienced				
Ozanimod	10	[redacted]	[redacted]	[redacted]
Ustekinumab	8	[redacted]	[redacted]	[redacted]
Vedolizumab[a]	6	[redacted]	[redacted]	[redacted]
BSC[b]	10	[redacted]	[redacted]	[redacted]

[a]Vedolizumab SC is only licensed for maintenance treatment.
[b]Patients will not experience an induction period related to BSC within the model however the clinical efficacy after induction is used to calculate the transition probabilities in the maintenance period for patients receiving BSC alone. The data for patients receiving placebo in the NMA was considered the best available data to inform these transitions.
Abbreviations: BSC: best supportive care; IV, intravenous SC: subcutaneous; TNFi: tumour necrosis factor alpha inhibitor; UC: ulcerative colitis.

Table 15: Clinical efficacy at the end of the maintenance period

Drug	Sustained Remission	Sustained Response
TNFi-naïve		
Ozanimod	[redacted]	[redacted]
Golimumab	[redacted]	[redacted]
Infliximab/biosimilar	[redacted]	[redacted]
Adalimumab/biosimilar	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]
Vedolizumab (IV)	[redacted]	[redacted]
Vedolizumab (SC)	[redacted]	[redacted]
BSC[a]	[redacted]	[redacted]
TNFi-experienced		
Ozanimod	[redacted]	[redacted]
Ustekinumab	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]
Vedolizumab (IV)	[redacted]	[redacted]
Vedolizumab (SC)	[redacted]	[redacted]
BSC[a]	[redacted]	[redacted]

[a] Patients will not experience an induction period related to BSC within the model however the clinical efficacy after induction is used to calculate the transition probabilities in the maintenance period for patients receiving BSC alone. The data for patients receiving placebo in the NMA was considered the best available data to inform these transitions.
Abbreviations: BSC: best supportive care; TNFi: tumour necrosis factor alpha inhibitor; UC: ulcerative colitis.

Table 16: Transition probabilities for loss of response in the maintenance period

Drug	Duration of maintenance period	Loss of Response	Loss of Response No Remission
TNFi-naïve			
Ozanimod	42	[redacted]	[redacted]
Golimumab	54	[redacted]	[redacted]
Infliximab/biosimilar	46	[redacted]	[redacted]
Adalimumab/biosimilar	44	[redacted]	[redacted]
Vedolizumab	46	[redacted]	[redacted]
BSC	42	[redacted]	[redacted]
TNFi-experienced			
Ozanimod	42	[redacted]	[redacted]
Ustekinumab	44	[redacted]	[redacted]
Vedolizumab	46	[redacted]	[redacted]
BSC	42	[redacted]	[redacted]

Abbreviations: BSC: best supportive care; TNFi: tumour necrosis factor alpha inhibitor.

Comparators

B9. Priority question. Clinical advice to the ERG suggests that there is consensus that tofacitinib is increasingly being used as a first-line treatment option in UK clinical practice for the appropriate target population because of its oral administration and rapid action. Clinical advice to the ERG also suggests that safety concerns, both in the first and second line, can and are clinically managed at an individual patient level. Therefore, please provide an updated model including tofacitinib as a comparator for both TNFi-naïve and TNFi-experienced populations.

Clinical consultation sought by the Company as part of the clarification process re-confirmed that owing to the significant safety concerns associated with tofacitinib, it is not routinely used in TNFi-naïve patients and its use in TNFi-experienced patients is typically restricted to later treatment lines. It was noted that whilst there may be growing use of tofacitinib amongst some clinicians, there is no UK wide consensus on the increasing use of tofacitinib.

Clinician feedback received as part of this appraisal noted that on rare occasions tofacitinib may be used as a first-line treatment option in TNFi-naïve patients. However, it was noted that due to the adverse events and safety concerns associated with tofacitinib its use is restricted to a small number of younger, healthier patients. For the majority of UC patients, the potential of adverse events and safety is a critical factor in the eventual treatment choice, with safety being highly prioritised in clinical decision making. Clinician feedback indicated that this has led clinicians to reserve use of tofacitinib to later in the treatment pathway.

The exclusion of tofacitinib as a relevant comparator in TA633 was accepted by the appraisal committee, which noted that tofacitinib is rarely used in clinical practice due to safety concerns.[9] Since the ustekinumab submission, there has been no downgrading in the European Medicines Agency (EMA) warnings and restrictions associated with tofacitinib. Rather, in June 2021 the Pharmacovigilance Risk Assessment Committee (PRAC) of the EMA recommended that tofacitinib should only be used if no suitable treatment alternative is available in patients over 65 years of age, patients who are current or past smokers, patients with other cardiovascular (CV) risk factors, and patients with other malignancy risk factors.[65] Very recently, the EMA's safety committee has started a review of the safety of JAK inhibitors in treating chronic inflammatory disorders including UC, initiated at the request of the European Commission (EC) under Article 20 of Regulation (EC) No 726/2004, highlighting ongoing safety concerns of this class of medication.[66]

In the USA, the Food and Drugs Administration (FDA) has also issued warnings associated with tofacitinib, noting increased risk of serious heart-related events, cancer, blood clots, and death.[67] In reference to these safety warnings the FDA has restricted approved use to tofacitinib in UC to only certain patients who are not treated effectively or who experience severe side effects with TNF inhibitors.

This highlights that despite more clinical experience with tofacitinib, there are increasing safety warnings and restrictions imposed by regulatory authorities in both the EU and USA, due to safety concerns, which serve to limit and restrict use of tofacitinib in clinical practice.

In the UK, the Medicines and Healthcare products Regulatory Agency (MHRA) issued guidance in October 2021 mirroring EMA guidance in recommending that tofacitinib should not be used in patients older than 65 years of age, people who are current or past smokers, or individuals with other cardiovascular (such as diabetes or coronary artery disease) or malignancy risk factors unless there are no suitable treatment alternatives.[68]

These cautions are likely to impact clinical decision making across a broad spectrum of the whole adult UC population. A recent UK epidemiological study indicates that up to 30% of the total adult population are 60 years or over.[69] Given that decision to start a UC treatment drug is taken with the view of long-term use this would represent a sizeable proportion of the UC population where tofacitinib would not be a standard comparator to other therapies based on safety. Approximately, 50% of the total adult UC patient population are aged 50 years or more,[69] of which a significant proportion are likely to have at least one cardiovascular risk factor (such as diabetes, hypertension or hyperlipidaemia).

Restrictions on the use of tofacitinib are not limited to older UC patients. Safety concerns would be important factor in choice of therapy for female IBD patients of child-bearing age; 30–40% of this population report using hormonal contraceptive methods for contraception, where treatment with tofacitinib would be associated with a higher risk of clots.[70, 71]

Due to the safety concerns associated with tofacitinib, and subsequent restriction to its use in a limited number of younger, healthier patients, the Company considers the exclusion of tofacitinib as a relevant comparator in the TNFi-naïve and TNFi-experienced settings appropriate. This view is in line with the committee's position in TA633 and feedback from clinical consultation sought as part of the clarification process.[9] As a result, tofacitinib has not been included as a comparator in the updated model provided alongside these responses.

Costs

B10. Excel model. Please explain what drives the difference in drug acquisition costs (provided in the model engines) between the two approaches adopted for the treatment regimen costs application in the model (that is, per model cycle or treatment cycle).

Applying costs per treatment cycle or per model cycle changes how costs are applied in the maintenance period in the model traces. If applied per model cycle, an average drug acquisition cost per cycle (2 weeks) is calculated and applied every model cycle. For example, the drug acquisition costs for an IV treatment administered once every 4 weeks in the maintenance period would be incurred by applying the average cost per cycle over two model cycles. This is the standard method of applying acquisition costs in economic models and is used in the base case.

The inclusion of per treatment cycle costs in the model was exploratory in nature. If applied per treatment cycle, acquisition costs are calculated for the full dose (e.g. 90 mg ustekinumab once every 12 weeks), and are incurred in line with the dosing schedule (e.g. once every 6 model cycles), assuming there are no deviations from this schedule across the entire patient cohort, rather than by applying an average per-cycle cost every model cycle. Please note that this method of applying treatment costs is only available for first line treatments, not subsequent treatments.

It is assumed that when costs are applied per treatment cycle a patient receives the full cost of the treatment upfront even if they discontinue treatment in subsequent model cycles. Whereas when costs are applied per model cycle, patients discontinuing treatment part-way through a treatment cycle accrue the proportional fraction of the full cost of the treatment cycle. If treatment cycles are greater in length than the model cycle (2 weeks) this leads to differences in the overall costs. In practice, there may be some deviations from the strict dosing schedules across the entire patient cohort, which would not be reflected in the per treatment cycle approach. Owing to this, applying costs per model cycle was deemed most appropriate. This approach was aligned with that taken for prior evaluations in the same indication (TA547).[10]

Utilities

B11. Document B, section B.3.4.1. From the TRUENORTH trial utility data presented in Tables 52 and 53, the improvement in 'Remission' and 'Response no remission' utilities between Week 10 (induction) and Week 52 (maintenance) is marginal (~[redacted]), whereas, for 'No response or remission (Active UC)', the improvement is greater (~[redacted]). Please explain any potential reasons for this difference.

The 'No response or remission (Active UC)' utility values at Week 10 and Week 52 were informed by distinct patient groups and cannot be directly compared.

Only patients who had responded to treatment in the induction period in TRUENORTH were included in the maintenance period efficacy analysis. As such, the patients informing the 'No response or remission (Active UC)' health state utility value at Week 52 were those who initially responded to treatment at Week 10 but subsequently lost response. The patients informing the 'No response or remission (Active UC)' health state utility value at Week 10 were those who did not achieve response at induction. These patients were subsequently excluded from efficacy analysis and do not inform the utility value at Week 52. Patients who had not responded in the induction period were optionally enrolled in an extension study and received open label ozanimod, and thus there are no 52-week data available for patients who did not achieve response or remission at Week 10.

Those patients having initially responded to treatment in the induction period, but who subsequently lost response in the maintenance period, may experience "residual" quality of life benefits from the initial response experienced following induction treatment. As such, these patients would likely report higher EQ-5D scores at Week 52 than would be observed for those who never achieved response, accounting for the difference in utility values for the 'No response or remission (Active UC)' state at Weeks 10 and 52.

Sensitivity analysis

B12. Document B, section B.3.8. The one-way sensitivity analysis (OWSA) tested key clinical parameters including sustained clinical response and remission at

maintenance. Please clarify why the probability of loss of response was not tested in the OWSA.

The probability of loss of response is calculated as 1 minus the sum of two probabilities: the probability of staying in remission and the probability of responding but not going into remission. Therefore, whilst this probability is not varied independently, it is varied when either of the other two probabilities it is calculated from are varied. Its effect on model outcomes is therefore tested implicitly as part of the sensitivity analysis of either the sustained remission or sustained response at maintenance.

B13. Document B, section B.3.8.1. Please clarify whether the probabilistic sensitivity analysis (PSA) accounts for joint uncertainty. If so, please describe how the correlation among the various input parameters (especially between the remission/response variables) has been captured in the PSA.

For the NMA parameters, for each treatment k, the raw CODA output of absolute probabilities of "no clinical response, no clinical remission" (PASI[1,k]), "clinical remission, no clinical response" (PASI[2,k]), and "clinical remission" (PASI[3,k]) from the MCMC iterations were used in the PSA. Each iteration of the MCMC sampling process represents an observation from the joint posterior distribution of all parameters specified in the NMA model, including the baseline treatment effect of placebo (parameter A), the treatment effects relative to placebo (d[k]), and the category effects of response (z[1]) and remission (z[2]). When the PSA is carried out in the cost-effectiveness analysis the absolute probabilities used in each PSA iteration are from a single iteration of the MCMC. Therefore, the correlation between each of these parameters is retained through the joint posterior distribution and in each iteration of the PSA.

Fully incremental analysis

B14. Document B, section B.3.7.1. The company has presented fully incremental analysis in Tables 70 and 71. However, the fully incremental analysis and the associated cost-effectiveness frontier are not available in the model. Please clarify.

The fully incremental analysis calculations and cost-effectiveness frontier have been incorporated into the revised model, attached alongside these responses. Please see the "Primary results" tab for details of the fully incremental analysis, and "Secondary results" tab for the cost-effectiveness frontier.

Half-cycle correction

B15. Document B, section B.3.2.2. The company submission states that a half-cycle correction has not been applied in the model owing to a shorter cycle length (2-weeks). However, it highlights that NICE technology appraisal 633 also had a shorter

cycle length (2-weeks) but applied a half-cycle correction. Please include an option in the model to allow for half-cycle correction.

A half cycle correction was independently applied to the Markov maintenance portion of the model (i.e. excluding the induction tunnel states, where one-off costs are applied and half-cycle correction is not appropriate) by manually calculating the half-cycle correction for the model engines. The results with the half-cycle correction are presented in Table 18 and Table 21. Given little difference was found between the results when a half cycle correction was applied (the change in incremental QALYs, incremental costs and ICERs at a £30,000 WTP threshold are of the order of a few percent as seen in Table 19 and Table 22) the company did not build the option to apply a half-cycle correction into the model as this would add significant complexity to the model, without substantial impact on the results.

The company also notes that for the model used in TA547, which had a cycle length of 8 weeks, a half-cycle correction was not applied. The reasoning given for this was that it is a relatively short duration of cycle length (Page 116 of the submission) which was accepted by the committee.[10]

Table 17: TNFi-naïve base case Results

	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER £ /QALY
Ozanimod	[redacted]	[redacted]	-	-	
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£28,686
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£167,024[a]
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£52,736[a]
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£71,023[a]

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Table 18: TNFi-naive half-cycle corrected results

	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER £ /QALY
Ozanimod	[redacted]	[redacted]	-	-	
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£28,912
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£168,417[a]
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£52,452[a]
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£70,502[a]

[a]SW quadrant ICER; costs saved per QALY forgone
Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Table 19: Percentage difference between half-cycle corrected and base case results in the TNFi-naïve model

	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER £ /QALY
Ozanimod	[redacted]	[redacted]			
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	0.79%
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	0.83%

Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	-0.54%
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	-0.73%

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Table 20: TNFi-experienced base case results

	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER £ /QALY
Ozanimod	[redacted]	[redacted]	-	-	
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£199,551[a]
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	-£33,725

[a]SW quadrant ICER; costs saved per QALY forgone

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Table 21: TNFi-experienced half-cycle corrected results

	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER £ /QALY
Ozanimod	[redacted]	[redacted]	-	-	
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£190,745[a]
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	-£33,962

[a]SW quadrant ICER; costs saved per QALY forgone

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Table 22: Percentage difference between the half-cycle corrected and base case results for the TNFi-experienced model

	Total costs	Total QALYs	Inc costs	Inc QALYs	ICER £ /QALY
Ozanimod	[redacted]	[redacted]			
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	-4.41%
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	0.70%

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Section C: Textual clarification and additional points

C1. Document B, section B.2.8.5.1. There seems to be an inconsistency in referencing of the appendix providing additional information on the assessment of inconsistency within the NMA. The company submission states that ‘An assessment of inconsistency was not possible due to the inconsistency model failing to converge (Appendix F.4.2).’ (p104) and ‘An assessment of inconsistency determined there to be little evidence of inconsistency between direct and indirect estimates for either model (Appendix F.4.2).’ (p99). Please confirm whether the reference should be to Appendix D.4.2. and, if so, correct this in the company submission.

The references in Section B.2.8.5.1 of the company submission should read Appendix D.4.2.

C2. Document B, section B.3.8.1. The company submission states that ‘A table containing a list of the inputs used in PSA is presented in .2.’. However, such a table is not available in the document. Please clarify and correct in the company submission, if necessary.

The reference in Section B.3.8.1 of the company submission should read Appendix J.2.

References

1. Cochrane Handbook for Systematic Reviews of Interventions version 6.1. In: Julian Higgins JT, Jacqueline Chandler, Miranda Cumpston, Tianjing Li, Matthew Page, Vivian Welch, ed. Volume 2020.
2. CRD's guidance for undertaking reviews in health care: University of York, 2009.
3. European Medicines Agency. Cimzia. Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/cimzia. [Accessed: 14/ February 2022].
4. Sandborn WJ, Feagan BG, D'Haens G, et al. Ozanimod as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med 2021;385:1280-1291.
5. Sandborn WJ, Feagan BG, Wolf DC, et al. Ozanimod Induction and Maintenance Treatment for Ulcerative Colitis. N Engl J Med 2016;374:1754-62.
6. NICE DSU TECHNICAL SUPPORT DOCUMENT 18: METHODS FOR POPULATION-ADJUSTED INDIRECT COMPARISONS IN SUBMISSIONS TO NICE. NICE DSU. Available at: http://nicedsu.org.uk/wp-content/uploads/2018/08/Population-adjustment-TSD-FINAL-ref-rerun.pdf [Last accessed: 24/07/21].
7. NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. Available at https://www.ncbi.nlm.nih.gov/books/NBK310366/pdf/Bookshelf_NBK310366.pdf. [Accessed 18 February 2022].
8. National Institute for Health and Care Excellence (NICE). TA633: Ustekinumab for treating moderately to severely active ulcerative colitis [ID1511]. Committee papers. Available at: https://www.nice.org.uk/guidance/ta633/documents/committee-papers-3 [Accessed 16 February 2022].
9. National Institute for Health and Care Excellence (NICE). TA633: Ustekinumab for treating moderately to severely active ulcerative colitis [ID1511]. Final appraisal document. Available at: https://www.nice.org.uk/guidance/ta633/documents/final-appraisal-determination-document-2 [Accessed 16 February 2022].
10. National Institute for Health and Care Excellence (NICE). TA547: Tofacitinib for treating moderately to severely active ulcerative colitis. Available at https://www.nice.org.uk/guidance/ta547. [Last accessed: 24/11/20].
11. Sands BE, Peyrin-Biroulet L, Loftus EV, Jr., et al. Vedolizumab versus Adalimumab for Moderate-to-Severe Ulcerative Colitis. N Engl J Med 2019;381:1215-1226.
12. Sandborn WJ, van Assche G, Reinisch Wea. Adalimumab in the Treatment of Moderate-to-Severe Ulcerative Colitis: ULTRA 2 Trial Results. Gastroenterology & hepatology 2013;9:317-320.
13. Sandborn WJ, Baert F, Danese S, et al. Efficacy and Safety of Vedolizumab Subcutaneous Formulation in a Randomized Trial of Patients With Ulcerative Colitis. Gastroenterology 2020;158:562-572.e12.

14. Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med 2017;376:1723-1736.
15. Sands BE, Sandborn WJ, Panaccione R, et al. Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med 2019;381:1201-1214.
16. Motoya S, Watanabe K, Ogata H, et al. Vedolizumab in Japanese patients with ulcerative colitis: A Phase 3, randomized, double-blind, placebo-controlled study. PLoS One 2019;14:e0212989.
17. 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). J Gastroenterol 2017;52:1101-1111.
18. 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. J Clin Gastroenterol 2015;49:582-8.
19. Suzuki Y, Motoya S, Hanai H, et al. Efficacy and safety of adalimumab in Japanese patients with moderately to severely active ulcerative colitis. J Gastroenterol 2014;49:283-94.
20. Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab maintains clinical response in patients with moderate-to-severe ulcerative colitis. Gastroenterology 2014;146:96-109.e1.
21. Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab induces clinical response and remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology 2014;146:85-95; quiz e14-5.
22. Sandborn WJ, Ghosh S, Panes J, et al. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. N Engl J Med 2012;367:616-24.
23. 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;142:257-65.e1-3.
24. Reinisch W, Sandborn WJ, Hommes DW, et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. Gut 2011;60:780-7.
25. Rutgeerts P, Sandborn WJ, Feagan BG, et al. Infliximab for induction and maintenance therapy for ulcerative colitis. N Engl J Med 2005;353:2462-76.
26. 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. J Gastroenterol 2016;51:241-51.
27. Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med 2013;369:699-710.
28. Milev S, DiBonaventura MD, Quon P, et al. An economic evaluation of tofacitinib for the treatment of moderately-to-severely active ulcerative colitis: modeling the cost of treatment strategies in the United States. J Med Econ 2019;22:859-868.
29. Scott FI, Luo M, Shah Y, et al. Identification of the Most Cost-effective Position of Vedolizumab Among the Available Biologic Drugs for the Treatment of Ulcerative Colitis. J Crohns Colitis 2020;14:575-587.

30. 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 Gastroenterol 2016;3:e000093.
31. Nguyen V, Carlson JJ, Bloudek L. PGI26 The Cost-Effectiveness of Pharmacologic Vs Surgical Intervention for Patients Hospitalized with Acute Severe Ulcerative Colitis in the United States. Value in Health 2020;23:S147.
32. Shaffer SR, Huang E, Patel S, et al. Cost-Effectiveness of 5-Aminosalicylate Therapy in Combination With Biologics or Tofacitinib in the Treatment of Ulcerative Colitis. Am J Gastroenterol 2021;116:125-133.
33. Ollendorf DA BL, Carlson JJ, Pandey R, Fazioli K, Chapman R, Bradt P, Pearson SD. ,. Effectiveness and Value; Evidence Report. Institute for Clinical and Economic Review. Available at https://icer.org/assessment/ulcerative-colitis-2020/. [Accessed 15 February 2022], 2020.
34. Bloudek LM, Pandey R, Fazioli K, et al. Optimal treatment sequence for targeted immune modulators for the treatment of moderate to severe ulcerative colitis. J Manag Care Spec Pharm 2021;27:1046-1055.
35. Wu B, Wang Z, Zhang Q. Cost-Effectiveness of Different Strategies for the Treatment of Moderate-to-Severe Ulcerative Colitis. Inflamm Bowel Dis 2018;24:2291-2302.
36. Ung V, Thanh NX, Wong K, et al. Real-life treatment paradigms show infliximab is cost-effective for management of ulcerative colitis. Clin Gastroenterol Hepatol 2014;12:1871-8.e8.
37. Chaudhary MA, Fan T. Cost-Effectiveness of Infliximab for the Treatment of Acute Exacerbations of Ulcerative Colitis in the Netherlands. Biol Ther 2013;3:45-60.
38. Trigo-Vicente C, Gimeno-Ballester V, López-Del Val A. Cost-effectiveness analysis of infliximab, adalimumab, golimumab, vedolizumab and tofacitinib for moderate to severe ulcerative colitis in Spain. Eur J Hosp Pharm 2020;27:355-360.
39. Trigo-Vicente C, Gimeno-Ballester V, Montoiro-Allué R, et al. Cost-effectiveness analysis of infliximab, adalimumab, golimumab and vedolizumab for moderate to severe ulcerative colitis in Spain. Expert Rev Pharmacoecon Outcomes Res 2018;18:321-329.
40. 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. Eur J Clin Pharmacol 2016;72:1319-1325.
41. 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:e0160444.
42. 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;36:472-81.
43. Petryszyn P, Ekk-Cierniakowski P, Zurakowski G. Infliximab, adalimumab, golimumab, vedolizumab and tofacitinib in moderate to severe ulcerative colitis: comparative cost-effectiveness study in Poland. Therap Adv Gastroenterol 2020;13:1756284820941179.
44. Hernandez L, Kuwabara H, Shah A, et al. Cost-Effectiveness Analysis of Vedolizumab Compared with Other Biologics in Anti-TNF-Naïve Patients with

Moderate-to-Severe Ulcerative Colitis in Japan. Pharmacoeconomics 2020;38:69-84.
45. Moradi N, Tofighi S, Zanganeh M, et al. Economic Evaluation of Infliximab for Treatment of Refractory Ulcerative Colitis In Iran: Cost-Effectiveness Analysis. Value in Health 2015;18:A628.
46. Beilman CL, Thanh NX, Ung V, et al. Real-Life Treatment Paradigms Show Adalimumab Is Cost-Effective for the Management of Ulcerative Colitis. Can J Gastroenterol Hepatol 2016;2016:5315798.
47. Xie F, Blackhouse G, Assasi N, et al. Cost-utility analysis of infliximab and adalimumab for refractory ulcerative colitis. Cost Eff Resour Alloc 2009;7:20.
48. Toor K, Druyts E, Jansen JP, et al. Cost per remission and cost per response with infliximab, adalimumab, and golimumab for the treatment of moderately-to-severely active ulcerative colitis. J Med Econ 2015;18:437-46.
49. Sruamsiri R. Cost-Effectiveness of Vedolizumab for the Management of Moderate-To-Severe Ulcerative Colitis in Thailand. Value in Health 2018;21:S2.
50. Menchén L, de Andrés-Nogales F, García S, et al. PGI16 Cost-Effectiveness of Tofacitinib for the Treatment of Moderate-to-Severe Ulcerative Colitis after Biologic Failure or Intolerance in Spain. Value in Health 2019;22:S619.
51. Taxonera C, de Andrés-Nogales F, García S, et al. PGI17 Cost-Effectiveness of Tofactinib for the Treatment of Moderate-to-Severe Ulcerative Colitis after Conventional Therapy in Spain. Value in Health 2019;22:S619.
52. Quon P, Sardesai A, Milev S, et al. PGI9 Cost-Effectiveness of Tofacitinib Compared with Infliximab, Adalimumab, Golimumab and Vedolizumab for the Treatment of Moderate to Severe Ulcerative Colitis in Germany. Value in Health 2019;22:S617.
53. De La Puente C, Hurtado V. PSY100 - Cost-Effectiveness Analysis of Second Line Treatment for Patients with Ulcerative Colitis. Value in Health 2018;21:S453.
54. Brito N, Mulinari E. PGI18 Cost-Effectiveness of Biologics for Moderate to Severe Ulcerative Colitis in Brazil. Value in Health 2020;23:S146.
55. Stern S, Ward AJ, Saint-Laurent Thibault C, 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. J Med Econ 2018;21:27-37.
56. Vellopoulou K, Stefanou G, Tzanetakos C, et al. Cost-effectiveness of tofacitinib for the treatment of moderate to severe active ulcerative colitis in Greece. Eur J Gastroenterol Hepatol 2021;33:325-333.
57. Taxonera C, de Andrés-Nogales F, García-López S, et al. Cost-effectiveness analysis of using innovative therapies for the management of moderate-to-severe ulcerative colitis in Spain. Expert Rev Pharmacoecon Outcomes Res 2022;22:73-83.
58. Sardesai A, Dignass A, Quon P, et al. Cost-effectiveness of tofacitinib compared with infliximab, adalimumab, golimumab, vedolizumab and ustekinumab for the treatment of moderate to severe ulcerative colitis in Germany. J Med Econ 2021;24:279-290.
59. Uda A, Eto Y, Li Y, et al. Evaluation of the Optimal Position for Vedolizumab in the Japanese Treatment Paradigm for Ulcerative Colitis Using Markov Modeling. Crohn's & Colitis 360 2020;2.

60. Sheng Y, Ma L, Chen J, et al. PGI9 Cost-Effectiveness Comparison of Vedolizumab, Infliximab and Conventional Therapy for the Treatment of Biologic-Naïve, Moderately to Severely ACTIVE Ulcerative Colitis Patients in China. Value in Health 2020;23:S534.
61. Canadian Agency for Drugs and Technologies in Health. Common Drug Review. Pharmacoeconomic Review Report. Adalimumab (Humira). Available at https://cadth.ca/sites/default/files/cdr/pharmacoeconomic/SR0450_Humira_UC_PE_Report.pdf. [Accessed 16 February 2022].
62. Canadian Agency for Drugs and Technologies in Health. Common Drug Review. Pharmacoeconomic Review Report. Tofactininb (Xeljanz). Available at https://www.cadth.ca/sites/default/files/cdr/pharmacoeconomic/sr0572-xeljanz-pharmacoeconomic-report.pdf. [Accessed 16 January 2022].
63. Canadian Agency for Drugs and Technologies in Health. Common Drug Review. Pharmacoeconomic Review Report. Ustekinumab (Stelara/Stelara IV). Available at https://www.cadth.ca/sites/default/files/cdr/pharmacoeconomic/sr0627-stelara-pharmacoeconomic-review-report.pdf. [Accessed 16 January 2022].
64. Canadian Agency for Drugs and Technologies in Health. Common Drug Review. Pharmacoeconomic Review Report. Vedolizumab (Entyvio SC). Available at https://www.cadth.ca/sites/default/files/cdr/pharmacoeconomic/sr0635-entyvio-pharmacoeconomic-review-report.pdf. [Accessed 16 January 2022].
65. European Medicines Agency. PRAC recommendations on signals. https://www.ema.europa.eu/en/documents/prac-recommendation/prac-recommendations-signals-adopted-7-10-june-2021-prac-meeting_en.pdf [Accessed 21 February 2022].
66. European Medicines Agency. EMA starts safety review of Janus kinase inhibitors for inflammatory disorders. Available at https://www.ema.europa.eu/en/news/ema-starts-safety-review-janus-kinase-inhibitors-inflammatory-disorders. [Accessed 21 February 2022].
67. Food and Drug Administration. FDA Drug Safety Communication. Available at https://www.fda.gov/media/151936/download. [Accessed 16 February 2022].
68. Medicines and Healthcare products Regulatory Agency. Drug safety update. Available at https://www.gov.uk/drug-safety-update/tofacitinib-xeljanzv-new-measures-to-minimise-risk-of-major-adverse-cardiovascular-events-and-malignancies. [Accessed 21 February 2022].
69. King D, Reulen RC, Thomas T, et al. Changing patterns in the epidemiology and outcomes of inflammatory bowel disease in the United Kingdom: 2000-2018. Aliment Pharmacol Ther 2020;51:922-934.
70. Gawron LM, Gawron AJ, Kasper A, et al. Contraceptive method selection by women with inflammatory bowel diseases: a cross-sectional survey. Contraception 2014;89:419-25.
71. Cotton CC, Baird D, Sandler RS, et al. Hormonal Contraception Use is Common Among Patients with Inflammatory Bowel Diseases and an Elevated Risk of Deep Vein Thrombosis. Inflamm Bowel Dis 2016;22:1631-8.
72. Sterne JAC, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. BMJ 2016; 355; i4919; doi: 10.1136/bmj.i4919.

73. Joanna Briggs Institute Critical Appraisal Tools: https://joannabriggs.org/critical-appraisal-tools.

Appendix 1: Non-randomised controlled trials SLR

Study selection

Studies were selected for inclusion in two stages: first, the titles and abstracts of the search results were reviewed for relevance using the eligibility criteria presented in Table 1; second, the full-texts of potentially relevant articles were screened in order to obtain the final list of included studies.

Abstract review

Each abstract was assessed for inclusion by two independent reviewers using the eligibility criteria (Table 1). Where the applicability of the inclusion criteria was unclear, the article was included at this stage to ensure that all potentially relevant studies were captured. The results of the two reviewers were compared and any disagreements were resolved by discussion until a consensus was met. If necessary, a third independent reviewer made the final decision.

Full-text review

Each full-text article was then assessed for inclusion by two independent reviewers using the eligibility criteria (Table 1). In cases where the article did not give enough information to be sure that it met the inclusion criteria, the article was excluded to ensure that only relevant articles were ultimately included in the SLR. The results of the two reviewers were compared and any disagreements were resolved by discussion until a consensus was met. If necessary, a third independent reviewer made the final decision.

Data extraction and quality assessment

For each included study, it was planned that data would be extracted by a single individual into a pre-specified data extraction table. It was planned that a second individual would then verify the extracted data, with any discrepancies resolved by a third independent reviewer.

Quality assessment strategy

It was planned that the quality of all included interventional studies would be assessed using the Risk Of Bias In Non-randomized Studies - of Interventions (ROBINS-I) critical appraisal tool[72] while the quality of all included observational studies would be assessed using the JBI critical appraisal tools developed by the Joanna Briggs Institute.[73] It was planned that the quality assessment would be completed by one independent reviewer and verified by a second independent reviewer. If necessary, any discrepancies would be resolved by a third independent reviewer.

Search results

A total of 252 records were retrieved by the electronic database searches. After de-duplication of results, 177 unique records were suitable for review. After title and abstract review, 6 records were selected to be reviewed at the full-text stage. Of these, 0 records were found to fulfil the eligibility criteria for inclusion in the SLR.

Supplementary searches of conferences, SLR bibliographies and clinical trials registries yielded 39 records. Of these, 0 records fulfilling the eligibility criteria were identified.

Studies included in the SLR

No studies were included in this SLR, therefore no data extractions or quality assessments were required.

Studies excluded from the SLR

A list of electronic database records excluded at the full-text review stage of the SLR is presented in Table 30, Electronic database records excluded at the full-text review stage of the SLR along with a brief rationale for exclusion.

Table 30: Electronic database records excluded at the full-text review stage of the SLR

#	Full Reference	Reason for Exclusion
1	Alimohammadi N, Koosha F, Rafeian-Kopaei M. Current, New and Future Therapeutic Targets in Inflammatory Bowel Disease: A Systematic Review. Curr Pharm Des. 2020;26(22):2668-2675.	SLR
2	Jairath V, Jeyarajah J, Zou G, Parker CE, Olson A, Khanna R, D'Haens GR, Sandborn WJ, Feagan BG. A composite disease activity index for early drug development in ulcerative colitis: development and validation of the UC-100 score. Lancet Gastroenterol Hepatol. 2019 Jan;4(1):63-70. doi: 10.1016/S2468-1253(18)30306-6.	RCT/modelling study
3	Lichert, F. Ulcerative colitis: Increased remission rates with 1 mg ozanimod daily. Zeitschrift fur Gastroenterologie. 2016;54(10):1114	Non-English language
4	Lucaciu LA, Seicean R, Seicean A. Small molecule drugs in the treatment of inflammatory bowel diseases: which one, when and why? - a systematic review. Eur J Gastroenterol Hepatol. 2020 Jun;32(6):669-677.	SLR
5	Singh S, Fumery M, Sandborn WJ, Murad MH. Systematic review with network meta-analysis: first- and second-line pharmacotherapy for moderate-severe ulcerative colitis. Aliment Pharmacol Ther. 2018 Jan;47(2):162-175.	SLR
6	Trigo-Vicente C, Gimeno-Ballester V, García-López S, López-Del Val A. Systematic review and network meta-analysis of treatment for moderate-to-severe ulcerative colitis. Int J Clin Pharm. 2018 Dec;40(6):1411-1419.	SLR

Abbreviations: RCT: randomised controlled trial; SLR: systematic literature review.

Appendix 2: NMA trials outcomes

The trial outcome results for clinical response and clinical remission included in the base case analysis are presented for each arm in each trial, along with the total number of patients included in each arm of each trial, in Table 31 to Table 33 (Appendix 2.1). The trial outcome results for clinical response and clinical remission included in the sensitivity analyses are presented in Table 35–Table 49 (Appendix 2.2).

Appendix 2.1: Base case NMA

Table 31: NMA inputs for induction period, TNFi-naïve population

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	8	Infliximab Pooled	159	243	65.4%	86	243	35.4%
		Placebo	45	121	37.2%	18	121	14.9%
ACT 2	8	Infliximab Pooled	161	241	66.8%	74	241	30.7%
		Placebo	36	123	29.3%	7	123	5.7%
GEMINI 1	6	IV Vedolizumab 300 mg	69	130	53.1%	30	130	23.1%
		Placebo	20	76	26.3%	5	76	6.6%
Jiang 2015	8	Infliximab Pooled	32	41	78.0%	22	41	53.7%
		Placebo	15	41	36.6%	9	41	22.0%
Kobayashi 2016	8	Infliximab Pooled	57	104	54.8%	21	104	20.2%
		Placebo	37	104	35.6%	11	104	10.6%
Motoya 2019	10	IV Vedolizumab 300 mg	42	79	53.2%	22	79	27.8%
		Placebo	15	41	36.6%	6	41	14.6%
OCTAVE 1 + 2	8	Placebo	43	110	39.1%	13	110	11.8%
		Tofacitinib 10 mg BID	284	440	64.5%	106	440	24.1%
PURSUIT-SC	6	Placebo	89	292	30.5%	20	292	6.8%
		SC Golimumab 200/100 mg	147	294	50.0%	52	294	17.7%
Study A3921063	8	Placebo	15	33	45.5%	NA	NA	NA
		Tofacitinib 10 mg BID	14	23	60.9%	NA	NA	NA
Suzuki 2014	8	Adalimumab 160/80/40 mg Q2W	45	90	50.0%	9	90	10.0%
		Placebo	34	96	35.4%	11	96	11.5%
TRUE NORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ULTRA 1	8	Adalimumab 160/80/40 mg Q2W	71	130	54.6%	24	130	18.5%
		Placebo	58	130	44.6%	12	130	9.2%
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	89	150	59.3%	32	150	21.3%
		Placebo	56	145	38.6%	16	145	11.0%
UNIFI	8	Placebo	56	158	35.4%	15	158	9.5%
		Ustekinumab Pooled	194	312	62.2%	60	312	19.2%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	151	305	49.5%	72	305	23.6%
		IV Vedolizumab 300 mg	213	304	70.1%	84	304	27.6%

Abbreviations: BID: twice a day; IV: intravenous; QD: once every day; Q2W: once every 2 weeks; SC: subcutaneous.

Table 32: NMA inputs for induction period, TNFi-experienced population

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	6	IV Vedolizumab 300 mg	32	82	39.0%	8	82	9.8%
		Placebo	13	63	20.6%	2	63	3.2%
Motoya 2019	10	IV Vedolizumab 300 mg	23	85	27.1%	8	85	9.4%
		Placebo	12	41	29.3%	4	41	9.8%
OCTAVE 1 + 2	8	Placebo	29	124	23.4%	1	124	0.8%
		Tofacitinib 10 mg BID	237	465	51.0%	53	465	11.4%
Study A3921063	8	Placebo	5	15	33.3%	NA	NA	NA
		Tofacitinib 10 mg BID	6	10	60.0%	NA	NA	NA
TRUE NORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	36	98	36.7%	9	98	9.2%
		Placebo	29	101	28.7%	7	101	6.9%

UNIFI	8	Placebo	44	161	27.3%	2	161	1.2%
		Ustekinumab Pooled	169	330	51.2%	40	330	12.1%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	26	81	32.1%	10	81	12.3%
		IV Vedolizumab 300 mg	44	79	55.7%	18	79	22.8%

Abbreviations: BID: twice a day; IV: intravenous; QD: once every day; Q2W: once every 2 weeks; SC: subcutaneous.

Table 33: NMA inputs for maintenance period, TNFi-naïve population

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	54	Infliximab Pooled	92	159	57.9%	53	159	33.3%
		Placebo	17	45	37.8%	10	45	22.2%
GEMINI 1	52	Placebo	21	79	26.6%	15	79	19.0%
		Vedolizumab Pooled	88	145	60.7%	68	145	46.9%
Motoya 2019	60	Placebo	10	28	35.7%	10	28	35.7%
		Vedolizumab Pooled	16	24	66.7%	13	24	54.2%
OCTAVE SUSTAIN	52	Placebo	27	109	24.8%	12	109	11.0%
		Tofacitinib Pooled	132	219	60.3%	94	219	42.9%
PURSUIT-J	54	Golimumab Pooled	18	32	56.3%	16	32	50.0%
		Placebo	6	31	19.4%	2	31	6.5%
PURSUIT-M	54	Golimumab Pooled	146	302	48.3%	101	302	33.4%
		Placebo	48	154	31.2%	34	154	22.1%
Suzuki 2014	52	Adalimumab 40 mg Q2W	50	82	61.0%	38	82	46.3%
		Placebo	12	34	35.3%	8	34	23.5%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	52	Adalimumab 40 mg Q2W	44	89	49.4%	34	89	38.2%
		Placebo	24	56	42.9%	15	56	26.8%
UNIFI	52	Placebo	44	87	50.6%	27	87	31.0%

		Ustekinumab Pooled	144	187	77.0%	91	187	48.7%
VISIBLE 1	52	Placebo	NA	NA	NA	7	37	18.9%
		Vedolizumab 108 mg Q2W SC	NA	NA	NA	36	67	53.7%
		Vedolizumab Pooled	NA	NA	NA	17	32	53.1%

Abbreviations: BID: twice a day; IV: intravenous; QD: once every day; Q2W: once every 2 weeks; SC: subcutaneous.

Table 34: NMA inputs for maintenance period, TNFi-experienced population

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI I	52	Placebo	6	38	15.8%	2	38	5.3%
		Vedolizumab Pooled	37	83	44.6%	30	83	36.1%
Motoya 2019	60	Placebo	5	14	35.7%	3	14	21.4%
		Vedolizumab Pooled	11	17	64.7%	10	17	58.8%
OCTAVE SUSTAIN	52	Placebo	13	89	14.6%	10	89	11.2%
		Tofacitinib Pooled	92	176	52.3%	54	176	30.7%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	52	Adalimumab 40 mg	15	36	41.7%	10	36	27.8%
		Placebo	6	29	20.7%	2	29	6.9%
UNIFI	52	Placebo	34	88	38.6%	15	88	17.0%
		Ustekinumab Pooled	98	161	60.9%	52	161	32.3%
VISIBLE 1	52	Placebo	NA	NA	NA	1	19	5.3%
		Vedolizumab 108 mg Q2W SC	NA	NA	NA	13	39	33.3%
		Vedolizumab Pooled	NA	NA	NA	6	22	27.3%

Abbreviations: BID: twice a day; IV: intravenous; QD: once every day; Q2W: once every 2 weeks; SC: subcutaneous.

Appendix 2.2: Sensitivity analyses NMA

Unpooled doses

Table 35: NMA inputs for induction period, TNFi-naïve population, unpooled doses

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	8	Infliximab 10 mg/kg	75	122	61.5%	39	122	32.0%
		Infliximab 5 mg/kg	84	121	69.4%	47	121	38.8%
		Placebo	45	121	37.2%	18	121	14.9%
ACT 2	8	Infliximab 10 mg/kg	83	120	69.2%	33	120	27.5%
		Infliximab 5 mg/kg	78	121	64.5%	41	121	33.9%
		Placebo	36	123	29.3%	7	123	5.7%
GEMINI 1	6	IV Vedolizumab 300 mg	69	130	53.1%	30	130	23.1%
		Placebo	20	76	26.3%	5	76	6.6%
Jiang 2015	8	Infliximab 5 mg/kg	32	41	78.0%	22	41	53.7%
		Placebo	15	41	36.6%	9	41	22.0%
Kobayashi 2016	8	Infliximab 5 mg/kg	57	104	54.8%	21	104	20.2%
		Placebo	37	104	35.6%	11	104	10.6%
Motoya 2019	10	IV Vedolizumab 300 mg	42	79	53.2%	22	79	27.8%
		Placebo	15	41	36.6%	6	41	14.6%
OCTAVE 1 + 2	8	Placebo	43	110	39.1%	13	110	11.8%
		Tofacitinib 10 mg BID	284	440	64.5%	106	440	24.1%
PURSUIT-SC	6	Golimumab 200/100 mg	147	294	50.0%	52	294	17.7%
		Placebo	89	292	30.5%	20	292	6.8%
Study A3921063	8	Placebo	15	33	45.5%	NA	NA	NA
		Tofacitinib 10 mg BID	14	23	60.9%	NA	NA	NA

Suzuki 2014	8	Adalimumab 160/80/40 mg Q2W	45	90	50.0%	9	90	10.0%
		Placebo	34	96	35.4%	11	96	11.5%
TRUE NORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 1	8	Adalimumab 160/80/40 mg Q2W	71	130	54.6%	24	130	18.5%
		Placebo	58	130	44.6%	12	130	9.2%
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	89	150	59.3%	32	150	21.3%
		Placebo	56	145	38.6%	16	145	11.0%
UNIFI	8	IV Ustekinumab 130 mg	90	156	57.7%	31	156	19.9%
		IV Ustekinumab 6 mg/kg	104	156	66.7%	29	156	18.6%
		Placebo	56	158	35.4%	15	158	9.5%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	151	305	49.5%	72	305	23.6%
		IV Vedolizumab 300 mg	213	304	70.1%	84	304	27.6%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 36: NMA inputs for induction period, TNFi-experienced population, unpooled doses

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	6	IV Vedolizumab 300 mg	32	82	39.0%	8	82	9.8%
		Placebo	13	63	20.6%	2	63	3.2%
Motoya 2019	10	IV Vedolizumab 300 mg	23	85	27.1%	8	85	9.4%
		Placebo	12	41	29.3%	4	41	9.8%
OCTAVE 1 + 2	8	Placebo	29	124	23.4%	1	124	0.8%
		Tofacitinib 10 mg OD BID	237	465	51.0%	53	465	11.4%
Study A3921063	8	Placebo	5	15	33.3%	NA	NA	NA

		Tofacitinib 10 mg BID	6	10	60.0%	NA	NA	NA
TRUE NORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	36	98	36.7%	9	98	9.2%
		Placebo	29	101	28.7%	7	101	6.9%
UNIFI	8	IV Ustekinumab 130 mg	74	164	45.1%	19	164	11.6%
		IV Ustekinumab 6 mg/kg	95	166	57.2%	21	166	12.7%
		Placebo	44	161	27.3%	2	161	1.2%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	26	81	32.1%	10	81	12.3%
		IV Vedolizumab 300 mg	44	79	55.7%	18	79	22.8%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 37: NMA inputs for maintenance period, TNFi-naïve population, unpooled doses

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	54	Infliximab 10 mg/kg	45	75	60.0%	25	75	33.3%
		Infliximab 5 mg/kg	47	84	56.0%	28	84	33.3%
		Placebo	17	45	37.8%	10	45	22.2%
GEMINI 1	52	Placebo	21	79	26.6%	15	79	19.0%
		Vedolizumab 300 mg Q4W	41	73	56.2%	35	73	47.9%
		Vedolizumab 300 mg Q8W	47	72	65.3%	33	72	45.8%
Motoya 2019	60	Placebo	10	28	35.7%	10	28	35.7%
		Vedolizumab 300 mg Q8W	16	24	66.7%	13	24	54.2%
OCTAVE SUSTAIN	52	Placebo	27	109	24.8%	12	109	11.0%
		Tofacitinib 10 mg BID	67	104	64.4%	46	104	44.2%
		Tofacitinib 5 mg BID	65	115	56.5%	48	115	41.7%
PURSUIT-J	54	Golimumab 100 mg Q4W	18	32	56.3%	16	32	50.0%

		Placebo	6	31	19.4%	2	31	6.5%
PURSUIT-M	54	Golimumab 100 mg Q4W	75	151	49.7%	51	151	33.8%
		Golimumab 50 mg Q4W	71	151	47.0%	50	151	33.1%
		Placebo	48	154	31.2%	34	154	22.1%
Suzuki 2014	52	Adalimumab 40 mg Q2W	50	82	61.0%	38	82	46.3%
		Placebo	12	34	35.3%	8	34	23.5%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	52	Adalimumab 40 mg Q2W	44	89	49.4%	34	89	38.2%
		Placebo	24	56	42.9%	15	56	26.8%
UNIFI	52	Placebo	44	87	50.6%	27	87	31.0%
		Ustekinumab 90 mg Q12W	78	102	76.5%	50	102	49.0%
		Ustekinumab 90 mg Q8W	66	85	77.6%	41	85	48.2%
VISIBLE 1	52	Placebo	NA	NA	NA	7	37	18.9%
		Vedolizumab 108 mg Q2W SC	NA	NA	NA	36	67	53.7%
		Vedolizumab 300 mg Q8W	NA	NA	NA	17	32	53.1%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 38: NMA inputs for maintenance period, TNFi-experienced population, unpooled doses

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI I	52	Placebo	6	38	15.8%	2	38	5.3%
		Vedolizumab 300 mg Q4W	17	40	42.5%	14	40	35.0%
		Vedolizumab 300 mg Q8W	20	43	46.5%	16	43	37.2%
Motoya 2019	60	Placebo	5	14	35.7%	3	14	21.4%
		Vedolizumab 300 mg Q8W	11	17	64.7%	10	17	58.8%
OCTAVE SUSTAIN	52	Placebo	13	89	14.6%	10	89	11.2%
		Tofacitinib 10 mg BID	55	93	59.1%	34	93	36.6%

		Tofacitinib 5 mg BID	37	83	44.6%	20	83	24.1%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	52	Adalimumab 40 mg	15	36	41.7%	10	36	27.8%
		Placebo	6	29	20.7%	2	29	6.9%
UNIFI	52	Placebo	34	88	38.6%	15	88	17.0%
		Ustekinumab 90 mg Q12W	39	70	55.7%	16	70	22.9%
		Ustekinumab 90 mg Q8W	59	91	64.8%	36	91	39.6%
VISIBLE 1	52	Placebo	NA	NA	NA	1	19	5.3%
		Vedolizumab 108 mg Q2W SC	NA	NA	NA	13	39	33.3%
		Vedolizumab 300 mg Q8W	NA	NA	NA	6	22	27.3%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

TRUENORTH 3-component Mayo data

Table 39: NMA inputs for induction period, TNFi-naïve population, TRUENORTH 3-component Mayo data[a]

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT	8	Infliximab Pooled	159	243	65.4%	86	243	35.4%
		Placebo	45	121	37.2%	18	121	14.9%
ACT 2	8	Infliximab Pooled	161	241	66.8%	74	241	30.7%
		Placebo	36	123	29.3%	7	123	5.7%
GEMINI 1	6	Placebo	20	76	26.3%	5	76	6.6%
		Vedolizumab Pooled	69	130	53.1%	30	130	23.1%
Jiang 2015	8	Infliximab Pooled	32	41	78.0%	22	41	53.7%
		Placebo	15	41	36.6%	9	41	22.0%
Kobayashi 2016	8	Infliximab Pooled	57	104	54.8%	21	104	20.2%
		Placebo	37	104	35.6%	11	104	10.6%
Motoya 2019	10	Placebo	15	41	36.6%	6	41	14.6%
		Vedolizumab Pooled	42	79	53.2%	22	79	27.8%
OCTAVE 1 + 2	8	Placebo	43	110	39.1%	13	110	11.8%
		Tofacitinib Pooled	284	440	64.5%	106	440	24.1%
PURSUIT-SC	6	Golimumab 200/100 mg	147	294	50.0%	52	294	17.7%
		Placebo	89	292	30.5%	20	292	6.8%
Study A3921063	8	Placebo	15	33	45.5%	NA	NA	NA
		Tofacitinib Pooled	14	23	60.9%	NA	NA	NA
Suzuki 2014	8	Adalimumab 160/80/40 mg Q2W	45	90	50.0%	9	90	10.0%
		Placebo	34	96	35.4%	11	96	11.5%
TRUENORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ULTRA 1	8	Adalimumab 160/80/40 mg Q2W	71	130	54.6%	24	130	18.5%
		Placebo	58	130	44.6%	12	130	9.2%
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	89	150	59.3%	32	150	21.3%
		Placebo	56	145	38.6%	16	145	11.0%
UNIFI	8	Placebo	56	158	35.4%	15	158	9.5%
		Ustekinumab Pooled	194	312	62.2%	60	312	19.2%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	151	305	49.5%	72	305	23.6%
		Vedolizumab Pooled	213	304	70.1%	84	304	27.6%

[a]As the sensitivity analysis focused on the inclusion of remission/response data based on the 3-component Mayo score from the TRUENORTH trial, only the values for response and remission for TRUENORTH differ from the base case values in this sensitivity analysis.
Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 40: NMA inputs for induction period, TNFi-experienced population, TRUENORTH 3-component Mayo data[a]

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	6	Placebo	13	63	20.6%	2	63	3.2%
		Vedolizumab Pooled	32	82	39.0%	8	82	9.8%
Motoya 2019	10	Placebo	12	41	29.3%	4	41	9.8%
		Vedolizumab Pooled	23	85	27.1%	8	85	9.4%
OCTAVE 1 + 2	8	Placebo	29	124	23.4%	1	124	0.8%
		Tofacitinib Pooled	237	465	51.0%	53	465	11.4%
Study A3921063	8	Placebo	5	15	33.3%	NA	NA	NA
		Tofacitinib 10 mg BID	6	10	60.0%	NA	NA	NA
TRUENORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	36	98	36.7%	9	98	9.2%

		Placebo	29	101	28.7%	7	101	6.9%
UNIFI	8	Placebo	44	161	27.3%	2	161	1.2%
		Ustekinumab Pooled	169	330	51.2%	40	330	12.1%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	26	81	32.1%	10	81	12.3%
		Vedolizumab Pooled	44	79	55.7%	18	79	22.8%

[a]As this sensitivity analysis focused on the inclusion of remission/response data based on the 3-component Mayo score from the TRUENORTH trial, only the values for response and remission for TRUENORTH differ from the base case values in this sensitivity analysis.
Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 41: NMA inputs for maintenance period, TNFi-naïve population, TRUENORTH 3-component Mayo data[a]

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	54	Infliximab Pooled	92	159	57.9%	54	159	34.0%
		Placebo	17	45	37.8%	10	45	22.2%
GEMINI 1	52	Placebo	21	79	26.6%	15	79	19.0%
		Vedolizumab Pooled	88	145	60.7%	68	145	46.9%
Motoya 2019	60	Placebo	10	28	35.7%	10	28	35.7%
		Vedolizumab Pooled	16	24	66.7%	13	24	54.2%
OCTAVE SUSTAIN	52	Placebo	27	109	24.8%	12	109	11.0%
		Tofacitinib Pooled	132	219	60.3%	94	219	42.9%
PURSUIT-J	54	Golimumab 100 mg Q4W	18	32	56.3%	16	32	50.0%
		Placebo	6	31	19.4%	2	31	6.5%
PURSUIT-M	54	Golimumab Pooled	146	302	48.3%	101	302	33.4%
		Placebo	48	154	31.2%	34	154	22.1%
Suzuki 2014	52	Adalimumab 40 mg Q2W	50	82	61.0%	38	82	46.3%
		Placebo	12	34	35.3%	8	34	23.5%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ULTRA 2	52	Adalimumab 40 mg Q2W	44	89	49.4%	34	89	38.2%
		Placebo	24	56	42.9%	15	56	26.8%
UNIFI	52	Placebo	44	87	50.6%	27	87	31.0%
		Ustekinumab Pooled	144	187	77.0%	91	187	48.7%
VISIBLE 1	52	Placebo	NA	NA	NA	7	37	18.9%
		Vedolizumab 108 mg Q2W	NA	NA	NA	36	67	53.7%
		Vedolizumab Pooled	NA	NA	NA	17	32	53.1%

[a]As this sensitivity analysis focused on the inclusion of remission/response data based on the 3-component Mayo score from the TRUENORTH trial, only the values for response and remission for TRUENORTH differ from the base case values in this sensitivity analysis.
Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 42: NMA inputs for maintenance period, TNFi-experienced population, TRUENORTH 3-component Mayo data[a]

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	52	Placebo	6	38	15.8%	2	38	5.3%
		Vedolizumab Pooled	37	83	44.6%	30	83	36.1%
Motoya 2019	60	Placebo	5	14	35.7%	3	14	21.4%
		Vedolizumab Pooled	11	17	64.7%	10	17	58.8%
OCTAVE SUSTAIN	52	Placebo	13	89	14.6%	10	89	11.2%
		Tofacitinib Pooled	92	176	52.3%	54	176	30.7%
TRUENORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	52	Adalimumab 40 mg Q2W	15	36	41.7%	10	36	27.8%
		Placebo	6	29	20.7%	2	29	6.9%
UNIFI	52	Placebo	34	88	38.6%	15	88	17.0%
		Ustekinumab Pooled	98	161	60.9%	52	161	32.3%
VISIBLE 1	52	Placebo	NA	NA	NA	1	19	5.3%
		Vedolizumab 108 mg Q2W	NA	NA	NA	13	39	33.3%

		Vedolizumab Pooled	NA	NA	NA	6	22	27.3%

[a]As this sensitivity analysis focused on the inclusion of remission/response data based on the 3-component Mayo score from the TRUENORTH trial, only the values for response and remission for TRUENORTH differ from the base case values in this sensitivity analysis.
Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Exclusion of treat-through trials at maintenance

Table 43: NMA inputs for maintenance period, TNFi-naïve population, exclusion of treat-through trials at maintenance

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	52	Placebo	21	79	26.6%	15	79	19.0%
		Vedolizumab Pooled	88	145	60.7%	68	145	46.9%
Motoya 2019	60	Placebo	10	28	35.7%	10	28	35.7%
		Vedolizumab Pooled	16	24	66.7%	13	24	54.2%
OCTAVE SUSTAIN	52	Placebo	27	109	24.8%	12	109	11.0%
		Tofacitinib Pooled	132	219	60.3%	94	219	42.9%
PURSUIT-J	54	Golimumab 100 mg Q4W	18	32	56.3%	16	32	50.0%
		Placebo	6	31	19.4%	2	31	6.5%
PURSUIT-M	54	Golimumab Pooled	146	302	48.3%	101	302	33.4%
		Placebo	48	154	31.2%	34	154	22.1%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
UNIFI	52	Placebo	44	87	50.6%	27	87	31.0%
		Ustekinumab Pooled	144	187	77.0%	91	187	48.7%
VISIBLE 1	52	Placebo	NA	NA	NA	7	37	18.9%
		Vedolizumab 108 mg Q2W	NA	NA	NA	36	67	53.7%
		Vedolizumab Pooled	NA	NA	NA	17	32	53.1%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 44: NMA inputs for maintenance period, TNFi-experienced population, exclusion of treat-through trials at maintenance

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	52	Placebo	6	38	15.8%	2	38	5.3%
		Vedolizumab Pooled	37	83	44.6%	30	83	36.1%
Motoya 2019	60	Placebo	5	14	35.7%	3	14	21.4%
		Vedolizumab Pooled	11	17	64.7%	10	17	58.8%
OCTAVE SUSTAIN	52	Placebo	13	89	14.6%	10	89	11.2%
		Tofacitinib Pooled	92	176	52.3%	54	176	30.7%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
UNIFI	52	Placebo	34	88	38.6%	15	88	17.0%
		Ustekinumab Pooled	98	161	60.9%	52	161	32.3%
VISIBLE 1	52	Placebo	NA	NA	NA	1	19	5.3%
		Vedolizumab 108 mg Q2W	NA	NA	NA	13	39	33.3%
		Vedolizumab Pooled	NA	NA	NA	6	22	27.3%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Exclusion of Asian trials

Table 45: NMA inputs for induction period, TNFi-naïve population, exclusion of Asian trials

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	8	Infliximab Pooled	159	243	65.4%	86	243	35.4%
		Placebo	45	121	37.2%	18	121	14.9%
ACT 2	8	Infliximab Pooled	161	241	66.8%	74	241	30.7%
		Placebo	36	123	29.3%	7	123	5.7%
GEMINI 1	6	Placebo	20	76	26.3%	5	76	6.6%
		Vedolizumab Pooled	69	130	53.1%	30	130	23.1%
OCTAVE 1 + 2	8	Placebo	43	110	39.1%	13	110	11.8%
		Tofacitinib Pooled	284	440	64.5%	106	440	24.1%
PURSUIT-SC	6	Golimumab 200/100 mg	147	294	50.0%	52	294	17.7%
		Placebo	89	292	30.5%	20	292	6.8%
Study A3921063	8	Placebo	15	33	45.5%	NA	NA	NA
		Tofacitinib Pooled	14	23	60.9%	NA	NA	NA
TRUE NORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 1	8	Adalimumab 160/80/40 mg Q2W	71	130	54.6%	24	130	18.5%
		Placebo	58	130	44.6%	12	130	9.2%
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	89	150	59.3%	32	150	21.3%
		Placebo	56	145	38.6%	16	145	11.0%
UNIFI	8	Placebo	56	158	35.4%	15	158	9.5%
		Ustekinumab Pooled	194	312	62.2%	60	312	19.2%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	151	305	49.5%	72	305	23.6%

		Vedolizumab Pooled	213	304	70.1%	84	304	27.6%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 46: NMA inputs for induction period, TNFi-experienced population, exclusion of Asian trials

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	6	Placebo	13	63	20.6%	2	63	3.2%
		Vedolizumab Pooled	32	82	39.0%	8	82	9.8%
OCTAVE 1 + 2	8	Placebo	29	124	23.4%	1	124	0.8%
		Tofacitinib Pooled	237	465	51.0%	53	465	11.4%
Study A3921063	8	Placebo	5	15	33.3%	NA	NA	NA
		Tofacitinib Pooled	6	10	60.0%	NA	NA	NA
TRUE NORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	36	98	36.7%	9	98	9.2%
		Placebo	29	101	28.7%	7	101	6.9%
UNIFI	8	Placebo	44	161	27.3%	2	161	1.2%
		Ustekinumab Pooled	169	330	51.2%	40	330	12.1%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	26	81	32.1%	10	81	12.3%
		Vedolizumab Pooled	44	79	55.7%	18	79	22.8%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 47: NMA inputs for maintenance period, TNFi-naïve population, exclusion of Asian trials

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	54	Infliximab Pooled	92	159	57.9%	54	159	34.0%
		Placebo	17	45	37.8%	10	45	22.2%
GEMINI 1	52	Placebo	21	79	26.6%	15	79	19.0%
		Vedolizumab Pooled	88	145	60.7%	68	145	46.9%
OCTAVE SUSTAIN	52	Placebo	27	109	24.8%	12	109	11.0%
		Tofacitinib Pooled	132	219	60.3%	94	219	42.9%
PURSUIT-M	54	Golimumab Pooled	146	302	48.3%	101	302	33.4%
		Placebo	48	154	31.2%	34	154	22.1%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	52	Adalimumab 40 mg Q2W	44	89	49.4%	34	89	38.2%
		Placebo	24	56	42.9%	15	56	26.8%
UNIFI	52	Placebo	44	87	50.6%	27	87	31.0%
		Ustekinumab Pooled	144	187	77.0%	91	187	48.7%
VISIBLE 1	52	Placebo	NA	NA	NA	7	37	18.9%
		Vedolizumab 108 mg Q2W	NA	NA	NA	36	67	53.7%
		Vedolizumab Pooled	NA	NA	NA	17	32	53.1%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 48: NMA inputs for maintenance period, TNFi-experienced population, exclusion of Asian trials

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	52	Placebo	6	38	15.8%	2	38	5.3%
		Vedolizumab Pooled	37	83	44.6%	30	83	36.1%
OCTAVE SUSTAIN	52	Placebo	13	89	14.6%	10	89	11.2%
		Tofacitinib Pooled	92	176	52.3%	54	176	30.7%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	52	Adalimumab 40 mg Q2W	15	36	41.7%	10	36	27.8%
		Placebo	6	29	20.7%	2	29	6.9%
UNIFI	52	Placebo	34	88	38.6%	15	88	17.0%
		Ustekinumab Pooled	98	161	60.9%	52	161	32.3%
VISIBLE 1	52	Placebo	NA	NA	NA	1	19	5.3%
		Vedolizumab 108 mg Q2W	NA	NA	NA	13	39	33.3%
		Vedolizumab Pooled	NA	NA	NA	6	22	27.3%

Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Inclusion of TOUCHSTONE

Table 49: NMA inputs for induction period, TNFi-naïve population, inclusion of TOUCHSTONE[a]

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	8	Infliximab Pooled	159	243	65.4%	86	243	35.4%
		Placebo	45	121	37.2%	18	121	14.9%
ACT 2	8	Infliximab Pooled	161	241	66.8%	74	241	30.7%
		Placebo	36	123	29.3%	7	123	5.7%
GEMINI 1	6	Placebo	20	76	26.3%	5	76	6.6%
		Vedolizumab Pooled	69	130	53.1%	30	130	23.1%
Jiang 2015	8	Infliximab Pooled	32	41	78.0%	22	41	53.7%
		Placebo	15	41	36.6%	9	41	22.0%
Kobayashi 2016	8	Infliximab Pooled	57	104	54.8%	21	104	20.2%
		Placebo	37	104	35.6%	11	104	10.6%
Motoya 2019	10	Placebo	15	41	36.6%	6	41	14.6%
		Vedolizumab Pooled	42	79	53.2%	22	79	27.8%
OCTAVE 1 + 2	8	Placebo	43	110	39.1%	13	110	11.8%
		Tofacitinib Pooled	284	440	64.5%	106	440	24.1%
PURSUIT-SC	6	Golimumab 200/100 mg	147	294	50.0%	52	294	17.7%
		Placebo	89	292	30.5%	20	292	6.8%
Study A3921063	8	Placebo	15	33	45.5%	NA	NA	NA
		Tofacitinib Pooled	14	23	60.9%	NA	NA	NA
Suzuki 2014	8	Adalimumab 160/80/40 mg Q2W	45	90	50.0%	9	90	10.0%
		Placebo	34	96	35.4%	11	96	11.5%
TOUCHTONE	8	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

TRUE NORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 1	8	Adalimumab 160/80/40 mg Q2W	71	130	54.6%	24	130	18.5%
		Placebo	58	130	44.6%	12	130	9.2%
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	89	150	59.3%	32	150	21.3%
		Placebo	56	145	38.6%	16	145	11.0%
UNIFI	8	Placebo	56	158	35.4%	15	158	9.5%
		Ustekinumab Pooled	194	312	62.2%	60	312	19.2%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	151	305	49.5%	72	305	23.6%
		Vedolizumab Pooled	213	304	70.1%	84	304	27.6%

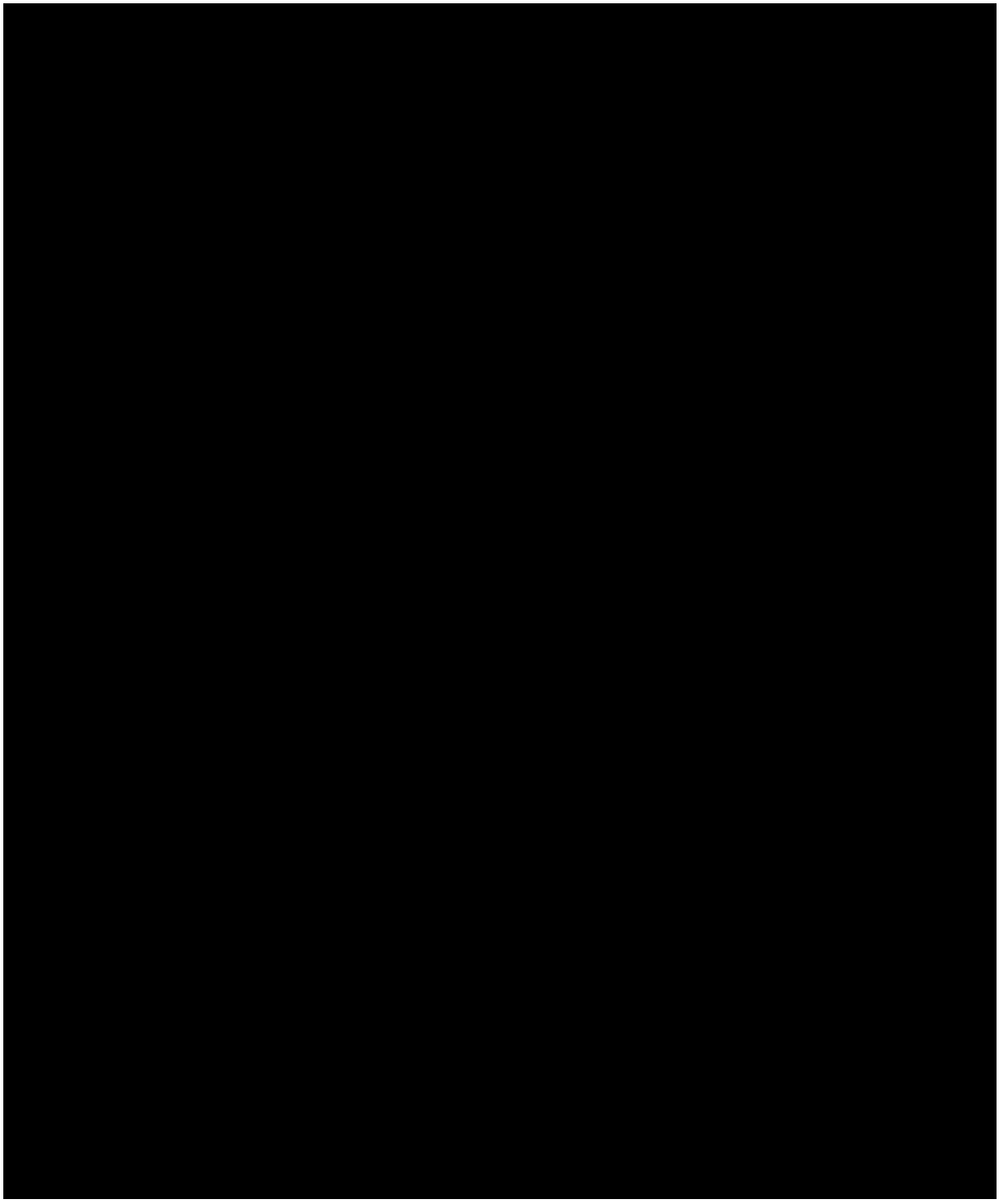
[a]As this sensitivity analysis focused on the inclusion of data from the TOUCHSTONE trial, all other values are unchanged from the base case.
Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Appendix 3: NMA JAGS code and summary results

JAGS code for the NMA are shown in Figure 11 and Figure 12 of Appendix 3.1 below. Lists of data inputs, value sets for any constants, and initial values and seeds used for the base case analysis are provided in Appendix 3.2. Summary results including the posterior means, medians, and 95% credible intervals for T and PASI are provided in Appendix 3.3.

Appendix 3.1: JAGS code

Figure 11: JAGS code used to generate fixed effect base case or sensitivity NMAs



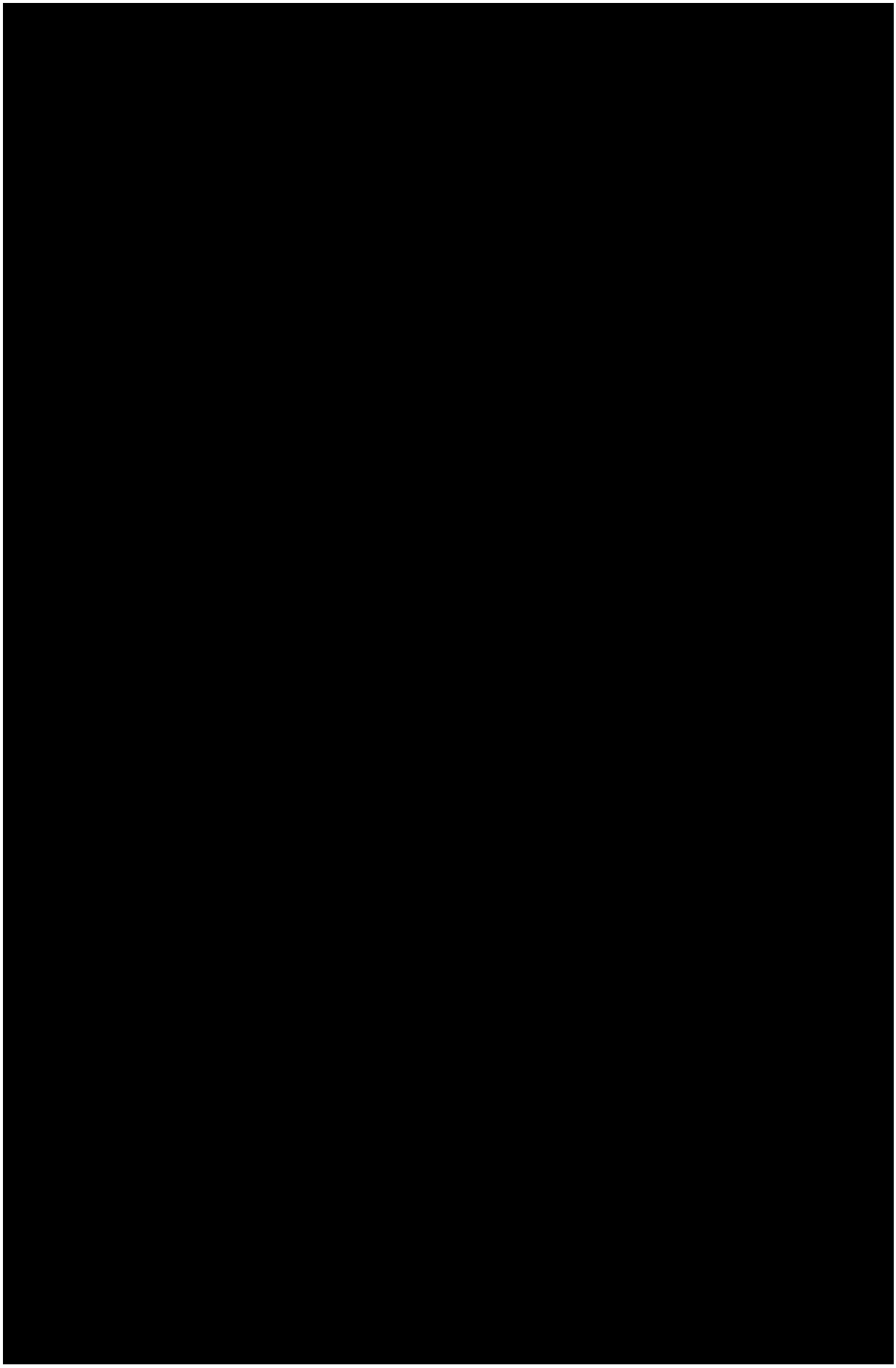
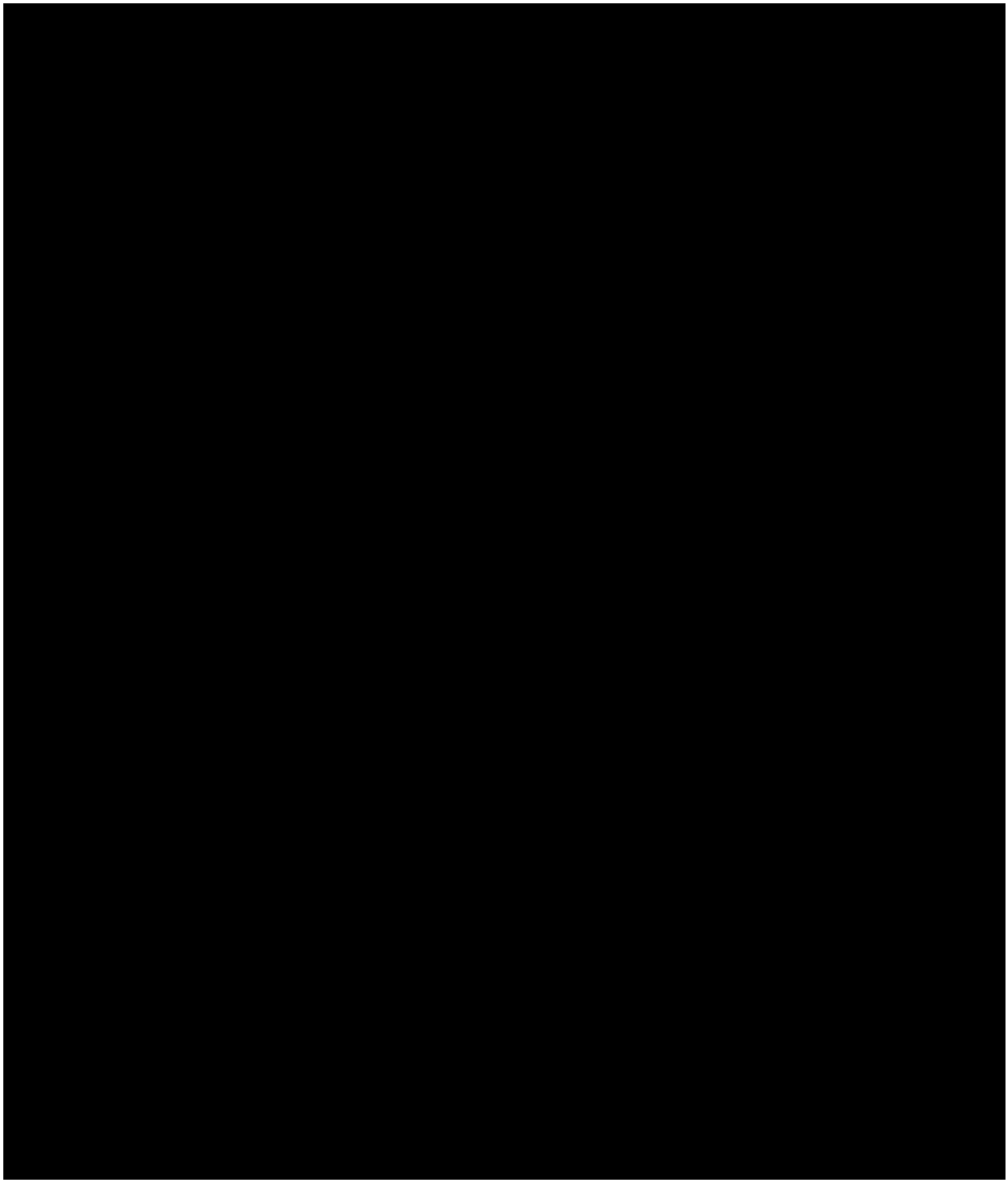
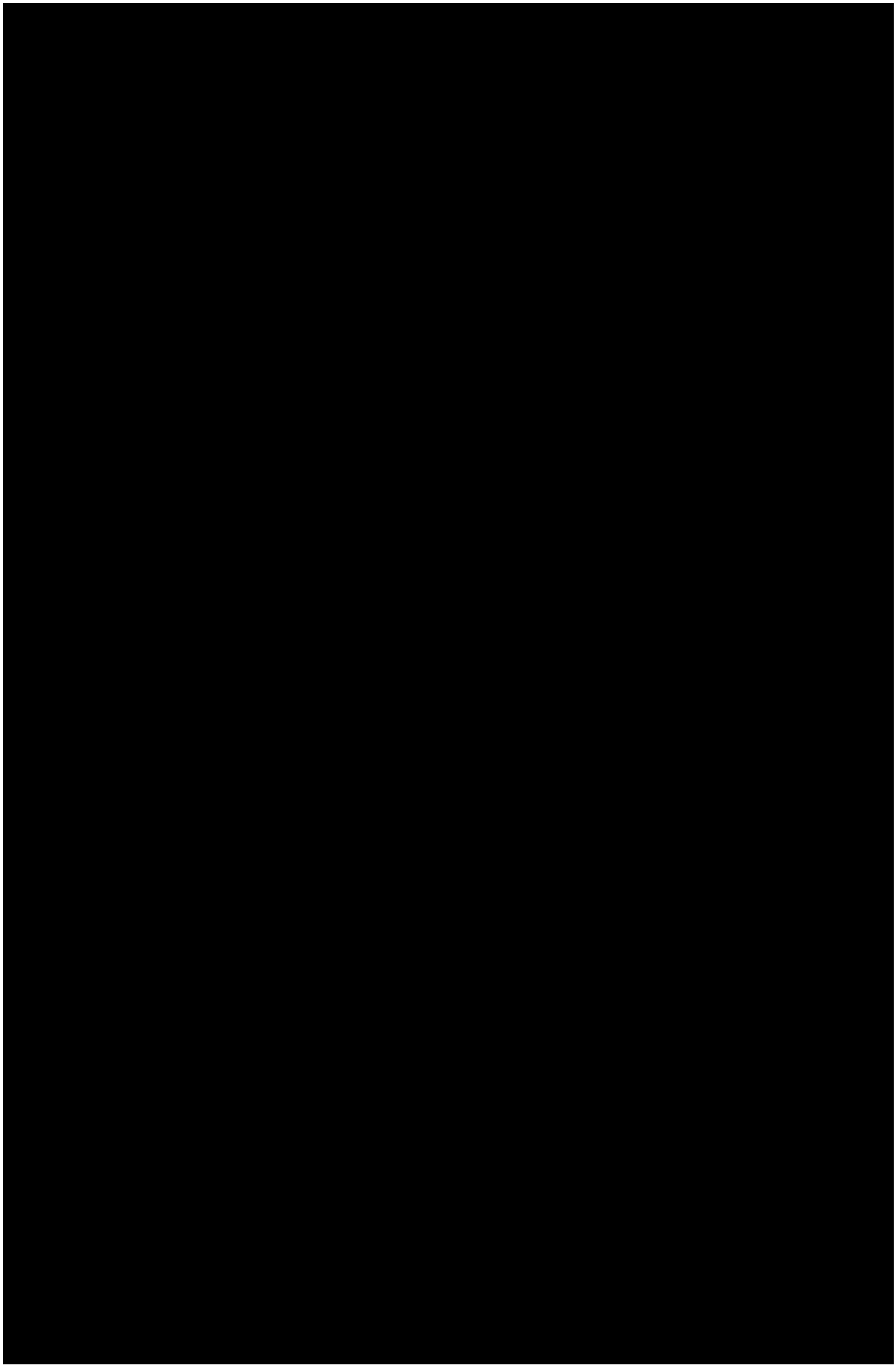


Figure 12: JAGS code used to generate random effects base case or sensitivity NMAs





[redacted]

Appendix 3.2: Data inputs, value sets, initial values, and seeds

Figure 13: Base case NMA: TNFi-naïve; induction

[redacted]

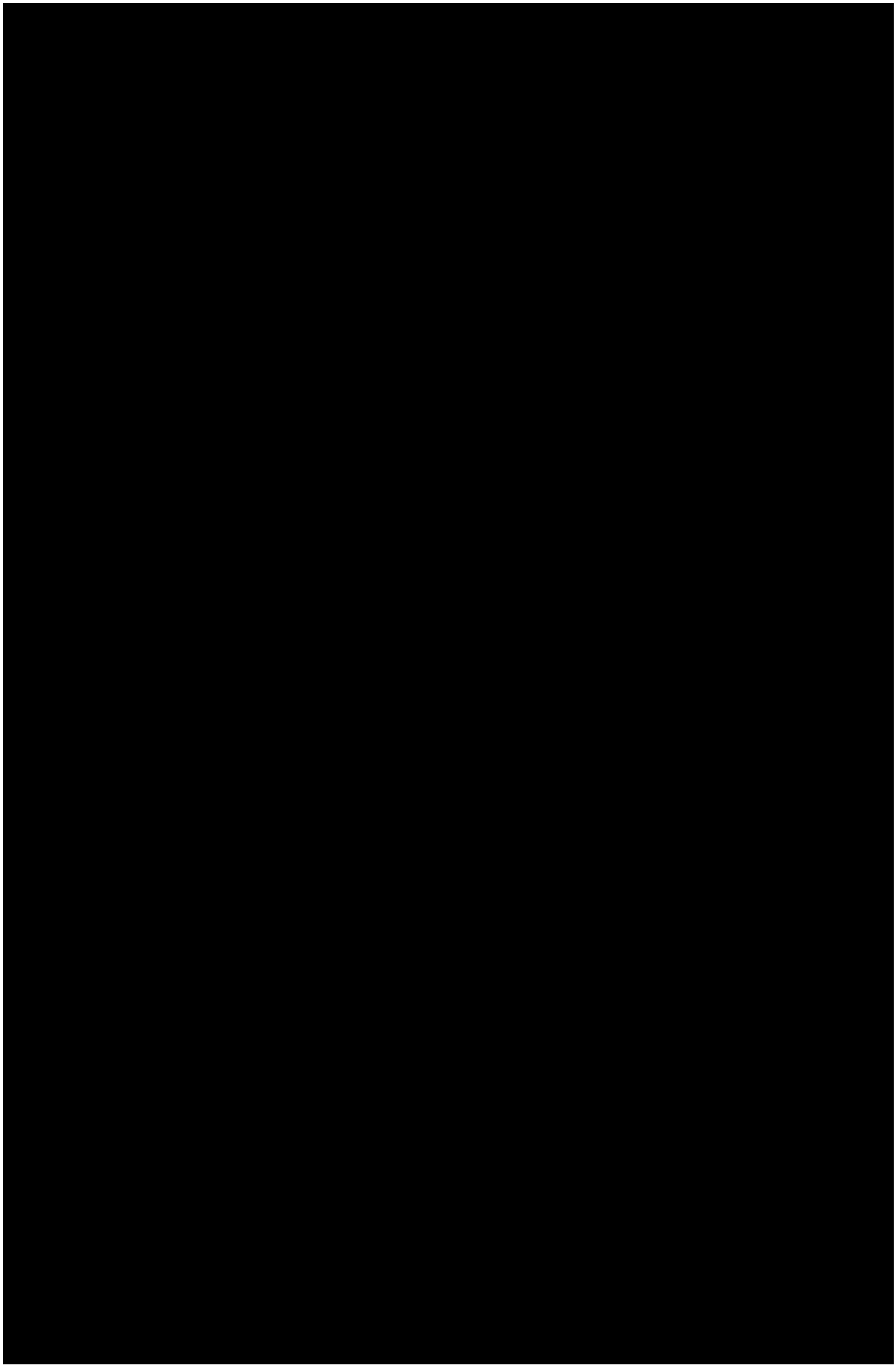
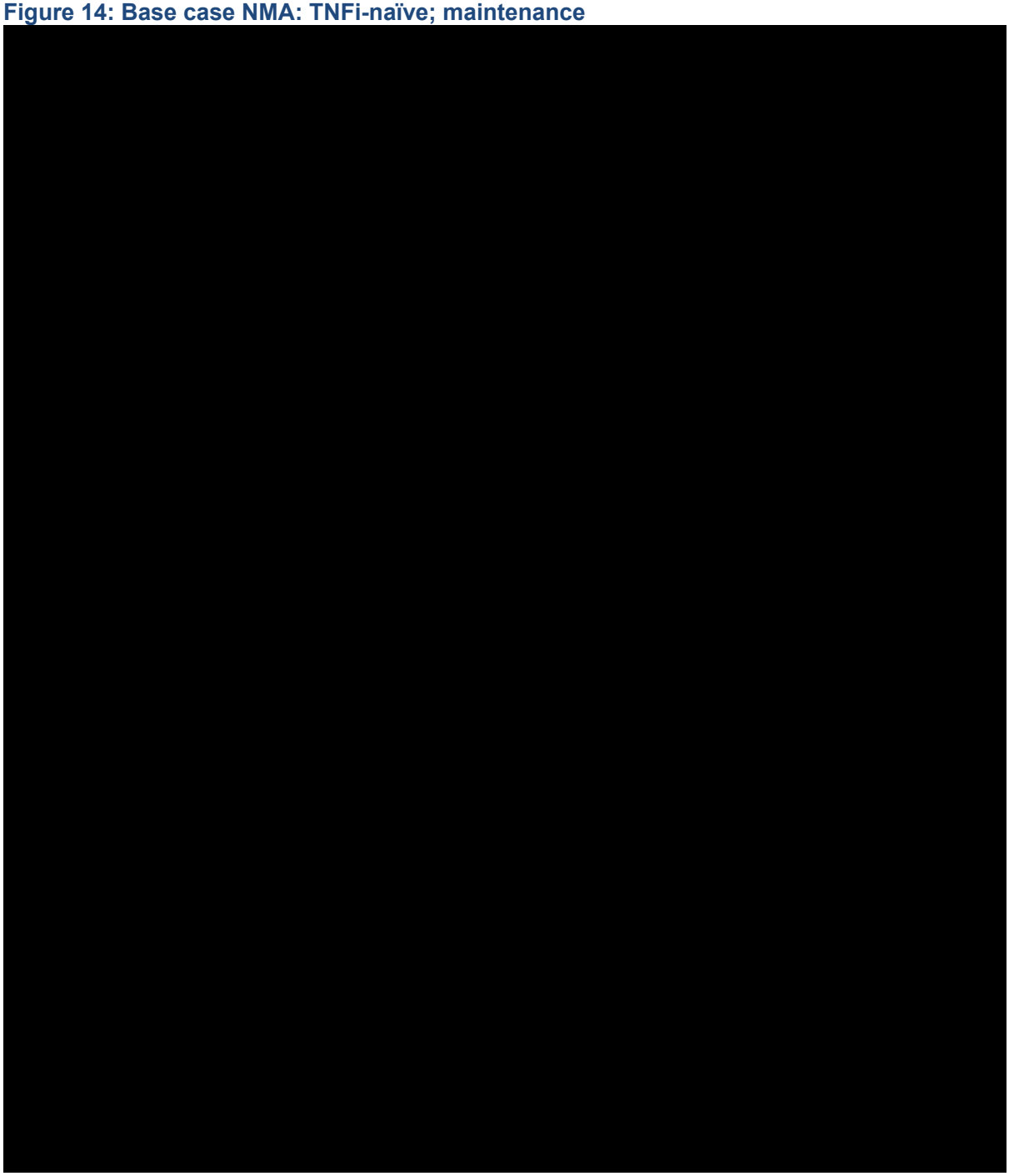
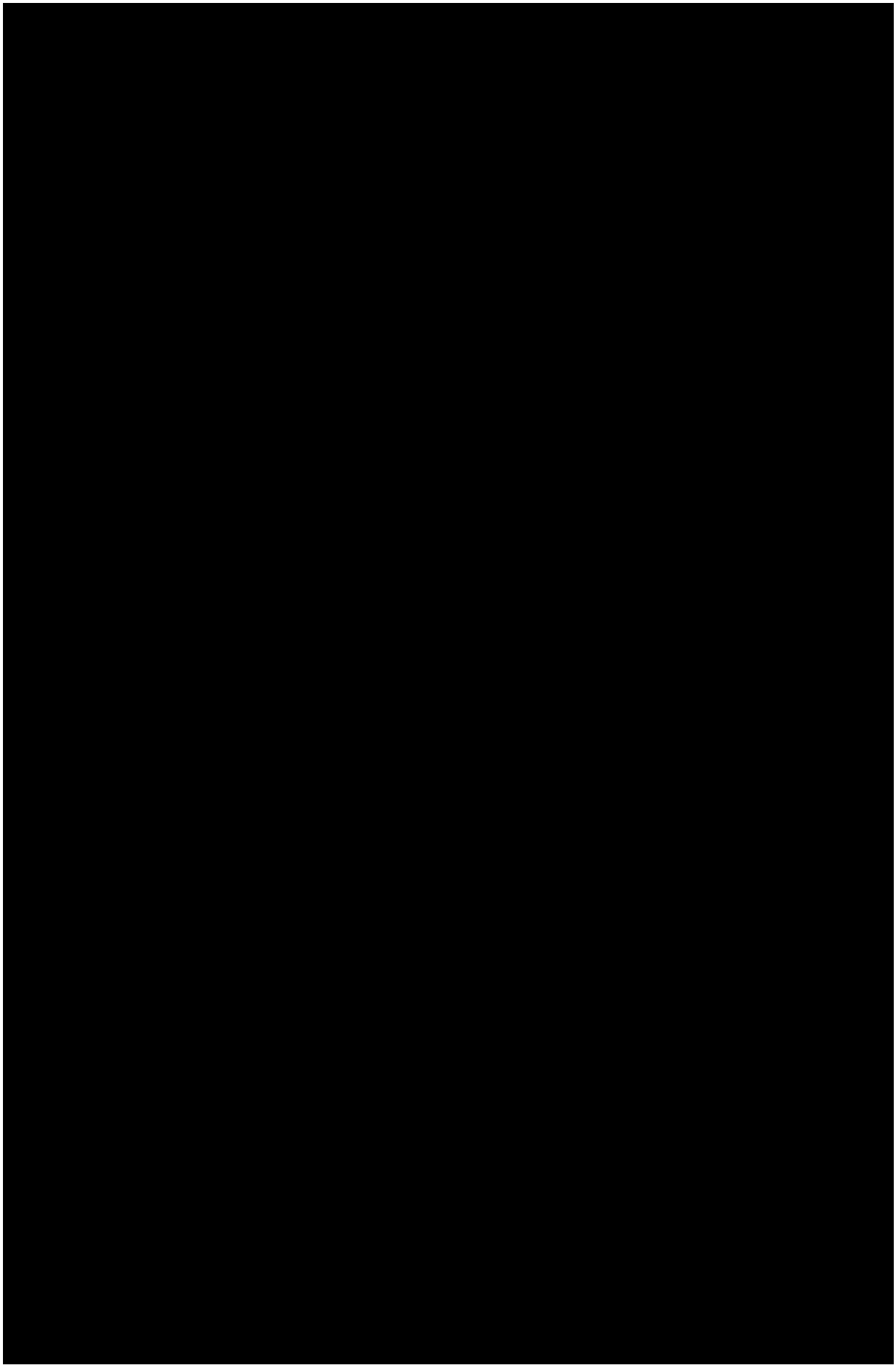
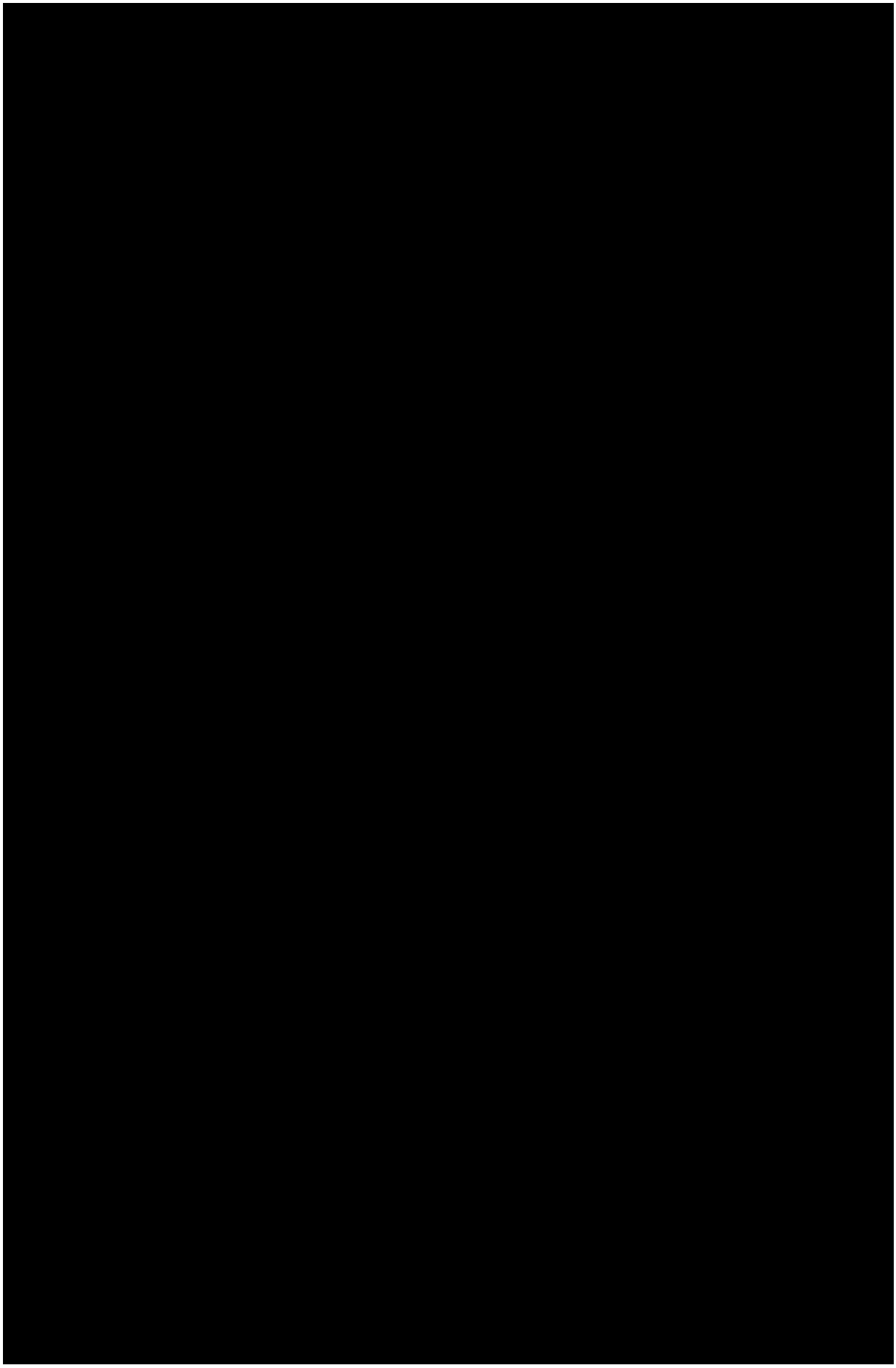


Figure 14: Base case NMA: TNFi-naïve; maintenance







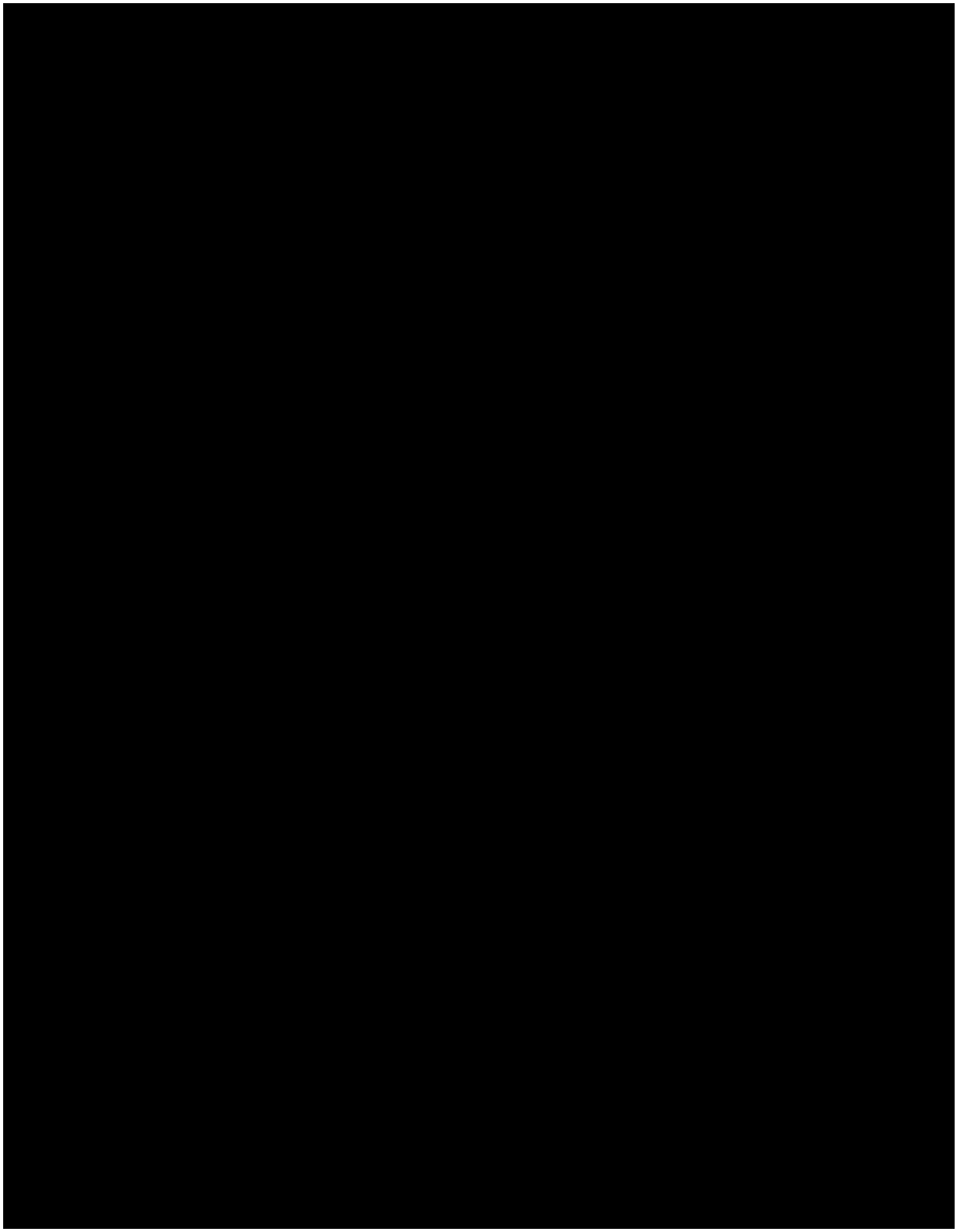
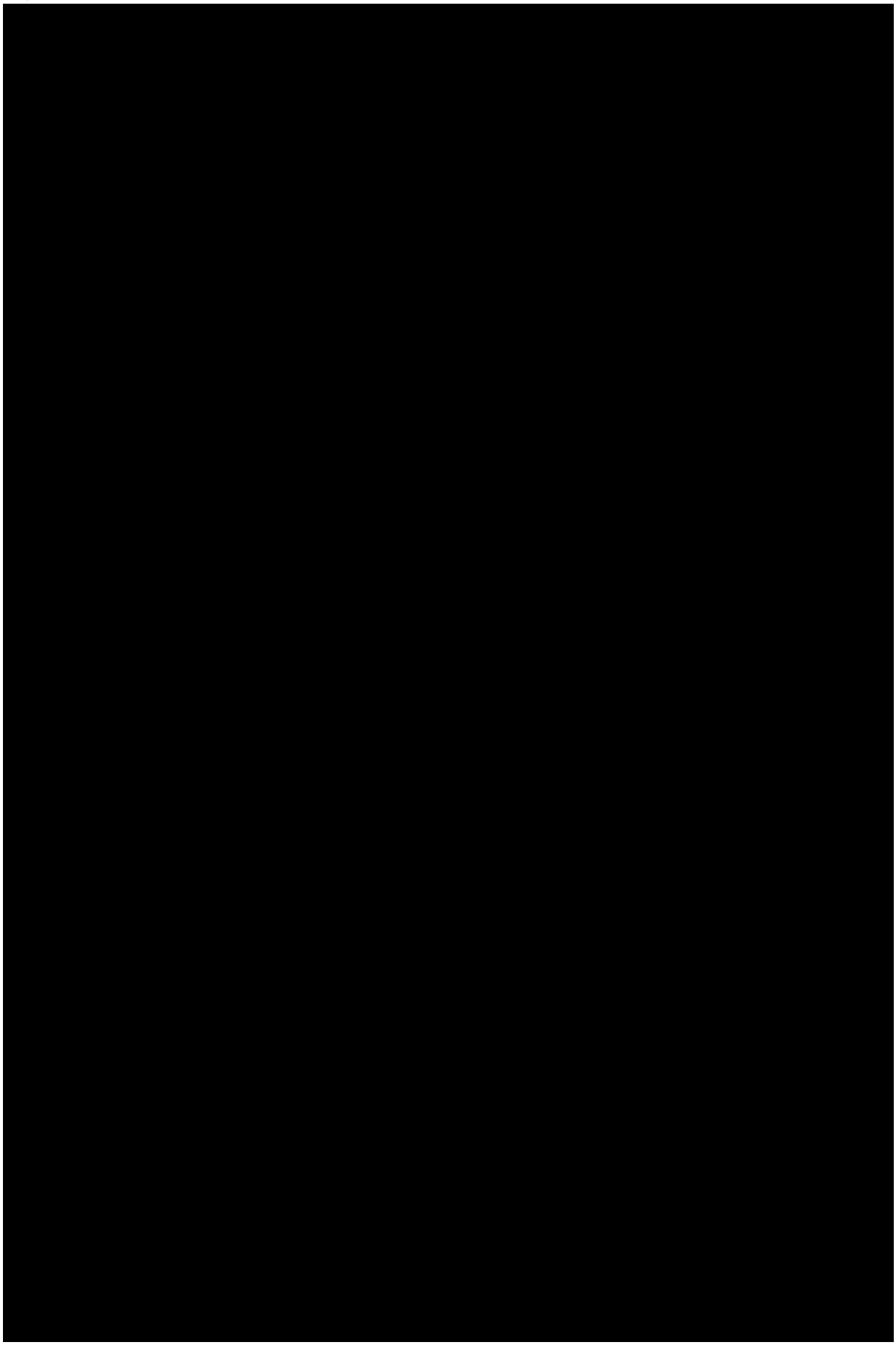
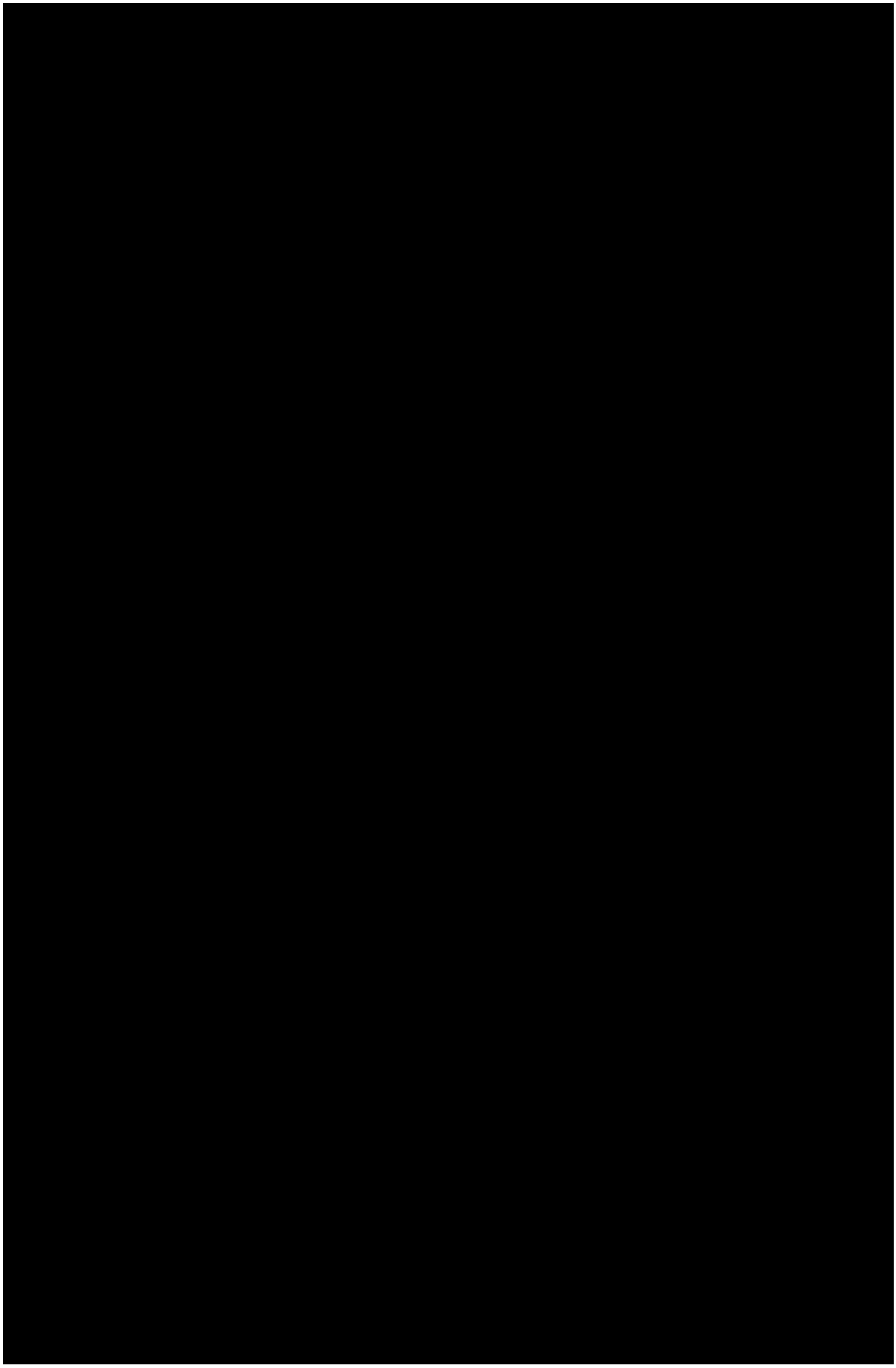


Figure 15: Base case NMA: TNFi-experienced; induction





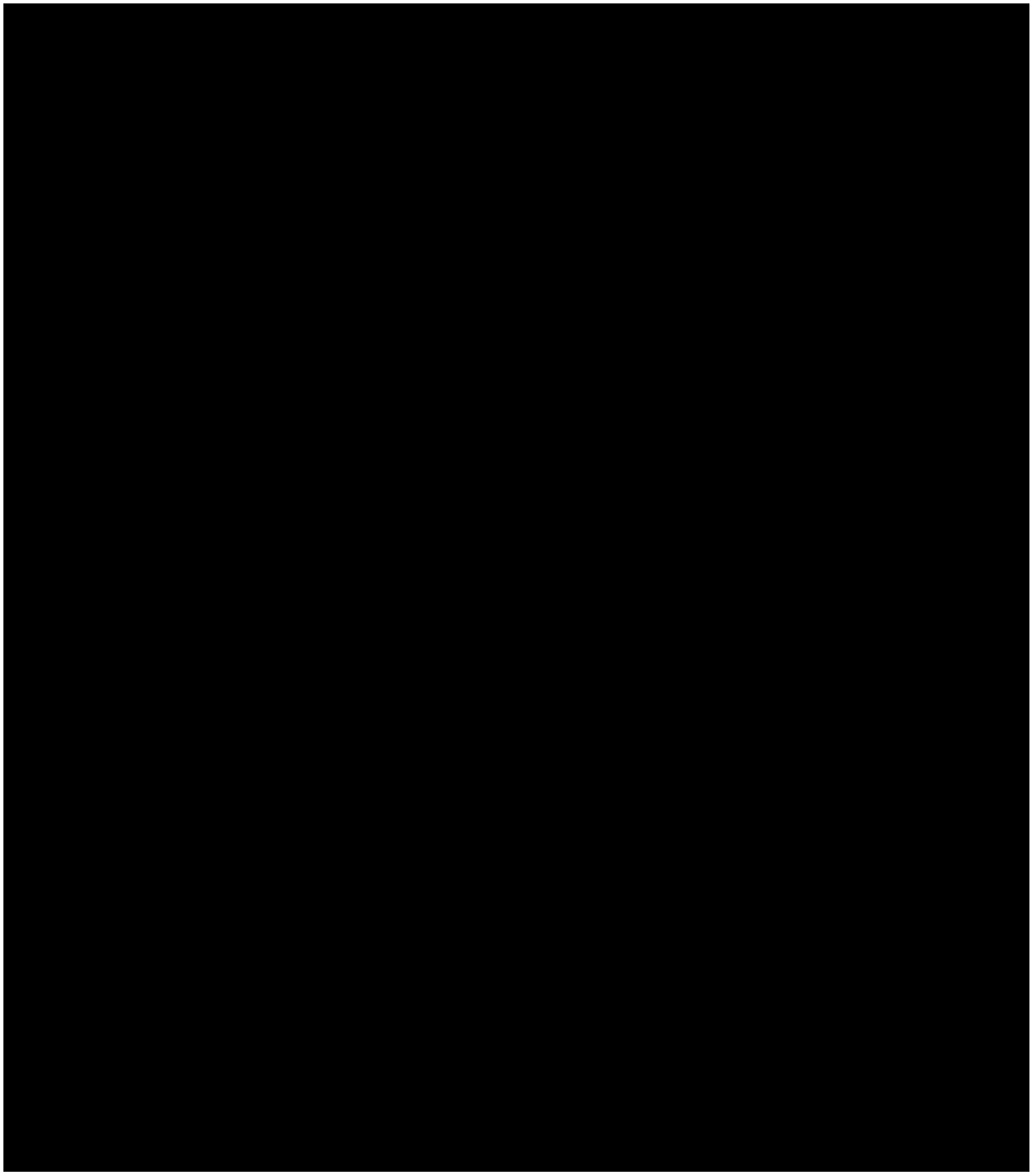
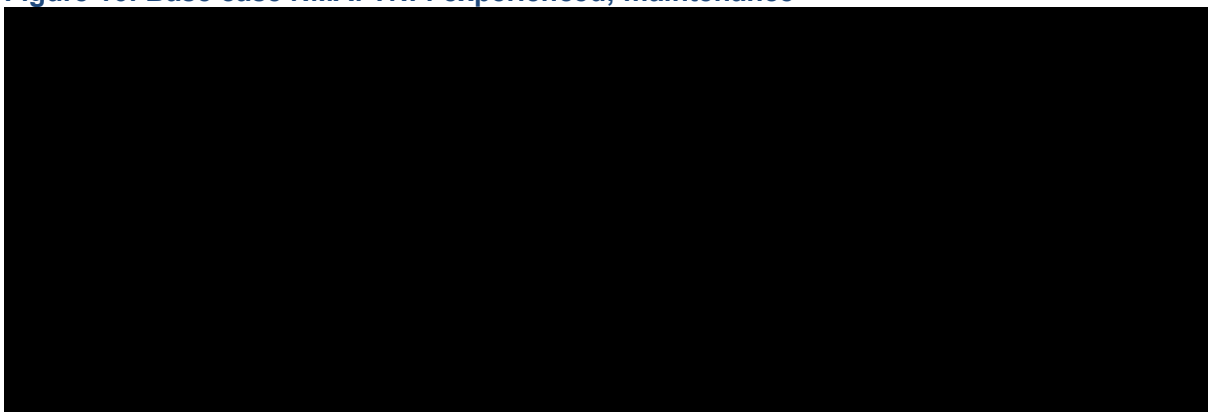
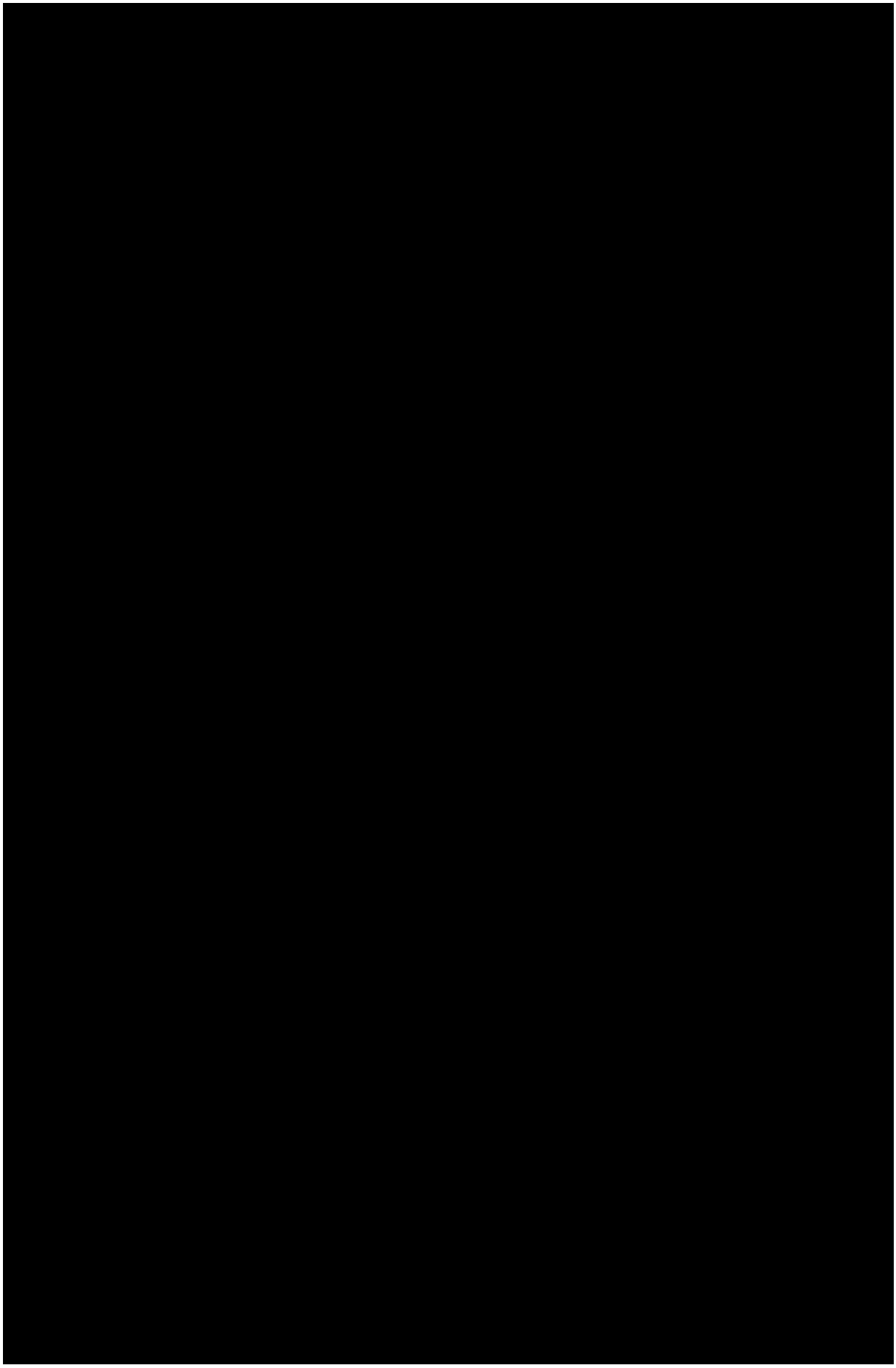
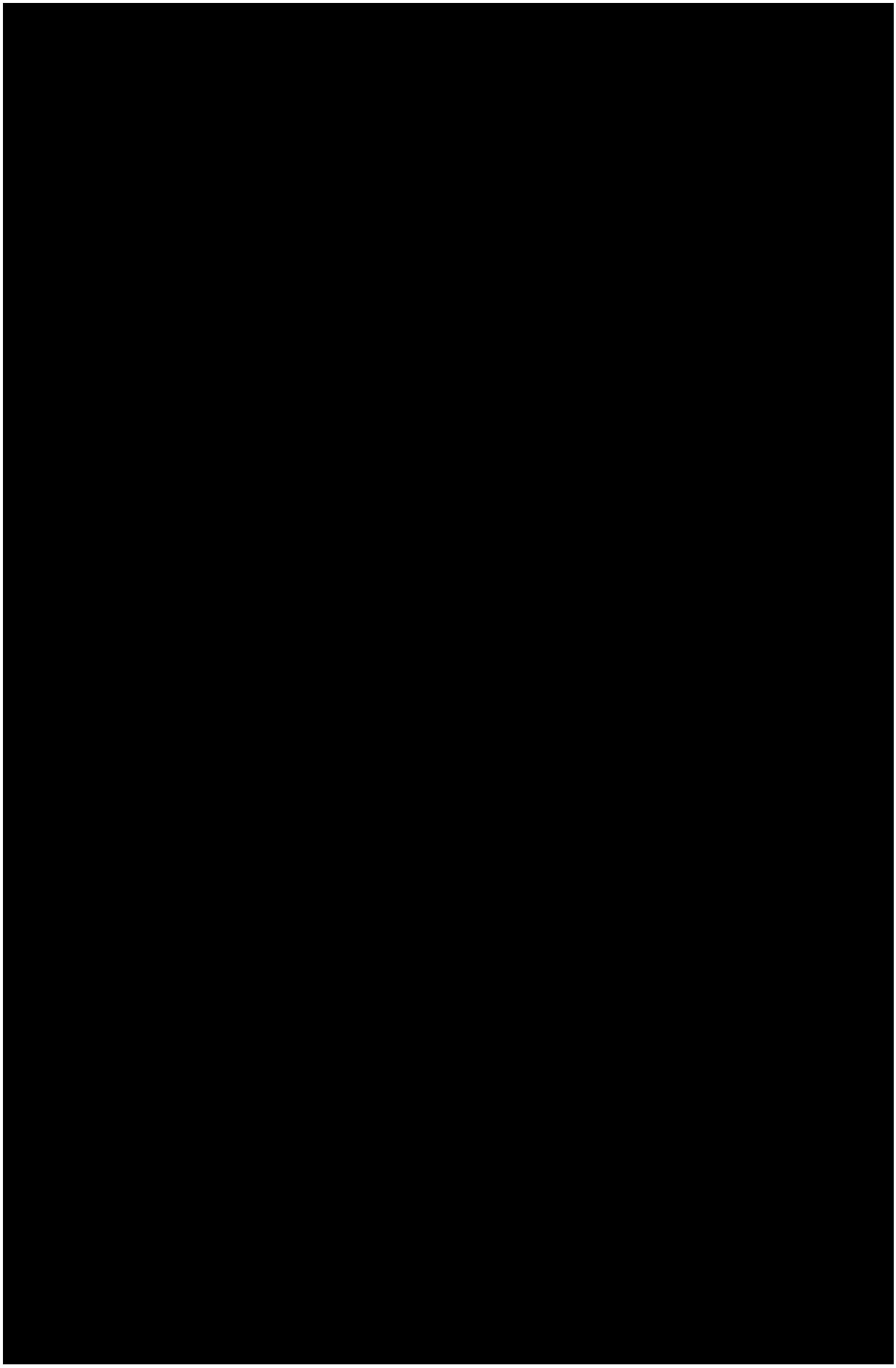


Figure 16: Base case NMA: TNFi-experienced; maintenance







Appendix 3.3: Summary results

Table 50: Summary results for key JAGS code parameters for TNFi-naïve induction base case NMA – posterior distribution means, medians, and 95% credible intervals

Parameter[a]	Mean	Q2.5	Q50 (Median)	Q97.5
PASI[1,1]				
PASI[2,1]				
PASI[3,1]				
PASI[1,2]				
PASI[2,2]				
PASI[3,2]				
PASI[1,3]				
PASI[2,3]				
PASI[3,3]				
PASI[1,4]				
PASI[2,4]				
PASI[3,4]				
PASI[1,5]				
PASI[2,5]				
PASI[3,5]				
PASI[1,6]				
PASI[2,6]				
PASI[3,6]				
PASI[1,7]				
PASI[2,7]				
PASI[3,7]				
PASI[1,8]				
PASI[2,8]				
PASI[3,8]				
T[1,1]				

T[2,1]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,2]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,2]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,3]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,3]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,4]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,4]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,5]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,5]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,6]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,6]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,7]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,8]	[redacted]	[redacted]	[redacted]	[redacted]

[a]Parameters named "PASI" are the absolute probability of no clinical response, no clinical remission (PASI[1,k]); clinical response, no clinical remission (PASI[2,k]); clinical remission (PASI[3,k]), for each treatment k. Parameters T[j,k] are the absolute probability of achieving at least clinical response (j = 1) or clinical remission (j = 2) on treatment k. Treatment k = 1 for placebo; k = 2 for ozanimod 1 mg QD; k = 3 for adalimumab 160/80/40 mg; k = 4 for golimumab 200/100 mg SC; k = 5 for infliximab pooled; k = 6 for tofacitinib 10 mg BID; k=7 for ustekinumab pooled; k=8 for vedolizumab 300 mg IV.
Abbreviations: BID: twice a day; IV: intravenous; QD: once-daily; Q2W: once every 2 weeks; SC: subcutaneous.

Table 51: Summary results for key JAGS code parameters for TNFi-experienced induction base case NMA – posterior distribution means, medians, and 95% credible intervals

Parameter[a]	Mean	Q2.5	Q50 (Median)	Q97.5
PASI[1,1]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,1]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,1]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,2]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,2]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,2]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,3]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,3]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,3]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,4]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,4]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,4]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,5]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,5]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,5]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,6]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,6]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,6]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,1]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,1]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,2]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,2]	[redacted]	[redacted]	[redacted]	[redacted]

T[1,3]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,3]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,4]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,4]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,5]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,5]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,6]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,6]	[redacted]	[redacted]	[redacted]	[redacted]

[a]Parameters named "PASI" are the absolute probability of no clinical response, no clinical remission (PASI[1,k]); clinical response, no clinical remission (PASI[2,k]); clinical remission (PASI[3,k]), for each treatment k. Parameters T[j,k] are the absolute probability of achieving at least clinical response (j = 1) or clinical remission (j = 2) on treatment k. Treatment k = 1 for placebo; k = 2 for ozanimod 1 mg; k = 3 for adalimumab 160/80/40 mg Q2W; k = 4 for tofacitinib 10 mg BID; k = 5 for ustekinumab pooled; k = 6 for vedolizumab 300 mg IV.
Abbreviations: BID: twice a day; IV: intravenous; Q2W: once every 2 weeks; SC: subcutaneous.

Table 52: Summary results for key JAGS code parameters for TNFi-naïve maintenance base case NMA – posterior distribution means, medians, and 95% credible intervals

Parameter[a]	Mean	Q2.5	Q50 (Median)	Q97.5
PASI[2,1]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,1]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,2]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,2]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,3]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,3]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,4]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,4]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,5]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,5]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,6]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,6]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,7]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,7]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,8]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,8]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,9]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,9]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,1]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,1]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,2]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,2]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,3]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,3]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,4]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,4]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,5]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,5]	[redacted]	[redacted]	[redacted]	[redacted]

T[1,6]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,6]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,7]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,7]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,8]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,8]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,9]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,9]	[redacted]	[redacted]	[redacted]	[redacted]

[a]Parameters named "PASI" are the absolute probability of no clinical response, no clinical remission (PASI[1,k]); clinical response, no clinical remission (PASI[2,k]); clinical remission (PASI[3,k]), for each treatment k. Parameters T[j,k] are the absolute probability of achieving at least clinical response (j = 1) or clinical remission (j = 2) on treatment k. Treatment k = 1 for placebo; k = 2 for ozanimod 1 mg QD; k = 3 for adalimumab 40 mg Q2W; k = 4 for golimumab pooled; k = 5 for infliximab pooled; k = 6 for tofacitinib pooled; k=7 for ustekinumab pooled; k= 8 for vedolizumab pooled; k=9 for vedolizumab 108 mg Q2W SC.
Abbreviations: QD: once daily; Q2W: once every 2 weeks; SC: subcutaneous.

Table 53: Summary results for key JAGS code parameters for TNFi-experienced maintenance base case NMA – posterior distribution means, medians, and 95% credible intervals

Parameter[a]	Mean	Q2.5	Q50 (Median)	Q97.5
PASI[1,1]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,1]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,1]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,2]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,2]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,2]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,3]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,3]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,3]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,4]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,4]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,4]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,5]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,5]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,5]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,6]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,6]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,6]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[1,7]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[2,7]	[redacted]	[redacted]	[redacted]	[redacted]
PASI[3,7]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,1]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,1]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,2]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,2]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,3]	[redacted]	[redacted]	[redacted]	[redacted]

T[2,3]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,4]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,4]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,5]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,5]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,6]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,6]	[redacted]	[redacted]	[redacted]	[redacted]
T[1,7]	[redacted]	[redacted]	[redacted]	[redacted]
T[2,7]	[redacted]	[redacted]	[redacted]	[redacted]

[a]Parameters named "PASI" are the absolute probability of no clinical response, no clinical remission (PASI[1,k]); clinical response, no clinical remission (PASI[2,k]); clinical remission (PASI[3,k]), for each treatment k. Parameters T[j,k] are the absolute probability of achieving at least clinical response (j = 1) or clinical remission (j = 2) on treatment k. Treatment k = 1 for placebo; k = 2 for ozanimod 1 mg; k = 3 for adalimumab 40 mg; k = 4 for tofacitinib pooled; k = 5 for ustekinumab pooled; k = 6 for vedolizumab pooled; k=7 for vedolizumab 108 mg Q2W SC.
Abbreviations: Q2W: once every 2 weeks; SC: subcutaneous.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Ozanimod for treating moderately to severely active ulcerative colitis [ID3841]

Patient Organisation Submission

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. [Please note that declarations of interests relevant to this topic are compulsory].

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	[redacted]

2. Name of organisation	Crohn's & Colitis UK
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Crohn's & Colitis UK is the UK's leading charity for everyone affected by Crohn's and Colitis. We're working to improve diagnosis and treatment, and to fund research into a cure; to raise awareness and to give people hope, comfort, and confidence to live freer, fuller lives. We want: • To drive world-class research that improves lives today and brings us closer to a world free from Crohn's and Colitis tomorrow • Everyone to understand Crohn's and Colitis • To support and empower everyone to manage their conditions • To drive high-quality and sustainable clinical care • Early and accurate diagnosis for all. Founded as a patients' association in 1979, we now have over 47,000 members across the UK. Our members include people living with the conditions, their families and friends, health professionals and others who support our work. We have 50 Local Networks which arrange educational meetings, generate publicity and organise fundraising. 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 Crohn's & Colitis UK's annual reports and accounts (crohnsandColitis.org.uk)
4b. Has the organisation received any funding from the manufacturer(s) of the	No

technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] If so, please state the name of manufacturer, amount, and purpose of funding.	
4c. 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?	We gather information about the experience of patients, carers and families through: • the Crohn's & Colitis UK helpline • local networks • calls for evidence via our website and social media • one to one discussion with people with IBD, clinicians, and the wider IBD community; and • research - our own and that of external organisations.

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. 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.[1] [2] *"Life with UC has been difficult, as I was constantly ill over a period of years, I had my relationship break down. I have been lucky that my previous line manager at work had a daughter of his own who suffered from UC, so any hospital stays weren't a problem and he allowed me to work from home on particularly bad days."* **Quote from a person living with Ulcerative Colitis.** Given that disease severity is wide-ranging, and while each person has their own individual experience, we would like to take this opportunity to describe the impact and experience of the specific cohort of patients with moderate to severe Ulcerative Colitis that this guidance is targeting. This cohort is likely to comprise of patients with Ulcerative Colitis who experience more severe flares, weight loss, fever and constitutional symptoms, and whose disease has not responded to or are unable to tolerate other treatments, and/or can benefit from this treatment in particular. Truelove and Witts define severe Ulcerative Colitis as six or more stools a day plus at least one of the features of systemic upset (marked with an *): visible blood; pyrexia*; pulse rate greater than 90 BPM*; erythrocyte sedimentation rate (mm/hour) * and anaemia.[3] The Mayo Score defines severe Colitis as more than five stools a day, blood passed without stool, obvious blood with stools in most cases and severe disease (spontaneous bleeding, ulceration).[4]

[1] Crohn's and Colitis UK (2018) Quality of Life Survey https://ibduk.org/ibd-standards
[2] IBD UK (2019) IBD Standards
[3] NICE (2019) NICE Guideline on Ulcerative Colitis: Management (NG130) https://www.nice.org.uk/guidance/ng130/chapter/Recommendations
[4] Dignass, A,. Second European evidence-based consensus on the diagnosis and management of Ulcerative Colitis Part 1: Definitions and diagnosis. Journal of Crohn's and Colitis Vol 6. Issue 10 https://www.sciencedirect.com/science/article/pii/S1873994612004047#t0020

	For this subgroup of patients with moderate to severe Ulcerative Colitis, the condition is more than challenging, but frequently overwhelming and detrimentally life-altering, as described below: *"I had 3 blood transfusions, multiple steroids, sleepless drained nights, cannula paracetamol, Iron deficiency, stomach ulcers and multiple drugs and many blood tests, not being able to eat and losing a huge amount of weight over 2 and a half stone in just 2 weeks wasn't expected out the blue in my life."* **Quote from a person living with Ulcerative Colitis.** **Mortality** There are risks and mortality associated with untreated and uncontrolled disease. NICE Guideline on Ulcerative Colitis states: 'Ulcerative Colitis is a lifelong disease that is associated with significant morbidity. It can also affect a person's social and psychological wellbeing, particularly if poorly controlled'.[5] This is echoed by BSG Guidelines that state that 'acute severe Colitis is a potentially life-threatening condition'.[6] Acute severe Colitis has a 1% mortality risk and a 29% chance of requiring emergency surgery to remove the inflamed bowel (colectomy).[7] Between 15-25% of patients with Ulcerative Colitis will need to be hospitalised due to an acute severe flare-up at some stage. Often this will be the first presentation of their disease.[8]

[5] NICE (2019) Guideline on Ulcerative Colitis: Management: Overview | Ulcerative Colitis: management | Guidance | NICE
[6] The British Society of Gastroenterology (2011) British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. https://gut.bmj.com/content/60/5/571.long
[7] Ibid
[8] Ibid

When a flare occurs in acute severe Colitis, deterioration can occur rapidly. Patients will require close monitoring and review by appropriate specialists. It's also vitally important to make decisions quickly to avoid severe complications.

The very real risks associated with acute severe Colitis include:

- Life-threatening haemorrhage
- Toxic megacolon - can occur in up to 1 in 40 people with Colitis[9]
- Perforation of the bowel[10]

Additional complications of chronic, uncontrolled, active Ulcerative Colitis also include:

- Osteoporosis and vitamin D deficiency. The major risk factors for osteoporosis complicating IBD are age, steroid use and disease activity[11]
- Anaemia[12].
- Increased risk of cancer[13]

Impact on emotional and mental health

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.[14] Stigma and lack of wider understanding of the condition exacerbate the impact.

Anxiety and depression are higher in people with IBD, with mood disorders at least in part a consequence of the IBD itself and its medical treatment (e.g., corticosteroid therapy), surgery, including specifically

[9] Parray, F. Q. et al. (2012). Ulcerative Colitis: a challenge to surgeons. Int. J. Prev. Med. 3, 749–63.
[10] IBDUK (2019) IBD Standards 2019: Homepage | IBD UK
[11] Mowat C, Cole A, Windsor A et al. (2011) Guidelines for the management of inflammatory bowel disease in adults. Gut, 60, 571-607.
[12] Crohn's and Colitis Foundation.(2020) Anaemia. https://www.crohnscolitisfoundation.org/sites/default/files/2020-03/anemia.pdf
[13] The British Society of Gastroenterology (2019) British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. https://www.bsg.org.uk/resource/bsg-consensus-guidelines-ibd-in-adults.html
[14] Cosnes J, et al., (2011). Epidemiology and natural history of inflammatory bowel diseases. Gastroenterology, 140 (6), 1785-94.

	colectomy and stoma formation.[15] Additionally, most reports indicate that stress may be involved in triggering relapse. "*The last 9 months have been really quite horrible for me dealing with my UC and I went through a really low point in my life, feeling very anxious and depressed. I took 5 months off work and only recently started a new job. My UC really affected my social life and confidence especially with getting out of the house and carrying out simple tasks."* **Quote from a person living with Ulcerative Colitis.** "*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."* **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*." **Quote from the parent of a person living with Ulcerative Colitis.** **Social functioning**

[15] Graff L. A. et al., (2009). Depression and anxiety in inflammatory bowel disease: a review of comorbidity and management. Inflamm Bowel Dis, 15 (7), 1105-18.

	Social functioning can be impaired - leading to an inability to work, attend school, participate in leisure activities, or have intimate relationships. *"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.** Research shows that young people aged 16-25 with IBD who have not yet entered full-time employment often feel that their condition has compromised their education and significantly limited their career aspirations. There is a clear associated "productivity loss" by health state, whereby the lowest score for health state (Visual Analogue Score 0-2.5) corresponds with a 71% productivity loss.[16]
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	There is unmet need amongst people with moderate to severe Ulcerative Colitis. Patients express dissatisfaction with many of the current treatment options. Many experience lack of response (primary or secondary) and/or adverse reactions. The effects of steroids are extremely unpleasant and long-term safety profile of other treatments, including biologics, are of some concern. "*When I am unwell, I struggle with extreme tiredness and extended periods in the bathroom which makes my working life very difficult. I work in construction so spend a lot of time away from toilets. Vedolizumab, when I first started, it was my wonder drug. It was difficult spending so much time in hospital but worth it to*

[16] Gay M et al. (2011) Crohn's, Colitis and Employment – from Career Aspirations to Reality. Crohn's and Colitis UK.

be completely symptom free. I was in remission for nearly 4 months.

I was then given Golimumab which was a lot more convenient, and I liked having the control of self-administering. This however never gave me remission and my CRP worsened over the period I was taking it. I am now being offered Tofacitinib but have been told this is my final option." **Quote from a person living with Ulcerative Colitis.**

"*I have suffered with UC for 13 years. It's always been moderate to severe. I have tried all drugs including all biologics. All failed after a while. The best was Infliximab, I had my first ever remission for 2 years. However, it came to an end in Aug 2017. I had 18 months of pain and blood, countless hospital admissions, yet I was still pushed to try yet another biologic, Vedolizumab then Golimumab. None of it worked. 6 weeks later I had an emergency op and my colon was removed. My recovery is slow as I was ill for quite some time before and I'm building up my stamina now."* **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.**

Steroids

"Corticosteroids have no proven efficacy in maintaining remission in IBD and should not be used for this purpose."[17] The BSG guidelines set out clear stipulations on the best practice of prescribing steroid therapies given their diminishing returns, harsh side effects and risk of dependency.[18]

Surgery

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.

	"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.** 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 by any pelvic surgery. *"I had severe Pan Ulcerative Colitis. I started my journey with an emergency admission in a very poor state (…). I spent 2 weeks in hospital while they tried to stop the frequency and bleeding, I came out on steroids, cyclosporine and Asacol. I was better for a little while but soon became very ill again and was off work. I was put on azathioprine but could not tolerate this, so I was switched to mercaptopurine. This put me in remission for 3 years, when this no longer worked I was put on Simponi. The initial double dose showed some promising results, but the single dose didn't keep me in remission. Following this I became dependent on steroids.* *My life was terrible quality. I missed out on opportunities at work, very rarely went anywhere and people would comment on my features from the steroids, and they said I looked a strange green-yellow colour.* *Finally, I had enough of being ill and hospital admissions and blood transfusions and requested surgery to remove my colon. My consultant told me if I was in any other country, they'd have taken it out much sooner. The surgeon said it disintegrated as he was taking it out it was in such a bad state. I now have a j-pouch and while life is a lot better it isn't the cure that was promised and it impacts on my life considerably."* **Quote from a person living with Ulcerative Colitis.**

[17] Barrett, K. (2018) Using corticosteroids appropriately in inflammatory bowel disease: a guide for primary care, British Journal of General Practice. 68 (675): 497-498. https://bjgp.org/content/68/675/497

[18] BSG (2019) British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. https://www.bsg.org.uk/resource/bsg-consensus-guidelines-ibd-in-adults.html

	Surgery has significant associated long- and short-term risks which include: - general anaesthetic complications - infections - adhesions - pouchitis - pouch leakage - abscesses - fistulae - small bowel obstruction - post-operative bleeding - sexual dysfunction - delayed wound healing - nerve damage.[19,20] Additionally, a meta-analysis has shown 'an approximate threefold increase (from 15% to 48%) in the risk of infertility in women with Ulcerative Colitis as a result of ileal pouch anal anastomosis (IPAA).[21] Johnson et al. reported the infertility rate in females who had pelvic pouch surgery was significantly higher compared to females who were managed medically (38.1 % compared with 13.3 %; $p < 0.001$).[22] We would also urge the Committee to consider the persistent quality of life issues that impact multiple domains, including psychological and sexual functioning. A 2015 study found 81% experienced problems in at least one of the following areas: depression, work productivity, restrictions in diet, body image, and sexual function. In the same study, amongst moderate to severe Ulcerative Colitis patients, post-colectomy, 27% of men and 28% of women reported that their sexual life was worse now than before surgery.[23]

[19] Ibid
[20] Brown, C. et al., (2015). Long-term outcomes of colectomy surgery among patients with Ulcerative Colitis. *Springerplus,* 4, 573.
[21] Waljee A, et al., (2006). Threefold increased risk of infertility: a meta-analysis of infertility after ileal pouch anal anastomosis in Ulcerative Colitis. Gut, 55 (11), 1575–1580.
[22] Johnson P, Richard C, Ravid A, Spencer L, Pinto E, Hanna M, Cohen Z, McLeod R. Female infertility after ileal pouch-anal anastomosis for Ulcerative Colitis. Dis Colon Rectum. 2004 Jul;47(7):1119-26. doi: 10.1007/s10350-004-0570-7.
[23] Brown, C. et al., (2015). Long-term outcomes of colectomy surgery among patients with Ulcerative Colitis. *Springerplus,* 4, 573.

8. Is there an unmet need for patients with this condition?	The range of options available for treating Ulcerative Colitis remain far from optimal for patients, a substantial number of whom experience lack of response (primary or secondary) and/or adverse reactions to biologic as well as conventional therapies. There are significant short and long-term side effects with corticosteroids, including opportunistic infections, steroid-induced psychosis, steroid dependence, diabetes and osteoporosis. Their use is also limited to induction of remission. *"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.** Up to one third of patients with IBD are intolerant to thiopurines and a further 10% are unresponsive to them. In the majority of patients who do respond, the benefits take three to six months to appear. Significant risks of thiopurines including non-Hodgkin's lymphoma (as high as 4-5-fold compared with unexposed IBD patients and further increased when used in combination with anti-TNFs). Other side effects include early hypersensitivity reactions such as fever and pancreatitis, bone marrow suppression and hepatotoxicity requiring frequent lab monitoring during treatment. Anti-TNFs are increasingly being used earlier in the treatment pathway and can have a significant and positive effect on quality of life for patients. However, up to 40% of patients treated with anti-TNF therapy do not respond to induction therapy. In the approximately one-third of patients who do achieve remission with anti-TNF therapy, between 10%-50% lose response over time.[24] Overall, there is a pressing need for additional treatment options which offer a different mode of action and the potential for people with Ulcerative Colitis to resume their lives and restore their quality of life.

[24] Roda, G. (2016). Loss of Response to Anti-TNFs: Definition, Epidemiology, and Management. *Clin Transl Gastroenterol*, 7 (1), e135.

Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	One of the key advantages is that Ozanimod is an oral therapy and would give patients a treatment option to be taken at home, which will allow people to be treated at home. Furthermore 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 would result in an associated reduction in NHS costs due to reduced infusions. Patients most likely to benefit from this drug are those for whom currently available therapies are ineffective, contraindicative or they develop an intolerance. In this group, it is likely that individuals, without further choice, will return to treatment/s which have already been established to be inadequate. This may include highly undesirable long-term steroid use or unproven unconventional therapy. It is also likely that patients in this group who exhaust all other treatment options would be forced to have a colectomy, either elective or as an emergency. *"I am well aware that these drugs have a very significant cost but without them, the last 12 years would have been very different for me. Even with them I have had to have 2 lots of surgery to remove scarred bowel but without them I think I would have had to have more extensive surgery and possibly not even be here to send this email. I am also well aware that I am on my last chance here with current available drugs having taken everything the NHS has to offer; if the vedo stops working then I have nowhere else to go with medication. New drugs and options for medication will be vital for my health going forward."* **Person with IBD, in which drug treatments have not been effective.**
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Prescription costs faced people living with long-term and chronic conditions, including Ulcerative Colitis, in England, are shown to contribute to economic disadvantage, which can impact adherence and lead to

	complications and increased cancer risks and cost to the NHS.[25] However, the disadvantage is not specific to Ozanimod, and the value of an additional treatment option will remain beneficial as it will reduce the risk of loss of response.
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. This would include young people wishing to complete studies and those for whom surgery would be considered unacceptable due to cultural or religious factors.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	For certain religious groups, the impact of active disease and the effects of surgery may interfere with religious practices and cause distress, which could be alleviated by an additional medical therapeutic option. Although not specific to Ozanimod, prescription costs 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.
- There is significant unmet need within the moderate to severe cohort. Current treatments remain far from optimal for patients, a substantial number of whom experience a lack of response (primary or secondary) and/or adverse reactions to medical treatments and may face the prospect of surgery with considerable anxiety.
- Ozanimod offers a novel and effective treatment option and increases choice for both clinicians and patients (in the context of shared decision making).
- Ozanimod may delay or prevent surgery in UC patients. This is particularly important for patients who have exhausted all over treatment options and wish to avoid or delay surgery (e.g. to complete studies.

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

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

NHS organisation submission (CCG and NHS England)

Ozanimod for treating moderately to severely active ulcerative colitis [ID3841]

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 10 pages.

About you	
1. Your name	Drs. Jimmy K. Limdi 1 , Rachel Cooney 2, Sreedhar Subramanian 3 on behalf of:
2. Name of organisation	British Society of Gastroenterology (Inflammatory Bowel Disease Committee)

3. Job title or position	1. Consultant Gastroenterologist & Head- Section of IBD The Northern Care Alliance NHS Foundation Trust, Manchester Hon. Professor of Gastroenterology, Manchester Academic Health Sciences, University of Manchester, Manchester UK 2. Consultant Gastroenterologist and IBD lead, Univeristy Hospitals Birmingham, Honorary Senior Clinical Lecturer, University of Birmingham 3. Consultant Gastroenterologist, Cambridge University Hospitals NHS Trust
4. Are you (please tick all that apply):	☐ commissioning services for a CCG or NHS England in general? ☐ commissioning services for a CCG or NHS England for the condition for which NICE is considering this technology? ☐ responsible for quality of service delivery in a CCG (for example, medical director, public health director, director of nursing)? ✓ an expert in treating the condition for which NICE is considering this technology? ☐ an expert in the clinical evidence base supporting the technology (for example, an investigator in clinical trials for the technology)? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	The British Society of Gastroenterology is a British professional organisation of gastroenterologists, surgeons, pathologists, radiologists, scientists, nurses, dietitians and others amongst its members, which number over 3,000. It was founded in 1937 and is a registered charity (Number:1149074). Its main activities include:

	1. Education and Training (annual scientific meeting, postgraduate training, clinical update meetings) 2. Supporting research into gut and liver disease (supporting academic development, promoting Gut) 3. Enhancing service standards (clinical service development, guidelines, sharing of best practice) 4. Supporting the gastrointestinal community 5. Promoting awareness of gastroenterology
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
Current treatment of the condition in the NHS	
6. Are any clinical guidelines used in the treatment of the condition, and if so, which?	Yes, there are national and international guidelines available. In the UK, we follow: 1. The British Society of Gastroenterology IBD guidelines mainly- citation as below: Lamb CA et al. British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. Gut. 2019 Dec;68(Suppl 3):s1-s106. doi: 10.1136/gutjnl-2019-318484. 2. NICE guidance for the management of ulcerative colitis(2019) https://www.nice.org.uk/guidance/ng130 accessed 24 December 2021 3. European Crohns and Colitis Organisation: most recent guidelines citation as below:

	Raine T et al. ECCO Guidelines on Therapeutics in Ulcerative Colitis: Medical Treatment. J Crohns Colitis. 2021 Oct 12
7. 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.)	The BSG guidelines recommend that "ulcerative colitis patients on maintenance therapy with high-dose mesalazine, who required two or more courses of corticosteroids in the past year, or who become corticosteroid-dependent or refractory, require treatment escalation with thiopurine, anti-TNF therapy), vedolizumab or tofacitinib. The choice of drug should be determined by clinical factors, patient choice, cost, likely adherence and local infusion capacity."(1) Since publication of the BSG guidance, Ustekinumab (a biological therapy targeting the p40 subunit, common to cytokines IL12 &IL-23 has been licensed and also approved by NICE (2): "as an option for treating moderately to severely active ulcerative colitis in adults when conventional therapy or a biological agent cannot be tolerated, or the disease has responded inadequately or lost response to treatment, only if a tumour necrosis factor-alpha inhibitor has failed (that is the disease has responded inadequately or has lost response to treatment) or a tumour necrosis factor-alpha inhibitor cannot be tolerated or is not suitable, and the company provides ustekinumab at the same price or lower than that agreed with the Commercials Medicines Unit". Despite well-defined pathways and updated guidelines, existing therapies have significant drawbacks, highlighting a major unmet therapeutic need for people living with moderate to severe ulcerative colitis. Specifically, primary failure of anti-TNF induction therapy occurs in 19–58% of patients in clinical trials (3-5). Among patients responsive to anti-TNF therapies, discontinuation due to secondary loss of response occurs in 17 to 22% of patients and approximately 40% required dose escalation to maintain treatment efficiency(3-5)(6). Treatment failure is even higher among patients undergoing second line TNF inhibitor therapy. In a meta-analysis the proportion of patients have discontinued treatment due to loss of response was 68-77% at 12 months and 82 -90% by the end of year 2(6). Diminishing efficacy stems in part from immunogenicity and the formation of antibodies against biologics (7). Evolution in our knowledge of disease pathophysiology and immune mechanisms led to development of other biologics, blocking the integrin α4β7 on leukocyte cells blocking lymphocyte trafficking to intestinal mucosa and IL12/23 inhibition with ustekinumab and also the small molecule (non-biological) Janus kinase inhibitor, Tofacitinib, but again, up to 50%

	of patients either do not respond or will have loss of response over time(8-10). In clinical practice, the high rate of non-response or incomplete response to ulcerative colitis medication indicates a need for newer therapeutic strategies. S1P modulators are among the most transformative new agents for the treatment of UC, with Ozanimod (“first in class”) , selectively targeting S1P1 and S1P5 demonstrating safety and efficacy in adults with moderate to severely active UC in the phase 3 True North study(11). Briefly, in this study, investigators randomly assigned adults (18-75 years) with moderate to severe UC to receive oral Ozanimod 1mg daily (n=429) after a 1-week dose escalation period or placebo (n=216) for a 10-week induction period. Patients were excluded from the trial if they had not had a response to induction therapy with at least two biologic agents approved for the treatment of ulcerative colitis, had a clinically relevant cardiac condition, or had a history of uveitis or macular edema. The incidence of clinical remission was significantly higher among patients who received ozanimod than among those who received placebo during both induction (18.4% vs. 6.0%, P<0.001) and maintenance (37.0% vs. 18.5% [among patients with a response at week 10], P<0.001). The incidence of clinical response was also significantly higher with ozanimod than with placebo during induction (47.8% vs. 25.9%, P<0.001) and maintenance (60.0% vs. 41.0%, P<0.001). All other key secondary end points were significantly improved with ozanimod as compared with placebo in both periods. The incidence of infection (of any severity) with ozanimod was similar to that with placebo during induction and higher than that with placebo during maintenance. Serious infection occurred in less than 2% of the patients in each group during the 52-week trial. On the basis of these safety and efficacy results, ozanimod received US Food and Drug Administration approval May 2021 (12). In addition to these data, and a novel mechanistic approach, ozanimod, as an oral small molecule offers added advantages in the form of no immunogenicity and a short half-life (19 hours), enabling a “start-stop” dosing strategy as clinically indicated. It is a welcome addition to the therapeutic armamentarium for moderate to severely active ulcerative colitis.

References

1. Lamb CA, Kennedy NA, Raine T, Hendy PA, Smith PJ, Limdi JK, et al. British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. Gut. 2019.
2. NICE. Ustekinumab for treating moderately to severely active ulcerative colitis. Technical appraisal guidance 633. wwwniceorguk/guidance/ta633. June 2020.
3. Rutgeerts P, Sandborn WJ, Feagan BG, Reinisch W, Olson A, Johanns J, et al. Infliximab for induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2005;353(23):2462-76.
4. 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. 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.
6. Gordon JP, McEwan PC, Maguire A, Sugrue DM, Puelles J. Characterizing unmet medical need and the potential role of new biologic treatment options in patients with ulcerative colitis and Crohn's disease: a systematic review and clinician surveys. Eur J Gastroenterol Hepatol. 2015;27(7):804-12.
7. Kennedy NA, Heap GA, Green HD, Hamilton B, Bewshea C, Walker GJ, et al. Predictors of anti-TNF treatment failure in anti-TNF-naive patients with active luminal Crohn's disease: a prospective, multicentre, cohort study. Lancet Gastroenterol Hepatol. 2019;4(5):341-53.
8. Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel JF, Sandborn WJ, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2013;369(8):699-710.
9. Sands BE, Sandborn WJ, Panaccione R, O'Brien CD, Zhang H, Johanns J, et al. Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2019;381(13):1201-14.
10. Sandborn WJ, Su C, Panes J. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2017;377(5):496-7.

	11. Sandborn W DHG, Wolf D et al. Ozanimod as induction therapy in moderate to severe ulcerative colitis-results from the phase 3 randomised double-blind placebo controlled True North study. United European Gastroenterol J. 2020;8:1258-75. 12. Squibb BM. https://news.bms.com/news/details/2021/U.S.-Food-and-Drug-Administration-Approves-Bristol-Myers-Squibbs-Zeposia-ozanimod-an-Oral-Treatment-for-Adults-with-Moderately-to-Severely-Active-Ulcerative-Colitis1/default.aspx. 2021.
8. What impact would the technology have on the current pathway of care?	Ozanimod will be an additional therapeutic option for UC patients who fail conventional or biological therapy or where such therapies are not tolerated or contra-indicated. There are no head to head studies comparing ozanimod to current available treatments. In the True North study patients who had failed two biologics were excluded from this study. Therefore we suggest Ozanimod should be made available to patients who have not failed two biologics
The use of the technology	
9. To what extent and in which population(s) is the technology being used in your local health economy?	N/A-ozanimod is currently not available to use for this indication
10. Will the technology be used (or is it already used) in	Ozanimod is approved for use in Multiple Sclerosis(MS) in Scotland (not England and Wales). It is not currently used in England and Wales but we anticipate that it will be used as standard care for UC patients in select situations.

the same way as current care in NHS clinical practice?	The dosage regimen used in UC and MS is the same.
• How does healthcare resource use differ between the technology and current care?	It will be an additional option to existing therapeutics. Screening for serious infection e.g. TB and exclusion of active malignancy are similar to other biologic and small molecule treatments in UC. Varicella screening and recommendation for vaccination if not immune, is already in place for patients with UC commencing tofacitinib. Patients who have a history of uveitis and macular oedema were excluded in the True North study. The European Medicines Authority recommends that patients with history of uveitis and macular oedema do not get this treatment. Patients with risk factors for macular oedema would need ophthalmology review prior to and following commencement of treatment. The British National Formulary (BNF) recommends: an ophthalmic examination in patients with diabetes , history of uveitis or retinal disease before initiation and periodically during treatment (with interrupt treatment if macular oedema occurs). Ozanimod is contraindicated in patients at risk of symptomatic bradycardia. The EMA recommends that an ECG should be obtained prior to treatment initiation with ozanimod to determine whether any pre-existing cardiac abnormalities are present in all patients. The BNF also recommends: ‘Patients with certain pre-existing cardiac conditions should be monitored for bradycardia for 6 hours after the first dose. An ECG should be obtained before dosing and after 6 hours—consult product literature for further information. Monitor blood pressure regularly during treatment.’

	References 1. OZANIMOD | Drug | BNF content published by NICE 2. https://www.ema.europa.eu/en/medicines/human/EPAR/zeposia
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Ozanimod would be used exclusively in secondary and tertiary care by gastroenterologists experienced in the care of people living with ulcerative colitis after and in line with NICE approval
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	.Specialist teams would adhere to prescribing information and related laboratory and clinical monitoring as is standard for other biological and newer small molecule therapies. Baseline ECG is required but is accessible to all outpatients so this would not require investment. If the BNF recommendation to monitor heart rate for 6 hours after first dose is followed this would have resource implications as the patient would need an area to sit whilst this monitoring is performed and the monitor may be unavailable for other patients. In the True North study none of the incidences of bradycardia were symptomatic and all were transient. The EMA does not have this recommendation; only a baseline ECG is recommended. If patients have risk factors for macular degeneration, e.g. diabetes, an ophthalmology review will also be required which could have implications and potentially affect starting this medication whilst awaiting review. BNF recommends that liver blood tests are performed before initiation of treatment and then at months 1, 3, 6, 9 and 12 and periodically thereafter. Treatment would need to be witheld if significant liver injury occurs: liver transaminases above 5 times the upper limit of normal). This is the same as currently done for azathioprine/mercaptopurine monitoring.

• If there are any rules (informal or formal) for starting and stopping treatment with the technology, does this include any additional testing?	There are no stipulated “rules” for starting and stopping treatment other than the assessment of ulcerative colitis disease activity at initiation and during monitoring of therapy as standard of care. These involve blood tests (full blood count and differential white cell count, standard blood biochemistry (renal, liver function), inflammatory markers (CRP) and stool marker of colitis activity (faecal calprotectin). Additionally, an endoscopic assessment (sigmoidoscopy/colonoscopy) is performed as standard of assessment upon initiation of biological/small molecule treatment and often for assessment of disease activity if considering stopping treatment.
11. What is the outcome of any evaluations or audits of the use of the technology?	N/A
Equality	
12a. Are there any potential equality issues that should be taken into account when considering this treatment?	None within a defined pathway of care for all individuals with moderate-severely active ulcerative colitis meeting criteria for prescription for Ozanimod.
12b. Consider whether these issues are different from issues with current care and why.	None, as stated above.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

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

Professional organisation submission

Ozanimod for treating moderately to severely active ulcerative colitis [ID3841]

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	**UK Clinical Pharmacy Association (UKCPA)**

3. Job title or position	[redacted]
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).	**UKCPA is a not-for-profit organisation, which invest all surplus back into the association in order to provide better services and benefits for members, and to support initiatives which improve patient care.**
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.]	**No**

If so, please state the name of manufacturer, amount, and purpose of funding.	
5c. 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 and maintain clinical remission in patients with moderately to severely ulcerative colitis Avoid major surgery (colectomy) which could have considerable impact on patient quality of life particularly if pouch or permanent stoma formation is necessary
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by	- Endoscopic improvement (intestine mucosal healing) - Corticosteroid free remission - Prevent relapse - Hospital and surgery avoidance - Improvement in IBDQ scores and quality of life

x cm, or a reduction in disease activity by a certain amount.)	
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Currently stratification of patients regarding response to different classes of drugs is not possible. A considerable number of patients do not currently respond to or lose response to available treatment. Hence, a novel mechanism of action is an additional option for these patients.
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Medical treatment: - Aminosalicylates, also known as 5-ASAs - Corticosteroids - Immunosuppressants - Ciclosporin - Biologic medicines - Tofacitinib Surgical treatment: - Colectomy - Ileostomy - Ileoanal pouch
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults https://www.bsg.org.uk/wp-content/uploads/2019/12/BSG-IBD-Guidelines-2019.pdf ECCO Guidelines on Therapeutics in Ulcerative Colitis: Medical Treatment https://academic.oup.com/ecco-jcc/advance-article/doi/10.1093/ecco-jcc/jjab178/6390052

	NICE: Ulcerative colitis management https://www.nice.org.uk/guidance/ng130
• 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.)	Currently within the NHS, Thiopurines (azathioprine or mercaptopurine) is used first if mesalazine alone failed to maintain remission. If thiopurine fails or if it's considered as an unsuitable treatment option due to contraindication or intolerance, biologic therapies would be the next step. Biologics that are available in the UK are tumour necrosis factor inhibitors (anti-TNFs e.g. infliximab, adalimumab, golimumab), integrin inhibitor (vedolizumab), interleukin inhibitor (ustekinumab). The choice of treatment is based on patient's past medical history, route of administration, side effect profile, patient's preference, drug cost. A non-biologic therapy (JAK inhibitor: tofacitinib) is another medical treatment option if failed all biologic therapies. The use of biologic and non-biologic therapy in ulcerative colitis is still a developing field as there are still a lot of unknown areas e.g. how genetic factors enhance prediction of response to anti-TNF therapy. Practice varies across the UK depending on local resources, access to treatment / laboratory tests.
• What impact would the technology have on the current pathway of care?	Additional medical treatment options before surgical intervention particularly for primary and secondary loss of response
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes, more likely 2nd or 3rd line therapy after thiopurines and biologics.
• How does healthcare resource use differ	Depends where this technology is positioned and the drug cost.

between the technology and current care?	
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Secondary care
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	None
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes, for the patient cohort who are not or no longer responding to current treatment options or are ineligible to avoid life changing surgery.
• Do you expect the technology to increase length of life more than current care?	Unsure (difficult to measure)
• Do you expect the technology to increase health-related quality of	Yes, for the patient cohort who are not or no longer responding to current treatment options or are ineligible to avoid life changing surgery.

life more than current care?	
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Patients with co-morbidities that are contraindicated will not be suitable for this treatment - Patients who in the last 6 months, experienced myocardial infarction, unstable angina, stroke, transient ischemic attack (TIA), decompensated heart failure requiring hospitalization, or Class III/IV heart failure or have the presence of Mobitz type II second-degree or third degree atrioventricular (AV) block, sick sinus syndrome, or sino-atrial block, unless the patient has a functioning pacemaker - Patients with severe hepatic dysfunction (child –pugh class C) - Patients taking a monoamine oxidase (MAO) inhibitor - During pregnancy or in women of childbearing potential (if not using effective contraception) - Patients with active malignancies - Patients with active infections (including chronic infection e.g. TB or viral hepatitis)
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	This technology will be less complicated than some other alternative agents as this is an oral medication. This removes the need for patient administration training. It also helps with capacity of the medical infusion units in hospital which are already stretched to capacity. This is also of benefit to the patient as they do not need to attend the hospital on a regular basis, potentially putting themselves at risk of infections. Most of the injectable agents used to treat UC need to be stored in the fridge; ozanimod does not have the same requirement so less chance of waste due to incorrect storage or temperature excursions.

clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	Cardiac assessment is required before starting treatment. An ECG is required to rule out any pre-existing conditions. This is not required for other treatment options. Patients would need to attend the hospital for this test to be conducted.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Patients will be escalated to a biologic agent or Janus kinase inhibitors or this technology if they have responded inadequately to conventional therapy including corticosteroids and mercaptopurine or azathioprine. Patients will be switched from one biologic to another if they are unable to tolerate it or they are not responding to it. This will be assessed by : • checking symptoms and symptom history • conducting tests such as faecal calprotectin, CRP, drug trough and antibodies level (if available) • ruling out other causes of symptoms such as infection • Primary or secondary loss of response Treatment will be stopped if patient is in deep remission clinically and endoscopically, or if patients develop serious side effects or adverse reactions secondary to treatment; or patients acquire new co-morbidities which place them into contra-indication category.

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?	Reduced hospital admissions and need for surgery which will impact beneficially on the health economy
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?	Treatment options for UC are limited in comparison to other diseases such as rheumatoid arthritis or psoriasis. Having an alternative agent with an alternative mode of action is of benefit to all patients, especially those whom have tried and failed other agents. Ozanimod was found to be significantly more effective at inducing and maintaining remission in moderate to severe UC in comparison to placebo. Novel mechanism of action in comparison to current medical treatment options in ulcerative colitis
• Is the technology a 'step-change' in the management of the condition?	Yes, it could be used as 2nd/3rd line treatment if current licensed biologic therapies are clinically inappropriate. It could be placed at the same position as tofacitinib in ulcerative colitis treatment pathway.

• Does the use of the technology address any particular unmet need of the patient population?	Additional medical treatment options with new mechanism of action for patients with moderate to severe ulcerative colitis before surgical intervention. Additional oral therapy which will be appropriate for patient who has no intravenous access or needlephobic
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?	Apart from the side effect profiles, the side effect profile for ozanimod (for MS) appears to be similar to other biologic agents used to treat UC. However a recent published systematic review by Lasa et al looked at the efficacy and safety of biologics or small molecule drugs used for induction or maintenance of remission for patients with moderate to severe UC and found that ozanimod ranked highest for serious adverse events compared to other biologic therapies and janus kinase inhibitors. Additional blood and cardiac monitoring may affect patient's quality of life and side effect(s) such as symptomatic bradycardia may require cardiologist input for further management and to determine the most appropriate monitoring strategy. Reference: Lasa J, Olivera P, Danese S, Peyrin-Biroulet L. Efficacy and safety of biologics and small molecule drugs for patients with moderate-to-severe ulcerative colitis: a systematic review and network meta-analysis. The Lancet Gastroenterology & Hepatology. 2021

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?	n/a
• What, in your view, are the most important outcomes, and were they measured in the trials?	Yes induction and maintenance of clinical remission, mucosal healing and corticosteroid-free remission are all measured in the trials. Ozanimod has significantly improved the outcome in all these fields.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	n/a
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	All highlighted in clinical trials

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 treatment(s) since the publication of NICE technology appraisal guidance TA633?	No
21. How do data on real-world experience compare with the trial data?	Consistent
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No

22b. Consider whether these issues are different from issues with current care and why.	N/A

Key messages

23. In up to 5 bullet points, please summarise the key messages of your submission.

- Novel oral medication as induction and maintenance therapy in patients with moderate to severe ulcerative colitis
- Significantly superior to placebo in induction of endoscopy improvement
- Significantly superior to placebo in induction of clinical remission
- Significantly superior to placebo in maintenance of remission
- Potentially serious adverse effect(s) or side effect(s) which require additional monitoring

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

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.

……

Ozanimod for treating moderately to severely active ulcerative colitis [ID3841]

A Single Technology Appraisal

Produced by	Peninsula Technology Assessment Group (PenTAG) University of Exeter Medical School South Cloisters St Luke's Campus Heavitree Road Exeter EX1 2LU
Authors	**Amanda Brand**[1] **Brian O'Toole**[1] **Madhusubramanian Muthukumar**[1] **Justin Matthews**[1] **Fraizer Kiff**[1] **Naomi Shaw**[1] **Edward CF Wilson**[1] **Jonathan Digby-Bell**[2] **Louise Crathorne**[1] **G.J. Melendez-Torres**[1] [1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter [2] Consultant Gastroenterologist, Royal Devon & Exeter Hospital
Correspondence to	Amanda Brand 3.09 South Cloisters, St Luke's Campus, Heavitree Road, Exeter, EX1 2LU; a.brand2@exeter.ac.uk
Date completed	22/03/2022
Source of funding	This report was commissioned by the NIHR Systematic Reviews Programme as project number 13/54/43.

Declared competing interests of the authors	Dr Digby-Bell received payment of delegate fee from AbbVie (a stakeholder in technology appraisal ID3841) for conference attendance at the British Society of Gastroenterology annual meeting 2021.
Acknowledgments	The authors acknowledge the administrative support provided by Mrs Sue Whiffin and Ms Jenny Lowe (both PenTAG) and clinical expert advice provided by Dr James Goodhand (consultant gastroenterologist, Royal Devon & Exeter Hospital). Dr Goodhand received grant funding over the last 18 months from Biogen and Celltrion (companies associated with the evaluation) for work completed assessing outcomes from COVID-19 infection and vaccination in patients treated with infliximab and vedolizumab.
Rider on responsibility for document	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	Brand, O'Toole, Muthukumar, Matthews, Kiff, Shaw, Wilson, Digby-Bell, Crathorne, Melendez-Torres. Ozanimod for treating moderately to severely active ulcerative colitis [ID3841] A Single Technology Appraisal. Peninsula Technology Assessment Group (PenTAG), 2022.
Copyright	© 2022, PenTAG, University of Exeter. Copyright is retained by Celgene (a Bristol-Myers Squibb company) for tables and figures copied and/or adapted from the company submission and other submitted company documents.

Author Contributions:	
Amanda Brand	Project lead, lead for the ERG's critical appraisal of the clinical evidence of the company submission, writing and editorial input
Brian O'Toole	Lead for the ERG's appraisal of the economic evidence, drafted economic sections of the report, writing and editorial input
Madhusubramanian Muthukumar	Critical appraisal of the economic evidence, checked and re-analysed the economic model, carried out ERG base case analyses and further scenario analyses, and drafted economic sections of the report
Justin Matthews	Critical appraisal of the clinical evidence, conducted additional clinical work on the NMA for the submission, and drafted sections of the report
Fraizer Kiff	Critical appraisal of the clinical evidence and drafted sections of the report
Naomi Shaw	Critical appraisal of the literature search strategies, conducted additional literature searching, and editorial review
Edward CF Wilson	Independent appraisal of the ERG model for errors and editorial review
Jonathan Digby-Bell	Clinical advice and review of draft report
Louise Crathorne	Critical appraisal of the company submission, writing and editorial input, and co-supervised the final report
G.J. Melendez-Torres	Critical appraisal of the company submission, writing and editorial input, and co-supervised the final report. Guarantor of the report

Abbreviations

ADA	adalimumab
AE	adverse event
AG	Assessment Group
ALT	alanine aminotransferase
AMSTAR	A MeaSurement Tool to Assess systematic Reviews
AST	aspartate aminotransferase
BID	*L. bis in die*/twice a day
β-hCG	beta-human chorionic gonadotropin
BMI	body mass index
BSC	best supportive care
CD	Crohn's disease
CEAC	cost-effectiveness acceptability curve
CI	confidence interval
CODA	Convergence Diagnosis and Output Analysis
CRD	Centre for Reviews and Dissemination
CrI	credible interval
CRP	C-reactive protein
cPAS	confidential patient access scheme
CS	company submission
CSR	clinical study report
CV	cardiovascular
CvT	conventional therapy
ECG	electrocardiogram
EMA	European Medicines Agency
EQ-5D(-5L)	European Quality of Life Five Dimension (Five Level)
ERG	Evidence Review Group
EU	European Union
FDA	Food and Drug Administration
FE	fixed effect
GOL	golimumab
HES	Hospital Episode Statistics
HRQoL	health-related quality of life
HTA	health technology assessment

ICER	incremental cost-effectiveness ratio
ICTRP	International Clinical Trials Registry Platform
IFX	infliximab
ITT	intention-to-treat
IV	intravenous
JAGS	Just another Gibbs sampler
JAK	Janus kinase
LYG	life years gained
MCID	minimally clinically important difference
MCMC	Markov chain Monte Carlo
MCS	mental composite summary score
MVH	Measurement in Valuation of Health
N/A	not applicable
NHB	net health benefit
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	network meta-analysis
NR	not reported
OLE	open-label extension
OR	odds ratio
OWSA	one-way sensitivity analysis
OZA	ozanimod
PBO	placebo
PCS	physical component summary
PSA	probabilistic sensitivity analysis
PSS	Personal Social Services
Q2W	*L. quaque* 2W/once every two weeks
QA	quality assessment
QALY	quality-adjusted life year
QD	*L. quaque die*/once a day
QoL	quality of life
RCT	randomised controlled trial
RBS	rectal bleeding sub-score
RE	random effects
SAE	serious adverse event

SC	subcutaneous
SD	standard deviation
SF-36	36-item Short Form Health Survey
SFS	stool frequency sub-score
SLR	systematic literature review
SmPC	Summary of Product Characteristics
SW	southwest
TA	technology appraisal
TDM	therapeutic drug monitoring
TEAE	treatment-emergent adverse event
TNFi	tumour necrosis factor inhibitor
TOF	tofacitinib
UC	ulcerative colitis
UCSS	ulcerative colitis symptom score
UST	ustekinumab
VAS	visual analogue scale
VEDO	vedolizumab
vs.	versus
WHO	World Health Organization
WTP	willingness-to-pay

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the evidence review group (ERG) as being potentially important for decision-making. It also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

- Section 1.1 provides an overview of the key issues and the differences in the assumptions of the company and the ERG in economic analysis.
- Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER.
- Sections 1.3 to 1.5 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main ERG report.
- Sections 1.6 and 1.7 provide an overview of the ERG's preferred base case and sensitivity analyses undertaken by the ERG.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1. Overview of the ERG's key issues

A brief overview of the key issues identified by the ERG in their appraisal of the company submission (CS) is provided in Table 1. Further detail of the issues is provided in Sections 1.3, 1.4 and 1.5.

Broadly speaking, the key clinical issues related to the exclusion of tofacitinib as a comparator as well as uncertainty surrounding the assumptions made in the analytical approach comparing ozanimod and its comparator treatments. This omission of tofacitinib and uncertainties around the network meta-analysis (NMA) have implications for the cost-effectiveness of ozanimod as well as for the positioning of ozanimod in a highly individualised treatment pathway. Furthermore, the ERG was of the opinion that random effects (RE) models should have been used to estimate clinical effectiveness in the NMAs, and that the use of fixed effect (FE) models may have resulted in inaccurate inputs of clinical effectiveness into the economic model. This, in turn, may have biased the results of the ICERs and increased the overall uncertainty of the cost-effectiveness evidence in the context of the decision problem. In terms of cost-effectiveness

issues, the ERG considered the exclusion of tofacitinib to have a high impact on cost-effectiveness and that its inclusion might result in considerably different fully incremental ICERs. The ERG also noted several concerns pertaining to the company's estimation of modelled transition probabilities and response rates for best supportive care (BSC) and uncertainty around the company's handling of subsequent treatments. Furthermore, the ERG did not consider the probabilistic sensitivity analysis (PSA) provided by the company to be helpful in decision-making, due to the exclusion of tofacitinib as a comparator and uncertainties around the NMA.

Table 1: Summary of key issues

ID	Summary of issues	Report sections
Key Issue 1	Tofacitinib was excluded as a comparator in TNFi-naïve and -experienced subgroups	Sections 1.3, 2.2.1, 2.3 and 6.1.1
The ERG reviewed the clinical effectiveness and safety evidence presented in the CS and identified the following key issues for consideration by the committee. Key Issue 2	Baseline risks for placebo anchors in the NMAs taken from the same trials those used for relative risk	Section 1.4, 3.3.2.4, 3.4.2.4 and 3.5.3.1
Key Issue 3	A RE model may be more appropriate for use in the maintenance phase NMAs	Section 1.4, 3.4.2.2 and 3.5.3.2
Key Issue 4	Modelled efficacy estimates for BSC in the post-active treatment phase	See Section 4.2.6, 4.2.6.3 and 6.3
Key Issue 5**Error! Reference source not found.**	There is uncertainty surrounding the handling of subsequent treatments/treatment sequencing within the model	See Sections 1.5 and 4.2.2.3
Key Issue 6	The PSA provided by the company was not considered helpful for decision making	See Sections 1.5 and 5.2.2

Abbreviations: NMA, network meta-analysis; PSA, probabilistic sensitivity analysis; RE, random effects; TNFi, tumour necrosis factor inhibitor

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are outlined in Table 2.

Table 2: Key differences between the company's preferred assumptions and ERG's preferred assumptions

Company's preferred assumption	ERG preferred assumption	Report Sections
To exclude tofacitinib as a comparator	The ERG's preference was to include tofacitinib as a relevant comparator in both the TNFi-naïve and TNFi experienced subgroups. However, it was not possible to include tofacitinib into the company's model.	4.2.4 and 6.1.1
Use of FE model in both TNFi-naïve and TNFi-experienced subgroups for the maintenance phase NMAs as well as FE model in the TNFi-experienced subgroup for the induction phase NMA	The ERG preferred the use of RE models for maintenance phase NMAs. The ERG was unable to produce RE models with sufficient convergence (without using an informative prior distribution) and were therefore unable to use a RE model as part of its preferred base case.	1.4, 3.4.2.2, 3.5.3.2 and 4.2.6.4
Baseline risks for placebo anchors included in the NMAs calculated from the identical set of trials used to calculate the relative treatment effect	The ERG preferred to use the placebo arm values from individual trials included in the NMA that were more generalisable to the UK context. This has been considered as part of the ERG preferred base case.	1.4, 3.3.2.4, 3.4.2.4, 3.5.3.1 and 6.3
Remission transition probabilities in the BSC arm were estimated via loss of overall response (including remission). Furthermore, for BSC, loss of response and loss of response (No remission) were based on pooled population estimates.	The ERG preferred revised post-active treatment transition probabilities for BSC which include an alternative means of estimating remission probabilities for BSC based on 'loss of remission' (directly from the sustained remission estimates) and different BSC response rates for the TNF-naïve and TNF-experienced populations. These changes were incorporated into the ERG's preferred base case.	4.2.6.3 and 6.3

Abbreviations: BSC, best supportive care; ERG, Evidence Review Group; FE, fixed effect; NMA, network meta-analysis; RE, random effects; TNFi, tumour necrosis factor inhibitor

1.2. Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Moving patients to the remission and response (no remission) health states at the end of the induction phase and by sustaining remission and response (no remission) for a

proportion of patients at the end of the maintenance phase. [redacted].

- Assuming loss of response to treatment whilst in remission and response (no remission) i.e. probability of loss of response during the maintenance phase is considered for all treatments.
- Considering discontinuation due to adverse events during the maintenance phase. Patients discontinuing treatment received best supportive care, comprising components of conventional therapy, and entered the 'Active' ulcerative colitis (UC) health state accruing costs and QALYs associated with this health state.

Overall, the technology is modelled to affect costs by:

- Having lower administration costs compared to IV comparators (see Section 4.2.8.4) and by having lower drug acquisition costs compared some comparator treatments. However, the ERG noted that when cPAS discounts were included, the cost effectiveness of ozanimod compared to comparator treatments changed considerably (see cPAS Appendix).
- Monitoring requirements were assumed to be similar for all treatments, however ozanimod was assumed to require an electrocardiogram (ECG) during induction.

The modelling assumptions that have the greatest effect on the ICER are:

- Alternative utility values for modelled health states, variation in dose escalation assumptions, % of patients receiving subcutaneous (SC) vedolizumab and the inclusion of extended induction (as evident by the company's scenario analyses).

1.3. The decision problem: summary of the ERG's key issues

The ERG reviewed the approach of the company to addressing the NICE decision problem for this appraisal and identified the following key issues for consideration by the committee.

Key Issue 1: Tofacitinib was excluded as a comparator in TNFi-naïve and -experienced subgroups

Report sections	Sections 1.3, 2.2.1, 2.3 and 6.1.1
Description of issue and why the ERG has identified it as important	The company excluded tofacitinib as a comparator to ozanimod in the TNFi-naïve subgroup, indicating that this treatment is not used at this line of treatment. Clinical advice to the ERG indicated that tofacitinib is used at this line of treatment in the UK clinical practice due to its rapid action and oral administration and confirmed that its use is increasing. Tofacitinib was also excluded as a comparator in the TNFi-experienced subgroup, with the company citing safety concerns. The company further indicated that the exclusion of tofacitinib was accepted by the committee for TA633, though the ERG noted that tofacitinib was included in the clinical effectiveness results and economics to provide the full picture of cost-effectiveness for ustekinumab. Clinical opinion to the ERG acknowledged these concerns but advised that these predominantly impact clinical practice in the US. Safety concerns regarding tofacitinib impact UK clinical practice far less, with concerns managed at the individual patient level. Furthermore, the ERG considered the treatment landscape to have evolved since TA633, with clinical experts advising that the use of tofacitinib is increasing. As a result, the ERG considered the exclusion of tofacitinib as a comparator in either subgroup to be an outstanding area of uncertainty.
What alternative approach has the ERG suggested?	During clarification, the ERG requested that tofacitinib be included as a relevant comparator in the model, for both subgroups, using the treatment efficacy estimates already derived from NMAs. However, this analysis was not provided to the ERG. For completeness, the ERG conducted a naïve cost comparison vs. tofacitinib, which assumed clinical equivalency between treatments in terms of efficacy. See Section 2.2.1, 2.3 and 6.1.1 for further discussion. Tofacitinib was found to be cost saving compared to ozanimod over lifetime of treatment, when considering the PAS price for tofacitinib.
What is the expected effect on the cost-effectiveness estimates?	The company did not provide the ERG with an additional analysis including tofacitinib as a relevant comparator. As such the base case results provided by the company should be interpreted with caution. The ERG expects the inclusion of tofacitinib to have a high impact on cost-effectiveness results. This may result in substantially different fully incremental ICERs.
What additional evidence or analyses might help to resolve this key issue?	The inclusion of clinical efficacy estimates for tofacitinib within the economic analysis would have sufficiently resolved this issue. The ERG was unable to amend the company's model to include tofacitinib, due to the lack of flexibility and time constraints.

Abbreviations: cPAS, confidential patient access scheme; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; NMA, network meta-analysis; TA, technology appraisal; TNFi, tumour necrosis factor inhibitor; vs., versus

1.4. The clinical effectiveness evidence: summary of the ERG's key issues

The ERG reviewed the clinical effectiveness and safety evidence presented in the CS and identified the following key issues for consideration by the committee.

Key Issue 2: Baseline risks for placebo anchors in the NMAs taken from the same trials those used for relative risk

Report sections	Section 1.4, 3.3.2.4, 3.4.2.4 and 3.5.3.1
Description of issue and why the ERG has identified it as important	The baseline risks for placebo anchors included in the NMAs (in this context, the probability of being in non-response and non-remission, under placebo) were calculated from the identical set of trials used to calculate the relative treatment effect. The ERG noted that this was contrary to NICE guidance (Dias et al 2013),[1] which recommends separate modelling and sources of information for relative treatment and baseline effects. Several trials included in the NMA did not match well to the decision problem due to diverse settings, demographic as well as clinical features of participants, and concomitant medication use as described in Sections 3.3.2.4 and 3.4.2.4.
What alternative approach has the ERG suggested?	The baseline risk in the placebo anchors could be estimated more accurately through the use of a baseline risk in the placebo arm(s) of a study or studies that match the decision problem more closely. Though not fully in line with NICE guidance, given the timeframe of the appraisal, the ERG used the placebo arm values from individual trials included in the NMA that were more generalisable to the UK context to generate estimates of clinical effectiveness for its base case. This approach is described in Section 3.4.2.4.
What is the expected effect on the cost-effectiveness estimates?	The ERG's base case NMA, utilising placebo baseline risks from a more UK-appropriate trial, resulted in lower response rates for placebo, as well as for ozanimod and most of the active treatment comparators, as discussed in Section 3.5.3.13.5.3.1. The impact of the revised baseline placebo risk on cost-effectiveness in conjunction with the alternative transition probabilities for BSC in the post-active treatment phase has been discussed in **Error! Reference source not found.** and Table 4.
What additional evidence or analyses might help to resolve this key issue?	Conducting the NMAs using baseline risk in the placebo arm, derived from a study that is highly generalisable to the UK context and was identified through a proper, protocol-driven systematic review would result in more generalisable estimations of treatment efficacy. The ERG could not conduct a comprehensive systematic review for highly generalisable evidence to inform its base case approach, as prescribed by NICE guidance, due to the timeframe of the appraisal, but used values from broadly representative single trials included in the NMA, as described in Section 3.4.2.4 and 3.5.3.1. As such, the ERG would like to highlight that there is residual uncertainty in its approach and recommends using placebo baseline risk values from a systematically identified trial that is highly specific to the decision problem to inform the NMA.

Abbreviations: BSC, best supportive care; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; NMA, network meta-analysis

Key Issue 3: A random effects model may be more appropriate for use in the maintenance phase NMAs

Report sections	Section 1.4, 3.4.2.2 and 3.5.3.23.5.3.2
Description of issue and why the ERG has identified it as important	Clinical efficacy parameters (clinical response and remission) for all treatments were derived from the induction and maintenance NMAs conducted by the company. In the base case analysis the company opted to use a FE model in both TNFi-naïve and TNFi-experienced subgroups for the maintenance phase NMAs, as well as for the TNFi-experienced subgroup during induction. The ERG acknowledged the company's rationale for using the FE model for these subgroups and phases (namely that the fit was reasonable, RE models did not converge or had highly uncertain posterior SD). However, due to the high degree of heterogeneity amongst the studies included in the NMA, the ERG considered FE models to be inappropriate.
What alternative approach has the ERG suggested?	The ERG suggests RE models to estimate clinical response and remission for both the TNFi-naïve and TNFi-experienced subgroups in the maintenance phase, as well as for the TNFi-experienced subgroup in the induction phase to address the heterogeneity in the evidence base, as discussed in Section 3.4.2.2. The ERG attempted to re-run the company NMAs using RE models with alternative baseline placebo risk; these also failed to converge (see Section 3.5.3.2). To address non-convergence, the ERG suggests the use of appropriate informative prior distributions from literature.
What is the expected effect on the cost-effectiveness estimates?	Using a RE model is likely to result in different clinical efficacy estimates for all treatments. This is anticipated to have an impact on the cost effectiveness results, however the directional impact of this analysis could not be determined in the timeframe of the appraisal.
What additional evidence or analyses might help to resolve this key issue?	The ERG noted that overall, there is uncertainty surrounding the clinical data used in the economic model, due to heterogeneity amongst studies and the lack of direct trial data. The ERG recommends that, because of heterogeneity, the NMA be run using RE models with informative prior distributions in the event of non-convergence.

Abbreviations: ERG, Evidence Review Group; FE, fixed effect; ICER, incremental cost-effectiveness ratio; NMA, network meta-analysis; RE, random effects; SD, standard deviation; TNFi, tumour necrosis factor inhibitor

1.5. The cost effectiveness evidence: summary of the ERG's key issues

The ERG reviewed the economic model and cost-effectiveness evidence presented in the CS and identified the following key issues for consideration by the committee.

Key Issue 4: Modelled efficacy estimates for BSC in the post-active treatment phase

Report sections	Section 4.2.6, 4.2.6.3 and 6.3
Description of issue and why the ERG has identified it as important	The ERG noted several concerns pertaining to the company's estimation of modelled transition probabilities and response rates for BSC (in the post active treatment phase), probability of remission, and the use the same pooled estimate for the BSC remission and response rates (for both TNF-naïve and TNFi-experienced subgroups). These concerns are discussed further in 4.2.6.3.
What alternative approach has the ERG suggested?	The ERG used an alternative approach to estimate remission state transition probabilities for BSC in the post active treatment phase i.e. these were calculated directly from the sustained remission estimates via 'loss of remission'. This approach has been incorporated into the ERG base case (see Section 6.3 for results). As noted in 4.2.6.3, the concerns surrounding the estimation of probability of remission, and the use of same pooled estimate for the BSC remission and response rates (for both TNF-naïve and TNF-experienced subgroups) were addressed as a result of using the alternative baseline placebo risk estimates which are different for TNF-naïve and TNF-experienced subgroups in the ERG's base case.
What is the expected effect on the cost-effectiveness estimates?	Due to the use of alternative placebo risk estimates derived by including only trials which are relevant to decision making, the overall response decreases across all treatments. For a complete description of the impact of these changes see Section 6.3.
What additional evidence or analyses might help to resolve this key issue?	The ERG base case analysis partly addressed this issue, however the uncertainty around the true remission estimates in the post active treatment phase remained.

Abbreviations: BSC, best supportive care; ERG, Evidence Review Group; QALY, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor; UC, ulcerative colitis

Key Issue 5: There is uncertainty surrounding the handling of subsequent treatments/treatment sequencing within the model

Report sections	Section 1.5 and 4.2.2.3
Description of issue and why the ERG has identified it as important	The company did not consider subsequent treatment use/treatment sequencing in the base case. The ERG noted that the company provided some scenario analyses for the TNFi-naïve subgroup which assumed that patients who do not respond to their initial TNFi treatment can go on to received either vedolizumab or ustekinumab. Results were not overly sensitive to this analysis, however the ERG considered the scenario analysis to be somewhat limited (See Section 4.2.2.3).
What alternative approach has the ERG suggested?	The ERG considered undertaking additional scenario analyses using various treatment sequencing strategies, including within

Report sections	Section 1.5 and 4.2.2.3
	class switching and step up/ step down approaches. However the model was not flexible enough to allow for this. As such there is some uncertainty surrounding the impact of treatment sequencing on the base case ICER. See Section 4.2.2.3 for further discussion.
What is the expected effect on the cost-effectiveness estimates?	It is anticipated that the inclusion of alternative treatment sequence options will have a moderate impact on total treatment costs, and a minor impact on total QALYs (See Section 4.2.2.3). Given the small differences in costs between the modelled treatments, ICERs may vary once treatment sequencing is considered.
What additional evidence or analyses might help to resolve this key issue?	Updating the economic model to allow for the consideration of various treatment sequencing options would help to further explore uncertainty.

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor

Key Issue 6: The PSA provided by the company was not considered helpful for decision making

Report sections	Section 1.5 and 5.2.2
Description of issue and why the ERG has identified it as important	Due to concerns relating to the NMA and the omission of tofacitinib as a relevant comparator in the economic model, the ERG considered that the PSA provided by the company is of limited use for decision-making and should be interpreted with caution.
What alternative approach has the ERG suggested?	Ideally the probabilistic analysis could have been done with the baseline risks of placebo arms from only the trials relevant to the decision problem and addressing the heterogeneity in the placebo arms adequately as this impacts the correlation of parameters. Further, tofacitinib could have also been included in the analysis presenting the true cost-effectiveness of relevant treatment alternatives both in the fully incremental analysis as well as with respect to the cost-effectiveness acceptability curve (CEAC).
What is the expected effect on the cost-effectiveness estimates?	The expected effect is that the probabilistic analysis would account for the correlation of treatment effects adequately and produce a CEAC including the relevant comparators presenting a true picture of cost-effectiveness of ozanimod closer to the reality.
What additional evidence or analyses might help to resolve this key issue?	Updating the economic model and re-running the PSA, addressing the concerns surrounding the NMA especially for the baseline risk estimates along with including tofacitinib as a relevant comparator would render the uncertainty analysis more suitable for decision making.

Abbreviations: CEAC, cost-effectiveness acceptability curve; ERG, Evidence Review Group; NMA, network meta-analysis; PSA, probabilistic sensitivity analysis

1.6. Summary of ERG's preferred assumptions and resulting ICER

The results below present the incremental and cumulative impact of the ERG's preferences. The ERG's preference would have been to include tofacitinib as a comparator within the economic analysis. However, due to the lack of model flexibility, it was not possible for the ERG to include tofacitinib in the economic model. As an exploratory analysis, the ERG has conducted a cost comparison versus tofacitinib (see Table 60 and Table 61 for results).

As part of the ERG preferred base case, the ERG used the following assumptions:

- Revised remission and response probability estimates for the treatments and BSC derived from the ERG run of the NMA using the alternative placebo baseline risks (as per Section 3.4.2.4)
- Revised post-active treatment transition probabilities for BSC which include an alternative means of estimating remission probabilities for BSC based on 'loss of remission' (directly from the sustained remission estimates) and different BSC response rates for the TNF-naïve and TNF-experienced populations, as opposed to an overall pooled estimate in the company's base case.

Table 3: Summary of ERG's preferred assumptions and ICER (TNFi-naïve subgroup)

Scenario	Total costs	Total QALYs	Incremental cost	Incremental QALYs	ICER: Ozanimod vs. comparators (£/QALY)
Company's base case					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£28,686
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£167,024*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£71,023*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£52,736*
ERG's preferred base case assumptions (applied incrementally over company's base case)					
Re-estimation of baseline placebo risks					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,479
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£169,098*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£82,608*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,298*
Revised modelled efficacy estimates for BSC in the post-active treatment phase					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,794
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£169,791*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£82,863*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,640*
Cumulative impact of ERG preferences (deterministic)					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,794
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£169,791*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£82,863*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,640*
Cumulative impact of ERG preferences (probabilistic)					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,842
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£1578721*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£87,452*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£68,470*

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; vs., versus

Note: * ICER in SW quadrant

Table 4: Summary of ERG's preferred assumptions and ICER (TNFi-experienced subgroup)

Scenario	Total costs	Total QALYs	Incremental cost	Incremental QALYs	ICER: Ozanimod vs. comparators (£/QALY)
Company's base case					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£33,725)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£199,551*
ERG's preferred base case (applied incrementally over company's base case)					
Re-estimation of baseline placebo risks					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£71,524)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£427,683*
Revised modelled efficacy estimates for BSC in the post-active treatment phase					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£70,807)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£436,080*
Cumulative impact of ERG preferences (deterministic)					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£70,807)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£436,080*
Cumulative impact of ERG preferences (probabilistic)					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£56,635)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£12,926)

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; vs., versus

Note: * ICER in SW quadrant

Including several labelling issues, the ERG noted a discrepancy between the CS Document B and the model in the distribution used for utilities in the PSA, as discussed in Section 5.2.2; however, it did not have any material impact on the results. Further, during clarification (clarification question B14) the ERG indicated that a fully incremental analysis with the associated CE frontier was missing, following which it was added to the model. Otherwise, no serious errors were found in the company's model that impacted the results.

1.7. Summary of exploratory and sensitivity analyses undertaken by the ERG

A summary of exploratory and sensitivity analyses undertaken by the ERG is provided in Table 5 (TNFi-naïve subgroup) and Table 6 (TNFi-experienced subgroup).

Table 5: ERG scenario analysis (TNFi-naïve subgroup)

	Section in ERG report	Incremental costs	Incremental QALYs	ICER £/QALY	% Change from company base case
Company base-case					
ozanimod	5.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£28,686	
infliximab		[redacted]	[redacted]	£167,024*	
golimumab		[redacted]	[redacted]	£71,023*	
vedolizumab		[redacted]	[redacted]	£52,736*	
Cost comparison with tofacitinib					
Incremental cost associated with ozanimod	6.1	[redacted]	-	-	-
Spontaneous remission (0.75% per model cycle)					
ozanimod	6.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£29,830	4%
infliximab		[redacted]	[redacted]	£169,731*	2%
golimumab		[redacted]	[redacted]	£72,123*	2%
vedolizumab		[redacted]	[redacted]	£53,983*	2%
Ozanimod AE discontinuation rate in maintenance phase (5% that of induction)					
ozanimod	6.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£29,790	4%
infliximab		[redacted]	[redacted]	£137,368*	-18%
golimumab		[redacted]	[redacted]	£65,285*	-8%

	Section in ERG report	Incremental costs	Incremental QALYs	ICER £/QALY	% Change from company base case
vedolizumab		[redacted]	[redacted]	£51,677*	-2%
Ozanimod AE rate in the maintenance phase (20% increase)					
ozanimod	6.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£28,750	0%
infliximab		[redacted]	[redacted]	£166,869*	
golimumab		[redacted]	[redacted]	£70,961*	
vedolizumab		[redacted]	[redacted]	£52,720*	
% patients receiving SC vedolizumab (80% after year 1)					
ozanimod	6.1	-	-	-	-
adalimumab		Not applicable			
infliximab					
golimumab					
vedolizumab		[redacted]	[redacted]	£44,204*	-16%
Treatment regimen costs applied per treatment cycle					
ozanimod	6.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£33,815	18%
infliximab		[redacted]	[redacted]	£188,210*	13%
golimumab		[redacted]	[redacted]	£71,528*	1%
vedolizumab		[redacted]	[redacted]	£53,501*	1%
Revised modelled efficacy estimates for BSC in the post-active treatment phase					
ozanimod	6.3	-	-	-	-
adalimumab		[redacted]	[redacted]	£28,797	0%
infliximab		[redacted]	[redacted]	£167,294*	0%
golimumab		[redacted]	[redacted]	£71,133*	0%
vedolizumab		[redacted]	[redacted]	£52,859*	0%

Abbreviations: AE, adverse events; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; SC, subcutaneous

Table 6: ERG scenario analysis (TNFi-experienced subgroup)

	Section in ERG report	Incremental costs	Incremental QALYs	ICER £/QALY	% Change from company base case
Company base-case					

	Section in ERG report	Incremental costs	Incremental QALYs	ICER £/QALY	% Change from company base case
ozanimod	5.1	-	-	-	-
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£33,725)	
vedolizumab		[redacted]	[redacted]	£199,551*	
Cost comparison with tofacitinib					
Incremental cost associated with ozanimod	6.1	[redacted]	Not applicable		
Spontaneous remission (0.75% per model cycle)					
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£34,594)	3%
vedolizumab		[redacted]	[redacted]	£198,146*	-1%
Ozanimod AE discontinuation rate in maintenance phase (5% that of induction)					
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£41,096)	22%
vedolizumab		[redacted]	[redacted]	£160,695*	-19%
Ozanimod AE rate in the maintenance phase (20% increase)					
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£33,689)	0%
vedolizumab		[redacted]	[redacted]	£199,367*	
% patients receiving SC vedolizumab (80% after year 1)					
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£33,725)	0%
vedolizumab		[redacted]	[redacted]	£161,152*	-19%
Treatment regimen costs applied per treatment cycle					

	Section in ERG report	Incremental costs	Incremental QALYs	ICER £/QALY	% Change from company base case
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£47,464)	41%
vedolizumab		[redacted]	[redacted]	£208,721*	5%
Revised modelled efficacy estimates for BSC in the post-active treatment phase					
ozanimod	6.3	-	-	-	-
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£33,354)	-1%
vedolizumab		[redacted]	[redacted]	£200,192*	0%

Abbreviations: AE, adverse events; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; SC, subcutaneous

2. INTRODUCTION AND BACKGROUND

2.1. Introduction

In this report, the Evidence Review Group (ERG) provides a review of the evidence submitted by Celgene, a Bristol Myers Squibb company, in support of ozanimod for treating moderately to severely active UC. The company provided an overview of the disease and burden of ulcerative colitis in the target population in Sections B.1.3.1 to B.1.3.3 in the CS.

UC is an inflammatory condition which affects the gastrointestinal tract.[2] The exact aetiology of the condition remains unknown, but the most popular hypothesis for its cause is centred around a complex interplay between genetic susceptibility, gastrointestinal microbiota, mucosal or generalised immune responses and environmental factors.[3] Ultimately, these factors cause chronic inflammation, which involves the degradation of the cells lining the lumen of the large intestine. As these cells are damaged, ulcers, which are the main cause of the symptoms associated with UC, form.

The symptoms of UC vary between people, depending largely on the extent and severity of their disease. The most common are characterised by symptoms related to an inflamed rectum and include bloody diarrhoea, abdominal pain, urgency and tenesmus.[4] However, the symptoms of UC are not limited to the GI tract as there can be extra-intestinal manifestations causing issues in the joints, eyes, bone, skin and liver, as well as anaemia and fatigue.[5-8] In addition, the condition is characterised by periods of remission interspersed by active disease relapses[4], with 50% of patients have at least one relapse per year.[9]

The degree of symptoms experienced by patients is largely dictated by the extent of their disease, although up to 25% will require hospitalisation at some point during the disease course.[10] The least severe category is ulcerative proctitis, where only the region closer to the rectum is affected. In cases of proctosigmoiditis the rectum and sigmoid colon are affected, whereas in left-sided colitis the rectum as well as sigmoid and descending colon are affected. The most extensive category is pancolitis, where the entire colon is inflamed.[11] This appraisal is focused on those with moderately to severely active UC. Mild-moderate UC is defined as less than six bowel movements per day with few systemic symptoms. Severe UC is more than six bowel movements per day with one or more of the following: a body temperature exceeding 37.8°C; pulse of more than 90 beats per minute; haemoglobin <10.5 g/dL or an erythrocyte sedimentation rate >30 mm/hour.[12]

Short-term symptoms are not the only concern associated with UC; there are also a number of potential longer-term complications. Some of these include bowel cancer, haemorrhage, perforation, strictures, abscesses, anorectal disease, primary sclerosing cholangitis, osteoporosis and toxic megacolon.[13,14] One of the main aims of treatment is to achieve remission in order to avoid the development of these longer-term conditions.

UC usually develops between the ages of 15 and 25 years, though there is a second peak between 55 and 65 years.[15] Historically, incidence has been highest in more economically developed countries, however more recently it has been increasing in developing countries while decreasing in western countries. In the UK specifically, UC affects 1 in 420 people, 52% of which have moderate to severe disease.[16] However, UK incidence rates are falling by 1.6% per year; this decrease is largely seen in the second peak while incidence in those under the age of 17 continues to increase.[2] In the UK, rates of UC are highest in the Northeast, East and Midlands.[2] UC is most common in black people and Caucasian people of European descent while it is less frequently seen in those from Asian communities. There is an equal split in prevalence between men and women,[2] although women with UC are at greater risk of relapse.

UC is initially diagnosed according to a patient's symptoms, in addition to a physical examination for anaemia, which can be confirmed with a blood test, and tenderness in the stomach. A stool sample can also be used to allow clinicians to rule out infections of the stomach or bowel, which can be mistaken for UC. Where UC is suspected, patients are referred for either an X-ray or CT scan to further rule out any serious complications. The diagnosis can then be confirmed with a sigmoidoscopy to determine the level and extent of the inflammation in the bowel, this may also involve a biopsy. If it is suspected that a greater portion of the large intestine is affected, a further colonoscopy may be carried out which can also involve a biopsy.[17]

Once a UC diagnosis is confirmed, the current treatment pathway is highly individualised. In the UK, patients will most often initially receive conventional therapies (CvTs) including corticosteroids, aminosalicylates and immunosuppressants. If CvTs are failing to manage a patient's condition, they will often progress to their first biologic treatment. In those in which they are suitable, tumour necrosis factor inhibitors (TNFi), also known as anti-TNF, treatment will be the first biologic. If patients become intolerant, or fail to respond to, their first biologic treatment, they will usually be treated with a second. The remaining treatments are biologics, vedolizumab and ustekinumab, and a small molecule drug, tofacitinib.[18] Which is used depends on many factors including comorbidities, rate of action and patient preference. If patients fail this

subsequent line of biologic treatment, they may require surgery to remove part of their large intestine, often leaving them in need of a stoma bag.

2.2. Background

2.2.1. Current treatment for ulcerative colitis

The company provides an overview of current treatment options for UC in Section B.1.3.4 of the CS.

The description of the current treatment pathway presented in the CS is broadly aligned with feedback from the ERG's clinical experts and a guidance algorithm of the NHS England (NSHE).[19] However, the CS lacks nuance in certain areas of the pathway and the use of tofacitinib within the NHS is underrepresented. Clinical advice to the ERG indicated that treatment of UC is highly individualised with factors such as comorbidities, contraindications and patient preference all relevant in establishing the optimal treatment for each patient. The CS largely describes the most common pathway and though it acknowledges the individualised nature of advanced treatment, it fails to account for even the most common patient-specific variations.

The CS suggested that CvTs are typically used as the first line treatment in UC patients. While this is the case in the majority of outpatients with UC, with most patients receiving CvT in the first line and moving on to their first biological treatment following relapse or contraindication, there are a small minority who receive a biological treatment in the first line. Clinical advice to the ERG indicated that, within this minority of patients, either vedolizumab or tofacitinib are most commonly used. The company did not consider that patients for whom CvT had failed would progress to treatment with another CvT non-concurrent with active treatment. As CvT was included as a comparator in TA633,[20] and TA547,[21] the ERG sought clinical advice which confirmed that, in UK clinical practice, CvT is not a relevant comparator in participants for whom CvTs have failed. The ERG therefore agreed with the company's exclusion of CvT as a comparator.

The company also indicated in the CS that the majority of patients will receive a TNFi as the first biological treatment. Clinical advice to the ERG concurred with this, stating that TNFis are unsuitable in 10-20 per cent of patients. The CS suggested that, in the minority of patients who do not receive a TNFi as the first biologic, vedolizumab will be prescribed. It does not, however, consider the use of tofacitinib despite its rapid mode of action and the convenience of

administration. Following relapse, the fast action of tofacitinib can reduce the need for steroid treatment in the interim and, being an oral treatment, it is often preferred by patients. Clinical advice to the ERG confirmed that the use of tofacitinib as a first treatment following CvT failure is increasing for the reasons described here.

Furthermore, the CS does not consider the use of a second TNFi following failure of a first TNFi to be routine practice. Clinical opinion to the ERG advised that standard practice is more nuanced than this simplified pathway. The CS accurately identified that TNFi therapeutic drug monitoring (TDM) is used to rationalise decision-making in this regard, with monitoring allowing clinicians to determine the reason for failure on a TNFi. Therefore, if discontinuation is due to immunogenicity, patients are eligible to receive a second TNFi. Clinical advice to the ERG indicated that TDM is much more commonplace in the UK than the US and, as a result, treatment with a second TNFi is more prevalent in the UK, though still fairly uncommon. Furthermore, the ERG noted the inclusion of adalimumab as a comparator in the subgroup of patients who had experienced failure of a biologic treatment in the appraisal of ustekinumab (TA633),[20] indicating that this TNFi is used in the second line. Clinical advice to the ERG further confirmed that, while uncommon, patients receiving a second TNFi would receive adalimumab following treatment failure with infliximab.

As stated previously, the CS does not consider tofacitinib a comparator in either the TNFi-naïve or -experienced patients. As a result of this exclusion, only vedolizumab and ustekinumab are considered relevant comparators following TNFi failure in the CS. The clinicians consulted by the ERG felt that the exclusion of tofacitinib from both subgroups did not reflect clinical practice. The CS states that tofacitinib is not routinely used in TNFi-experienced patients due to its safety profile, citing concerns raised in TA633.[20] The ERG noted, however, that tofacitinib was included as a comparator in TA633 to enable a full picture of cost-effectiveness for ustekinumab; though it further noted that the committee for this appraisal agreed with its subsequent exclusion as a comparator. Clinical experts consulted by the ERG did acknowledge the complex safety profile associated with tofacitinib, but also indicated that the treatment can be very beneficial for some patients. The CS refers to safety warnings regarding tofacitinib from the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA). Monitoring of safety warnings regarding tofacitinib by the Medicines and Healthcare products Regulatory Agency (MHRA) were also referenced by the company in its response to ERG clarification question B.9. Clinical advice to the ERG mentioned the more conservative approach to the safety of tofacitinib in the US and, notably, that use of tofacitinib is increasing in the UK,

driven largely by patients' preference for an oral treatment and its fast-acting nature, and estimated the use of tofacitinib to be approximately 5% in the first line and 25% in the second line in the Royal Devon and Exeter NHS Foundation Trust. Based on these factors, the ERG considered the treatment landscape for people with moderately to severely active UC to be changing, and the exclusion of tofacitinib as a comparator in either subgroup to be an area of particular uncertainty. Finally, real-world evidence from a recent multicentre UK cohort study[22] reported that adverse events requiring curtailment of the treatment were uncommon in the studied population, with no occurrence of thromboembolic events; the authors concluded that tofacitinib was well-tolerated. The ERG therefore requested during clarification that tofacitinib be added to the comparative cost-effectiveness evidence through its inclusion in the model; the company maintained its position regarding safety and opted not to include tofacitinib (company clarification response, question B9). Clinical advice to the ERG did not agree with the company's argument that tofacitinib would not be used in patients over the age of 65 years, those who are past or current smokers, or those who have cardiovascular or malignancy factors, instead indicating that tofacitinib may be offered to such patients following patient-involved decision-making.

Though golimumab is excluded from the decision problem table presented in the CS (Document B, Table 1, p.12), the ERG noted the inclusion of golimumab as a comparator for both TNFi-naïve and TNFi-experienced patients in the NMA and the company model. Clinical advice to the ERG indicates that, while golimumab is currently used in practice, its use is extremely limited. In most cases where it is currently used, it is predominantly in patients with comorbidities for which golimumab is appropriate. The ERG, therefore, agreed with the company that the use of golimumab is limited in UK practice, though it considered its inclusion in the NMA and model for the sake of completeness to be appropriate.

Finally, at the time of the appraisal, filgotinib (GID-TA10600) was under appraisal by NICE as a treatment for patients with moderately to severely active UC who have had an inadequate response, loss of response or were intolerant to a previous biologic agent or conventional therapy. It was not clear where in the treatment pathway filgotinib would be positioned if recommended.

2.2.2. The technology

The CS provided an overview of the mechanism and dosage of ozanimod (Zeposia®) in Section B.1.2; the company also presented the proposed positioning of the treatment in clinical practice in Section B.1.3.4.1 of the CS.

Ozanimod is a sphingosine 1-phosphate (S1P) receptor modulator that is hypothesised to sequester lymphocytes in lymph nodes by binding with high affinity to G protein-coupled S1P receptors 1 and 5 ($S1P_1$ and $S1P_5$).[23,24] Through this receptor binding, it is thought to prevent lymphocyte trafficking via the periphery to, inter alia, the intestine; thereby inhibiting inflammation of the area.[25] As a result of binding to $S1P_1$, also found in cardiac muscle and smooth arterial muscle tissue, ozanimod may have safety considerations related to the heart, in particular bradyarrythmias, as well as blood pressure increases.[25] Its affinity for $S1P_1$ also increases the risk of macular oedema, though this mechanism is more poorly understood. Further safety considerations include increased susceptibility to infections, related to the sequestration of lymphocytes; transient increases in liver enzymes; reduction in in forced expiratory volume; possible foetal harm and, in rare cases, posterior reversible encephalopathy syndrome (PRES).[25] The latter, however, was reported in a patient treated with ozanimod for multiple sclerosis.[26]

The company indicated in the CS that ozanimod is an orally administered medication taken at a dose of 1 mg daily, following an up-titration regimen of 0.25 mg on Days 1 to 4, 0.5 mg on Days 5 to 7, and a 1 mg maintenance dose thereafter. This is in line with the Summary of Product Characteristics (SmPC) for the treatment.

In the CS, the company proposed that ozanimod may be used to treat people with moderately or severely active UC, whether they had prior exposure to TNFis or not. As the line of treatment for the target population is highly individualised, the appropriate positioning of this treatment is dependent on the clinician's perspective on its efficacy and safety relative to comparators, as well as the patient's preference. In this regard, the company indicated that ozanimod satisfies an unmet need through its novel mechanism of action, tolerable safety profile and oral route of administration. Given the individualised nature of the treatment landscape for this condition, the ERG did not consider there to be a fixed position for ozanimod in the treatment pathway.

2.3. Critique of company's definition of decision problem

The company statement regarding the decision problem is presented in Section B.1.1 of the CS. The company position and the ERG response is provided in Table 7 below.

Table 7: Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	Adults with moderately to severely active UC who are intolerant of, or whose disease has had an inadequate response, or loss of response to previous biologic therapy (a tumour necrosis factor-alpha inhibitor TNFi, ustekinumab or vedolizumab), a JAK inhibitor (tofacitinib), or CvT (oral corticosteroids and/or immunomodulators)	Adults with moderately to severely active UC who have had an inadequate response, lost response, or were intolerant to either CvT or a biologic agent This comprises two mutually exclusive sub-populations: •TNFi-naïve: patients who have not previously received a TNFi •TNFi-experienced: patients who have previously received a TNFi and experienced treatment failure due to intolerance, lack of treatment efficacy or loss of response	The population addressed in the submission is in line with the final scope. TNFis are typically used as the first biologic treatment in patients who are intolerant or have had an inadequate response, or loss of response to CvT.1 As a result, exposure to TNFis forms the basis for clinical decision-making, with treatment options differing in two distinct sub-populations: TNFi-naïve and TNFi-experienced. This is reflected in the NICE restriction on the use of ustekinumab and is in line with the current use of other biologic treatments in UK clinical practice.[1]	The ERG considered the overall population included in the company scope to be broadly appropriate. While the ERG agreed with the company's decision to stratify its analyses by subpopulations related to treatment experience, it considered the stratification to be inconsistent with the NICE scope in that TNFi experience does not provide an absolute distinction between the first and second line following CvT.
Intervention	Ozanimod	Ozanimod[a]	N/A	The intervention in the company's main trial, TRUENORTH,[27,28] matches the scope and licence for ozanimod. The company's phase 2 trial compared the licensed dose of 1 mg daily with a lower dose; the ERG appraisal of this trial is restricted to the licensed dose.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Comparator (s)	Current clinical management including: • TNFi (infliximab, adalimumab and golimumab) • Vedolizumab • Ustekinumab • Tofacitinib • Conventional therapies (aminosalicylates, oral corticosteroids and/or immunomodulators), without biological treatments	The submission population has been split into two distinct subpopulations: TNFi-naïve and TNFi-experienced. The relevant comparators differ in these two populations: •TNFi-naive: o Infliximab (and associated biosimilars) o Adalimumab (and associated biosimilars) o Vedolizumab •TNFi-experienced: o Vedolizumab o Ustekinumab	The SmPC for ozanimod states that patients must have failed CvT or a biologic. As biologics are only offered after failure on CvT in clinical practice, CvT is not viewed as a relevant comparator to ozanimod in either population. **TNFi-naïve:** • Following failure with CvT the majority of patients are initially treated with TNFis • As a result, whilst the NICE recommendation for vedolizumab and tofacitinib do not restrict their use in patients who have failed, cannot tolerate or are unsuitable for TNFis, neither tofacitinib nor vedolizumab are typically used first line in TNFi-naïve patients. This was supported by feedback received from clinical consultation conducted as part of this appraisal • TNFis are not suitable for all patients and vedolizumab may be used in a small proportion of TNFi-naïve patients who are contraindicated to TNFis or have specific safety concerns surrounding their use • TNFis and vedolizumab have therefore been considered as relevant comparators in the TNFi-naïve population	The ERG noted that the comparators included in the submission were not consistent with the NICE final scope. The ERG considered the exclusion of tofacitinib as a comparator from both the TNFi-naïve and –experienced subgroups, as an outstanding area of uncertainty, and misaligned with UK clinical practice. The ERG further noted the exclusion of TNFis in the TNFi-experienced subgroup, though clinical advice to the ERG indicated that within-class treatment switching does occur if TNFi failure is due to immunogenicity. The exclusion of adalimumab was of particular concern as it was included as a comparator in the biologic failure subgroup in TA633,[20] indicating that one TNFi may be prescribed following the failure of another. This was confirmed by clinical advice to the ERG. The company excluded CvT as a comparator. Based on clinical advice to the ERG, this was considered to be an appropriate exclusion for the target population. The ERG accepted that the exclusion of golimumab from the company's scope was an omission made in error, due to clarification

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
			TNFi-experienced: • In line with the NICE final scope both ustekinumab and vedolizumab were considered relevant comparators in the TNFi-experienced populations • Neither tofacitinib or TNFis were considered relevant comparators in the TNFi-experienced population • Tofacitinib was not viewed as a relevant comparator as, in line with the opinion of clinicians consulted in TA633,[20] feedback from clinical consultation received as part of this appraisal noted that whilst tofacitinib may be effective for some patients, concerns regarding its safety profile mean it is not typically used as a first line treatment option in TNFi-experienced patients. There has been no downgrading in the EMA warnings and restrictions associated with tofacitinib since the ustekinumab submission.2 The restricted use of tofacitinib combined with concerns of its safety profile negates it as a standard comparator to ozanimod in this population (Section B.1.3.4) • TNFis were not considered relevant comparators in the TNFi-experienced population as TNFi switching is no longer routine clinical practice. As a result, receiving a second TNFi is only clinically relevant in a small	and its inclusion in both the NMA and as a comparator in the economics.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
			proportion of TNFi-experienced patients. The exclusion of TNFis is in line with the accepted assumption in TA633.[20]	
Outcomes	Outcome measures include: • Mortality • Measures of disease activity • Rates of and duration of response, relapse and remission • Rates of hospitalisation • Rates of surgical intervention • Endoscopic healing • Mucosal healing (combined endoscopic and histological healing) • Corticosteroid-free remission • Adverse effects of treatment • Health-related quality of life	Outcome measures include: • Measures of disease activity; change in the 3-component Mayo score • Rates of and duration of response, relapse and remission • Endoscopic healing • Mucosal healing (combined endoscopic and histological healing) • Corticosteroid-free remission • Adverse effects of treatment • Health-related quality of life	Mortality, rates of hospitalisation and rates of surgical intervention were not primary or secondary endpoints in TRUENORTH.[27,28] Data were therefore only collected on these events when assessing adverse events.	The outcomes reported by the company for the trial TRUENORTH[27,28] are relevant to the NICE scope, and clinically meaningful for evaluating the efficacy of treatments for UC. The ERG noted the omission of mortality, rates of hospitalisation and rates of surgical intervention as primary or secondary outcomes from the company scope. Clinical advice to the ERG indicated that mortality and rates of hospitalisation are broadly invariant with respect treatment with biologics or small molecules. Based on further clinician input, rates of surgery are likely similarly unchanged, though uncertainty remains as to whether the use of these treatments may result in a reduction in surgery. The ERG considered that this outcome could have been included in the NMA and used subsequently in the economic modelling, given its importance in the treatment pathway and disease course. However, the ERG acknowledges that rates of surgery would likely be little changed between treatments and are unlikely to impact the base case ICER, given that the key drivers of

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
				cost-effectiveness relate to treatment effect which, to some extent, incorporate surgical outcomes. The economic model captured treatment discontinuation, disutility and costs associated with serious adverse events only. HRQoL data were included in the economic model. The ERG noted that HRQoL data were collected in the pivotal TRUENORTH[27,28] trial using the EQ-5D-5L instrument; however, these data were not used in the company's base case. Instead, the company used published literature (from Woehl et al.[29] and Arsenau et al.[30]) and assumption to derive health state utility values.
Economic analysis	• The cost-effectiveness of treatments is expressed in terms of incremental cost per quality-adjusted life year • The time horizon for estimating cost-effectiveness was set at a lifetime horizon to sufficiently reflect any differences in costs or outcomes between the technologies being compared	As per final scope and NICE reference case	In line with the NICE final scope	The company submitted a cost utility analysis which used ICERs and QALYs as appropriate. A lifetime horizon was used in the base case. The ERG considered this to be reasonable (see Section 4.2.5). Costs were considered from an NHS and Personal Social Services perspective, in line with NICE guidance. Overall, the ERG considered that the economic analysis provided by the company was aligned with NICE's preferred reference case with respect to time horizon, perspective and outcomes.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	• Costs are considered from a NHS and Personal Social Services perspective • The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account			
Subgroups	If the evidence allows the following subgroups will be considered: • People who have been previously treated with one or more biologic • People who have not received prior biologic therapy	Clinical consultation conducted as part of this appraisal indicated that exposure to TNFis forms the basis for clinical decision-making, with treatment options differing in two distinct sub-populations: • TNFi-naïve • TNFi-experienced	• Economic analyses were conducted for ozanimod for sub-populations based on prior TNFi exposure owing to the relevant comparators differing between these sub-populations. These analyses informed the base case cost-effectiveness analysis for comparisons versus infliximab, adalimumab, golimumab and vedolizumab (in TNFi-naïve patents) and vedolizumab and ustekinumab (in TNFi-experienced patients) • Subgroup analyses were informed by the Network Meta-analysis (NMA). The efficacy of ozanimod in the NMA was based on the subgroups of TRUENORTH stratified by TNFi exposure.	In the economic analysis, results have been presented for two distinct subgroups i.e., TNFi-naïve and -experienced patients. The ERG noted that final scope issued by NICE stated that subgroups should be stratified according to those who have been treated previously with biologics and those who have not received biologic treatment. The ERG noted that the previous UC appraisal for ustekinumab (TA633) was for a treatment licensed for patients with moderately to severely active UC who had an inadequate response or lost response to previous biologic therapy.
Special considerations including	None	The company did not identify any equity or equality concerns in the scope	N/A	The ERG agreed that there are no equity or equality concerns to be considered in this appraisal.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
issues related to equity or equality				

Abbreviations: cPAS, confidential patient access scheme; CvT, conventional therapy; EMA, European Medicines Agency; EQ-5D-5L, European Quality of Life Five Dimension Five Level; ERG, Evidence Review Group; HRQoL, health-related quality of life; ICER, incremental cost-effectiveness ratio; JAK, Janus kinase; N/A, not applicable; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; NMA, network meta-analysis; SmPC, Summary of Product Characteristics; TA, technology appraisal; TDM, therapeutic drug monitoring; TNFi(s), tumour necrosis factor inhibitor(s); UC, ulcerative colitis

Note: [a] Ozanimod presents in three distinct capsule strengths each with two reportable weights (ozanimod hydrochloride 0.25 mg,0.50 mg, and 1.0 mg, which are equivalent to ozanimod 0.23 mg, 0.46 mg, and 0.92 mg, respectively).

3. CLINICAL EFFECTIVENESS

3.1. Critique of the methods of review(s)

The Company undertook a systematic literature review (SLR) to identify randomised controlled trials (RCTs) providing evidence for ozanimod (summarised in Section 3.2) and comparators to ozanimod. These were used to inform their indirect treatment comparison (Sections 3.3 and 3.4) in people with moderately to severely active UC. An additional SLR was conducted to identify any non-randomised trials of ozanimod, but yielded no results. An overview of the methods used in the SLRs is provided in Table 8 below.

Table 8: Summary of ERG's critique of the methods implemented by the company to identify evidence relevant to the decision problem

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix D.2.1.	The searches of bibliographic databases and grey literature sources are considered broadly appropriate; however, the ERG noted that specific searches for adverse reactions were not conducted. The search methods for the additional SLR to identify non-randomised trials were provided in response to clarification question A1.
Inclusion criteria	Appendix D.2.2., Table 7 (pages 50-51)	The inclusion criteria for the clinical effectiveness review are considered broadly appropriate to the decision problem. The ERG noted that the subgroup 'biologic treatment-failure and biological treatment non-failure with and without prior corticosteroid use' is fully aligned with the population detailed in the NICE scope, but not the company scope defining subgroups by TFNi experience; as highlighted in Table 7. The ERG further noted the inclusion of certolizumab as a comparator, though this treatment is listed in neither the NICE scope nor the decision problem addressed by the CS. The ERG accepted the company's clarification that certolizumab was included in error. The ERG disagreed with the company's decision to exclude phase 4 trials from the NMA, though it notes that such trial data are not currently available and therefore did not investigate this further.
Screening	Appendix D.2.2. (page 49)	Screening was conducted to appropriate standards to minimise selection bias, with duplicate screening at title/abstract and full-text stages. Arbitration by a third reviewer is also described, though the ERG noted that it was

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		not explicitly stated whether this was done at both screening stages.
Data extraction	Appendix D.2.3. (page 51)	Data extraction was conducted to appropriate standards to minimise selection bias, with single reviewer extractions validated by a second reviewer. Though data extraction was not done independently and in duplicate, the ERG noted that data validation by a second reviewer is permissible with the AMSTAR 2 critical appraisal tool,[31] and further concluded that arbitration conducted by a third reviewer, if necessary, would minimise potential error or bias.
Tool for quality assessment of included study or studies	Document B, Section B.2.4., Table 19 (page 60); Appendix D.2.3. (page 51), Appendix D.4.4 (page 129), Appendix D.6. (pages 169-170)	The risk of bias assessment of TRUENORTH[27,28] in Document B of the CS was reported according to the Centre for Reviews and Dissemination (2009)[32] tool. The tool was also used to assess the risk of bias of all RCTs included in the company's NMA. The ERG considered this method appropriate, though it noted that the updated Cochrane risk of bias 2 tool[33] is generally preferred. No risk of bias assessment was reported for the long-term trial extension to TRUENORTH. The ERG considered this acceptable, given the ongoing nature of this trial.
Evidence synthesis	Appendix D.4.1. (pages 122-123), Appendix D.4.2. (pages 124-126), Appendix 4.3. (pages 127-129)	No synthesis of trials investigating ozanimod was conducted, as there is only one trial per comparison available. The ERG considered this reasonable. The company conducted several NMAs to evaluate the comparative efficacy of ozanimod with other available treatments within the TNFi-naïve and –experienced subgroups; these were further stratified by induction and maintenance phases for each subgroup. The ERG considered that further outcomes, particularly adverse events or treatment discontinuations, could have been evaluated in the NMAs, however, the company did not report their feasibility assessment with regards to outcomes and therefore it is not possible to determine if these outcomes were considered but found not feasible for analysis. The methods used in the NMAs were appropriate, though the ERG highlighted concerns about heterogeneity in the networks and the paucity of evidence, which both contributed to uncertainty in the results.

Abbreviations: AMSTAR, A MeaSurement Tool to Assess systematic Reviews; CRD, Centre for Reviews and Dissemination; CS, Company submission; ERG, Evidence Review Group; NMA, network meta-analysis; TNFi, tumour necrosis factor inhibitor

3.2. Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

The company presented evidence for ozanimod from one Phase 3 RCT (TRUENORTH) including two cohorts: one placebo-controlled cohort and a second enrichment cohort; with responders from both cohorts re-randomised following an induction period. Further evidence came from a key supporting phase 2 placebo-controlled dose-ranging RCT in participants with endoscopically-confirmed UC (TOUCHSTONE). An overview of the methods used in these studies is presented in Sections 3.2.1 to 3.2.5.

3.2.1. Study design

The company's primary evidence for ozanimod is derived from TRUENORTH,[27,28] a multicenter, phase 3 study with a 10-week induction phase followed by a 42-week maintenance phase. The trial enrolled a total of 1,012 participants: some had no prior experience to TNFi; others had been treated with TNFi before. Eligible participants were either randomised in a 1:2 ratio to placebo or 1 mg ozanimod (called 'cohort 1') or included in an open-label enrichment cohort which was also allocated 1 mg ozanimod ('cohort 2') during the induction phase. Following induction, responders to ozanimod from both cohorts were re-randomised to receive placebo or 1 mg ozanimod during the maintenance phase, while responders to placebo continued placebo in the maintenance phase. Induction non-responders in both arms, as well as those who had relapsed during the maintenance phase, had the option of entering the open-label extension (OLE) trial. The trial measured a broad range of clinical efficacy, quality of life and safety outcomes up to 52 weeks and the ERG considered the large trial to be well conducted, though some methodological concerns that could bias results are described in Section 3.2.4.3. Despite these concerns, the ERG agreed with the company's decision to use data from this trial as the primary clinical effectiveness evidence.

The key supporting trial, TOUCHSTONE,[34] is a multicenter, phase 2 dose-finding study with an 8-week induction and 24-week maintenance phase. The study enrolled 199 participants who were randomised to receive 0.5 mg ozanimod, 1 mg ozanimod or placebo. Participants with clinical improvement during the induction phase continued their blinded regimen during maintenance; induction non-responders and those who relapsed during the maintenance phase had the option of entering the OLE trial. The ERG agreed that data from this trial are suitable as supporting evidence, given that the trial was well conducted and reported on outcomes within the NICE scope of this appraisal.

Following the completion of either TOUCHSTONE or TRUENORTH, participants had the option of continuing in the single-arm 1 mg ozanimod OLE trial. The study had, at the time of writing, reached its primary completion date and is expected to report maximal follow-up to six years, with its primary outcomes related to the safety of ozanimod. The ERG is of the opinion that the long-term safety evidence from this trial would reduce the uncertainty around the safety of ozanimod for moderately to severely active UC. An overview of the trial designs in provided in Table 9.

Table 9: Overview of ozanimod trial designs

Study name and acronym	Study design	Phase	Intervention / Comparator	Study objectives	Population
TRUENORTH (NCT02435992)	Multicentre, placebo-controlled study. 1-week dose titration within a 10-week induction. Induction period had 2 cohorts, one randomised and double blind and one open label enrichment cohort. Responders to ozanimod re-randomised to 42-week maintenance period.	3	1 mg ozanimod hydrochloride daily / placebo	Safety and efficacy	N=1012. Adults aged 18 to 75 with moderately to severely active UC. (N=526 for maintenance period)
TOUCHSTONE/ TRUENORTH OLE – Ongoing (NCT02531126)	Multicenter, single group assignment, OLE.	3	1 mg ozanimod	Safety and efficacy	N=878. Adults aged 18 to 75 who had participated in either TRUENORTH or TOUCHSTONE.
TOUCHSTONE (NCT01647516)	Multicentre, randomised, double-blind, placebo-controlled study. 1-week dose titration within a 9-week induction. Responders re-randomised to	2	1 mg ozanimod / 0.5 mg ozanimod / placebo	Safety and efficacy	N=199. Adults ages 18 to 75 with moderately to severely active UC.

Study name and acronym	Study design	Phase	Intervention / Comparator	Study objectives	Population
	24-week maintenance period.				

Abbreviations: OLE, open-label extension; UC, ulcerative colitis

3.2.2. Trial populations

3.2.2.1. Eligibility criteria

Key inclusion and exclusion criteria used in the two included trials and their open-label extension are summarised in Table 10 below. Potential participants were identified through endoscopically confirmed UC of moderate to severe activity, defined by a Mayo score of 6 to 12. Participants in TRUENORTH[27,28] were additionally required to receive aminosalicylate or corticosteroids and could have had prior treatment with immunosuppressants, though the use of these needed to be stopped prior to randomisation. The ERG considered the age, definition of the condition and other inclusion criteria to be appropriate for the target population.

The TRUENORTH trial[27,28] excluded potential participants with a physician-judged likelihood of receiving colectomy or ileostomy within 12 weeks of baseline, recent evidence of serious UC symptoms, diagnosis of Crohn's disease (CD) or other types of colitis, cardiovascular (CV) conditions, or a history of certain eye conditions; as well as excluding participants who are pregnant or lactating. Potential participants in TOUCHSTONE[34] were excluded for current use of TNFis. As there were no explicit exclusion criteria related to TNFi experience, the ERG agreed that both TNFi-naïve and -experienced participants would be included in the trial populations and that data from these trials align with the proposed positioning of ozanimod as a first- or second-line treatment.

For the open-label extension, all participants who participated in either TOUCHSTONE[34] or TRUENORTH[27,28] were eligible for inclusion. For this long-term extension, participants were excluded if they were treated with breast cancer resistance protein inhibitors, had clinically relevant CV conditions, or had liver function impairment. The ERG considered these exclusions appropriate given the long-term safety concerns of S1P receptor modulators, but noted that this would limit the generalisability of any conclusions on the safety of ozanimod, though it acknowledged that participants with similar conditions would likely not be prescribed ozanimod in UK practice.

Table 10: Eligibility for the included trials

Study	Inclusion criteria	Exclusion criteria
TRUENORTH (NCT02435992)	Aged 18 to 75 years (at screening for Cohort 1 and 2) UC confirmed on endoscopy Moderately to severely active UC (Mayo score 6-12) Currently receiving treatment with aminosalisylate, prednisone, or budesonide Can be receiving azathioprine, mercaptopurine, or methotrexate, but treatment will be stopped prior to randomisation	Physician judgment that the patient is likely to require colectomy or ileostomy within 12 weeks of baseline. Current or recent (within 3 months) evidence of fulminant colitis, toxic megacolon, or bowel perforation. Diagnosis of CD, indeterminate colitis, or the presence of fistula consistent with CD or microscopic colitis, radiation colitis, or ischemic colitis Clinically relevant cardiovascular conditions or other relevant diseases that could impact the implementation or interpretation of the trial, or put the patient at risk History of uveitis or unknown macular edema Pregnancy, lactation, or a positive serum beta-human chorionic gonadotropin (β-hCG) measured during screening
TOUCHSTONE (NCT01647516)	18 Years to 75 Years UC confirmed on endoscopy Moderately to severely active UC (Mayo score 6-12)	Current use of anti-TNF agents
OLE (NCT02531126)	Aged 18 to 75 years Previously participated in a trial of ozanimod and meets the criteria for participation in the open-label extension as outlined in the prior trial (i.e. non-responders after induction or relapse/completion of maintenance phase)	Receiving treatment with breast cancer resistance protein inhibitors Clinically relevant cardiovascular conditions Liver function impairment

Abbreviations: β-hCG, beta-human chorionic gonadotropin; CD, Crohn's disease; OLE, open-label extension; TNF, tumour necrosis factor; UC, ulcerative colitis

3.2.2.2. Baseline characteristics

The baseline characteristics of the participants in the TOUCHSTONE[34] and TRUENORTH[27,28] trials are presented in Table 11, alongside comparative characteristics from a cross-sectional, retrospective UK cohort dataset presented by King et al. (2020).[35]

Table 11: Baseline characteristics of the intention-to-treat populations of the included trials at induction, and their comparability with a cross-sectional and retrospective UK cohort study[35]

Characteristic	TRUENORTH			TOUCHSTONE			King et al.[35]
	Ozanimod (cohort 1)	Placebo	Ozanimod (cohort 2)	Placebo	Ozanimod 0.5 mg	Ozanimod 1 mg	
Mean age (years) (SD)	41.4 (13.54)	41.9 (13.64)	42.1 (13.72)	41.9 (12.3)	38.8 (12.1)	41.8 (11.0)	51 (37-65)[a]
Female	42.9%	33.8%	41.7%	46%	51%	28%	50.1%
White race	[redacted]	[redacted]	[redacted]	94%	91%	93%	
Mean weight (kg) (SD)	[redacted]	[redacted]	[redacted]	72.6 (14.9)	72.3 (16.9)	77.4 (16.3)	
BMI (kg/m^2) (SD)	25.40 (5.492)	25.11 (4.477)	25.88 (5.796)	NR	NR	NR	<25 – 38.49% 25-30 – 28.21% >30 – 14.55% Unknown – 18.76%
Tobacco/nicotine use	[redacted]	[redacted]	[redacted]	Current – 5%	Current – 6%	Current – 6%	Current – 12.33% Former – 26.36% Never – 54.22% Unknown – 7.08%
Region	[redacted]	[redacted]	[redacted]	NR	NR	NR	

Characteristic	TRUENORTH			TOUCHSTONE			King et al.[35]
4-component Mayo score	8.9 (1.47)	8.9 (1.35)	9.1 (1.49)	8.6 (1.5)	8.3 (1.5)	8.5 (1.6)	
Median C-reactive protein (mg/L) (range)	4.0 ([redacted])	5.0 ([redacted])	5.0 ([redacted])	4.9 (0.20-141.4)	3.9 (0.10-131.2)	4.3 (0.10-82.5)	
Median faecal calprotectin (µg/g) (range)	1079.48 ([redacted])	1349.79 ([redacted])	1259.85 ([redacted])	1272 (30-8380)	1477 (66-11,108)	1238 (10-10,511)	
Median lactoferrin (µg/g) (range)	NR	NR	NR	29.0 (1.4-1049)	30.6 (1.4-483)	29.9 (1.4-586)	
Disease extent	Left side of colon – 62.5% Extensive – 37.5%	Left side of colon – 62.0% Extensive – 38.0%	Left side of colon – 64.6% Extensive – 35.4 %	Left side of colon – 63% Extensive – 37%	Left side of colon – 63% Extensive – 37%	Left side of colon – 61% Extensive – 39%	
Concomitant medication	Glucocorticoid – 27.7% Aminosalicylate – 87.2%	Glucocorticoid – 32.4% Aminosalicylate – 84.3%	Glucocorticoid – 33.8% Aminosalicylate – 85.8%	Glucocorticoid – 37% Aminosalicylate – 88%	Glucocorticoid – 34% Aminosalicylate – 82%	Glucocorticoid – 40% Aminosalicylate – 79%	
Previous medication	TNFi – 30.3%	TNFi – 30.1%	TNFi – 43.4%	Immunosuppressant – 26% TNFi – 15%	Immunosuppressant – 37% TNFi – 20%	Immunosuppressant – 33% TNFi – 19%	
Mean age at onset/diagnosis (years) (SD)	[redacted]	[redacted]	[redacted]	35.8 (13.0)	33.1 (11.3)	35.2 (12.1)	
Mean years since diagnosis (SD)	6.9 (6.61)	6.8 (7.04)	7.91 (7.365)	6.1 (5.5)	5.9 (5.4)	6.7 (6.8)	

Abbreviations: BMI, body mass index; NR, not reported; SD, standard deviation; TNFi, tumour necrosis factor inhibitor

Note: [a] King et al. only reported the median age with interquartile range, as shown

Comparability of trial arms

As shown in Table 11, the baseline characteristics of the participants included in the ITT populations of the TRUENORTH[34] and TOUCHSTONE[34] trials were balanced across trial arms. The ERG noted that randomisation had been stratified by prior corticosteroid use and prior TNFi exposure. The company did not provide baseline characteristics per trial arm for the corticosteroid use stratum; these were, however, reported separately by TNFi experience in the CS (Document B, Table 15, p.54). Demographic and anthropometric characteristics between the placebo and ozanimod trial arms were comparable, though the ERG noted slightly higher proportions of male participants in the placebo arms of both the TNFi-naïve and –experienced strata. The ERG further noted a higher proportion of participants from Europe, as well as lower proportions of participants from other regions and with extensive disease in the ozanimod Cohort 2, when compared with the ozanimod Cohort 1 and placebo arms.

Relevance of trial populations to the target population

The overall characteristics of participants in the trials appear broadly comparable with those of the UK cohort dataset, with the only exceptions being that the average age of participants in the UK dataset is approximately 10 years older than in the ozanimod trials, and that there is a smaller proportion of current smokers included in the trials. The comparative data available for the relevant population in the UK are limited, however, and no comparisons are possible for a number of baseline characteristics; in particular across the range of biomarkers reported in the ozanimod trials. The ERG acknowledges that such unknown imbalances in respect of the UK population may exist in the trial populations, but considered the comparability between demographic and anthropometric characteristics to be reassuring. In addition, consultation with clinical experts indicated that the populations in TRUENORTH[27,28] and TOUCHSTONE[34] broadly reflect the characteristics of people with moderate to severe UC in the UK.

No comparative characteristics for the TNFi-naïve and –experienced strata could be found in published literature. The generalisability of evidence from these subgroups to the corresponding UK populations is an area of uncertainty in this appraisal.

3.2.3. Intervention characteristics

The characteristics of interventions delivered during the TOUCHSTONE[34] and TRUENORTH[27,28] trials, as well as their open-label extension, are summarised in

Table 12. Ozanimod is delivered through oral administration of 1 mg (0.93 mg ozanimod hydrochloride) capsules once a day, following an up-titration regimen of 0.25 mg for the first four days, 0.5 mg for Days 5 to 7 and 1 mg thereafter. The company did not provide an explicit rationale for the up-titration of ozanimod in the CS, but the ERG noted that this approach is identical to treatment described in patients with multiple sclerosis, with the rationale being attenuation of first-dose heart rate and atrioventricular conduction effects.[36] Furthermore, up-titration seems to be associated with the use of S1P modulators in general. As a result, the ERG considered this step to be appropriate.

A dose of 1 mg daily was selected following the completion of the phase 2 TOUCHSTONE[34] dose-finding trial, which also evaluated a lower maintenance dose (0.5 mg daily) of ozanimod and found slightly higher occurrences of clinical remission and clinical response and lowered lymphocyte counts with the higher dose.[34] No reductions or increases in dose, with the exception of up-titration, were permitted during TRUENORTH or indicated in the license for ozanimod.

The ERG noted that the use of concomitant treatment in TRUENORTH[27,28] (Document B, Table 14, p.52) and TOUCHSTONE[34] (Appendix L.1.3., Table 73, p.297) was balanced between the ozanimod cohort 1, ozanimod cohort 2 and placebo arms for TRUENORTH as well as the between the ozanimod 0.5 mg, ozanimod 1 mg and placebo arms for TOUCHSTONE[34] at induction. The most commonly used concomitant medication was aminosalicylates, followed by glucocorticoids and immunomodulators (reported for TRUENORTH only). The ERG further noted that the placebo responders arm during the maintenance phase of TRUENORTH comprised a far larger proportion of participants on aminosalicylates and lower proportion of participants on immunomodulators compared to those who has received ozanimod during induction: this was considered to be a function of the type of CvTs placebo responders were receiving at the time of response.

Table 12: Intervention characteristics of the included trials

Trial		Treatment
TRUENORTH induction (Cohort 1)	Ozanimod	Up-titration of 0.25 mg for Days 1 to 4, 0.5 mg for Days 5 to 7 and 1 mg thereafter with daily double-blinded oral administration 10 weeks
	Placebo	Matched double-blind oral placebo administered daily 10 weeks
TRUENORTH induction (Cohort 2)	Ozanimod	Up-titration of 0.25 mg for Days 1 to 4, 0.5 mg for Days 5 to 7 and 1 mg thereafter with daily double-blinded oral administration 10 weeks
TRUENORTH maintenance (re-randomised)	Ozanimod	Dose of 1 mg, daily double-blinded oral administration 42 weeks, up to study duration of 52 weeks
	Placebo	Matched double-blind oral placebo administered daily 42 weeks, up to study duration of 52 weeks
TOUCHSTONE/ TRUENORTH OLE		Dose of 1 mg, daily open-label oral administration Up to 6 years, or upon discontinuation from the sponsor
TOUCHSTONE	Ozanimod 0.5 mg	Up-titration of 0.25 mg for Days 1 to 4, and 0.5 mg from Day 5 onwards with daily double-blinded oral administration 32 weeks
	Ozanimod 1 mg	Up-titration of 0.25 mg for Days 1 to 4, 0.5 mg for Days 5 to 7 and 1 mg thereafter with daily double-blinded oral administration 32 weeks
	Placebo	Matched double-blind oral placebo administered daily 32 weeks

Abbreviation: OLE, open-label extension

3.2.4. Clinical effectiveness results

An overview of the clinical outcomes specified by NICE i.e., whether they were reported in the trials, how they were defined and how they were measured is provided in Section 3.2.4.1, along with limitations of these means of ascertainment captured where necessary.

3.2.4.1. NICE-scoped outcomes

Mortality

Mortality was not assessed as an outcome in TRUENORTH[27,28] or TOUCHSTONE.[34] The company did indicate that mortality is captured to a certain extent in TRUENORTH through its

reporting of adverse events. Though the ERG did not consider adverse events to be sufficiently specific to make any conclusions relating to the effect of ozanimod on mortality in the population of interest, clinical advice to the ERG did confirm that mortality is broadly invariant with respect to treatment with biologics or small molecules. As a result, the ERG did not consider the omission of mortality from the company submission to be highly problematic, though it noted the uncertainty around the effect of ozanimod on this outcome.

Measures of disease activity

Both TOUCHSTONE[34] and TRUENORTH[27,28] report disease activity. TRUENORTH measures disease activity through the three-component Mayo score, consisting of three sub-scores; rectal bleeding, stool frequency and mucosal appearance through endoscopy. By comparison, TOUCHSTONE reports disease activity using the 4-component score however the CS refers to the 3-component score. The primary outcomes of both studies included patients achieving clinical remission according to their three-component Mayo score. In addition, secondary outcomes included change in Mayo score and clinical response according to this score at weeks 8 and 32 in TOUCHSTONE; and weeks 10 and 52 in TRUENORTH. The company indicated that the Mayo scoring system is the most widely used, the ERG noted mention of the Truelove and Witts' severity index system and the UC symptom score (UCSS). Clinical advice to the ERG confirmed that the use of the Mayo scoring system is broadly appropriate.

Rates of and duration of response, relapse and remission

Both the TOUCHSTONE[34] and TRUENORTH[27,28] trials reported the number of patients achieving remission or clinical response to treatment. TOUCHSTONE defined clinical remission as a four-component Mayo score <2, with none of the individual sub-scores >1. This was recorded as a primary endpoint at week 8 and a secondary endpoint at week 32. TRUENORTH defined remission according to both the overall and the sub-scores of the three- and four-component Mayo scores. Clinical remission per the three-component Mayo score was defined as a rectal bleeding sub-score (RBS) of 0, with both the stool frequency sub-score (SFS) and endoscopy sub-score ≤1; clinical remission per the four-component Mayo score was defined the same as for TOUCHSTONE. The trial reported rates of remission at both 10 and 52 weeks as primary outcomes. It also reported the number of patients who remained in remission from week 10 to week 52. The TRUENORTH study also included a further group of those in 'durable clinical remission' at the 52-week time point.

Clinical response was defined in TRUENORTH[27,28] using both the overall and constituent sub-scores of both the three- and four-component Mayo score. The four-component definition was a reduction from baseline in the overall score of ≥3 points and ≥30%, and a reduction from baseline in the RBS of ≥1 point or an absolute RBS of ≤1 point. Similarly, the three-component definition was a reduction from baseline in the overall score of ≥2 and ≥35%, a reduction from baseline in RBS of ≥1 point and an absolute RBS of ≤1 point. In TOUCHSTONE,[34] clinical response was defined as a decrease in Mayo score of ≥3 points and ≥30% and a decrease in RBS of ≥1 point or an absolute RBS of ≤1.

Rates of hospitalisation

The company reported that rates of hospitalisation were not assessed as an outcome in TRUENORTH[27,28] or TOUCHSTONE,[34], though the ERG noted that rates of hospitalisation were listed as an outcome of TRUENORTH in the CS (Document B, Section B.2.2, Table 8). Instead of reporting this outcome, the company indicated that hospitalisation is captured to a certain extent in TRUENORTH through its reporting of adverse events. Though the ERG did not consider adverse events to be sufficiently specific to make any conclusions relating to the effect of ozanimod on hospitalisation in the population of interest, clinical advice to the ERG did confirm that this outcome is broadly invariant with respect treatment with biologics or small molecules. As a result, the ERG did not consider the omission of hospitalisation rates from the company submission to be highly problematic, though it noted the uncertainty around the effect of ozanimod on this outcome.

Rates of surgical intervention

Surgical intervention rates were not assessed as an outcome in TRUENORTH[27,28] or TOUCHSTONE.[34] The company did indicate that surgeries are captured to a certain extent in TRUENORTH through its reporting of adverse events. Though the ERG did not consider adverse events to be sufficiently specific to make any conclusions relating to the effect of ozanimod on surgery rates in patients with moderately to severely active UC, clinical advice to the ERG did indicate that rates of surgery are likely unchanged by treatment with biologics and small molecules. However, some uncertainty remains as to whether the use of these treatments may result in a reduction in surgery. The ERG considered that this outcome could have been included in the NMA and used subsequently in the economic modelling, given its importance in the treatment pathway and disease course.

Endoscopic healing

TOUCHSTONE[34] and TRUENORTH[27,28] both report endoscopic findings in addition to the endoscopic sub-score within the four-component Mayo score. However, the ERG noted that only TRUENORTH pre-specified the percentage of patients with 'endoscopic improvement', also defined as a Mayo endoscopy sub-score of ≤1, at the end of induction and maintenance phases (10 and 52 weeks respectively) as an outcome. The ERG further noted an inconsistency in TOUCHSTONE, where an endoscopy sub-score ≤1 at the end of the induction and maintenance phases (weeks 8 and 32 respectively) was also pre-specified as an outcome, however, this was termed 'mucosal healing'.

Mucosal healing (combined endoscopic and histological healing)

The ERG noted that mucosal healing was defined in TRUENORTH as a combination of the endoscopic healing outcome as well as histological healing, defined as a Geboes score <2.0. The latter is achieved when there are no neutrophils in the epithelial crypts or lamina propria and none of the following: increased eosinophils; crypt destruction; or erosions, ulceration or granulation of the tissue. TRUENORTH[27,28] reported the percentage of patients with mucosal healing at weeks 10 and 52, which was defined as 'endoscopic improvement with histological remission'; this also included a Mayo endoscopy sub-score of ≤1 and Geboes score <2.0. The TOUCHSTONE[34] trial, however, pre-specified endoscopy sub-scores as an outcome, as described in the section above, and called this 'mucosal healing'. In addition, this trial pre-specified 'histological remission', defined as a Geboes score <2.0, suggesting that TOUCHSTONE did not consider histological remission to be a component of mucosal healing. The ERG noted the company's acknowledgement of the stricter definition of mucosal healing in TRUENORTH when compared to other trials in UC.

Corticosteroid-free remission

TRUENORTH[27,28] reported the percentage of patients in corticosteroid-free remission. This was defined as those who had not received corticosteroids more than 12 weeks at week 52 of the trial. The company indicated in the CS that relapse within 12 weeks of corticosteroid discontinuation demonstrates steroid-dependent remission.

Adverse effects of treatment

TOUCHSTONE[34] separately reported the number of patients with treatment emergent adverse events (TEAEs) in the induction period and the maintenance period. A TEAE was classed as any event beginning on or after the first dose or an ongoing event that became more severe after the first dose, or up to 90 days after the last dose. An adverse event (AE) was described as serious if it resulted in death; was life threatening; required hospitalisation or elongation of a hospital stay; caused persistent disability/incapacity; was a congenital anomaly; or constituted an important medical event. The severity of AEs was assessed by the investigator according to their impact of patients' normal activities.

TRUENORTH[27,28] also reported the incidence, severity and relationship between the following TEAEs, serious AEs, TEAEs leading to discontinuation of ozanimod and TEAEs of special interest. In addition, changes from baseline in clinical laboratory measures, vital signs, ECG and pulmonary function tests were measured.

Health-related quality of life

Change in health-related quality of life (HRQoL), from baseline to week 10, was assessed in the TRUENORTH[27,28] study using both the SF-36 and EQ-5D five-level (5L) version, using both a summary index score and the patient's self-rated health status using a graduated visual analogue scale (VAS). The company reported that EQ-5D-5L data were cross-walked to EQ-5D-3L index scores using the algorithm included in van Hout et al. (2012);[37] the weighted average across treatment arms was used to inform health states. The ERG noted that this is the approach preferred by NICE.

3.2.4.2. Trial outcomes

Table 13 and Table 14 list the outcomes measured in the two trials providing the primary and key supporting evidence (TRUENORTH[27,28] and TOUCHSTONE,[34] respectively). Outcomes corresponding to NICE-scoped outcomes are also indicated.

Table 13: Outcomes per treatment phase reported in TRUENORTH

Outcome	NICE-scoped
Proportion in clinical remission; three- and four-component Mayo score (induction)	✓
Proportion with clinical response; three- and four-component Mayo score (induction)	✓
Proportion with endoscopic improvement (induction)	✓ (endoscopic healing)
Proportion with mucosal healing (induction)	✓
Changes in three-, four- and partial Mayo scores (induction)	✓ (measures of disease activity)
Proportion in histologic remission (induction)	× (only as part of mucosal healing)
Proportion with clinical response, clinical remission or endoscopic improvement in patients with prior TNFi experience (induction)	✓
Change in SF-36 and EQ-5D (induction)	✓
Proportion in clinical remission; three- and four-component Mayo score (maintenance)	✓
Proportion with clinical response; three- and four-component Mayo score (maintenance)	✓
Proportion with endoscopic improvement (maintenance)	✓ (endoscopic healing)
Proportion with maintenance of remission (maintenance)	×
Proportion with corticosteroid-free remission (maintenance)	✓
Proportion with mucosal healing (maintenance)	✓
Proportion with durable clinical remission (maintenance)	×
Changes in three-, four- and partial Mayo scores (maintenance)	✓ (measures of disease activity)
Proportion in histologic remission (maintenance)	× (only as part of mucosal healing)
Proportion with clinical response, clinical remission or endoscopic improvement in patients with prior TNFi experience (maintenance)	✓
Change in SF-36 and EQ-5D (maintenance)	✓
Health resource utilisation (maintenance)	×
Work productivity (maintenance)	×

Abbreviations: EQ-5D, European Quality of Life Five Dimension; NICE, National Institute for Health and Care Excellence; SF-36, 36-item Short Form Health Survey

Table 14: Outcomes per treatment phase reported in TOUCHSTONE

Outcome	NICE-scoped
Proportion with clinical response (induction)	✓
Changes in Mayo scores (induction)	✓ (measures of disease activity)
Proportion with mucosal healing (induction)	✓ (different definition to TRUENORTH)
Proportion with TEAE (induction)	✓
Proportion with clinical response (maintenance)	✓
Proportion in clinical remission (maintenance)	✓
Changes in Mayo scores (maintenance)	✓ (measures of disease activity)
Proportion with mucosal healing (maintenance)	✓ (different definition to TRUENORTH)
Proportion with durable clinical remission (maintenance)	×
Proportion with TEAE (maintenance)	✓

Abbreviation: NICE, National Institute for Health and Care Excellence; TEAE, treatment-emergent adverse event

3.2.4.3. Critical appraisal of the design of the studies

The company's approach to the critical appraisal of included trials was reported in the CS (Appendix D.2.3., p.51). The critical appraisal of published evidence from the key supporting trial, i.e. the TOUCHSTONE (Sandborn et al. 2016)[34] study, according to the University of York CRD[32] criteria was reported in Appendix D.6. (p.169). Published and unpublished evidence from the pivotal TRUENORTH study (Sandborn et al. 2021[28] and TRUENORTH CSR,[27] respectively) was also critically appraised using the University of York CRD criteria; the results of this appraisal were reported in Document B, Section B.2.4., Table 19 (p.60). No risk of bias assessment was reported for the long-term trial extension to TRUENORTH, but the ERG considered this acceptable, given the ongoing nature of this trial.

As noted in Table 8, the ERG considered the CRD criteria to be appropriate for the appraisal of these studies, though it noted that the Cochrane risk of bias tool was updated in 2019[33] and that Cochrane risk of bias 2 is generally the preferred tool for appraising risk of bias in RCTs.

TOUCHSTONE

The company appraised this trial as having no methodological concerns, aside from some uncertainty around blinding of providers, participants and outcome assessors. Assessments were made at the trial level and not the individual outcome level. The ERG agreed broadly with

the assessments of the company according to the domains of the tool, though some domains were less obvious than others. The ERG noted that there was no explicit description of allocation concealment and considered 'Unclear' to be a more appropriate judgment than 'Yes' for this domain. While the ERG agreed that participants in the two groups of interest, placebo and 1 mg ozanimod, were similar in terms of prognostic factors it did note some considerable differences; with more men included in the 1 mg ozanimod arm, a lower lactoferrin range on average in the 1 mg ozanimod group, and a lower proportion of participants in the placebo arm with previous medication use. These differences were not large or numerous enough to consider the randomisation and balancing of known and unknown prognostic factors to have failed. The ERG also noted that NCT01647516 does not list the change in Mayo score at week 32 as an outcome, even though it is listed and reported in the trial publication (Sandborn et al 2016);[34] it did take cognisance that this was an exploratory secondary outcome. Finally, it was noted that while an intention-to-treat analysis was conducted for the primary analysis this was not done using an imputation technique, but rather an assumption of non-response in participants with missing data. While this is a conservative assumption which would bias results to the null, the ERG considered 'Unclear' to be a more appropriate response for this domain.

TRUENORTH

The company appraised this trial as having no methodological concerns, with the assessment made at the trial level. No differential judgments were made by outcomes. The ERG agreed broadly with the assessments of the company according to the domains of the tool, though some domains were less obvious than others. While the ERG agreed that participants in the two groups of interest, placebo and 1 mg ozanimod in cohort 1, were similar in terms of prognostic factors it did note considerable differences; less men were included, median faecal calprotectin was lower on average, and the range of C-reactive protein (CRP) was lower on average in the ozanimod arm. These differences were not large or numerous enough to consider the randomisation and balancing of known and unknown prognostic factors to have failed. Furthermore, the ERG noted a discrepancy in the manner in which blinding was assessed between this trial and the TOUCHSTONE[34] trial. Given that blinding was reported similarly in the two trial publications (Sandborn et al 2016[34] and Sandborn et al 2021[28]) as well as the corresponding trial registries (NCT01647516 and NCT02435992, respectively), combined with Cohort 2 receiving ozanimod open-label, the ERG considered 'High' to be a more appropriate judgment for this domain. The receipt of open-label ozanimod in Cohort 2 is of particular concern for the induction phase, with participants self-reporting QoL and stool frequency

outcomes whilst unblinded; the ERG noted that such participants accounted for approximately 46% of participants treated with ozanimod in the ITT population. The ERG did not consider the company's judgment on imbalances in dropouts between groups and consequent analytical approach to be fully appropriate: it noted high and differential attrition between the ozanimod and placebo arms; indicating that the value of the outcome was unlikely to be independent of the missingness. While this may not have been unexpected, this may have influenced the validity of the non-response assumption for missing participants in the intention-to-treat analysis. Given that this assumption if conservative, and that sensitivity analyses using multiple imputation indicated that the results of primary analyses are robust, the ERG considers 'Unclear' to be a more reasonable judgment for the domain dealing with analytical approach.

TOUCHSTONE/TRUENORTH OLE

The company did not conduct a quality assessment for the OLE; it also did not comment on potential sources of bias present in this study. The ERG considered the study to be at high risk of bias. As the study did not have a control group, it is not possible to determine whether any observed changes are due to treatment with ozanimod, or natural disease progression over time. Furthermore, the open-label design may have resulted in ascertainment bias, with self-reported sub-scores of the Mayo score (e.g., stool frequency and rectal bleeding) particularly prone to this over-estimation of treatment effect.

3.2.5. Description and critique of the results of the studies

3.2.5.1. Clinical effectiveness results

The primary goal of treatment for UC is to induce remission. During periods of remission, patients' symptoms are minimal and the inflammation of the colon is reduced. In turn, this improves long-term prognoses by reducing the likelihood of developing complications such as colorectal cancer. In addition to the requirement of inducing remission, UC drugs must be able to maintain this state for as long as possible without relapse, ideally without the need for concomitant corticosteroids. This is particularly important since treatment options for UC are limited, with the result that each failed line of treatment takes a patient closer to a surgical last resort. The direct effect of these sub-scores on patients can also be identified through measures of quality of life, this is particularly pertinent given that UC is a chronic disease without a known cure.

Measures of disease activity

The company presented results from TRUENORTH[27,28] for the change from baseline in three-component Mayo score, and reported a greater reduction in the ozanimod group than in the placebo group at week 10 (LS mean (SE) change from baseline [redacted] for ozanimod and [redacted] for placebo, [redacted]. The company also reported a greater reduction in the three-component Mayo score in patients treated with ozanimod compared to those treated with placebo during maintenance at the 52-week time point (LS mean (SE) change from baseline [redacted] for ozanimod and [redacted] for placebo, [redacted].

In TOUCHSTONE[34], a significantly greater reduction in the three-component Mayo score was reported in the 1 mg ozanimod group when compared to placebo following induction up to week 10 (mean (SD) change from baseline -3.4 (2.79) for 1 mg ozanimod and -2.0 (2.52) for placebo, p=0.0042). A significantly greater reduction was also observed in the 1 mg ozanimod group when compared to placebo after maintenance at the 32-week time point (mean (SD) change from baseline -3.4 (2.93) for 1 mg ozanimod and -1.6 (2.72) for placebo, p=0.0004).

Clinical remission

Achievement of clinical remission with the three-component Mayo score (as defined in Section 3.2.4.1) was the primary endpoint during both the induction and maintenance phases of the TRUENORTH[27,28] study. During the induction phase, a significantly greater proportion achieved clinical remission at week 10 in the ozanimod arm versus placebo (18.4% vs. 6.0%, p<0.0001; OR (95% Wald confidence interval [CI]) 3.59 (1.94 to 6.63)). This was also reflected in the results from the maintenance phase at 52 weeks (37.0% vs. 18.5%, p<0.0001; OR (95% Wald CI) 2.76 (1.767 to 4.294)).

During the maintenance phase, the company provided further characterisation of patients' remission states with secondary endpoints measuring maintenance of remission and durable clinical remission. Maintenance of clinical remission, defined as the proportion of patients in clinical remission at the end of the maintenance period (52-week timepoint) in the subset of patients in clinical remission at the end of the induction period (10-week time point), was significantly higher in those receiving ozanimod when compared with placebo (51.9% vs. 29.3%, p=0.0025; OR (95% Wald CI) 2.88 (1.45 to 5.74)). Similarly, durable clinical remission, defined as those achieving remission at the end of both induction and maintenance periods, was

significantly higher in those receiving ozanimod than in the placebo arm (17.8% vs. 9.7%, p=0.003; OR (95% Wald CI) 2.65 (1,39 to 5.06)).

By comparison, the rates of clinical remission (defined as a Mayo score ≤2, with no subscore >1) at week eight were also greater in the ozanimod arm of the TOUCHSTONE[34] study (16.0% vs. 6.0%, p=0.048). Though the ERG noted a slight discrepancy in the reporting of the p-value presented in the CS appendices (Appendix L.3.2., p.302 and Table 80, p.303) in the maintenance phase at week 32, a significantly greater proportion (by either value) of those in the 1 mg ozanimod arm also achieved clinical remission than in the placebo arm (21% vs. 6%, p=0.01).

Clinical response

Clinical response in TRUENORTH[27,28] is presented by the company according to the three-component Mayo score. At week 10, a significantly greater proportion of patients receiving ozanimod achieved clinical response than those receiving placebo (47.8% vs. 25.9%, p<0.0001; OR (95% Wald CI) 2.67 (1.86 to 3.84)). This was also reflected in the maintenance phase at week 52 (60.0% vs. 41.0%, p<0.0001; OR (95% Wald CI) 2.27 (1.542 to 3.33)).

At the end of the induction phase of the TOUCHSTONE [34] study, there was also a significantly higher proportion of patients in the 1 mg ozanimod arm achieving clinical response than in the placebo arm (57% vs. 37%, p=0.02). At the end of the maintenance phase, there remained a greater proportion of the 1 mg ozanimod group with clinical response compared to placebo (51% vs. 20%, p<0.001) at week 32.

Hospitalisation

Though the company reported that rates of hospitalisation were not assessed as an outcome in TRUENORTH[27,28], the ERG noted that rates of hospitalisation were listed as an outcome of TRUENORTH and that some information on hospitalisations was provided in the CS (Document B, p.72). A very low overall rate of hospitalisations was reported with no accompanying test for significance between rates for ozanimod versus placebo ([redacted] vs. [redacted]).

Endoscopic improvement

While included in the overall four-component Mayo score used to quantify remission rates, endoscopic improvement was also reported as a separate secondary endpoint. It is unclear how endoscopic improvement is defined in the CS, however, endoscopic healing is defined in the

TRUENORTH[27,28] study as an endoscopic sub-score ≤1. The ERG considered that the terms 'improvement' and 'healing' may have been used interchangeably in the CS.

The company presented endoscopic improvement data as a secondary endpoint in both the induction and maintenance phases of TRUENORTH[27,28]. Endoscopic improvement was significantly greater in the ozanimod arm than in the placebo arm at week 10 of the induction phase (27.3% vs. 11.6%, p<0.001; OR (95% Wald CI) 2.88 (1.80 to 4.60)) and week 52 of the maintenance phase (45.7% vs. 26.4%, p<0.001; OR (95% Wald CI) 2.48 (1.65 to 3.72)).

Mucosal healing

Mucosal healing was also reported as secondary endpoint for both 10- and 52-week time points for the induction and maintenance phases. Mucosal healing can be defined as a lack of endoscopic or histological activity, or a combination of these. The CS, as in the TRUENORTH[27,28] study, defined mucosal healing as a Mayo endoscopy sub-score ≤1 and a Geboes index score <2.0. Again, this is presented for both the induction and maintenance phases. At week 10 of the induction phase, a significantly greater proportion of those in the ozanimod arm showed mucosal healing than those in the placebo arm (12.6% vs. 3.7%, p<0.001; OR (95% Wald CI) 3.77 (1.76 to 8.07)). This was similar at week 52 of the maintenance phase (29.6% vs. 14.1%, p<0.001; OR (95% Wald CI) 2.64 (1.64 to 4.26)).

Mucosal healing was defined differently in TOUCHSTONE[34], when compared to the stricter definition in TRUENORTH[27,28], as a Mayo endoscopy subscore of ≤1. This was also greater in the 1 mg ozanimod arm of the study compared to placebo at both the 8-week induction phase (34% vs. 12%, p=0.002) and 32-week maintenance phase (33% vs. 12%, p=0.005) time points. TOUCHSTONE[34] also reported histological remission separately, and defined this as a Geboes score of <2.0. Rates of histological remission were not significantly different between the 1 mg ozanimod and placebo arms at week eight (22% vs. 11%, p=0.07). Following maintenance at the 32-week time point, however, a statistically significantly greater proportion of the 1 mg ozanimod arm had achieved histological remission (31% vs. 8%, p<0.001).

Corticosteroid-free remission

During the maintenance phase, the company provided further characterisation of patients' remission states with a secondary endpoint measuring corticosteroid-free remission. A significantly greater proportion of patients treated with ozanimod achieved corticosteroid-free remission, defined as having been in remission without the need for corticosteroids for the prior

≥12 weeks during the maintenance phase, than those receiving placebo at week 52 (31.7% vs. 16.7%, p<0.001; OR (95% Wald CI) 2.56 (1.60 to 4.09)). The company considered this 12-week threshold to be clinically meaningful since relapse within 12 weeks is considered to be an indicator of steroid dependence in UC patients.

Adverse effects

The safety population reported in the CS includes all patients who received at least one dose of ozanimod. Treatment emergent adverse events (TEAEs) were defined as any AE with onset or worsening on or after the date of the first dose. TEAEs which occurred beyond the 90-day follow-up period were excluded. The company provide an overview of adverse events in Table 37 of document B of the CS.

Rates of treatment emergent adverse events (TEAEs) were similar between the ozanimod and placebo arms of cohort 1 (40.1% vs. 38% respectively). Rates of severe TEAEs, serious TEAEs, related serious TEAEs and those leading to interruption or discontinuation were also similar during the induction phase. However, in the maintenance period specifically, rates of TEAEs were higher in the ozanimod arm than the placebo arm (49.1% vs. 36.6% respectively). TEAEs suspected to be related to treatment were also higher in the ozanimod arm than placebo arm during the maintenance phase ([redacted] respectively)

In addition, a health resource utilisation questionnaire was used to collect data on hospitalisations, doctor visits and emergency room visits during both the induction and maintenance phases. During the induction phase, hospitalisation rates were [redacted] and [redacted] for ozanimod and placebo respectively. During the maintenance phase, hospitalisation rates were similarly low at [redacted] and [redacted] for ozanimod and placebo respectively. Conversely, in those re-randomised to ozanimod, rates of serious TEAEs and TEAEs leading to discontinuation were slightly lower than in those re-randomised to placebo.

One death occurred during the induction period in cohort 2. However, this was considered unrelated to ozanimod.

TOUCHSTONE also reported AE data for all three arms. The ERG noted the proportion of patients affected by AEs in the placebo, 0.5 mg ozanimod and 1 mg ozanimod arms were similar (40% vs. 40% vs. 39%, respectively). Serious AEs also occurred in similar proportion across the three arms (9% vs. 2% vs. 4% respectively. The most common AEs were UC flares, [redacted]

The ERG noted that the company only included serious infection AEs in the model; the company justified this approach by citing its high associated cost. This approach was accepted in TA633,[20] therefore the ERG considered it broadly appropriate. The TRUENORTH CSR[27] reported the incidence rates of serious infections during the induction phase as [redacted] and [redacted] for ozanimod and placebo, respectively. During the maintenance phase, those re-randomized to ozanimod had incidence rates of [redacted] compared to [redacted] in those re-randomised to placebo. The timeframe of reporting these results, specifically for the maintenance phase, was not clear to the ERG, i.e. it was not clear whether these results were annualised. As a result, the ERG was not able to validate the two-week cycle probability of serious infections as reported in the CS (Document B, Section B.3.3.9, Table 50). Furthermore, the ERG noted the data provided within the CSR are limited: Tables 14.3.2.1A and 14.3.2.1B were cited, but neither were made available to the ERG.

Health-related quality of life

Health related quality of life was presented for both the induction and maintenance phases through the EQ-5D-5L and the SF-36 in TRUENORTH[27,28]. However, the reporting of both the overall and component sub-scores was incomprehensive. The elements of the EQ-5D and SF-36 that were presented in the CS are described below.

For the induction phase, the physical component summary (PCS) score of the SF-36 was significantly improved in those treated with ozanimod compared to placebo [redacted], with a significantly greater proportion of patients treated with ozanimod compared to placebo achieved a minimally clinically important difference (MCID) for this score ([redacted] vs. [redacted], [redacted]). The mental composite summary score (MCS) score, however, showed no significant difference ([redacted] vs.[redacted] [redacted]) between ozanimod and placebo, respectively, and no difference in the proportion of patients who achieved MCID. The company reported that there were certain domains of the MCS score which showed significant improvement in the ozanimod group compared to placebo, including vitality ([redacted]), social functioning ([redacted]) and mental health ([redacted]). Scores on the SF-36 global health were also significantly improved with ozanimod compared to placebo ([redacted]), as were health utility scores ([redacted]). In terms of the EQ-5D summary index score, those in the ozanimod arms had a significantly greater mean change from baseline than those receiving the placebo ([redacted] vs.[redacted], [redacted]). Similarly, the mean change from baseline in the VAS, representing the self-reported health status, was significantly greater in the ozanimod arm than the placebo arm ([redacted] vs. [redacted], [redacted]).

For the maintenance phase, the company reported that SF-36 scores generally improved for patients randomised to both ozanimod and placebo. The PCS score, however, was most significantly improved in those receiving ozanimod compared to those receiving placebo ([redacted]). In addition, a significantly greater proportion of patients treated with ozanimod compared to placebo achieved MCID for this score ([redacted] vs. [redacted], [redacted]). There was, however, no difference in the proportion of patients achieving MCID in the SF-36 MCS scores at the 52-week time point. Also at week 52, the company reported that there was no significant difference between the EQ-5D summary index between placebo and ozanimod groups, though the ERG noted an inconsistency in reporting: [redacted] in text (Document B, p.70) versus [redacted] (Document B, Figure 19, p.71). VAS scores in those receiving ozanimod relative to those receiving placebo were significantly improved ([redacted] vs. [redacted], [redacted]).

Subgroup analyses

The CS presents two subgroups, defined as those that are TNFi-naïve and those that are TNFi-experienced.

During the induction phase, rates of clinical remission were higher in the ozanimod groups compared to placebo for both the TNFi-naïve ([redacted] vs. [redacted], respectively; [redacted] and TNFi-experienced ([redacted] vs. [redacted], respectively; [redacted] subgroups. These results were reflected in the maintenance phase at week 52 with higher rates of remission in the ozanimod groups compared to placebo for both the TNFi-naïve ([redacted] vs. [redacted], respectively; [redacted] and TNFi-experienced ([redacted] vs. [redacted], respectively; [redacted] subgroups. These results broadly reflect those from the overall population. The ERG noted and agreed with the company's position that lower rates of remission in the TNFi-experienced subgroup is explained by the fact that those with prior TNFi exposure are more challenging to treat.

This pattern also extended into the secondary outcomes, as shown in Table 15 below.

Table 15: Efficacy outcomes by TNFi-exposure subgroup and treatment arm

	TNFi-naïve		TNFi-experienced	
	Ozanimod	**Placebo**	**Ozanimod**	**Placebo**
Induction				
Clinical remission	[redacted]	[redacted]	[redacted]	[redacted]
Clinical Response	[redacted]	[redacted]	[redacted]	[redacted]
Endoscopic Improvement**	[redacted]	[redacted]	[redacted]	[redacted]
Mucosal Healing	[redacted]	[redacted]	[redacted]	[redacted]
Maintenance				
Clinical remission	[redacted]	[redacted]	[redacted]	[redacted]
Clinical Response	[redacted]	[redacted]	[redacted]	[redacted]
Endoscopic Improvement	[redacted]	[redacted]	[redacted]	[redacted]
Mucosal Healing	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviation: TNFi, tumour necrosis factor inhibitor

Note: * Significant difference at the 5% level between ozanimod and placebo; ** Different value for endoscopic improvement in TNFi-naïve placebo arm during induction in text ([redacted]; Document B, p.74) and Figure 22 ([redacted]; Document B, p.75)

3.3. Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

3.3.1. Search strategy

Two search strategies were used, one to identify RCTs and another to identify non-randomised trials, evaluating the efficacy and safety of ozanimod and its comparators for moderately to severely active UC for the company submission. The methods of these searches are described in Section 3.1.

3.3.2. Feasibility assessment

The company conducted a feasibility assessment for the NMA included in its submission, as described in CS Section 2.8.3 (Document B, pp.86-95). The company motivated its subgrouping of populations based on TNFi experience for the NMA by citing differential efficacy of treatment between first and second line biologics, with the former resulting in higher response rates and fewer patients requiring dose escalations than the latter. The ERG noted that subgrouping

based on degree of biologic experience is in line with the approach required in the NICE scope, with differences in efficacy between patients treated with first and second line treatments highly biologically plausible.

The ERG did not, however, consider the company's justification based on lower efficacy of treatments in the second line to be necessarily appropriate: literature cited included clinician surveys[38] and small studies (including less than 100 patients each) in participants with CD,[39-41] all describing comparative efficacy of TNFis (all considered first line according to the company decision problem). The ERG did not find these studies to be generalisable to the target population and considered that their publication dates reflect a clinical treatment landscape that is different to the current context. In contrast, a recent NMA[42] conducted in patients with moderately to severely active UC, and including most of the trials included in the company's NMA, generally found higher rates of clinical remission and endoscopic improvement for second-line treatments (ustekinumab and tofacitinib) versus placebo, when compared to first-line treatments (vedolizumab and TNFis) versus placebo. This was especially true for patients with prior TNFi exposure, in whom second-line treatments are typically used.[42]

3.3.2.1. Trial design

Maintenance trial design

The company submission also detailed the management of differences in trial design for the management phase; differentiated by the use of a 'treat-through' approach, i.e. once-off randomisation to treatment or placebo at baseline, or a 're-randomised' approach, i.e. randomising responders to a treatment again following an induction period during which participants were randomised to treatment or placebo. In addition to the information described in the CS (Document B, pp.86-88 and pp.97-98), further details on this approach are also provided in the appendices (Appendix D.4.1.) and in the response to clarification question A18.

Briefly, the approach allowed a comparison of like with like in terms of remitters in the maintenance phase who had responded by the end of an induction phase; this was achieved by estimating the number of clinical remitters among induction responders by applying a responders-to-remitters ratio from a comparable treatment arm from a similar trial investigating the same treatment or treatment class. For example, the responders-to-remitters ratio for TNFi-naïve participants of ULTRA 2[43] (comparing adalimumab with placebo) was applied to the total

number of remitters in ACT1[44] (comparing infliximab with placebo in a TNFi-naïve population) to estimate the number of remitters in ACT1 who had been induction responders.

The ERG considered it sensible to account for this source of heterogeneity, and noted the precedent for an approach to re-calculating data from different designs to allow comparisons in TA547[21] and TA633.[20] In TA633, the appraisal committee preferred the converse approach (i.e. recalculating re-randomised trials to approximate treat-through trials). The ERG acknowledged, however, the considerable uncertainty in recalculating re-randomised trials to approximate treat-through trials as reconfiguring these numbers to mimic the results of re-randomised trials may have biased relative measures of effectiveness. In particular, this approach assumes that there are no systematic differences between the baseline characteristics of induction responders and non-responders. If this is not the case, potential imbalances in treatment effect modifiers may have biased the results to an unknown extent, though the ERG accepted that the results of sensitivity analyses excluding treat-through trial designs (Appendix D.4.5.3.) demonstrated very little difference in point estimates, therefore the ERG did not consider the inclusion of data from these trials to be inappropriate. The assessment of the differences between the NMA base case and sensitivity analyses excluding these trials is described in more detail in Section 3.5.2.

Time point of assessment

The ERG considered the company's decision to restrict time points of assessment for the induction and maintenance phases of treatment within each subgroup to be a sensible approach in dealing with heterogeneity introduced by varying time points between trials. It was noted, however, that no clinical basis was provided for the selected restrictions, i.e. 6 to 14 weeks for induction and 52 to 60 weeks for maintenance. The ERG considered that the choice of these time points may have been led by maximal available data rather than clinical guidance, particularly in the induction phase; where only the trial by Sands et al. (2001)[45] was excluded.

3.3.2.2. Eligibility criteria

The company described trial eligibility criteria as a potential source of heterogeneity, a position with which the ERG concurred. The company report comparable inclusion criteria across trials for age, time since diagnosis, Mayo score and endoscopic sub-score and prior experience with CvT. Though the ERG acknowledged that these are important potential sources of heterogeneity that have been addressed in the company submission, it did note that UC SUCCESS[46] included participants aged 21 years and older while the Suzuki[47] trial listed 15

years and older as an age inclusion criterion. In addition, a study by Macaluso et al. (2018),[48] investigating factors affecting clinical and endoscopic outcomes in placebo arms of trials for UC treatments, indicated that concomitant steroids use, no prior TNFi experience, endoscopic central reading and duration of disease at baseline all affected these outcomes differentially. While the ERG considered TNFi experience to have been addressed through the stratification of analyses by prior TNFi exposure, other factors listed were not addressed through subgrouping or explored in sensitivity analysis.

The ERG noted the exclusion of trials conducted in exclusively Asian participants in a sensitivity analysis. The rationale for this was not clear or well-described in the CS, with the ERG noting no reported difference in pharmacokinetics between Japanese and Caucasian patients according to the SmPC. In addition, while inflammatory bowel disease (IBD) is modelled to be an emerging epidemic in Asia[49] this phenomenon is hypothesised to be related to changes in lifestyle, particularly the westernisation of diet,[50] rather than physiological differences in response. The ERG further considered that Asian patients are treated in the NHS and that the NICE scope did not exclude this population, therefore it agreed with the company's decision to include these trials in the base case NMA.

3.3.2.3. Subgroup definitions

TNFi versus biologic experience

The company's decision to separate participants into two mutually exclusive subgroups based on whether they were TNFi naïve or experienced, was considered a departure from the NICE scope; indicating subgrouping based on experience with biologics. The ERG noted that, of the trials included in the NMAs that reported on TNFi-experienced participants or a mix of TNFi-naïve and -experienced participants, only half listed TNFi experience as an explicit criterion. Four trials included participants with experience of 'biologics' or 'investigational' treatments, with two trials explicitly including those experienced with tofacitinib and vedolizumab (Motoya[51] and UNIFI,[52] respectively). The company justified this approach by indicating the following:

- TNFis are used almost exclusively in the first line.
- With the exception of UNIFI, across all trials included in the NMA, including TRUENORTH, subgroups were stratified by TNFi experience rather than biologic experience. This terminology is therefore a more accurate classification of the subgroups in which efficacy results are available.

- This approach is in line with a previous NMA in UC (TA547).

With respect to the first of these points, clinical advice to the ERG presented a more complex situation in UK clinical practice, as described in Section 2.2.1.

The ERG considered the heterogeneity in the TNFi-experienced subgroup to be a source of uncertainty in respect of the results of the NMA. The company reported that an overwhelming majority of participants (98.8%) in the UNIFI trial in the 'biologic failure' subgroup had experienced failure with at least one TNFi and justified their approach based on these numbers. The ERG could not find similar proportions for the study by Motoya et al. (2019)[51] (including biologic experience and describing required tofacitinib clearance period) or A3921063[53] (stipulating clearance periods for 'investigational' treatments). Furthermore, the ERG noted that, though the company described subgroups as being defined by TNFi experience, TOUCHSTONE[34] also specified clearance periods for 'biologic' or 'investigational' experience. As a result, the ERG is of the opinion that subgrouping by prior TNFi experience may limit generalisability of the results to the NICE scope, but is in line with the method of stratification used by the majority of the trials included in the NMA.

TNFi experience versus failure

The ERG accepted the company's explanation of heterogeneity within the TNFi-experienced subgroup in respect of failure, intolerance or inadequate response – particularly as Table 26 of the CS (Document B, p.91) showed that this information was not available in at least half of the trials. While the uncertainty caused by this heterogeneity is noted, the ERG further considered that this was not something the company could address as exclusion of these trials would result in sparse NMA networks. Finally, the company's approach is in line with the NICE scope, which specified inadequate response, loss of response or intolerance to biologic therapy.

3.3.2.4. Baseline characteristics

The company provided baseline characteristics of participants entering the induction and maintenance phases of trials included in the NMA in appendices to the CS (Appendix D.4., Tables 13 and 14, pp.113-121; though the ERG noted an erroneous reference to Appendix D.4.1.). These characteristics were reported as 'broadly similar' by the company, though the ERG noted large variations in C-reactive protein (CRP) levels and the proportion of patients with extensive disease, as well as some variation in concomitant steroid use and years since diagnosis. While the company also acknowledged this variation in the CS, it indicated that this

heterogeneity in baseline characteristics had been accepted in TA633[20] and TA547.[21] The ERG noted this precedent, and agreed to an extent that there are no alternative approaches which would improve the certainty around the estimates generated by the NMA, given that it is not known whether these characteristics are effect modifiers for UC treatment.

The ERG did, however, consider that excluding trials with outlier values for potential effect modifiers in sensitivity analyses, or running meta-regressions based on the values of such characteristics, could provide additional certainty around effect estimates generated by the NMA, but found there to be a paucity of published literature on the identification and cut-off levels of effect modifiers for UC.

3.3.2.5. Outcomes

The company presented the availability of outcome data by TNFi subgroups for the induction and maintenance phases by trials included in the NMA in the CS (Tables 27 and 28, respectively; pp.92-93), though the data itself were not presented. In cases where studies only reported clinical remission, but not clinical response, the company reported leveraging an ordinal response-remission NMA to retain studies – this approach is described in further detail in Section B.2.8.4 of the CS.

The ERG also noted the variation reported by the company in respect of the measurement of outcomes. This was considered to introduce some heterogeneity through the use of a three- or four-component Mayo score; the endoscopy sub-score of this measurement introduced further heterogeneity through local or central readings, with the latter providing greater objectivity than the former. The ERG further noted that the company restricted outcome measurement to the Mayo score, with trials using the UC Symptom and modified Truelove and Witts scoring systems excluded from the NMAs. This was considered to be an appropriate step in managing heterogeneity in outcome measurement: while this could have been further managed by restricting to the three- or four-component Mayo score, or to endoscopy scores read centrally, the ERG appreciated that this may have resulted in very sparse networks and highly imprecise effect estimates. Furthermore, a sensitivity analysis including three-component Mayo score data instead of four-component data from TRUENORTH[27,28] showed very little change in the effect estimates of the NMAs (comparing Appendix D.4.5.4. and base case league tables provided by the company in its clarification response to question A13).

The ERG did, however, consider the use of outcome data from an unweighted average of the placebo arms to be contrary to NICE guidance on this topic, which suggests using the evidence

sources that are generalisable to the decision problem to inform the baseline model (Dias et al. 2013).[1] It considered the approach taken by the company to increase the uncertainty and considerably reduce the generalisability of the findings of the NMA, and therefore recommends that the use of placebo rates from a single, generalisable trial (or possibly multiple generalizable sources of evidence) would yield results that are more aligned with the NICE scope.

3.3.3. Study selection criteria

The selection criteria used by the company are described in the CS appendices, with specific selection criteria presented in Appendix D.2.2. (Table 7, p.50-51). The ERG considered these criteria to be broadly appropriate, and noted specifically the inclusion of tofacitinib as a comparator treatment, resulting from the company's decision to include all treatments specified in the NICE final scope in the NMA.

As discussed in Section 3.3.2, the ERG considered the stratification of the company's analyses by prior TNFi experience to be a departure from NICE scope which may limit the generalisability, both to populations which are naïve to and have experience of biologics, but is in line with the method of stratification used by the majority of the trials included in the NMA.

The company chose to exclude phase 4 trials from the submission, which the ERG did not consider appropriate as such evidence could have been used to inform other links in the networks. In particular, real-world evidence for ozanimod could provide additional insights into the long-term safety and efficacy of the treatment. To determine the potential impact of this exclusion, the ERG conducted a search for phase 4 trial evidence for ozanimod and its comparators in the submission. The results of this search yield are reported in Section 3.5.1.

Following the completion of its screening, the company imposed additional exclusion criteria as part of its feasibility assessment. While this typically considered to introduce potential bias, the ERG agreed with the company's exclusions based on unlicensed doses and ineligible comparisons, as well as trials with substantially different follow-up time points. Notably, tofacitinib was not excluded from the search yield at this stage and was included in the company's NMAs. Furthermore, the ERG noted the absence of UC-SUCCESS[46] from the TNFi-naïve induction phase evidence network, even though it had not been excluded for any reasons stated in the feasibility assessment. The ERG did not consider the exclusion of this ostensibly eligible source of evidence to be appropriate and regarded it as decreasing the confidence in the results of the NMA.

3.3.4. Included studies

The flow of studies identified for the NMAs was reported clearly in a PRISMA diagram (Appendix D.3., Figure 1). The company reported including 28 trials in the qualitative synthesis, of which 22 trials were included in the quantitative synthesis, i.e. NMA. This discrepancy resulted from three trials, namely OASIS (etrasimod),[54] HICKORY[55] and EUCALYPTUS[56] (both etrolizumab), being excluded due to the treatments of interest not having FDA or European Medicines Agency (EMA) approval at the time of the appraisal. The ERG noted that this approach was in line with TA547[21] and TA633[20] and considered the exclusions appropriate. It was not clear to the ERG at which stage the PURSUIT-IV trial,[57] investigating intravenous golimumab, was excluded. This trial was reportedly excluded on the basis of the treatment not having FDA or EMA approval, bringing the total number of trials excluded for this reason to four. Along with the trials reported in Table 24 of the CS (Document B, pp.84-86), this increases the number of trials that should have been included in the qualitative synthesis to 29.

A further three trials were excluded due to one comparing an approved and unapproved dose of adalimumab (SERENE-UC)[58] and the remaining two not using the Mayo clinic score to assess outcomes: the study by Probert et al. (2003)[59] used the UC Symptom scoring system; the study by Sands et al. (2001)[45] used the modified Truelove and Witts scoring system (Sands). The ERG agreed with the exclusion of the SERENE-UC[58] trial as no placebo arm was included in the study, and therefore no eligible comparison was available to inform links in the networks. Very little information was available in the CS, primary studies or in the published literature on what the modified Truelove and Witts and UC Symptom scoring system comprised. The ERG considered that the inclusion of these trials would have exacerbated outcome-related heterogeneity in the NMAs and found the exclusions broadly appropriate.

For the induction phase of treatments, a total of 18 trials reported at least one outcome in the TNFi-naïve and/or TNFi-experienced subgroup. For the TNFi-naïve subgroup, 15 trials reported on clinical response and 14 on clinical remission. For the TNFi-experienced subgroup, eight trials reported on clinical response and seven on clinical remission. For the maintenance phase of treatments, a total of 12 trials reported at least one outcome in at least one of the subgroups related to TNFi experience. Of the trials reporting on TNFi-naïve populations, 10 reported on clinical response and 12 on clinical remission. Of the trials reporting outcomes for TNFi-experienced populations, six reported on clinical response and eight on clinical remission.

The majority of trials (n=20) included in the NMA had a placebo arm, though many trials included multiple arms investigating different doses of the same treatment (n=12). A total of three trials included a head-to-head comparison between active treatments, namely UC-SUCCESS[46] (comparing azathioprine and infliximab), VARSITY[60] (comparison vedolizumab and adalimumab) and GEMINI 1[61] (comparing infliximab with various arms treated with vedolizumab). Of these head-to-head trials, only GEMINI 1 was placebo-controlled.

The number of included RCTs for each comparator treatment were as follows: adalimumab, n=4; azathioprine, n=1 (not a relevant comparator in this appraisal); golimumab, n=3; infliximab, n=6; ozanimod, n=2; tofacitinib, n=4; ustekinumab, n=1; vedolizumab, n=4; and placebo, n=20.

Trials included in the NMA were conducted between dates ranging from 2002 to 2021, according to the trial registries of these studies. The trials were conducted across a range of geographic locations and healthcare settings, with the majority conducted in multiple countries (n=17). Four trials were conducted in Japan only, and one trial was conducted in China. Time points of assessment following the induction phase ranged from six to 14 weeks; follow-up after induction and maintenance ranged from 30 to 60 weeks.

The eligibility criteria of the trials in the company NMA are reported in the appendices to the CS (Appendix D.3.1., Table 9, pp.56-62). These criteria showed very little between-trial variation in diagnostic criteria, with only UC-SUCCESS[46] not specifying an endoscopic sub-score of the Mayo score and TRUENORTH[27,28] specifying additional criteria for the rectal bleeding and stool frequency sub-scores. A number of trials did not specify age as an eligibility criterion, indicating the possibility of including paediatric patients. The ERG considered this to be unlikely, given that UC typically does not present before 15 years of age, and was of the opinion that the heterogeneity introduced and departure from the NICE-scoped population by including a few paediatric patients would be negligible. There was some heterogeneity in the trials reporting age inclusion criteria: most specified participants aged 18 or older, with some indicating an upper age limit; the studies by Motoya et al. (2019)[51] and Suzuki et al. (2014)[47] recruited patients from 15 years, while UC-SUCCESS[46] amended its minimum age criteria from 18 years to 21 years. The ERG did not consider this variation in age inclusion to meaningfully affect heterogeneity, given the high background heterogeneity between the included trials. Finally, previous experience with CvT and active treatment was stipulated for a number of trials – the ERG was of the opinion that this heterogeneity was address through the subgrouping of the overall

population by TNFi experience, though it did not consider the approach to be exactly aligned to the NICE scope, as discussed in Section 3.3.2.3.

3.3.5. Quality assessment of studies included in indirect treatment comparison

The company reported using the University of York CRD[32] criteria for assessing risk of bias for the trials included in the NMA. The ERG noted that the domains used in the assessment of the trials were appropriate for the CRD criteria. The judgments were summarised in Appendix D.6. (Table 15, pp.169-170). Overall, the company assessed most studies included in the NMA to have had appropriate randomisation and, to a somewhat lesser extent, adequate concealment of treatment allocation. The ERG noted that all trials at unclear risk of bias for appropriate randomisation were also at unclear risk for allocation concealment, representing a serious potential risk of baseline imbalance for confounders and effect modifiers. A number of studies, some of those with potential risk of selection bias, were considered to have an unclear risk of bias related to baseline imbalance for prognostic factors. These factors were identified by the ERG as potential factors that could increase the uncertainty in the NMA estimates. The blinding of care providers, participants and outcome assessors was mostly assessed as unclear across trials; the ERG considered that this may have systematically biased results in favour of active treatments, particularly given that outcomes of interest were at least partially self-reported, i.e. stool frequency and rectal bleeding as part of the Mayo score. Studies included in the NMAs were generally considered not to have quality issues related to unexpected imbalance in attrition between trial arms, or for selective outcome reporting; most trials were also judged as having conducted intention-to-treat analyses, and for doing so appropriately.

The ERG noted, however, that the company did not provide justifications for their quality assessments, which made it difficult to determine whether these were reasonable. The company did report that these assessments were made by a single reviewer, with validation by a second reviewer and, where necessary, resolution by a third reviewer. This was considered to be appropriate. Within the timeframe of the appraisal it was not possible for the ERG to conduct independent assessments of the quality of trials included in the NMA. The ERG did, however, compare these judgments with a comprehensive assessment of quality appraisal done in TA633.[20] This assessment suggested that all trials were considered to be a low risk of bias ('Yes' according to CRD criteria) for the randomisation and allocation concealment domains – this is contrary to the assessment in this appraisal, with the SERENE-UC,[62] Probert et al. (2003)[59] and Sands[45] trials considered 'Unclear' for these domains. These assessments in

terms of randomisation were considered by the ERG to be appropriate, as very little information on SERENE-UC was available in the public domain at the time of submission; the studies by Probert[59] and Sands[45] did not provide clear methodology around randomisation. The ERG noted that none of these trials were included in the NMAs following the feasibility assessment. In terms of allocation concealment, the ERG noted that PURSUIT J[63] and UNIFI[52] were additionally discrepant from the assessment in TA633[20] as they were also considered at 'Unclear' risk. The ERG also agreed with these assessments in the current appraisal, as all five trials reported insufficient information to enable an assessment.

TA633[20] acknowledged heterogeneity between trials for balance in prognostic factors, but broadly considered an assessment of low risk ('Yes') to be appropriate, though it identified the Study A3921063 and ACT 1 trial to have the greatest within-trial variability. This was broadly reflected in the assessments of the current submission, with Study A3921063[53] as well as ACT1[44] judged as having unclear risk. The ERG noted that ACT 2,[44] SERENE-UC,[62] Sands, Motoya,[51] OCTAVE 2[64] and OCTAVE SUSTAIN[64] and were additionally judged as having unclear risk, but could not verify this as the prognostic factors used in the assessment were not clearly identified in the CS.

The largest discrepancy between the assessment done in TA633[20] and the current submission is with regards to blinding, with the former indicating most trials were at low risk of bias while the latter assessed most trials as posing an unclear risk of bias. In addition, TA633[20] indicated that 'No' for all trials except ULTRA 1 is an appropriate assessment for imbalance in dropouts – the ERG noted that several trials were judged as 'Unclear' or 'High' in this submission, with ULTRA 1[65] not among these. The assessment of intention-to-treat analyses also yielded some discrepancies between the assessment in TA633 and the current appraisal – the OCTAVE trials are assessed in the CS as having conducted appropriate analyses, in line with opinion in TA633; on the other hand, the CS judged Probert[59] and PURSUIT-M[66] as 'Unclear' and 'Yes' for this domain while TA633[20] considered that intention-to-treat analyses were not reported for these trials. The ERG considered that this may be due to systematic differences in assessing these domains, with assessment of attrition and appropriate analysis further exacerbated by uncertainty around the time of assessment (induction versus maintenance), and noted this as an area of potential uncertainty.

3.4. Critique of the indirect comparison and/or multiple treatment comparison

The following sections contain an appraisal of the company's NMA methods and results, as conducted by the ERG. The ERG considered that the model applied by the company followed recommended practice and that logical steps had been taken to address some potential effect modifiers (primarily prior TNFi treatment and differences in trial design). Overall, the ERG considered that the company could have selected a more generalisable approach to assessing baseline risk in the placebo arm (in this context, the probability of being in non-response and non-remission, under placebo) of the NMAs, thereby providing effect estimates that are more applicable to the UK context. As a result, the ERG conducted scenario analyses using more UK-appropriate alternate values for the placebo arm of the NMAs. The ERG was also of the opinion that the placebo arms of trials included in the NMAs showed considerable variability in baseline characteristics, and may differ in respect of potential effect modifiers. Furthermore, given the heterogeneity in the evidence base used to conduct the NMAs, the ERG considered RE modelling to be a more appropriate choice; an approach that was not applied to all company analyses. Though the ERG noted the company's assertion that RE modelling was not done throughout due to failure of the model to converge, it considered that RE modelling could have been conducted using an informative prior distribution from a relevant context, e.g. those reported in Turner et al. (2012)[67] or Turner et al. (2015).[68]

3.4.1. Summary of analyses undertaken

The company carried out four NMAs representing the combinations of TNFi-naïve and -experienced during induction and maintenance periods. The NMA models were based on a multinomial model with probit link, described further in NICE TSD2,[69] and the analysis was carried out in a Bayesian framework using JAGS. The company explained that the underlying JAGS code was 'in line with' the WinBUGS code presented in Example 6 of the appendix to NICE TSD 2.[69]

The company favoured the default use of RE models, but selected FE models in three of its four settings (induction in the TNFi-experienced subgroup, and maintenance phases for both TNFi experience subgroups), following an assessment of convergence with the Gelman-Rubin convergence statistic. With RE models they found non-convergence, or overly large variances in estimates, in three settings, so FE models were applied instead.

Doses were pooled in analyses, that is, grouped as the same treatment where individual doses of the same active agent had the same method of administration (summarised in Document B, Table 31). A sensitivity analysis was carried out without pooling of doses; the results of which were broadly comparable to the base case (see Section 3.4.4.5).

The trials included were both 'treat-through' and 're-randomised' in design. The company used a procedure to make the treat-through trials emulate re-randomised trials (see 3.4.2.3). A sensitivity analysis was carried out in which treat-through trials were excluded (see Section 3.4.4.5) and, given the potential uncertainty introduced by this approach, as highlighted in Section 3.3.2.1, the ERG undertook a head-to-head comparison of the results of the base case and sensitivity analysis. The results of this comparison demonstrated very little difference between the two approaches, as described in Section 3.5.2; the ERG therefore considered the approach to be appropriate.

The company did not include basic results in the CS, i.e. numbers or proportions partially responding or remitting by trial arm. These were supplied in Appendix 2 of the company's clarification response, and are reproduced with reformatting in Table 16 to Table 19 below. Estimates of the proportions responding/remitting under the company's NMAs are given in Appendix 3.3 of the company's clarification response. Complete sets of pairwise estimates of odds ratio from the NMAs are provided in Figures 3 to 10 of the company's clarification response.

Table 16: Company's base case NMA inputs for TNFi-naïve subgroups during induction; provided during clarification

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	8	Infliximab Pooled	159	243	65.4%	86	243	35.4%
		Placebo	45	121	37.2%	18	121	14.9%
ACT 2	8	Infliximab Pooled	161	241	66.8%	74	241	30.7%
		Placebo	36	123	29.3%	7	123	5.7%
GEMINI 1	6	Vedolizumab 300 mg IV	69	130	53.1%	30	130	23.1%
		Placebo	20	76	26.3%	5	76	6.6%
Jiang 2015	8	Infliximab Pooled	32	41	78.0%	22	41	53.7%
		Placebo	15	41	36.6%	9	41	22.0%
Kobayashi 2016	8	Infliximab Pooled	57	104	54.8%	21	104	20.2%
		Placebo	37	104	35.6%	11	104	10.6%
Motoya 2019	10	Vedolizumab 300 mg IV	42	79	53.2%	22	79	27.8%
		Placebo	15	41	36.6%	6	41	14.6%
OCTAVE 1 + 2	8	Placebo	43	110	39.1%	13	110	11.8%
		Tofacitinib 10 mg BID	284	440	64.5%	106	440	24.1%
PURSUIT-SC	6	Placebo	89	292	30.5%	20	292	6.8%
		Golimumab 200/100 mg SC	147	294	50.0%	52	294	17.7%
Study A3921063	8	Placebo	15	33	45.5%	NA	NA	NA
		Tofacitinib 10 mg BID	14	23	60.9%	NA	NA	NA
Suzuki 2014	8	Adalimumab 160/80/40 mg Q2W	45	90	50.0%	9	90	10.0%
		Placebo	34	96	35.4%	11	96	11.5%
TRUE NORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 1	8	Adalimumab 160/80/40 mg Q2W	71	130	54.6%	24	130	18.5%
		Placebo	58	130	44.6%	12	130	9.2%
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	89	150	59.3%	32	150	21.3%
		Placebo	56	145	38.6%	16	145	11.0%
UNIFI	8	Placebo	56	158	35.4%	15	158	9.5%
		Ustekinumab Pooled	194	312	62.2%	60	312	19.2%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	151	305	49.5%	72	305	23.6%
		Vedolizumab 300 mg IV	213	304	70.1%	84	304	27.6%

Abbreviations: BID, twice a day; IV, intravenous; NMA, network meta-analysis; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

Table 17: Company's base case NMA inputs for TNFi-naïve subgroups during maintenance; provided during clarification

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
ACT 1	54	Infliximab pooled	92	159	57.9%	53	159	33.3%
		Placebo	17	45	37.8%	10	45	22.2%
GEMINI 1	52	Placebo	21	79	26.6%	15	79	19.0%
		Vedolizumab pooled	88	145	60.7%	68	145	46.9%
Motoya 2019	60	Placebo	10	28	35.7%	10	28	35.7%
		Vedolizumab pooled	16	24	66.7%	13	24	54.2%
	52	Placebo	27	109	24.8%	12	109	11.0%

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
OCTAVE SUSTAIN		Tofacitinib pooled	132	219	60.3%	94	219	42.9%
PURSUIT-J	54	Golimumab pooled	18	32	56.3%	16	32	50.0%
		Placebo	6	31	19.4%	2	31	6.5%
PURSUIT-M	54	Golimumab pooled	146	302	48.3%	101	302	33.4%
		Placebo	48	154	31.2%	34	154	22.1%
Suzuki 2014	52	Adalimumab 40 mg Q2W	50	82	61.0%	38	82	46.3%
		Placebo	12	34	35.3%	8	34	23.5%
TRUE NORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	52	Adalimumab 40 mg Q2W	44	89	49.4%	34	89	38.2%
		Placebo	24	56	42.9%	15	56	26.8%
UNIFI	52	Placebo	44	87	50.6%	27	87	31.0%
		Ustekinumab pooled	144	187	77.0%	91	187	48.7%
VISIBLE 1	52	Placebo	NA	NA	NA	7	37	18.9%
		Vedolizumab 108 mg Q2W SC	NA	NA	NA	36	67	53.7%
		Vedolizumab pooled	NA	NA	NA	17	32	53.1%

Abbreviations: NMA, network meta-analysis; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

Table 18: Company's base case NMA inputs for TNFi-experienced subgroups during induction; provided during clarification

Trial name	Induction period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	6	Vedolizumab 300 mg IV	32	82	39.0%	8	82	9.8%
		Placebo	13	63	20.6%	2	63	3.2%
Motoya 2019	10	Vedolizumab 300 mg IV	23	85	27.1%	8	85	9.4%
		Placebo	12	41	29.3%	4	41	9.8%
OCTAVE 1 + 2	8	Placebo	29	124	23.4%	1	124	0.8%
		Tofacitinib 10 mg BID	237	465	51.0%	53	465	11.4%
Study A3921063	8	Placebo	5	15	33.3%	NA	NA	NA
		Tofacitinib 10 mg BID	6	10	60.0%	NA	NA	NA
TRUE NORTH	10	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	8	Adalimumab 160/80/40 mg Q2W	36	98	36.7%	9	98	9.2%
		Placebo	29	101	28.7%	7	101	6.9%
UNIFI	8	Placebo	44	161	27.3%	2	161	1.2%
		Ustekinumab Pooled	169	330	51.2%	40	330	12.1%
VARSITY	14	Adalimumab 160/80/40 mg Q2W	26	81	32.1%	10	81	12.3%
		Vedolizumab 300 mg IV	44	79	55.7%	18	79	22.8%

Abbreviations: BID, twice a day; IV, intravenous; NMA, network meta-analysis; Q2W, every two weeks; QD, once a day; TNFi, tumour necrosis factor inhibitor

Table 19: Company's base case NMA inputs for TNFi-experienced subgroups during maintenance; provided during clarification

Trial name	Maintenance period (weeks)	Treatments	Clinical response			Clinical remission		
			n	N	%	n	N	%
GEMINI 1	52	Placebo	6	38	15.8%	2	38	5.3%
		Vedolizumab pooled	37	83	44.6%	30	83	36.1%
Motoya 2019	60	Placebo	5	14	35.7%	3	14	21.4%
		Vedolizumab pooled	11	17	64.7%	10	17	58.8%
OCTAVE SUSTAIN	52	Placebo	13	89	14.6%	10	89	11.2%
		Tofacitinib pooled	92	176	52.3%	54	176	30.7%
TRUENORTH	52	Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ULTRA 2	52	Adalimumab 40 mg	15	36	41.7%	10	36	27.8%
		Placebo	6	29	20.7%	2	29	6.9%
UNIFI	52	Placebo	34	88	38.6%	15	88	17.0%
		Ustekinumab pooled	98	161	60.9%	52	161	32.3%
VISIBLE 1	52	Placebo	NA	NA	NA	1	19	5.3%
		Vedolizumab 108 mg Q2W SC	NA	NA	NA	13	39	33.3%
		Vedolizumab pooled	NA	NA	NA	6	22	27.3%

Abbreviations: NMA, network meta-analysis; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

The same trials supplied data for both the baseline risk estimates and the relative risk model. The baseline risk model made use of all placebo arms from these trials.

3.4.2. Critique of assumptions used in the indirect treatment comparison

3.4.2.1. General methodology

The company made use of a multinomial model with probit link for the counts of trial participants exceeding a set of thresholds. This modelling approach followed recommended practices (see TSD2[69]) and has precedent in UC submissions (e.g. TA547[21]).

The multinomial with probit modelling approach allows trials to utilise alternative thresholds within the pooled analysis. The set of thresholds across trials represent varying definitions of trial response and remission (see Document B, Tables 29 and 30). Although the inclusion criteria restricted trial outcomes to those involving reductions in Mayo scores, between-trial variation remained, primarily in the extent of Mayo score reduction constituting a response category. The modelling also accounts for inherent correlation between outcomes (i.e. between counts in different response categories), which has an advantage over a previous UC submission (TA633[20]) in which separate NMAs were carried out for remission and response, thereby disregarding the correlation.

The ERG noted that company network diagrams (Document B, Figures 31, 34, 37 and 40) are star-shaped, or nearly so, with a central node representing placebo. The networks have very few loops, indicating a lack of 'indirect' evidence in the networks. There are also few replicates of trials between the nodes, so difficulties in estimating heterogeneity precisely (if at all) may be anticipated. The Bayesian approach to modelling, as used by the company, seems apt if prior information can be justified and utilised, though this was not the case.

The company analysis took steps to account for heterogeneity, including:

- conducting separate NMAs for TNFi-experienced/TNFi-naïve combined with induction/maintenance phases,
- processing of 'treat-through' trials to emulate 're-randomised' trials (see Section 3.4.2.3 below) to remove or at least modulate a major heterogeneity from trial design,
- restriction of trials to those reporting outcomes determined using Mayo scores only, and

- restricting the time points of assessment in trials to 6 to 14 weeks for induction, and 52 to 60 weeks for maintenance.

The company pooled information of 'doses of the same active agent that had the same method of administration' for the base case (Document B, section B.2.8.4). Pooling of doses is contrary to Dias et al. (2018)[70] recommendation 'The default network meta-analysis … treats every intervention, every dose and every treatment combination a separate treatment' [p15]. On the other hand, the doses were at licensed levels and the ERG was advised by clinicians that the doses matched clinical practice (i.e. not implausibly low/high).

3.4.2.2. Choice of model (fixed effect or random effects)

The ERG noted that the underlying trials informing the NMA displayed considerable heterogeneity in respect of setting, inclusion criteria and baseline characteristics. As such, the ERG considered the RE model to be more appropriate than the FE model, as the latter assumes no between-trial heterogeneity is present. Heterogeneity in UC trials is known to be high, based on published literature (Macaluso et al. 2018).[48] It is also evident in the decision problem (variation in TNFi experience, see Table 7) and data (see baseline characteristics, CS Appendix, Tables 13 and 14; and outcome data, Table 16-Table 19 of Section 3.4.1). RE models are, therefore, better suited to NMAs of this condition due to the highly heterogeneous nature of studies; the company acknowledged this in the CS and reported that these were favoured over FE by default (Document B, Section B.2.8.4,p.96).

The ERG noted, however, that when attempting to fit RE models, the company reported model non-convergence in both TNFi-naïve and TNFi-experienced maintenance settings; and highly uncertain posterior standard deviation of between-trial variation within treatments in the setting for TNFi-experienced patients in the induction phase. As a result, the FE model was selected and applied by the company in all but the setting for TNFi-naïve patients in the induction phase.

The ERG noted that the aforementioned problems can often be remedied in RE models by using a more informative prior distribution on the variance parameter. The company did not report trying this remedy. The ERG acknowledged that in comparison to a FE model the likely results under the more defensible RE model would be similar point estimates, but an increase in the width of the credible intervals.

3.4.2.3. Choice of trial design ('treat-through' versus 're-randomised')

The company attempted to reduce a major potential source of heterogeneity in study design by carrying out a procedure to translate 'treat-through' trials to 're-randomised'. The ERG agreed with the principle but found the explanation poor in CS – it was clarified in response to questions A17 & A18. This process or similar appears to have been applied previously in e.g. TA547,[21] while the reverse approach (translating re-randomised to treat-through) was preferred in e.g. TA633.[20]

The procedure involved assuming those responding at the end of induction then enter the maintenance phase (as would happen in a 're-randomising' trial). This is discussed in more detail in Section 3.3.2.1 of this document. The assumptions are that the ratio of response during maintenance and response during induction is similar between similar trials, and that there is no delayed response, i.e. no participants respond after induction and then enter the maintenance phase. The ERG considered that there may be problems with this approach: there is evidence that these assumptions may be violated, and further there may be systematic differences between 're-randomised' and 'treat-through' trials due to differences in the trial process following randomisation. However, a sensitivity analysis conducted by the company indicated that the modification of treat-through trials is not very influential (see Section 3.5.2), and therefore the ERG did not investigate further.

3.4.2.4. Baseline risk

The sources for the baseline risk (the probability in the placebo arm of remaining in UC / not in response or remission) in the CS are the placebo arms, where they exist, of exactly the same set of trials used to estimate relative treatment effects. This runs contrary to TSD5,[71] which recommends separate modelling of relative treatment effects and baseline effects, and potentially separate sources of evidence, with the latter not restricted to randomized trials.

The studies informing the baseline model in the CS have not undergone a separate search process oriented to the baseline setting ("Investigators should identify evidence sources to inform the baseline model based on a protocol-driven systematic search ... " Dias et al. (2018)[70] (p.157)).

Furthermore, there has been no filtering of these studies towards the baseline setting. The ERG noted there is high variability in baseline characteristics across placebo arms of included trials (Appendix D.4, Tables 13 and 14). High heterogeneity in baseline variables may weaken

external validity: it may be that only a subset of these trials will match the decision problem population, or even none at all (in which case other sources of information e.g. registry data would be essential).

The ERG recommends selecting sources closest to UK clinical practice, with most appropriate choices for factors identified as important determinants of outcome, i.e. concomitant steroid use, duration of disease, prior TNFi, endoscopic central reading (Macaluso et al. 2018).[48] Clinical advice to the ERG confirmed that these factors are important to consider, in addition to severity of disease defined by a modified Mayo score of 9 or 10; endoscopy comprises one of the parts of this score. For its revised base case analysis, the ERG chose the following sources as best representing baseline risk: for the induction phase, PURSUIT SC[72] was selected for the TNFi-naïve subgroup, as it includes a similar gender split and roughly the same age as a large UK cohort[35] and OCTAVE 2[64] was selected for the TNFi-experienced subgroup, for the same reason; for the maintenance phase, PURSUIT M[66] was selected for the TNFi-naïve subgroup, and GEMINI1[61] for the TNFi-experienced subgroup, as the placebo arms of these trials still matched the age and gender split of the UK cohort most closely. In addition, these trials were all conducted in populations not exclusively including Asian participants; all trials were also assessed as having low risk of selection bias and were considered balanced in terms of prognostic factors at baseline (Appendix D.6., Table 15, pp.169-170). In these four studies, trial-level placebo arm average age since diagnosis was between 6.0 and 7.8 years, and average concomitant steroid use was between 42.9% and 57%. The characteristics of participants in the placebo arms of the selected trials are summarised in Table 20.

The company highlighted that remission or response data for the ERG's selected baseline trial for TNFi-experienced participants during induction (OCTAVE 2) were only available when pooled with results from the OCTAVE 1 trial. The baseline values for OCTAVE 1 are therefore also supplied in Table 20, where it can be seen that there is generally good correspondence, but that compared with OCTAVE 2 the percentage of males is about 13% higher, and the percentage with TNFi exposure about 5% less. The need to pool the ERG's selected trial (OCTAVE 2) with a similar trial (OCTAVE 1) is a limitation of the ERG's exploratory analysis.

Table 20: Baseline characteristics of participants in the placebo arms of trials selected for the ERG's placebo baseline risk NMA scenario

Characteristic (mean, unless otherwise specified)	TNFi-naïve (induction phase)	TNFi-experienced (induction phase)		TNFi-naïve (maintenance phase)	TNFi-experienced (maintenance phase)
	PURSUIT SC	OCTAVE 1	OCTAVE 2	PURSUIT M	GEMINI 1
Age, years (mean)	39.0	41.8	40.4	40.2	40.3
Male (%)	52.9	63.1	49.1	48.1	54.8
CRP (mg/L) (mean)	10.7	4.7	5.0[a]	9.6	NR
Years since diagnosis (mean)	6.0	6.0	6.2[a]	6.9	7.8
Mayo score (mean)	8.3	9.1	8.9	8.3	8.4
Left-sided disease (%)	57.0	30.3	35.1	NR	42.1
Extensive disease (%)	43.0	54.1	50.5	NR	13.5
Concomitant steroid use (%)	42.9	47.5	49.1	56.4	57
Biologic (TNFi) exposure (%)	NA	53.3	58.0	NA	37
Prior TNFi failure (%)	NA	52.5	53.6	NA	30.2

Abbreviations: CRP, C-reactive protein; NMA, network meta-analysis; TNFi, tumour necrosis factor inhibitor

Note: [a] median values

The ERG acknowledges the limitations of using these trials for placebo baseline risk, given the unsystematic selection of these based on limited information related to demographics, settings and methodological quality. This approach was selected due to time constraints within the appraisal and should therefore be seen as an attempt at improving the generalisability of results, vis-à-vis that of an unweighted average of all placebo arms, albeit with its own uncertainty. Clinical advice to the ERG confirmed that the baseline risk values used from the selected trials are broadly acceptable and representative of the relevant population by TNFi experience, as well as the treatment phase, but cautioned that no patients with proctitis were explicitly included. These patients are estimated by clinical experts to the ERG as representing approximately 20% of patients treated in the Royal Devon and Exeter NHS Foundation Trust. As a result, clinical advice to the ERG indicated that participants included in trials generally have more severe disease and increased use of steroids when compared to the 'general' population of patients with moderately to severely active UC in the UK. The ERG noted this caution and

considered that baseline placebo risk from the selected trials may not be fully generalisable to the target population, and recommends that a proper protocol-driven systematic review procedure as described in NICE guidance (Dias et al 2013)[1] is followed in respect of estimating baseline placebo risk, including non-RCT sources where available.

The baseline data (response or remission proportion; remission proportion in the placebo arm) for the selected trials are shown in Table 21 along with the estimates from the ERG's updated NMA. These estimated proportions are smaller than those used in the company base case where all trials with placebo arms were used. More information on the results of the ERG's updated NMA is given in Section 3.5.3.1.

Table 21: Comparison of response/remission proportions from data with estimates from NMA with updated baseline selection

Setting	Trial supplying baseline risk	Data source table[a]	Response or remission	NMA estimate	Remission	NMA estimate
Induction/ naïve	PURSUIT SC	31	89/292 (0.30)	0.30	20/292 (0.07)	0.07
Induction/ experienced	OCTAVE 1 + 2	32	29/124 (0.23)	0.22	1/124 (0.008)	0.03
Maintenance/ naïve	PURSUIT M	33	48/154 (0.31)	0.31	34/154 (0.22)	0.17
Maintenance/ experienced	GEMINI 1	34	6/38 (0.16)	0.15	2/38 (0.05)	0.06

Abbreviation: NMA, network meta-analysis

Note: [a] Clarification response, Appendix 2

3.4.2.5. Effect modification

The NMAs are to some extent protected from bias with respect to select potential effect modifiers by the company's approach. In the case of trial design (re-randomised versus treat-through), the company made a statistical adjustment to the treat-through trials (though the ERG notes some issues with this process). In the case of prior TNFi, the NMAs are conditioned on this factor i.e. they are analysed separately and prior TNFi is held fixed within each analysis (though the division between these levels (naïve/ experienced) may be somewhat blurred, see 3.3.2.3). Some other measures taken by the company are listed in Section 3.4.2.1.

The CS recorded baseline characteristics in Appendices Tables 13 and 14. This includes information on known prognostic or effect modifying factors mentioned by clinicians advising the

ERG, such as steroid use, duration of disease and age. There is wide variation between trials in some characteristics that may be effect modifiers, for example, extent of disease is plausibly related to treatment effect. A comment on the variation was requested in clarification A8. Placebo arm extensive disease in the induction phase ranges from 7.1% (VISIBLE1[73]) to 80.8% (Kobayashi et al. 2016[74]). In response, the company indicated that the range was reduced (7.1% to 56.2%) when excluding trials which only recruited Asian participants; the ERG noted that this remains a considerable range, a position clinical advice to the ERG confirmed. In relation to placebo arm CRP (mg/L) ranges from 3.2 (ULTRA1[65]) to 35.1 (Jiang et al. 2015[75]) in the induction phase, the company responded that clinicians in a previous appraisal described CRP measurements as 'non-specific' inflammatory marker. Clinical advice to the ERG confirmed that CRP has limited utility as marker outside of the acute severe UC context not covered in this appraisal, however, clinical experts noted that CRP levels <5 mg/L in the UNIFI[52] trial were predictive of response. Duration of disease appears to be fairly consistent when reported for the maintenance phase (ranging from 5.4 to 8.7 years) but much less so in the induction phase (ranging from 3.8 to 14.6 years). Concomitant steroid use is also relatively consistent during maintenance (ranging from 28% to 58%) but not the induction phase (ranging from 27% to 100%) (Document B, p.91; full trial-level data in CS Appendix D.4, Tables 13 and 14).

The ERG carried out random-effects meta-analyses of the response (no remission) and remission proportions in the placebo arms (data supplied in Clarification Response, Tables 31 to 34). The results are shown in Figure 1 to Figure 8, and the estimated I^2 from these analyses in Table 22.

Table 22: Estimates of heterogeneity (I^2) from placebo arms of the trials included in the CS, as calculated by ERG

	Estimate of I^2 (%) for remission	Estimate of I^2 (%) for response (no remission)
Induction/TNFi-naïve	44.5	28.8
Induction/TNFi-experienced	56.8	27.9
Maintenance/TNFi-naïve	50.6	65.0
Maintenance/TNFi-experienced	3.3	65.3

Abbreviation: TNFi, tumour necrosis factor inhibitor

Heterogeneity is substantial in many settings, most notably in the maintenance setting for response no remission. It is more moderate in the induction setting for response no remission,

and very low in the maintenance/experienced setting for remission. Unexplained heterogeneity in outcomes between trials might signal the influence of effect modifiers (known or unknown) and therefore potential bias in the NMA. Overall, there appears to be some reduction in heterogeneity compared to an earlier analysis of placebo arm outcomes in UC (Macaluso et al. 2018),[48] which could be attributed to the measures taken by the company (e.g. restricting time point of assessment, conditioning on TNFi experience, etc.)

Figure 1 : Placebo-arm, trial-specific response (no remission) proportions for the TNFi-naïve during induction NMA setting

Abbreviation: RE, random effects

Note: Data obtained from the company's clarification response, Appendix 2, Tables 31

Figure 2: Placebo-arm, trial-specific response (no remission) proportions for the TNFi-experienced during induction NMA setting

Abbreviation: RE, random effects

Note: Data obtained from the company's clarification response, Appendix 2, Tables 32

Figure 3: Placebo-arm, trial-specific response (no remission) proportions for the TNFi-naïve during maintenance NMA setting

Abbreviation: RE, random effects

Note: Data obtained from the company's clarification response, Appendix 2, Tables 33

Figure 4: Placebo-arm, trial-specific response (no remission) proportions for the TNFi-experienced during maintenance NMA setting

Abbreviation: RE, random effects

Note: Data obtained from the company's clarification response, Appendix 2, Tables 34

Figure 5: Placebo-arm, trial specific remission proportions for the TNFi-naïve during induction NMA setting

Abbreviation: RE, random effects

Note: Data obtained from the company's clarification response, Appendix 2, Tables 31

Figure 6: Placebo-arm, trial specific remission proportions for the TNFi-experienced during induction NMA setting

Abbreviation: RE, random effects

Note: Data obtained from the company's clarification response, Appendix 2, Tables 32

Figure 7: Placebo-arm, trial specific remission proportions for the TNFi-naïve during maintenance NMA setting

Abbreviation: RE, random effects

Note: Data obtained from the company's clarification response, Appendix 2, Tables 33

Figure 8: Placebo-arm, trial specific remission proportions for the TNFi-experienced during maintenance NMA setting

Abbreviation: RE, random effects

Note: Data obtained from the company's clarification response, Appendix 2, Tables 34

3.4.3. Relevance to the target population

The ERG considered the company's analyses to be broadly appropriate for the populations of interest, though it raised concerns about the generalisability of results stratified by prior TNFi experience, given the reality of the current treatment pathway is more complex. As such, the ERG is of the opinion that analyses stratified by biologic experience would have been more appropriate. The ERG also did not agree with the exclusion of tofacitinib, given its prominent role in the UK treatment landscape, as per clinical advice to the ERG.

There is a lack of published literature on effect modifiers in UC; as a result, the ERG considered that high variability in baseline characteristics of placebo arms of included trials may have represented imbalances in unknown effect modifiers, though it is known that treatment efficacy

varies widely between individuals based on demographic, medication use and clinical characteristics[48]. Furthermore, in line with NICE guidance[1], the ERG was of the opinion that an unweighted average of outcomes reported in the placebo arms of trials included in the NMA, the approach confirmed by the company in clarification response A15, was not appropriate. The ERG considered the use of placebo outcomes from studies that are highly generalisable to the UK-specific context and are identified through a proper, protocol-driven systematic review to be the most appropriate. However, given time constraints in this appraisal, the ERG selected placebo baseline risk from a single trial per TNFi experience and treatment setting dyad from the trials included in the NMA which matched the relevant dyad most closely for its base case.

3.4.4. Results of the indirect treatment comparison

The results of the company's base case NMAs are provided in the following sections, according to subgroups by prior TNFi experience and stratification by the induction and maintenance phases of the treatment.

3.4.4.1. TNFi-naïve participants (induction phase)

A summary of the results of the company's base case NMA for TNFi-naïve participants during the induction phase, comparing ozanimod to comparators and placebo, is presented in Table 23. Rank data taken from league tables provided during clarification indicated that ozanimod ranked 3rd (of 8) for both clinical response and clinical remission in this subgroup during induction; the ERG interpreted this with caution as no confidence intervals were reported around these ranks.

Furthermore, the results suggested that ozanimod was more likely to achieve clinical response and clinical remission during treatment induction of TNFi-naïve participants when compared to pooled doses of ustekinumab, tofacitinib, golimumab, adalimumab and placebo; the ERG noted that this effect was only statistically significant versus placebo. Ozanimod was out-performed by pooled doses of infliximab and vedolizumab for both outcomes, though the ERG noted that these effects were also non-significant.

Table 23: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-naïve participants during the induction phase

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Infliximab	Pooled	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Golimumab	200/100 mg SC	[redacted]	[redacted]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

Notes: [a] random effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Due to the wide credible intervals reported around the relative effect estimates for comparators, the ERG considered all comparators versus placebo as effects were expected to be more precise as a result of the weight and proximity of evidence relative to placebo. This comparison reflected ozanimod versus other treatments: pooled doses of infliximab and vedolizumab resulted in greater relative effect against placebo when compared to ozanimod, all other treatments had smaller relative effects against placebo; all results against placebo were statistically significant. The results of this comparison are reported in Table 24.

Table 24: NMA outcomes for comparators versus placebo in TNFi-naïve participants during the induction phase

Treatment	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Infliximab	Pooled	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Golimumab	200/100 mg SC	[redacted]	[redacted]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]

Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

Notes: [a] random effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour treatment, grey cells favour placebo.

3.4.4.2. TNFi-naïve participants (maintenance phase)

A summary of the results of the company's base case NMA for TNFi-naïve participants during the maintenance phase, comparing ozanimod to comparators and placebo, is presented in Table 25. Rank data taken from league tables provided during clarification indicated that ozanimod ranked 7th (of 9) for both clinical response and clinical remission in this subgroup during maintenance; the ERG interpreted this with caution as no confidence intervals were reported around these ranks.

The results suggested that ozanimod was out-performed by all treatments, with the exception of adalimumab and placebo, in terms of clinical response. Notably, pooled doses of vedolizumab as well as vedolizumab 108 mg resulted in significantly higher clinical response than ozanimod. Ozanimod was also out-performed by all treatments except adalimumab and placebo for clinical remission as an outcome. The ERG noted that pooled doses of vedolizumab and pooled doses of tofacitinib resulted in significantly higher clinical remission than ozanimod. The results of comparisons with all other active treatments for the two outcomes were statistically non-significant.

Table 25: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-naïve participants during the maintenance phase

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Infliximab	Pooled	[redacted]	[redacted]
Golimumab	Pooled	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

Abbreviations: BID, twice a day; CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

Notes: [a] fixed effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Due to the large number of non-significant relative effect estimates for comparators, the ERG considered all comparators versus placebo as effects were expected to be more precise as a

result of the weight and proximity of evidence relative to placebo. All treatments, with the exception of adalimumab, resulted in greater relative effect against placebo when compared to ozanimod; all results against placebo were statistically significant. The results of this comparison are reported in Table 26.

Table 26: NMA outcomes for comparators versus placebo in TNFi-naïve participants during the maintenance phase

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Infliximab	Pooled	[redacted]	[redacted]
Golimumab	Pooled	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]

Abbreviations: BID, twice a day; CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

Notes: [a] fixed effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour treatment, grey cells favour placebo.

3.4.4.3. TNFi-experienced participants (induction phase)

A summary of the results of the company's base case NMA for TNFi-experienced participants during the maintenance phase, comparing ozanimod to comparators and placebo, is presented in Table 27. Rank data taken from league tables provided during clarification indicated that ozanimod ranked 2nd (of 6) for both clinical response and clinical remission in this subgroup during induction; the ERG interpreted this with caution as no confidence intervals were reported around these ranks.

Results suggested that ozanimod was more likely to achieve clinical response and clinical remission during treatment induction of TNFi-experienced participants when compared to pooled doses of ustekinumab, vedolizumab, adalimumab and placebo; the ERG noted that this effect was only statistically significant versus adalimumab and placebo. Ozanimod was out-performed by tofacitinib for both outcomes, though the ERG noted that these effects were non-significant.

Table 27: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-experienced participants during the induction phase

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; TNFi, tumour necrosis factor inhibitor

Notes: [a] fixed effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Due to a number of non-significant relative effect estimates for comparators and wide credible intervals for adalimumab, the ERG considered all comparators versus placebo as effects were expected to be more precise as a result of the weight and proximity of evidence relative to placebo. All treatments resulted in greater relative effect against placebo when compared to ozanimod; all results against placebo were statistically significant with the exception of adalimumab. The results of this comparison are reported in Table 28.

Table 28: NMA outcomes for comparators versus placebo in TNFi-experienced participants during the induction phase

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]

Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; TNFi, tumour necrosis factor inhibitor

Notes: [a] fixed effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour treatment, grey cells favour placebo.

3.4.4.4. TNFi-experienced participants (maintenance phase)

The results of the company's base case NMA for TNFi-experienced participants during the maintenance phase, comparing ozanimod to comparators and placebo, are summarised in Table 29. Rank data taken from league tables provided during clarification indicated that

ozanimod ranked 4th (of 7) for both clinical response and clinical remission in this subgroup during maintenance; the ERG interpreted this with caution as no confidence intervals were reported around these ranks.

Results suggested that ozanimod was more likely to achieve clinical response and clinical remission during treatment maintenance of TNFi-experienced participants when compared to pooled doses of ustekinumab, adalimumab and placebo; the ERG noted that this effect was only statistically significant versus placebo. Ozanimod was out-performed by pooled doses of tofacitinib, pooled doses of vedolizumab and vedolizumab 108 mg for both outcomes, though the ERG noted that these effects were non-significant.

Table 29: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-experienced participants during the maintenance phase

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Ustekinumab	Pooled	[redacted]	[redacted]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

Notes: [a] fixed effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Due to the large number of non-significant relative effect estimates for comparators, the ERG considered all comparators versus placebo as effects were expected to be more precise as a result of the weight and proximity of evidence relative to placebo. Pooled doses of vedolizumab as well as tofacitinib and vedolizumab 108 mg resulted in greater relative effect against placebo when compared to ozanimod, with pooled doses of ustekinumab and adalimumab resulting in smaller relative effects against placebo when compared to ozanimod. The ERG noted that all results against placebo were statistically significant, but that large imprecision was still observed. The results of this comparison are reported in Table 30.

Table 30: NMA outcomes for comparators versus placebo in TNFi-experienced participants during the maintenance phase

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Ustekinumab	Pooled	[redacted]	[redacted]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

Notes: [a] fixed effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour treatment, grey cells favour placebo.

3.4.4.5. Sensitivity analyses

Sensitivity analyses reported by the company comprised an assumption of not pooling doses of the same active treatment if it had the same method of administration, the exclusion of trials with a treat-though design, using the three-component Mayo score instead of the four-component Mayo score in the TRUENORTH trial, the exclusion of trials conducted exclusively in Asian populations, and the inclusion of TOUCHSTONE[34] in the TNFi-naïve analysis.

The ERG generally agreed with the company's assessment that the results of sensitivity analyses demonstrated robustness of the base case NMA for the factors that were explored, but noted that wide confidence intervals in the base case, in particular for the TNFi-experienced subgroup, made it difficult to identify meaningful differences across analyses. Of note, the sensitivity analyses using three-component Mayo scores for TRUENORTH indicated large shifts in the point estimates for ozanimod relative to placebo, particularly in the TNFi-experienced subgroup. The ERG also noted higher point estimates, on average, for vedolizumab, infliximab, golimumab and adalimumab when Asian trials were excluded – this was considered to be an expected effect given the investigation of these treatments by the five trials including Asian participants only, i.e. Motoya,[51] Kobayashi,[74] Jiang,[75] PURSUIT J[63] and Suzuki.[47]

The ERG also did not consider all uncertainties in the NMA to have been addressed by the sensitivity analyses, though it did note an exploration of the effect of re-calculated data from treat-through trials, which was identified as a source of considerable uncertainty. It conducted a

thorough comparative assessment of the sensitivity analysis excluding data from treat-through trial designs, reported in Section 3.5.2.

3.4.5. Conclusions on the indirect treatment comparison

The company carried out NMAs in four settings combining trial phase (maintenance/induction) with prior TNFi treatment (experienced/naïve) using a modelling approach (multinomial with probit link) that was very appropriate in the ERG's view. A further strength of the submission was in the form of measures the company took to counter or reduce potential effect modification from factors including trial phase, prior TNFi treatment, trial design (re-randomised/treat-through), outcome definition and timepoint of assessment. However, evidence of further variation in potential effect modifiers is apparent in baseline characteristics (e.g. extensive disease, see Section 3.4.2.5) and placebo arm responses (see Figure 1 through Figure 8Figure 5). The company also mainly used FE models when RE models are more appropriate (given heterogeneity), as the latter would not converge – the ERG recommends using informative priors as a potential remedy as FE models do not account adequately for the observed between-trial heterogeneity. Furthermore, the ERG considered the company's approach in using all available trials to inform the baseline placebo risk to be a limitation, and recommends instead a systematic review to select the most appropriate sources for the UK context and rerun the NMAs.

Results of the NMA indicated that ozanimod was a ranked in the top three treatments for the induction phase in both TNFi-naïve and -experienced patients, though only placebo was significantly outperformed in both subgroups; adalimumab was additionally significantly outperformed in the TNFi-experienced subgroup during the induction phase. Comparison with all other active treatments yielded non-significant results; in particular, the ERG noted that tofacitinib was non-significantly outperformed by ozanimod in the TNFi-naïve subgroup, and non-significantly outperformed ozanimod in the TNFi-experienced subgroup. These results indicate that ozanimod is a moderately effective treatment, that considerable uncertainty exists around its relative treatment effect and that tofacitinib is a comparable treatment in terms of efficacy in the induction phase (approximately 6 to 14 weeks), regardless of prior TNFi experience.

In the maintenance phase, results of the NMA show that ozanimod was ranked in the lowest three for the TNFi-naïve subgroup and middle of the range for the TNFi-experienced subgroup. In the former, ozanimod only significantly outperformed placebo and was significantly

outperformed by pooled doses of vedolizumab (for clinical response and remission), vedolizumab 108 mg (for clinical response only) and tofacitinib (for clinical remission only). Comparisons with all other active treatments were non-significantly in favour of the comparator. These results indicate that ozanimod may be a less efficacious treatment for the maintenance of TNFi-naïve patients, though considerable uncertainty exists about its relative treatment effect. The ERG also noted that results suggested tofacitinib may be a more efficacious treatment compared to ozanimod in this setting. In the TNFi-experienced subgroup, ozanimod significantly outperformed placebo and non-significantly outperformed ustekinumab and adalimumab, with all other comparators non-significantly outperforming ozanimod. The results suggest that considerable uncertainty exists around the relative efficacy of ozanimod against comparators for the maintenance of TNFi-experienced patients.

3.5. Additional work on clinical effectiveness undertaken by the ERG

3.5.1. Additional searches

As described in Sections 3.1 and 3.3.3, the ERG did not consider the company's exclusion of evidence from phase 4 studies from the submission to be appropriate. Consequently, the ERG carried out additional searches for phase 4 trials reporting on ozanimod and its comparators for moderately to severely active UC. Searches for phase 4 studies of ozanimod were conducted in Ovid MEDLINE, Ovid Embase, clinicaltrials.gov, the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) and the European Union (EU) Clinical Trials Register; Ovid MEDLINE and Embase were searched from 1946 and 1974, respectively (Appendix A). These searches yielded no results, with only one trial investigating ozanimod for multiple sclerosis identified.

Searches for Phase 4 studies of relevant comparators were conducted in Ovid MEDLINE and Ovid Embase, from 1946 and 1974, respectively (Appendix A). The searches yielded 28 potentially eligible records, though screening of this yield identified no studies reporting real-world evidence for adalimumab, tofacitinib or ustekinumab, filgotinib, etrasimod or etrolizumab. The ERG noted that filgotinib is currently under appraisal by NICE for moderately to severely active UC, and that etrasimod and etrolizumab do not currently have FDA or EMA approval – it therefore considered that these were not comparators of interest for this appraisal and that it would not have been possible for evidence on these treatments to inform links in NMA networks.

A total of 13 publications reporting results of phase 4 trials investigating golimumab were identified. Nine records, all related to the GO-COLITIS trial (NCT02092285), reported results for TNFi-naïve populations; four records reported results for participants with mixed TNFi experience, with two of these related to the GORE-UC trial,[76] one related to GO-LEVEL (NCT03124121)[77] and another related to a trial by Yu et al. (2021).[78] These trials were all considered ineligible as they described single-group assignment to golimumab.

Ten publications reporting phase 4 results for infliximab were identified; six were related to the NOR-SWITCH trial (n=4; NCT02148640) and its open label extension (n=2), two related to the NOR-DRUM trial (NCT03074656), one to the SECURE[79] trial and one to the trial by Park et al. (2015).[80] All trials were assumed to describe results in participants with mixed TNFi experience, though none stated this explicitly. The trials were all considered ineligible by the ERG: NOR-SWITCH and the SECURE trial both compared infliximab with a biosimilar; NOR-DRUM compared infliximab plus standard of care with infliximab plus TDM and the trial by Park et al. (2015)[80] described single-group assignment to an infliximab biosimilar.

Five publications reporting phase 4 results for vedolizumab were obtained during the searches. Two publications were related to a trial by Coletta et al. (2020)[81] (Eudract number 2015-003270-32) and one each to the trials by Danese et al. (2021)[82] (NCT02743806), Osterman et al. (2020)[83] and Vermeire et al. (2020).[84] As for infliximab, all trials were assumed to describe results in participants with mixed TNFi experience, though this was not explicit. All trials were considered ineligible as the trial by Coletta et al. (2020)[81] and Danese et al. (2021)[82] described single-group assignment, the trial by Vermeire[84] investigated single-group de-escalation of vedolizumab dosing and the trial by Osterman et al. (2020)[83] compared serum vedolizumab concentration in responders and non-responders.

As such, the ERG did not consider that the exclusion of phase 4 evidence from the CS meaningfully changed the results of the NMA or conclusion of the submission, though the methodological bias this approach could introduce is reiterated.

3.5.2. Validation of robustness of NMAs including treat-through trial data

As described in Section 3.3.2, the ERG considered the combination of re-calculated treat-through trial data with data from re-randomised trials to be a potential source of heterogeneity and bias in the base case NMA. Therefore, the sensitivity analysis of each comparator relative

to placebo was compared with the corresponding base case result; these are summarised in Table 31 and Table 32.

Table 31: Comparison of treatment effect relative to placebo between the NMA base case and sensitivity analysis excluding treat-through trial data: TNFi-naïve population in the maintenance phase

Comparator vs. placebo	Clinical response estimate; OR (95% CrI)		Clinical remission estimate; OR (95% CrI)	
	Base case [a]	Sensitivity analysis [a]	Base case [a]	Sensitivity analysis [a]
Tofacitinib 10 mg BID	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab 108 mg Q2W SC	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab pooled	[redacted]	[redacted]	[redacted]	[redacted]
Ustekinumab pooled	[redacted]	[redacted]	[redacted]	[redacted]
Golimumab pooled	[redacted]	[redacted]	[redacted]	[redacted]
Infliximab pooled	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]
Adalimumab 40 mg Q2W	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BID, twice a day; CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor; vs., versus

Notes: [a] fixed effects NMA. Blue cells represent statistically significant results

Table 32: Comparison of treatment effect relative to placebo between the NMA base case and sensitivity analysis excluding treat-through trial data: TNFi-experienced population in the maintenance phase

Comparator vs. placebo	Clinical response estimate; OR (95% CrI)		Clinical remission estimate; OR (95% CrI)	
	Base case [a]	Sensitivity analysis [a]	Base case [a]	Sensitivity analysis [a]
Vedolizumab 108 mg Q2W SC	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib pooled	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab pooled	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1 mg QD	[redacted]	[redacted]	[redacted]	[redacted]
Adalimumab 40 mg	[redacted]	[redacted]	[redacted]	[redacted]

Comparator vs. placebo	Clinical response estimate; OR (95% CrI)		Clinical remission estimate; OR (95% CrI)	
	Base case [a]	Sensitivity analysis [a]	Base case [a]	Sensitivity analysis [a]
Ustekinumab pooled	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor; vs., versus

Notes: [a] fixed effects NMA; blue cells represent statistically significant results

The ERG considered the comparison between the NMA base case and the sensitivity analysis excluding treat-through data to indicate that re-calculated treat-through data did not meaningfully bias the relative point estimates, with the majority of relative risks differing by less than 0.2 between the base case and sensitivity analysis in both subgroups related to prior TNFi experience. The only exception was for comparisons of golimumab with placebo in the TNFi-naïve subgroup, where base case estimates were 0.6 lower compared to estimates derived from sensitivity analyses.

Furthermore, the ERG noted that the 95% credible intervals around relative effect estimates were considerably wider for sensitivity analyses conducted in the TNFi-naïve subgroup, and fewer results were nominally significant as a result. The ERG found this plausible and likely attributable to the fact that three of the four treat-through trials, i.e. ACT1,[44] ULTRA2[43] and Suzuki (2014),[47] were conducted in exclusively TNFi-naïve populations; their exclusions therefore resulting in a very sparse network and considerable imprecision around the point estimate. Given that only the VARSITY[60] trial reported treat-through data for the TNFi-experienced population, the ERG noted that base case precision and nominal significance was generally retained.

3.5.3. Validation of company NMAs

The company supplied JAGS model code in CS Appendix D.4.3, but no accompanying data (either coded or within the CS). In clarification question A6 the ERG requested code for data setup and execution of the NMA. Although in response the company supplied data and setup information, this was not provided in an executable form. It was therefore necessary for the ERG to finalise coding itself, including reconfiguring response data and initial values provided as a printout or list. The procedure to obtain stable estimates as described in the CS (Document B, p.97) was not fully specified and the ERG could not implement it. In addition, the ERG noted that the code supplied as the RE model (company's clarification response, Figure 12) was for

the FE case. As a result, implementation of JAGS code by the ERG was time-consuming, resulting in limited latitude for revisions of the base case NMA or further NMA scenarios. In spite of this, coding was largely successful, and the ERG were able to closely, but not identically, replicate the NMA results provided by the company.

3.5.3.1. ERG base case NMA using alternative baseline placebo risk values

As discussed in Section 3.4.2.4, the ERG considered the use of a more generalisable study to the UK context to inform the placebo baseline risk in the NMA to be a more appropriate approach than the unweighted average of all placebo arms used by the company. The ERG included the NMA estimates generated using these values in its revised base case, though the limitations associated with this approach are acknowledged in Section 3.4.2.4. Time constraints precluded the ideal approach, i.e. using highly generalisable studies identified through a protocol-driven systematic review. The trials selected for the ERG base case NMA, as well as the baseline characteristics of participants in the placebo arm of each trial, are summarised in Table 20.

For its base case, the ERG modified the company JAGS code to select the placebo arm of a single trial in each setting (whereas the company's code averaged all placebo arms) as data for estimation of baseline risk. The process for estimation of relative effects was unchanged and estimates of OR were similar between the ERG and company base case. The pattern of convergence was similar to that reported by the company, and the same model choices were applied, namely RE for the induction setting in TNFi-naïve participants, and FE otherwise. Estimated probabilities of being in each response category by treatment and setting are affected; the results of the ERG base case NMAs, and comparative ERG-replicated company NMA results, are summarised in Figure 9 and Table 33.

Figure 9: Visualisation of results of ERG base case NMAs with revised placebo baseline risk compared to ERG-replicated results of company base case NMAs

Abbreviations: ADA, adalimumab; CS, company submission; ERG, evidence review group; GOL, golimumab; IFX, infliximab; ind/naïve, TNFi-naïve subgroup during induction; ind/exp, TNFi-experienced subgroup during induction; maint/naïve, TNFi-naïve subgroup during maintenance; maint/exp, TNFi-experienced during maintenance; no_resp, no response; OZA, ozanimod; partial_resp, partial response (response no remission); PBO, placebo; TOF, tofacitinib; UST, ustekinumab; VEDO, vedolizumab; VEDO 108, vedolizumab 108 mg Q2W SC

Table 33: Numerical results of ERG base case NMAs with revised placebo baseline risk compared to ERG-replicated results of company base case NMAs

Treatment	Setting (TNFi experience/treatment phase)	Response		Remission		No response	
		Company	ERG	Company	ERG	Company	ERG
PBO	Naïve/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PBO	Experienced/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PBO	Naïve/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PBO	Experienced/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OZA	Naïve/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OZA	Experienced/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OZA	Naïve/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OZA	Experienced/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	Naïve/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	Experienced/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	Naïve/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	Experienced/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	Naïve/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	Naïve/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX	Naïve/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX	Naïve/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF	Naïve/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF	Experienced/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF	Naïve/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOF	Experienced/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
UST	Naïve/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
UST	Experienced/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
UST	Naïve/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
UST	Experienced/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VEDO	Naïve/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VEDO	Experienced/induction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VEDO	Naïve/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VEDO	Experienced/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VEDO 108	Naïve/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
VEDO 108	Experienced/maintenance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ADA, adalimumab; ERG, evidence review group; GOL, golimumab; IFX, infliximab; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab; VEDO, vedolizumab; VEDO 108, vedolizumab 108 mg Q2W SC

The comparison of results from the ERG base case NMAs with results from the company NMAs indicate that the company approach estimated a higher response for most treatments, with the 'no response' outcome higher for placebo and almost all active treatments in the ERG NMAs. The ERG noted that these findings are indirectly validated by clinical advice to the ERG, which suggested that remission and response in the placebo arms of trials included in the company NMA are higher than expected. The ERG concluded that using placebo baseline risks from more generalisable studies represented a more conservative base case.

3.5.3.2. Re-estimation of company NMA using random effects modelling

After closely replicating the NMA results provided by the company as described in Section 3.5.3.1, the ERG attempted to re-run the NMA using RE models with alternative baseline placebo values for the three analyses where FE were used (see Section 3.4.2.2). Clear non-convergence, broadly in line with what was observed in the company base case, was detected for the odds ratios of these analyses. As described in Section 3.4.2.2, the ERG considered that the use of an appropriate informative prior distribution, e.g. those reported in Turner et al. (2012)[67] or Turner et al. (2015),[68] could resolve this problem.

3.6. Conclusions of the clinical effectiveness section

Based on the evidence presented in the CS for the pivotal TRUENORTH[27,28] trial and the supplementary TOUCHSTONE.[34] trial, as summarised in Section 3.2.5.1, the ERG concluded that ozanimod has a significant effect on the outcomes of clinical remission and clinical response in both the induction and maintenance phases in the overall population, when compared to placebo. Furthermore, ozanimod resulted in significant improvements in other categories of remission (maintenance of remission, durable remission and corticosteroid-free remission), endoscopic healing, mucosal healing and measures of disease activity compared to placebo. The results of the effect of ozanimod on HRQoL were more variable, with some domains showing no significant change compared to placebo. Furthermore, the ERG noted that the proportion of various adverse events were higher for ozanimod compared to placebo, though no formal tests of significance are reported. As a result, the ERG concluded that ozanimod is an effective treatment compared to placebo, though its effect on quality of life and its safety are uncertain. These results were mostly reflected in the clinical results reported for subgroup analyses by TNFi experience, though the ERG noted that effects were smaller in the TNFi-experienced subgroup. The company posited that this was due to TNFi-experienced patients being more difficult to treat, a position the ERG agreed with.

As discussed in Section 3.4.5, the results of the indirect treatment comparison showed that ozanimod is a moderately effective treatment during the induction phase of treatment, regardless of TNFi experience, but that considerable uncertainty exists around its relative treatment effect. The ERG further noted that tofacitinib is a comparable treatment to ozanimod in terms of efficacy in the induction phase. The results of the indirect treatment comparison for the maintenance phase further indicated that ozanimod may be a less efficacious treatment for the maintenance of TNFi-naïve patients, though considerable uncertainty exists about its relative treatment effect, and that tofacitinib may be a more efficacious treatment compared to ozanimod in this setting. Furthermore, in the TNFi-experienced subgroup, results suggested that considerable uncertainty exists around the relative efficacy of ozanimod against comparators for the maintenance of patients with prior TNFi experience.

4. COST-EFFECTIVENESS

4.1. ERG comment on company's review of cost-effectiveness evidence

The company undertook a SLR to identify evidence for outcomes relevant to the cost-effectiveness, HRQoL, healthcare resource use (HCRU) and cost of ozanimod and comparator treatments for the treatment of moderate to severe UC.

Table 34. Summary of ERG's critique of the methods implemented by the company to identify cost-effectiveness evidence

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix G.2.	The ERG noted the following limitations: a recognised filter for identifying cost-effectiveness studies was not used; database searches applied limits that excluded conference abstracts from search results, however, hand searching and database searches of known conference proceedings may have mitigated this issue. Missing search strategies were provided in response to clarification question B1.
Inclusion criteria	Appendix G.3	The ERG notes that studies reporting 'primarily clinical outcomes' are excluded. It is not clear whether this may involve the exclusion of studies reporting both clinical and cost-effectiveness evidence. Despite this, the ERG considered the inclusion criteria to be broadly appropriate to encompass the cost-effectiveness evidence for all the relevant comparators to this appraisal.
Screening	Appendix G.3	Title and abstract screening was conducted by two reviewers, with a third available to resolve discrepancies. The same procedure was followed at full-text screening. Where studies gave insufficient information at full-text screening, they were excluded. It is not clear how this was applied and whether this may have led to relevant studies being excluded. In addition, updates were made to the SLR during screening. While the impact of these factors is unclear, the ERG considers

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		the screening methods to be broadly appropriate.
Data extraction	Appendix G.4	It is unclear why detailed data extractions were done in US studies while those from other countries underwent less extensive extraction. Despite this, the ERG considers the data extraction to be broadly appropriate.
QA of included studies	Appendix G.4	QA was completed using the Drummond checklist, as recommended by NICE. Therefore the ERG considers the QA to be appropriate.

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; HRQoL, health-related quality of life; NICE, National Institute for Health and Care Excellence; QA, quality assessment

Table 35. Summary of ERG's critique of the methods implemented by the company to identify health related quality of life

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix H.2.	The ERG noted the following limitations: a recognised filter for identifying health-related quality of life studies was not used; database searches applied limits that excluded conference abstracts from search results, however, hand searching and database searches of known conference proceedings may have mitigated this issue. Missing search strategies were provided in response to clarification question B1.
Inclusion criteria	Appendix H.3.	Studies with <50 patients of interest are excluded though it is unclear why this is the case and whether this may have excluded relevant studies. Besides this, the ERG considers the inclusion criteria to be broadly appropriate to capture HRQoL studies relevant to UC.
Screening	Appendix H.3.	Title and abstract screening was conducted by two reviewers, with a third available to resolve discrepancies. The same procedure was followed at full-text screening. Where studies gave insufficient information at full-text

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		screening, they were excluded. It is not clear how this was applied and whether this may have led to relevant studies being excluded. In addition, updates were made to the SLR during screening. While the impact of these factors is unclear, the ERG considers the screening methods to be broadly appropriate.
Data extraction	Appendix H.4.	Data from all studies was extracted in detail regardless of geographic region. The ERG considers the methods to be appropriate to have extracted all relevant data.
QA of included studies	Appendix H.4.	Methods of QA of the included studies are not described.

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; HRQoL, health-related quality of life; QA, quality assessment; SLR, systematic literature review; UC, ulcerative colitis

Table 36. Summary of ERG's critique of the methods implemented by the company to identify healthcare resource use and costs

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix G.2.	Searches for cost and resource use studies and cost-effectiveness studies were conducted in a single SLR. The ERG noted the following limitation: database searches applied limits that excluded conference abstracts from search results. The ERG conducted additional searches of Ovid MEDLINE (reported in Appendix A) to retrieve studies that may have been missed. Following screening of additional search results the ERG was satisfied all relevant studies had been identified. Missing search strategies for bibliographic database searches were provided in response to clarification question B1.
Inclusion criteria	Appendix G.3.	The ERG notes that studies reporting 'primarily clinical outcomes' are excluded. It is not clear whether this may involve the exclusion of studies reporting both clinical and resource use and cost evidence. Despite this, the ERG considered the inclusion criteria to

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		be broadly appropriate to encompass the resource use and cost evidence for all the relevant comparators to this appraisal.
Screening	Appendix G.3.	Title and abstract screening was conducted by two reviewers, with a third available to resolve discrepancies. The same procedure was followed at full-text screening. Where studies gave insufficient information at full-text screening, they were excluded. It is not clear how this was applied and whether this may have led to relevant studies being excluded. In addition, updates were made to the SLR during screening. While the impact of these factors is unclear, the ERG considers the screening methods to be broadly appropriate.
Data extraction	Appendix G.4.	It is unclear why only US and UK studies were extracted in detail, particularly for resource use data. Besides this, the data extraction is considered acceptable by the ERG.
QA of included studies	Appendix G.4.	The description of the methods for QA of resource use and costs is not completely clear, however the CS states that each study was compared to the NICE reference case, suggesting that QA will have been conducted appropriately.

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; HRQoL, health-related quality of life; NICE, National Institute for Health and Care Excellence; QA, quality assessment; SLR, systematic literature review

4.2. Summary and critique of company's submitted economic evaluation by the ERG

4.2.1. NICE reference case checklist

The NICE reference case checklist with regards to aspects of the appraisal, as well as the ERG's comment on the company's submission, is summarised in Table 37.

Table 37: NICE reference case checklist

Attribute	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	QALYs were estimated for patients. Carer disutility was not included in the analysis.
Perspective on costs	NHS and PSS	NHS and PSS as appropriate.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	The company submitted a cost utility analysis and presented both pairwise results and a fully incremental analysis.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	A lifetime horizon was used in the base case analysis. The ERG considered this to be appropriate.
Synthesis of evidence on health effects	Based on systematic review	Clinical data used in the economic model for both the treatment naïve and treatment experienced subgroups were primarily derived from the induction and maintenance NMAs conducted by the company. For the extended induction scenario analysis, clinical efficacy data were based on individual trial arms.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	QALYs were used as appropriate.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Utility values used in the base case were derived from published literature (Woehl et al.[29] and Arsenau et al.[30]). The ERG noted that QoL data were collected in the TrueNorth study using the EQ-5D-5L (which were cross-walked to EQ-5D-3L). These values were used in a company scenario analysis.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	The paper by Dolan (1997)[85] was used and was considered to be appropriate.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	There were no equity concerns.

Attribute	Reference case	ERG comment on company's submission
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	NHS reference costs, previous NICE appraisals and published literature were used to estimate costs and resource use.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Costs and outcomes were discounted at 3.5% as appropriate.

Key: EQ-5D, European Quality of Life Five Dimension; HRQoL: health-related quality of life; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; PSS, Personal Social Services; QALY, quality-adjusted life year; QoL, quality of life; TA, technology appraisal

4.2.2. Model structure

The company submitted a de novo hybrid decision analytic model which consisted of two parts (active treatment and post active treatment). The ERG noted that the model is a Markov-based model that utilised tunnel states within the Markov cohort trace for the induction period. Based on a review of previous UC appraisals i.e. ustekinumab TA633[20] and tofacitinib TA547[21], models were characterised by the use of decision trees (for the induction period) and a Markov (for the maintenance/post active treatment periods). Conceptually, a decision tree can be modelled within a Markov model framework as a set of tunnel states. In addition, the company stated (p141 of the CS) that the use of tunnel states has the added benefit *'of allowing patients to enter the maintenance phase at any cycle, therefore enabling the variable length of induction periods between treatments'*. Furthermore, the use of tunnel states was stated to capture '*the effective decision tree at the end of the induction period'* thereby determining the initial health state distribution for the maintenance period. As such, the tunnel state approach is equivalent to a decision tree and, as it explicitly includes a time dimension, allows for more accurate modelling of the varying induction lengths of different treatments. For completeness, the company provided additional rationale for the use of tunnel states upon request from the ERG (see as B5 in the clarifaction document). The ERG considered the company's justifcation to be reasonable and that the approach did not appear to introduce any bias into the analysis.

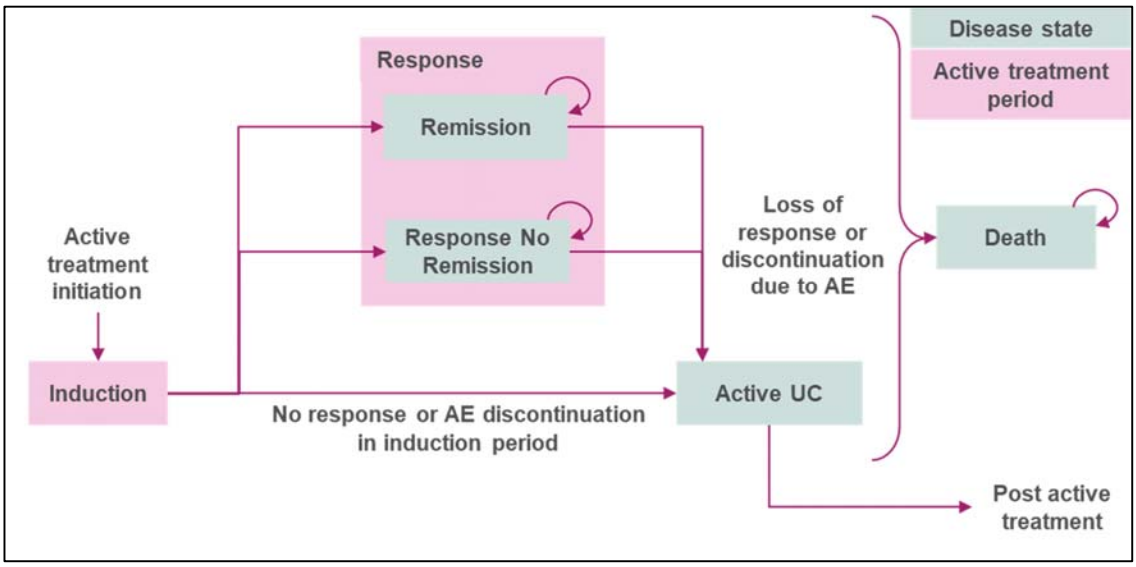
4.2.2.1. Active treatment period

The active treatment portion of the model consisted of both an induction and maintenance phase. All patients entered the model by initiating active treatment and progressed through a series of tunnel states (which reflected the specific induction period for each treatment). Once patients reached the final induction tunnel state, they were distributed into one of three health

states, 'Remission', 'Response (No remission)' and 'Active UC' (Figure 10). The probability of transitioning between key health states was derived from the NMAs outlined in Section 3.4.4 . See Section 4.2.6 for further discussion.

During induction patients could discontinue treatment due to serious adverse events or be absorbed by the 'Death' state. Patients distributed into the 'Remission', 'Response (No remission)' health states were assumed to remain in these states until they lose their initial response, discontinue due to adverse events or die. This is active treatment maintenance phase.

Figure 10: Active treatment model structure (without extended induction)



Abbreviation: AE, adverse events

4.2.2.2. Post active treatment phase

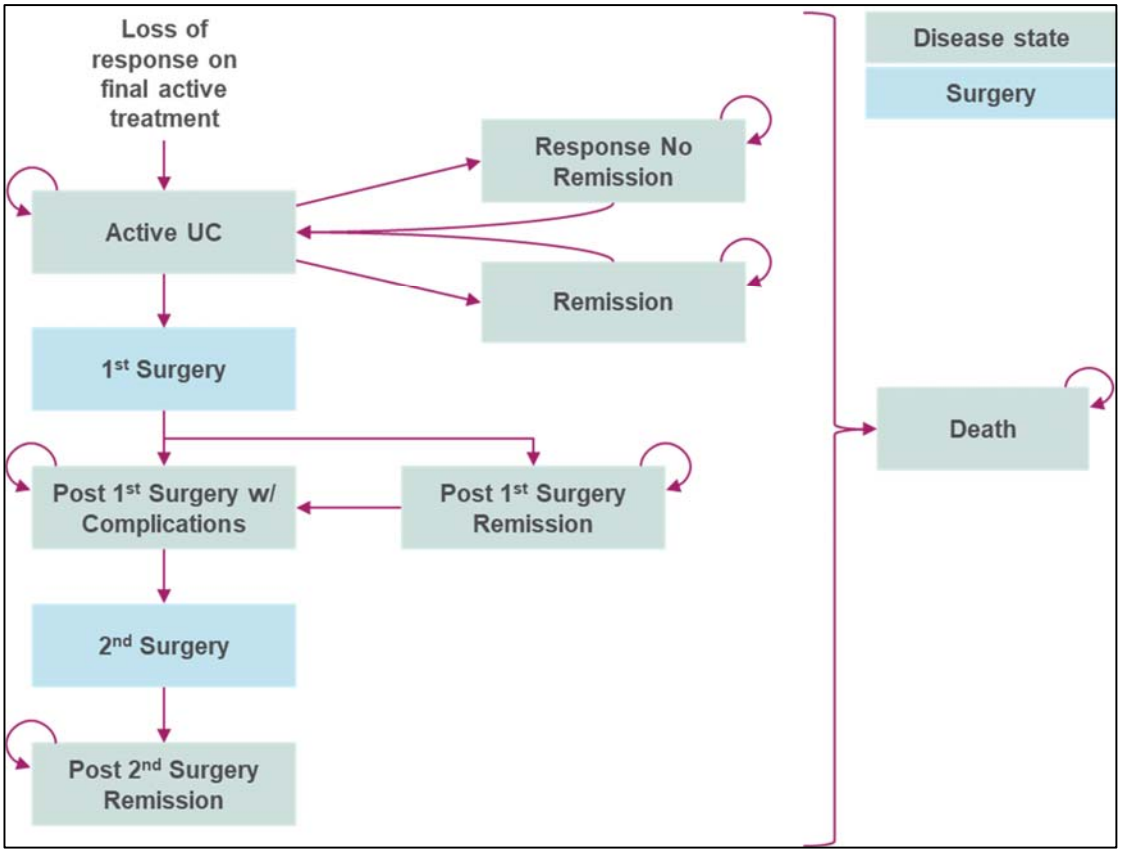
Patients that discontinued active treatment due to AEs, loss of response, or failure to achieve response entered the post active treatment component of the model and were assumed to initially enter the 'Active UC' health state. This part of the model consisted of 9 health states including 'Remission', 'Response (No remission)', 'Active UC', '1st Surgery', 'Post 1st Surgery Remission', 'Post 1st Surgery Complications', '2nd Surgery', 'Post 2nd Surgery Remission' and 'Death'. The ERG confirmed that these health states have been used and accepted in previous ustekinumab TA633[20] and are considered to accurately reflect the nature of the condition and key clinical events. In TA633,[20] the ERG had highlighted that a major limitation with the model structure was the non-inclusion of remission and response health states in the post active

treatment phase as not all patients follow an active form of the disease. This limitation has been addressed in the current model structure submitted by the company, as it considers remission and response health states following the 'Active UC' health state. In addition, the model allows for the spontaneous response and remission as well in the 'Active UC' health state.

The ERG noted that the modelled spontanous remission rate of 0.5% per model cycle (12% per year), was not based on clinical data, but was an arbitrary value chosen by the company to align with NICE committee preferences in TA633.[20] In TA633, the NICE committee stated that *'there is likely to be a small number of people who improve without treatment'* and generally preferred a low spontaneous remission rate, closer to the company's original modelled estimate of 0%. In order to validate the spontaneous remission rate, the ERG sought clinical input. Based on clinical expert responses, spontaneous remission was considered plausible for patients who no longer received active treatment. The rate of spontaneous remission in clinical practice, was somewhat variable i.e. between 5% to 30% per flare of active disease. In order to explore uncertainty, the company conducted scenario analyses in which 0% and 1% rates of spontaneous response were tested. The ERG has additionally conducted a scenario analysis which used a higher rate of spontaneous response compared to the company's base case estimate (reflective of 0.75% per model cycle). It should be noted that this analysis may lead to implausibly high spontaenous remission rates over the modelled time horizon and therefore should be interpreted with caution. See Section 6.1.5 for results.

Finally, it should be noted that in the post active model component, patients were assumed to receive best supportive care, comprising components of conventional therapy (See Section 4.2.8.3 for the list of CvT treatments provided).

Figure 11. Post active treatment model structure



Abbreviation: UC, ulcerative colitis

4.2.2.3. Subsequent treatments/Treatment sequencing

The base case analysis did not include subsequent treatments. Although the company acknowledged that it is possible for patients in the TNFi-naïve and TNFi-experienced subgroups to receive subsequent treatment (based on clinical expert feedback), subsequent treatments were not considered in the base case due to the lack of robust efficacy data available and uncertainty surrounding the treatments patients are likely to receive. The ERG noted that the exclusion of subsequent treatments is in line with TA633[20] (although the impact of including subsequent treatment was tested in a scenario analysis by the ERG). In TA547,[21] the base case did not consider subsequent treatments, however the model was flexible enough to allow the ERG to conduct a scenario analyses in the TNFi-naïve subgroup whereby various treatment sequencing strategies were explored including within class switching amongst TNFis as well as step up and step down approaches (outside class). The ERG noted in this appraisal that the results were not especially sensitive to this analysis. Overall, the ERG considered the impact of

treatment sequencing to have a moderate impact on costs, but minor impact on QALYs gained, as outlined in a published study by Wu et al. (2018)[86], which assessed the cost effectiveness of TNFis for the treatment of UC (from a UK and Chinese perspective). In this study, which compared 14 different treatment sequencing strategies, the total QALY gained for each strategy were largely similar (ranging between 10.49 to 12.37).

Within the current appraisal for ozanimod, the company provided limited scenario analysis which allowed for subsequent treatment usage in the TNFi-naïve subgroup (this was not conducted for the TNF-experienced subgroup as the company stated that there was a lack of available data to inform efficacy and clinical opinion to the company noted that treatments provided after failure on multiple biologics were likely to be patient dependent and variable). The scenario analysis allowed for the modelling of either vedolizumab or ustekinumab as plausible subsequent treatment options (after having received ozanimod or TNFis first line). The ERG noted that this did not have a meaningful impact on base case cost-effectiveness results. For completeness, the ERG considered undertaking additional scenario analyses using alternative subsequent treatments/sequencing options, however the model lacked the flexibility to conduct this. The ERG acknowledged that the exclusion of subsequent treatments from the base case analysis is an area of structural uncertainty, as it was not possible to adequately test via scenario analysis. Given the small differences in costs between treatments, incorporating treatment sequencing may have considerable impact on the base case results.

4.2.2.4. Extended induction

The ERG noted that extended induction was not considered as part of the company's base case analysis on the basis that it is not standard clinical practice in the UK for all treatments and further noted limitations associated with using trial data to inform the patient distributions into the health states (see p.154 of the CS). However, the company did provide a scenario analysis which explored the impact of including extended induction based on the SmPC for each treatment (See Section 5.2.3). Clinical input to the ERG, confirmed that extended induction does occur in clinical practice, albeit it is treatment dependent and therefore highly variable, i.e. if patients do not respond to treatment by the end of induction as per the SmPC, they can receive further treatment up to a total of 12 to 16 weeks before switching. The ERG noted that a delayed response phase was included in the TA633[20] model and was generally accepted, however, the ERG broadly agreed with the company's approach to exclude extended induction in the current appraisal, as the assumption that all treatments will require extended induction is highly unlikely. Furthermore, the ERG noted that for the scenario analysis which included

extended induction, patient distribution into the Remission, Response (No remission) and Active UC health states were derived directly from individual clinical trials (as opposed to the NMAs).

4.2.3. Population

For the economic analysis, the company submitted two distinct subgroup analyses for TNFi-naïve and TNFi- experienced patients, which was considered to be broadly consistent with the decision problem in Table 7. Modelled patient characteristics were taken from the TrueNorth[27,28] study for both subgroups, which was a multicentered, international study. Clinical expert opinion to the ERG confirmed that mean weight and mean age were generally representative of patients in the UK (albeit there may be more of an equal distribution of male and female patients in both subgroups). Patient characteristics were also found to be broadly similar to those used in previous NICE TAs for moderately to severe UC including ustekinumab TA633[20] and tofacitinib TA547.[21] Overall, the ERG considered the modelled patient characteristics presented in Table 38 to be appropriate.

Table 38: Modelled patient characteristics

Characteristic	Population	
	TNFi-naïve	TFNi-experienced
Mean weight, kg	[redacted]	[redacted]
Proportion of female, %	[redacted]	[redacted]
Mean age, years	[redacted]	[redacted]

Abbreviations: TNFi, tumour necrosis factor inhibitor

4.2.4. Interventions and comparators

In the TNFi-naïve subgroup the primary comparators were TNFis including infliximab (biosimilar), adalimumab (biosimilar), golimumab and the biologic treatment vedolizumab. The company stated that ustekinumab was not considered as a relevant comparator within this subgroup, given that NICE guidance states that ustekinumab is restricted to patients who have failed CvT or a biologic AND who have failed a TNFi or for whom a TNFi cannot be tolerated or is unsuitable. The ERG acknowledged that this restriction is in place for ustekinumab and that the company's rationale to exclude ustekinumab from this subgroup seemed reasonable. Furthermore, clinical opinion to the ERG noted that ustekinumab is not used as a first line treatment.

In the TNFi-experienced subgroup, comparators were ustekinumab and vedolizumab. The company stated that TNFis were not appropriate comparators for this subgroup, as TNFi switching is no longer routine clinical practice. Based on clinical expert opinion to the ERG, treatment switching amongst TNFis may occur in UK clinical practice. One expert stated that further TNFis are used, though the choice of subsequent TNFi is dependent largely on why patients did not respond to initial treatment. For example if the patient failed due to immunogenicity, a second TNFi would be tried. A second clinical expert noted that switching is uncommon, however patients could switch to adalimumab if they do not respond to infliximab. The ERG also noted that adalimumab was included as a relevant comparator in TA633.[20] The ERG considered undertaking a scenario analysis which included adalimumab as a relevant comparator in the TNFi-experienced subgroup, however the model does not allow a flexible selection of comparators interchangeably between the subgroups and so it was not possible conduct this analysis.

The ERG noted that the company excluded tofacitinib as a comparator from both the TNFi-naïve and TNFi- experienced subgroups stating that there are significant safety concerns associated with treatment. The ERG did not consider the company's rationale to be sufficient, given that tofacitinib (TA547)[21] has been recommended for use by NICE as a viable treatment option. Furthermore, based on clinical opinion to the ERG, tofacitinib safety concerns were considered to be clinically managed at an individual patient level. Clinician input also confirmed that tofacitinib is used in UK clinical practice for treating TNFi-naive patients and treatment experienced patients with moderately to severe UC, as it is a fast-acting treatment and reduces the need for corticosteroid use. As such, the ERG subsequently asked the company to provide a revised analysis including tofacitinib as a relevant comparator within both subgroups. The company did not provide this analysis, stating that clinical opinion to the company confirmed that although there may be use of tofacitinib, it is not considered routine practice (refer to B9 of the company's clarification response). In contrast, the ERG noted that a recent multicentre real-world cohort study conducted in the UK by Honap et al. (2020)[22] has found that adverse events requiring curtailment of the treatment were uncommon with no occurrence of thromboembolic events and further concluded that tofacitinib was well-tolerated. The company also stated that in TA633,[20] the committee agreed that the exclusion of tofacitinib was appropriate. Whilst the ERG noted this observation, it should be acknowledged that in TA633,[20] the company had included tofacitinib as a relevant comparator within their model. The ERG were unable to alter the company's model to include a cost effectiveness comparison with tofacitinib (due to inflexibility),

however as an exploratory analysis, a cost comparison was undertaken to determine the comparative difference in drug costs, monitoring costs and adverse event costs between treatments (see 6.1 and 6.1.5 for results).

As a minor point, the ERG acknowledged that in both tofacitinib TA547 and ustekinumab TA633,[20,21] conventional therapy was included as a comparator (in both the biologic naïve and biologic experienced patient populations). However, within this current appraisal for ozanimod, the company did not consider conventional therapy as an active comparator, on the basis that patients were specifically those who have not responded to conventional therapy. The ERG sought clinical expert opinion to comment on the appropriateness of this assumption. Based on input to the ERG, it was considered reasonable to exclude conventional therapy from the analysis on the basis that patients have already failed conventional therapy.

4.2.5. Perspective, time horizon and discounting

The ERG did not identify concerns surrounding discounting. Costs and benefits were discounted at 3.5% which reflects NICE guidance. Furthermore, costs and outcomes were estimated from an NHS and PSS perspective, as appropriate.

The company used a lifetime horizon (58 and 60 years in the TNFi-naïve and TNFi-experienced populations, respectively) in the base case analysis and justified this on the basis that a lifetime horizon has been used in previous UC appraisals including ustekinumab TA633[20] and toficitinib TA547.[21] The ERG noted that using a lifetime horizon is consistent with both TA633[20] and TA547,[21] however in older UC appraisals i.e. vedolizumab TA342[87] and infliximab, adalimumab and golimumab TA329,[88] shorter time horizons (10 years) have been used. Overall, the ERG considered a lifetime horizon to be appropriate as UC is chronic condition characterised by remission and loss of response, thereby affecting patients over the duration of their lifetime.

The ERG noted that a two-week cycle length was used in the model. The company justified this on the basis that it captured the variety of treatment regimens. This is consistent with the cycle length used in ustekinumab TA633,[20] however an 8 week cycle length has been used previously in tofacitinib TA547[21] and vedolizumab TA342.[87] Clinical opinion to the ERG was somewhat mixed regarding the appropriateness of the modelled cycle length, noting that 8 weeks is broadly reasonable for assessing response to treatment, however 2 weeks is also used (particularly with respect to tofacitiib). The ERG did not conduct a scenario analysis using an 8 week cycle length, on the basis that this model parameter was not programmable in the

company's model.Overall, the ERG considered the higher resolution of the 2 week versus 8 week cycle, to allow for greater flexibility within the model, and was therefore reasonable.

Finally, the ERG noted that the model did not incorporate a half-cycle correction. The company justified this on the basis that the model uses a short two-week cycle length, and that a half-cycle correction was not applied in TA547 despite the submitted model having an eight-week cycle length. However, given that TA633[20] included a half-cycle correction (and used a two-week cycle length), the ERG asked the company to include an option in the model that allowed for a half-cycle correction. Based on clarification response B15, including half-cycle correction did not have a meaningful impact on results. Overall, the ERG considered that the company's decision to exclude half-cycle correction did not meaningfully impact results.

4.2.6. Treatment effectiveness and extrapolation

4.2.6.1. Induction period and maintenance period transition probabilities

The proportion of patients achieving 'Remission' and 'Response (No remission)' at the end of the induction phase (for both subgroups) was taken from the NMAs, discussed in Section 3.4.4. The company stated (CS, Document B, p153), that mean absolute probabilities were based on NMA outputs including baseline anchor, response effect, remission effect and standardised mean difference versus baseline for a given treatment in the induction period (Table 39). For the maintenance phase the probability of achieving sustained remission and sustained response were estimated based on the maintenance NMA (See Table 40 below for the mean absolute values).

Table 39: Clinical efficacy at the end of the induction period

Treatment	Induction length (weeks)	Remission	Response (No remission)	No response (Active UC)
TNFi-naïve				
Ozanimod	10	[redacted]	[redacted]	[redacted]
Golimumab	6	[redacted]	[redacted]	[redacted]
Infliximab (biosimilar)	8	[redacted]	[redacted]	[redacted]
Adalimumab (biosimilar)	8	[redacted]	[redacted]	[redacted]
Vedolizumab	6	[redacted]	[redacted]	[redacted]
TNFi-experienced				
Ozanimod	10	[redacted]	[redacted]	[redacted]
Ustekinumab	8	[redacted]	[redacted]	[redacted]
Vedolizumab	6	[redacted]	[redacted]	[redacted]

Treatment	Induction length (weeks)	Remission	Response (No remission)	No response (Active UC)
BSC	10	[redacted]	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; TNFi, tumour necrosis factor inhibitor; UC, ulcerative colitis

Table 40: Clinical efficacy in the maintenance period

Treatment	Sustained remission	Sustained response
TNFi-naïve		
Ozanimod	[redacted]	[redacted]
Golimumab	[redacted]	[redacted]
Infliximab/biosimilar	[redacted]	[redacted]
Adalimumab/biosimilar	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]
Vedolizumab (IV)	[redacted]	[redacted]
Vedolizumab (SC)	[redacted]	[redacted]
TNFi-experienced		
Ozanimod	[redacted]	[redacted]
Ustekinumab	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]
Vedolizumab (IV)	[redacted]	[redacted]
Vedolizumab (SC)	[redacted]	[redacted]
BSC	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; IV, intravenous; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

4.2.6.2. Loss of response

In line with TA633,[20] the model assumed that a proportion of patients lose response over time. This was assumed to be a constant loss of response that extended beyond the trial duration. In TA633,[20] the ERG accepted this assumption given the absence of longer-term follow-up data outlining how absolute or relative loss of response changes over time. Loss of response was estimated using an equation provided in Appendix N, provided by the company.

Table 41: Transition probabilities for loss of response

Treatment	Duration of maintenance period	Loss of response	Loss of response (No remission)
TNFi-naïve			
Ozanimod	42	[redacted]	[redacted]
Golimumab	54	[redacted]	[redacted]
Infliximab/biosimilar	46	[redacted]	[redacted]
Adalimumab/biosimilar	44	[redacted]	[redacted]
Vedolizumab	46	[redacted]	[redacted]
TNFi-experienced			
Ozanimod	42	[redacted]	[redacted]
Ustekinumab	44	[redacted]	[redacted]
Vedolizumab	46	[redacted]	[redacted]
BSC	42	[redacted]	[redacted]

Abbreviations: BSC, best supportive care; TNFi, tumour necrosis factor inhibitor

4.2.6.3. Post active treatment transitions for BSC

In the company model (post active treatment phase), the modelled cohort progress to the 'Active UC' health state where some may continue to receive best supportive care, comprising components of conventional therapy, and still continue to experience 'Remission' or 'Response No Remission' since UC is a relapse-remitting disease. The company submission stated that the transitions among the 'Active UC', 'Remission', and 'Response No Remission' health states for BSC were informed using the 'Loss of Response' and 'Loss of Response No Remission' derived from the pooled placebo arm (from the RCTs included in the NMA) estimates across the subgroups for sustained remission and sustained response.

However, the ERG noted that the loss of overall response (including remission) was used to inform remission transition probabilities for BSC i.e., remission equals overall response (through loss of response) which differed from the approach taken for active treatments (where remission state membership was derived as: overall response (through loss of response) – response no remission (through loss of response no remission). The ERG considered that the remission probabilities for BSC could be calculated through 'loss of remission' (i.e. calculated directly from the sustained remission estimates), as deriving it from overall response slightly overestimates the remission probability. This approach has been incorporated into the ERG base case (see Section 6.3 for results).

Additionally, the ERG noted that the loss of response and loss of response (no remission) estimates were noticeably different between the non-biologic failure and biologic failure subgroups in TA633[20] (Table 43 and Table 44 TA633 committee papers), and that the company had used the data for the TNFi-experienced group in both populations, given patients receiving BSC in the model (regardless of the population selected) do so in the post-active treatment setting, and thus have failed at least one active treatment by definition. The ERG considered that the approach used in TA633 was more appropriate, on the basis that available subgroup data were used to inform loss of response. As a result of using the alternative baseline placebo risk estimates in the ERG's revised base case, loss of response and loss of response (no remission) were based on TNFi-naïve and TNFi-experienced subgroup estimates, as appropriate.

4.2.6.4. Uncertainty surrounding clinical effectiveness estimates

In the base case analysis the company opted to use a FE model in both TNFi-naïve and TNFi-experienced subgroups for the maintenance phase NMAs, as well as the TNFi-experienced subgroup for the induction phase. The ERG acknowledged the company's rationale for using the FE model for the maintenance phase, i.e. that the fit was reasonable and that the RE model did not converge; it also noted the highly uncertain posterior SD in the induction phase NMA for TNFi-experienced participants. However, due to the high degree of heterogeneity amongst the studies included in the NMA, the ERG considered that FE models were inappropriate. As noted in Section 3.5.3.2 the ERG was unable to produce RE models with sufficient convergence (without using an informative prior distribution) and were therefore unable to use a RE model as part of its preferred base case. Furthermore, as noted in Section 3.4.2.4, the ERG identified concerns surrounding the baseline estimation of placebo risk in the NMA. In order to generate estimates of clinical effectiveness for its base case, the ERG used the placebo arm values from individual trials included in the NMA that were more generalisable to the UK context, the results of which are reported in Section 3.5.3.1. The ERG noted that re-running the NMA using the alternative means of estimating baseline placebo risk resulted in fewer total QALYs for all treatments. The ERG used these estimates to inform the ERG's preferred modelling assumptions, as described in Section 6.3.

With respect to treatment discontinuation due to serious adverse events, modelled per cycle probabilities for ozanimod were taken from the TRUENORTH[27,28] study (see Table 47 on p157 of the CS). For vedolizumab, golimumab, infliximab and adalimumab, treatment discontinuation rates were derived from GEMINI 1,[61] VISIBLE 1,[73] PURSUIT M,[66] Kobayshi et al. (2016)[74] and

ULTRA 1[65] and 2[43] respectively. The ERG noted that the discontinuation rate for ozanimod was considerably lower than comparator treatments, particularly in the maintenance phase. The ERG were unclear whether these rates would be reflective of clinical practice and therefore conducted a scenario analysis using an alternative treatment discontinuation rate for ozanimod. See Section 6.1.

4.2.6.5. Validation of model outcomes

In order to assess the validity of the company's base case estimation of QALYs, the ERG reviewed previous economic models (which predominantly considered lifetime horizons), including ERG analyses for previous NICE technology appraisals in UC and other relevant published literature. Though the ERG understands and acknowledges the differences in terms of model structure and methodology across these economic models, the ERG's view is that such a comparison would still serve as a useful means of cross-validating model outcomes (especially the QALYs), irrespective of the differences.

The ERG compared the model generated QALYs of ozanimod (as outlined in the company submission), to that of NICE TA329,[88] TA342[87] and Wu et al.(2018)[86], as summarised in Table 42. It should be noted that the QALY results from tofacitinib (TA547)[21] and ustekinumab (TA633)[20] were not available, as they were commercial in confidence. However, it was noted that in TA633 ERG's validation highlighted that the ustekinumab company model QALY estimates were lower than those from other lifetime models. In general, ERG noted that the QALY estimates from ozanimod company model (though slightly lower) were mostly comparable to the previous appraisals and publications. The difference in the total QALYs between the company's model and previous appraisals could be due to the consideration of additional response health states for the BSC in the post-active treatment phase (which is a change in this model structure compared to the ustekinumab appraisal (TA633)[20]).

Table 42: Comparison of modelled QALYs

Study name (time horizon)	QALYs	
	TNFi-naïve	**TNFi-experienced**
Ozanimod company model (lifetime)	Oza: [redacted]	Oza: [redacted]
	Ada: [redacted]	Ved: [redacted]
	Inf: [redacted]	Ust: [redacted]
	Ved: [redacted]	
	Gol: [redacted]	
TA342 (lifetime, ERG preferred base case)	Ada:12.39	Ved:11.84
	Ved:12.37	CvT:11.28
	Gol:12.05	Surgery:14.60
	Inf: 12.01	
	Surgery:14.60	
	CvT: 11.73	
TA329 (lifetime, AG model)	**Moderate to severe UC who failed at least 1 prior therapy**	
	Ada:10.82	
	Inf: 10.81	
	Gol: 10.63	
	CvT: 10.47	
Wu et al. (lifetime)	**Moderate to severe UC**	
	CvT:10.49	
	Ved→CvT: 11.48	
	Tof→CvT: 11.51	
	Inf→CvT: 10.87	
	Gol→CvT:10.89	
	Ada→CvT: 10.71	
	Ved→Tof→CvT: 12.37	
	Inf→Tof→CvT:11.81	
	Gol→Tof→CvT:11.83	
	Ada→Tof→CvT:11.67	
	Tof→Ved→CvT:12.37	
	Tof→Inf→CvT:11.84	
	Tof→Gol→CvT:11.86	

Abbreviations: Ada, adalimumab; AG, Assessment Group; CvT, conventional therapy; ERG, Evidence Review Group; Gol, golimumab; Inf, infliximab; Oza, ozanimod; QALY, quality-adjusted life year; TA, technology appraisal; TNFi, tumour necrosis factor inhibitior; Tof, tofacitinib; UC, ulcerative colitis; Ust, ustekinumab; Ved, vedolizumab

4.2.6.6. Probability of surgery and complications

A proportion of patients in the Active UC health state (of the post-active treatment phase) were assumed to require surgery each model cycle. In the company's base case analysis, annual

probabilities were converted into per cycle probabilities (see Table 48 and 49 on p159 of the CS). The company derived the probability of 1st surgery from a published study by Misra et al. (2016),[89] which was a large retrospective analysis of UK Hospital Episode Statistics (HES) for UC, based on 71,966 patients. The ERG noted this study has been used previously in TA633[20] and TA547[21] to estimate the probability of 1st surgery. Based on a review of TA633 and TA547,[20,21] the ERG identified several other alternative published studies which could be used to inform the probability of 1st surgery, including a UK study by Chhaya et al. (2015)[90] However, in both appraisals, the ERG agreed with the company's selection of Misra et al. (2016)[89] in the base case.

The proportion of patients who had complications after having a 1st surgery (1st surgery complications) was derived from a national clinical audit of inpatient care (for adults with UC).[91] This was estimated to be 33.5% (based on a weighted average of 32% elective and 35% non-elective surgeries). The proportion of patients with complications following post-surgery remission (1st surgery remission) was derived from a published UK study by Segal et al. (2018)[92] which assessed long term outcomes of prepouch ileitis in 31 patients. The ERG noted that both of these sources were used previously in TA633.[20] In TA547,[21] the proportion of patients with complications following post-surgery remission (1st surgery remission) was based on a Belgian study by Ferrante et al. (2008).[93] The company and ERG undertook scenario analyses using other published literature sources, however results were not sensitive to this.

In the current appraisal for ozanimod, the company made several simplifying assumptions with respect to modelled surgery rates, including the assumption that the probability of patients requiring a 2nd surgery is the same as the 1st surgery. The ERG noted that this assumption was consistent with TA633[20] and was generally accepted by the ERG. Overall, the ERG considered that the modelled surgery rates were not key drivers of cost effectiveness in this appraisal due to the small proportion of patients who transitioned into the surgical health states.

Table 43: Modelled per cycle probability of surgery

From	To	Per cycle probability
Active UC (post active treatments)	1st Surgery	0.00018
1st Surgery	1st Surgery Complications	0.33500
1st Surgery	1st Surgery Remission	0.66500
1st Surgery Remission	1st Surgery Remission	0.99876
1st Surgery Remission	1st Surgery Complications	0.00124
1st Surgery Complications	2nd Surgery	0.00018

Abbreviation: UC, ulcerative colitis

4.2.6.7. Mortality

The model included all-cause mortality i.e. patients could die in any health state, based on age and gender adjusted background mortality (using UK lifetables). The company assumed that UC is not associated with an additional mortality risk, therefore a standardised mortality ratio of 1 was used for key UC health states (Table 44). Based on clinician input to the ERG, UC was not considered to result in excess mortality. As such the company's assumption appeared reasonable. As per TA633[20] and TA547,[21] the company used a study by Jess et al. (2007)[94] to estimate a 30% mortality risk associated with surgery. This additional mortality risk was also assumed to apply to second surgery. Overall, the ERG considered the company's modelled mortality estimates to be acceptable. Mortality was not considered to be a driver of cost effectiveness, as there is no difference in LY gains between treatments.

Table 44: Standardised mortality ratio by health state

Health state	Standardised Mortality Ratio
Remission	1.0
Response (No remission)	1.0
Active UC	1.0
1st surgery	1.3
1st surgery remission	1.0
1st surgery complications	1.0
2nd surgery	1.3
2nd surgery remission	1.0

Abbreviation: UC, ulcerative colitis

4.2.7. Health-related quality of life

4.2.7.1. Health state utility values

The company conducted a systematic literature review to identify plausible health state utility values for inclusion in the model. On p162 of the CS, the company stated that 27 unique studies which reported HRQoL were identified in patients with moderately to severely active UC and 9 HTA appraisals. The ERG noted that the company did not use a recognised filter for HRQoL studies and restricted the bibliographic database searches to exclude conference abstracts. The company stated that these were separately hand-searched and excluded to avoid double counting. Additional searches in Ovid MEDLINE (Appendix A) conducted by the ERG identified other UK and non UK studies which reported HRQoL data in patients with moderate to severe

UC. However, after screening these studies, the ERG considered the company's search and study identification to be broadly reasonable. In the base case analysis, the company opted to use health state values from ustekinumab TA633.[20] The ERG noted that these values were derived from published literature sources i.e. Woehl et al. (2008)[29] was selected for the remission, response (no remission), active UC and post 1st surgery remission health states. Arseneau et al. (2006)[30] was used for 1st surgery and post 1st surgery complications, as these values were not reported in Woehl et al.[29] The ERG noted that the utility value for 1st surgery was based on a weighted average of the utilities for ileostomy (0.57) and J-Pouch surgery (or ileal pouch-anal anastomosis) (0.68), with weights of 60% and 40% respectively.

The company assumed that the utility associated with the second surgery and post-second surgery remission were equal to the first surgery (due to the lack of published data surrounding second surgery HRQoL values). The ERG found this assumption to be consistent with ustekinumab TA633,[20] where the ERG accepted these values (see discussion below). Based on a review of tofacitinib TA547[21] and vedolizumab TA342,[87] second surgery was not considered, therefore these models did not include second surgery utilities. Health state utilities were adjusted appropriately for age and gender using a published equation by Ara and Brazier (2010)[95] to account for the natural decline in QoL as a result of aging. See Table for the health state utility values used in the company's base case.

The ERG noted Woehl et al. (2008)[29] to be a UK study, which collected HRQoL data on 180 patients with active UC in the UK (using the EQ-5D), whilst Arseneau et al. (2006)[30] collected HRQoL data on 48 US patients from the University of Virginia Health System and Duke University Medical Centre. The ERG considered Woehl et al. (2008)[29] to be generalisable and broadly appropriate (albeit the study is somewhat dated). However, the use of Arseneau et al. (2006)[30] raises some generalisability concerns given the small sample size and participant characteristics i.e. a mean age of 45 years (thereby likely underrepresenting the second disease peak), overwhelmingly Caucasian (96%) and predominantly male (62%), a mean disease duration of 9.8 years (thereby likely underrepresenting the first disease peak) and very little participants who had undergone colectomy (21%). Despite these limitations, Arseneau et al has been considered a reasonable source for use in TA633 (see commentary below).

Table 45: Modelled health state utility values

Health state	Utility	Source
Remission	0.87	Woehl et al. (2008)[29]
Response (no remission)	0.76	Woehl et al. (2008)[29]
Active UC	0.41	Woehl et al. (2008)[29]
1st Surgery	0.61	Arseneau et al. (2006)[30]
Post 1st surgery remission	0.72	Woehl et al. (2008)[29]
Post 1st surgery complications	0.34	Arseneau et al. (2006)[30]
2nd Surgery	0.61	Assumption (as per TA633)[20]
Post 2nd surgery remission	0.72	Assumptions (as per TA633)[20]

Abbreviation: UC, ulcerative colitis

4.2.7.2. Utility value sources used in previous UC appraisals

Ustekinumab TA633 (2020)

The ERG noted that the use of Woehl et al. (2008)[29] and Arseneau et al. (2006)[30] was in line with the recent UC appraisal for ustekinumab TA633,[20] where the ERG considered the values reported in Woehl et al. (2008)[29] and Arseneau et al. (2006)[30] to be 'generally reasonable'. In TA633,[20] the ERG further agreed with the company's decision to not use direct HRQoL data from the pivotal study UNIFI,[52] as they were 'inconsistent with the values used in previous NICE appraisals for UC'. It was not possible to validate this statement as UNIFI utility values were marked as CIC in TA633.[20]

Tofacitinib TA547 (2018)

Based on a review of tofacitinib TA547,[21] Woehl et al. (2008)[29] was also used to derive health state utilities all health states. Trial based utilities from the OCTAVE studies were not considered appropriate for use in the base case due to the re-randomisation design and the lack of intermediate assessment of clinical response and remission between week 8 and week 52. The ERG considered Woehl et al. (2008)[29] to be the most appropriate source for base case utility parameters, and used values reported in Swinburn et al. (2012)[96] as a scenario analysis.

Vedolizumab TA342 (2015)

In the company's base case, health state utility values were derived from the pivotal study GEMINI I,[61] whereby QoL values for the Remission, Mild disease and Moderate to severe disease were estimated based on EQ-5D data. As utility data were not collected for the surgery

health states (post-surgery remission and post-surgery complications), the company used published literature from Punekar and Hawkins (2010)[97] (stated to be an epidemiology and costs study of CD). Although the ERG considered that using trial-based utilities in the base case was appropriate, it was noted that the value for post-surgical remission was lower than the value for moderate to severe UC. This was considered to lack plausibility, as it did not capture any benefit from surgery. The committee agreed that quality of life may be improved after 1st surgery (compared to having moderate to severe UC), although the magnitude of difference was uncertain. Two alternative sources were identified by the ERG and used in scenario analyses i.e. Woehl et al. (2008)[29] and Swinburn et al. (2012).[96] The ERG considered that values from Woehl et al. (2008)[29] (for patients who had surgery) were higher than those reported in Punekar and Hawkins et al.[97] The committee considered that Woehl et al. (2008)[29] and Swinburn et al. (2012)[96] had some important limitations i.e. small patient numbers and uncertainty regarding generalisability to UK practice. However, in the TNFi-experienced population, the committee expressed a preference for using both of these sources.

Infliximab, adalimumab and golimumab TA329 (2015)

The ERG considered Woehl et al. (2008)[29] and Swinburn et al. (2012)[96] to be the most useful sources of utility values in the model as they were UK-based, included a large number of patients (n=180 and n=230 respectively) and reported EQ-5D utility values for most modelled health states. In TA329,[88] the ERG selected utility values by Woehl et al. (2008)[29] to inform their base case analysis and used Swinburn et al. (2012)[96] as a scenario analysis.

4.2.7.3. The availability of direct HRQoL data from TRUENORTH

For ozanimod, quality of life data were available from patients directly in the TRUENORTH[27,28] study (from cohort 1 and cohort 2). Utility values were collected using the EQ-5D-5L at baseline, the end of induction (10 weeks) and the end of maintenance (52 weeks), see Table 46 and Table 47. Values were then cross-walked to EQ-5D-3L values using an appropriate published algorithm by Van Hout et al. (2012)[37] and UK value set from Dolan et al. (1997).[85] The ERG noted that health state values in the induction and maintenance phases were based on the weighted average across placebo arm and ozanimod arms i.e. utility values were health state dependent as opposed to treatment dependent. The company justified this approach on the basis that placebo and ozanimod values were broadly similar across health states.

The ERG noted that despite the availability of direct trial data, the company did not use QoL data from TRUENORTH[27,28] in the base case analysis due to limitations. The company outlined key limitations with the TRUENORTH utility data on p.162 of the CS. These included the following;

- In the induction phase, utility values for the Active UC health state (No response or remission at week 10) may be somewhat overestimated, as Active UC patients in TRUENORTH were receiving ozanimod. However, the modelled Active UC health state assumes that no further treatment would be received. Similarly, in the maintenance phase, patients continued to receive ozanimod in the Active UC health state (No response or remission at week 52). However, the modelled Active UC health state assumes that no further treatment would be received. The QoL of patients in the TRUENORTH Active UC health state was therefore not considered to be reflective modelled patients.
- Maintenance phase utility values were based on small patient numbers and are therefore subject to uncertainty (See Table 53 on p161 of the CS).
- Length of trial data considered too short and may not accurately capture the change in utility over time.
- QoL data for surgical health states were not captured.

In addition to the limitations highlighted by the company, the ERG further noted that in the maintenance phase, the utility value for placebo patients in the Response (no remission) at week 52 health state was higher than re-randomised patients in the same health state who received ozanimod in the induction phase and ozanimod in the maintenance phase ([redacted] versus [redacted] respectively). This result appeared somewhat counterintuitive, as the ERG expected that patients receiving ozanimod in both trial phases would have a higher QoL than those who initially received placebo during induction and then continued to receive placebo in the maintenance phase.

Table 46: TRUENORTH utility data (induction phase)

	Cohort 1		Cohort 2	Weighted average
Health state	**Ozanimod**	**Placebo**	**Ozanimod**	
Baseline (Active UC)	[redacted]	[redacted]	[redacted]	[redacted]
Remission at week 10	[redacted]	[redacted]	[redacted]	[redacted]
Response (No remission) at week 10	[redacted]	[redacted]	[redacted]	[redacted]
No response or remission at week 10 (Active UC)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: UC, ulcerative colitis

Table 47: TRUENORTH utility data (maintenance phase)

	Re-randomised		Placebo	Weighted average
Health state	**Ozanimod/Placebo**	**Ozanimod/Ozanimod**		
Remission at Week 52	[redacted]	[redacted]	[redacted]	[redacted]
Response (No remission) at Week 52	[redacted]	[redacted]	[redacted]	[redacted]
No response or remission at Week 52 (Active UC)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: UC, ulcerative colitis

Based on cross-validation, TRUENORTH[27,28] utility values for active UC (No response or remission) and Response (No remission) were considerably higher compared with published literature sources noted in Section 4.2.7.2. As such, using TRUENORTH values in the base case could potentially bias the analysis in favour of treatments with relatively poorer clinical effectiveness estimates, as a high percentage of patients transition to the active UC health state. Overall, the ERG agreed with the company that TRUENORTH utility values were subject to limitation and the use of these values in the base case may have introduced further uncertainty. For completeness, the company provided scenario analyses using alternative sources including TRUENORTH,[27] TA342[87] and TA547[21] (See Section 5.2.3 for results).

4.2.7.4. Disutility associated with adverse events

The base case analysis included disutility associated with serious infection only, which is consistent with previous appraisals including utekinumab TA633[20] and tofacitinib TA547.[21] As per TA633, the company elicited the utility decrement for serious infection from a published systematic review and economic evaluation by Stevenson et al. (2016),[98] which assessed the impact of treatments on rheumatoid arthritis. Modelled disutility associated with a serious adverse event was estimated to be 0.156 and symptoms were assumed to last for 4 weeks (28

days). The ERG noted that the duration of symptoms was considered reasonable and is in line with TA329.[88]

4.2.8. Resources and costs

4.2.8.1. Treatment acquisition costs

Medicine acquisition costs were included in the analysis for active treatments (with the exception of tofacitinib, which was excluded as a relevant comparator by the company) and concomitant treatments. Unit costs were derived from appropriate sources including the British National Formulary (BNF) and the Drugs and Pharmaceutical Electronic Market Information Tool (eMIT). Where more than one formulation of treatment with similar strength was available, the company selected the cheapest for use in the model (see Table 60, p.166 of the CS for the full list of treatments and unit costs). Overall, the ERG considered this approach to be reasonable and likely conservative. However, based on cross validation of the company's medicine acquisition costs with those reported in the BNF and TA633,[20] the ERG noted that the company's cost for adalimumab (£633.60 for 40 mg/0.8 mL) represented the solution for injection pre filled syringes. Another formulation was available for use i.e. solution for injection vials (£316.93 for 40 mg/0.8 mL), which is stated to be for hospital use only. However, based on clinical expert opinion to the ERG, pre-filled syringes are predominently used in practice.

Treatment costs were based on the cost per pack for each treatment and the dosing regimen, as outlined in the SPC for each treatment and/or clinical trials (see Tables 58, 59 and 60 in the CS for the treatment doses used in the model). For each subgroup (TNFi-naïve and TNFi-experienced), the company modelled treatment costs seperately for the induction phase, extended induction phase (scenario analysis only) and the maintenance phase in order to account for variance in dosing and duration. For the induction phase, costs were applied as one off costs at the start of induction. The ERG considered this approach to be reasonable. Clinical opinion to the ERG confirmed that the dosing used in both the induction and maintenance phases were broadly appropriate, albiet there may be some variation in clinical practice with respect to dosing frequency and escalation. The company has provided scenario analysis testing the impact of dose escalation on base case results (see Section 5.2.3). The ERG noted that the company estimated treatment costs on a per model cycle basis which was considered to be appropriate and in line with TA633.[20] However, given that the company's model allowed for the estimation of costs on a per treatment cycle basis, the ERG conducted a scenario analysis using this approach to determine the impact on the ICER. See Section 6.1.6.

Two formulations were available for vedolizumab (SC and IV). In the base case analysis the company assumed that the proportion of patients receiving SC and IV as maintenance therapy was 'evenly distributed' i.e. 50% received SC and 50% received IV. In order to validate the company's base case assumption, the ERG elicited clinical expert opinion. Based on clinical responses, 50% of patients receiving SC vedolizumab for maintenance therapy was considered to be largely reasonable and reflective of current clinical practice. However the proportion is likely to increase over time as one expert noted that not many patients are expected to remain on IV vedolizumab after one year. The company provided scenario analysis which varied the proportion of patients receiving SC vedolizumab. For completeness, the ERG conducted an additional scenario analysis to capture the opinions of clnical experts. See Section 6.1.5.

4.2.8.2. Dose escalation

In the base case analysis, the company assumed that 30% of patients would require dose escalation in the maintenance period i.e. 70% of patients would recive standard dosing (see Table 60 in the CS for). This assumption was applied to all treatments apart from vedolizumab SC and ozanimod. The company stated that dose escalation was not considered for these treatments, as per information contained in their respective SPCs. Based on clinical opinion to the ERG, dose escalation for biologics is common in clinical practice (with between 30%-40% of patients on infliximab receiving an escalated dose). Clinical experts stated that the proportion of patients requiring dose escalation would vary depending on treatment received, however the company's base case assumption of 30% may be somewhat low (with figures more aligned to 40%-50%).

Overall, the ERG considered the company's handling of dose escalation in the base case to be reasonable and in line with TA633,[20] whereby the ERG noted 30% dose escalation was reflective of data within a published study by Lindsay et al. (2017).[99] In order to test uncertainty surrounding dose escalation, the company provided scenario analyses which assumed 0% and 50% of patients required dose escalation. See Section 5.2.

4.2.8.3. Concomitant treatment and conventional therapy costs

Whilst on active treatment, patients received concomitant treatment with conventional therapy. Conventional therapy costs were also applied to patients entering the post active treatment phase of the model i.e. patients in the Active UC health state. As noted in Table below, the per cycle cost (per average patient) was estimated based on the weighted proportion of patients receiving each treatment. The proportion of patients receiving conventional therapy were taken

from previous UC appraisals (ustekinumab TA633[20] and vedolizumab TA342[87]). As stated by the company, patients receiving ozanimod were contraindicated to azathioprine, 6-mercaptopurine and methotrexate, therefore the costs of these treatments were not included in the ozanimod treatment arm. The ERG considered this to be reasonable.

Based on a review of TA547,[21] the ERG noted that alternative conventional therapy proportions were used i.e. these were taken from a national audit of the Royal College of Physicians (RCP) on IBD.[91] The ERG highlighted several concerns surrounding these proportions, namely that it was inappropriate to assume equal usage for the four aminosalicylic (5ASA) drugs, as most patients received mesalazine. As such, the ERG considered the proportions from TA633[20] and TA342[87] to be reasonable.

Table 48: Modelled conventional therapy treatments and proportions

Drug	Dose description	Patient usage (Ozanimod)	Patient usage (other treatments)
Balsalazide	1.5 g twice daily	0.0%	0.0%
Mesalazine	1.2 g/day (divided doses)	13.0%	13.0%
Olsalazine	500 mg twice daily	0.0%	0.0%
Sulfasalazine	500 mg 4 times daily	0.0%	0.0%
Prednisolone	20.0 mg/day for two weeks	36.0%	36.0%
Hydrocortisone	20 mg/day	0.0%	0.0%
Azathioprine	2.5 mg/kg/day	0.0%	39.0%
6-mercaptopurine	1.5 mg/kg/day	0.0%	15.0%
Methotrexate	17.5 mg/week	0.0%	9.0%
Budesonide	3.0 mg/3xday for eight weeks	1.0%	1.0%

4.2.8.4. Administration and monitoring costs

The company's base case analysis included administration costs for all IV treatments only. The cost per IV administration was £186.36 reflecting the average of a consultant and non consultant led face to face attendance. Costs were based on 2019/2020 NHS reference costs, which was considered to be an appropriate source. As per ustekinumab TA633,[20] the company assumed that there to be no cost involved with administering SC treatment, as most patients self administer. Based on clinical opinion to the ERG, most patients would be able to self administer SC treatment, however a small proportion (2%) may require assistance. The ERG noted that the inclusion of administration costs for such a small proportion of patients would not

have an meaningful impact on results and therefore considered the company's base case assumption to be acceptable.

One-off nurse training to teach patients how to self administer was assumed to be incurred by the manufacturer. Based on a review of TA633,[20] the ERG acknowledged that patient education and home delivery is provided by biologic manufacturers. Ozanimod was assumed to incur no administration cost as it is an oral treatment. The ERG considered this assumption to be reasonable.

With respect to monitoring costs, for ozanimod the company inluded the cost of a single ECG during induction which was estimated to be £61.80. This was included to reflect guidance within the SmPC for ozanimod. The cost was derived from 2019/20 NHS reference costs[100] as appropriate. The company assumed that all other monitoring requirements were similar between treatments (as per previous appraisals TA633,[20] TA547[21] and TA342[87]). Based on clinical expert opinion to the ERG, this was cosidered to be a reasonable assumption.

4.2.8.5. Health state costs

The company's analysis included disease management costs and health state specific costs, which applied to all treatments (see **Error! Reference source not found.** below for a complete list). Resource use estimates were mostly derived from a published study by Tsai et al. (2008)[101], which estimated annual resource use for each modelled health state based input from a panel of UK gastroenterologists. The ERG noted Tsai et al to be a UK cost effectiveness study which assessed a scheduled maintenance treatment of infliximab in moderate to severe UC. Although the study was somewhat dated, Tsai et al. (2008)[101] has been used and accepted as an appropriate source for resource use estimates in previous UC appraisals including TA633[20]. The ERG noted that Tsai et al. (2008)[101] did not report resource use estimates for surgery health states, as such the company assumed that resource use for 1st surgery and 2nd surgery were the same resource use in the active UC health state. This assumption is in line with TA633.[20]

Table 49: Modelled health state resource use

Resource item	Unit cost	Remission	Remission (no response)	Active UC	1st/2nd Surgery	Post 1st/2nd surgery remission	Post 1st surgery complications
Outpatient							
Consultant visit	£183.43	2	4.5	6.5	6.5	1.5	1.75
Blood test	£1.81	3.25	3.90	6.5	6.5	1.5	3.25
Inpatient							
Emergency endoscopy	£814.46	0	0.25	0.75	0.75	0.50	0.13
Elective endoscopy	£330.51	0.20	0.50	2	2	1.25	0.65
Care without colectomy	£2,301.47	0	0	0.15	0.15	0	3.25
Stoma care (post-colectomy)	£541.75	-	-	-	1	-	-

Abbreviations: UC, ulcerative colitis

Unit costs were based on 2018/2019 NHS reference costs values as appropriate. The cost of stoma care costs (post colectomy), was based on TA547 [21], which appeared reasonable. The model included acute costs associated with 1st and 2nd surgeries. The ERG noted that the costs associated with 1st and 2nd surgeries were estimated to be £14,309.51 and £10,438.22 respectively. These costs were elicited from expert opinion to the company were broadly in line TA633,[20] which reported these to be £15,311 and £10,998 respectively.

Finally, in the economic model, resource use costs were estimated based on a per cycle basis (see **Error! Reference source not found.**). Overall, the ERG considered the company's handling of health state resource use to be reflective of prior UC appraisals and therefore appropriate.

Table 50: Total per cycle health state costs

Health state	Total cost per cycle
Remission	£16.82
Response (No remission)	£46.05
Active UC	£108.13
1st and 2nd Surgery	£128.90
Post 1st and 2nd surgery (Remission)	£42.09
Post 1st and 2nd Surgery (Complications)	£311.52

Abbreviations: UC, ulcerative colitis

5. COST-EFFECTIVENESS RESULTS

5.1. Company's base case cost-effectiveness results

The company presented both pairwise and fully incremental results for consideration. Pairwise results reported by the company are shown in Table 51 and Table 52 and the fully incremental results are provided in Table 53 and Table 54. As previously highlighted by the ERG, the company has been asked to provide a revised analysis which includes tofacitinib as a comparator in both subgroups. Unfortunately, as this analysis was not provided, the cost effectiveness of ozanimod compared to tofacitinib is unknown. Furthermore, comparator PAS (cPAS) discounts were not included in the company's base case results. These are provided in a confidential appendix.

5.1.1. TNFi-naïve subgroup results (pairwise)

Based on the pair-wise analysis, ozanimod was cost effective compared to adalimumab at a conventional willingness-to-pay threshold of £30,000, resulting in an ICER of £28,686, based on an incremental QALY gain of [redacted] and an incremental cost of [redacted]. Compared to infliximab, golimumab and vedolizumab, ozanimod resulted in southwest (SW) ICERs i.e. ozanimod was less costly and less effective. [redacted].

Table 51: Company (Pairwise) base case results: TNFi-naïve subgroup (Discounted)

	Total costs	LYG	Total QALYs	Incremental costs	Inc. LYG	Inc. QALYs	Cost/ QALY gained
Company deterministic base case							
Ozanimod	[redacted]	[redacted]	[redacted]	-	-	-	-
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£28,686
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£167,024*
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£71,023*
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£52,736*
Company probabilistic base case							
Ozanimod	[redacted]	NR	[redacted]	-	-	-	-
Adalimumab	[redacted]		[redacted]	[redacted]	-	[redacted]	£28,934
Infliximab	[redacted]		[redacted]	[redacted]	-	[redacted]	£155,144*
Golimumab	[redacted]		[redacted]	[redacted]	-	[redacted]	£71,945*

	Total costs	LYG	Total QALYs	Incremental costs	Inc. LYG	Inc. QALYs	Cost/ QALY gained
Vedolizumab	[redacted]	-	[redacted]	[redacted]	-	[redacted]	£63,862*

Abbreviations: LYG, life years gained; NR, not reported; QALY, quality-adjusted life year

Note: * ICER in SW quadrant

5.1.2. TNFi-experienced subgroup (pairwise)

Based on the pair-wise analysis provided by the company, ozanimod was considered less costly and less effective compared to vedolizumab, resulting in a SW ICER of £199,551 [redacted]. Compared to ustekinumab, ozanimod was dominant i.e. less costly and more effective. It should be noted that the probabilistic results presented below are based on the ERG's re-run of the PSA, as the company did not provide these values in the CS.

Table 52: Company (Pairwise) base case results: TNFi-experienced subgroup (Discounted)

	Total costs	LYG	Total QALYs	Incremental costs	Inc. LYG	Inc. QALYs	Cost/ QALY gained
Company deterministic base case							
Ozanimod	[redacted]	[redacted]	[redacted]	-		-	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£199,551*
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Ozanimod dominant
Company probabilistic base case							
Ozanimod	[redacted]	NR	[redacted]	-		-	-
Vedolizumab	[redacted]		[redacted]	[redacted]	-	[redacted]	£1,324,054*
Ustekinumab	[redacted]		[redacted]	[redacted]	-	[redacted]	Ozanimod dominant

Abbreviations: LYG, life years gained; NR, not reported; QALY, quality-adjusted life year

Note: * ICER in SW quadrant

5.1.3. TNFi-naïve subgroup results (fully incremental)

Based on the fully incremental analysis provided by the company, ozanimod was considered the most cost effective treatment compared to adalimumab, resulting in an ICER of £28,686 (based on an incremental QALY gain of [redacted] and an incremental cost of [redacted]. Infliximab was

[redacted] by golimumab, and golimumab was [redacted] by vedolizumab. Vedolizumab resulted in an ICER of £52,736.

Table 53: Company (fully incremental) base case results: TNFi-naïve subgroup (Discounted)

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Cost per QALY gained
Adalimumab	[redacted]	[redacted]	-	-	-
Ozanimod	[redacted]	[redacted]	[redacted]	[redacted]	£28,686
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£52,736

Abbreviations: QALY, quality-adjusted life year

5.1.4. TNFi-experienced subgroup results (fully incremental)

Based on the fully incremental analysis provided by the company ustekinumab was dominated by ozanimod, resulting in an incremental QALY loss of [redacted] and an incremental cost of [redacted]. Compared to ustekinumab, vedolizumab resulted in an ICER of £199,551.

Table 54: Company (fully incremental) base case results: TNFi-experienced subgroup (Discounted)

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Cost per QALY gained
Ozanimod	[redacted]	[redacted]			
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£199,551

Abbreviations: QALY, quality-adjusted life year

5.2. Company's sensitivity analyses

5.2.1. One-way sensitivity analysis

One-way sensitivity analysis (OWSA) was provided and model parameters were varied by +/- 20% (95% CI were used where standard errors of the mean were available). The company presented results based on Net Health Benefit (NHB) and presented results using tornado diagrams (CS, Document B, Section B.3.8.2) as such the impact on ICER was not reported in the CS. Overall, the ERG considered the company's OWSA to be useful in deteriming the sensitivity of model parameters to variation, however the results were be of limited use for

decision making/interpretation as most parameters were varied by an abitrary percentage, and cPAS results were not included for comparator treatments.

5.2.2. Probabilistic sensitivity analysis

The company PSA which varied the model parameters simultaneously to determine the impact on the ICER. The results of the company's PSA (cost-effectiveness plane scatterplots and CEACs) were presented in the CS (Document B, Section B.3.8.1) and results can also be found in Table 51 and Table 52). As there are scatterplots for ozanimod vesus each comparator for both TNF-naïve and TNF-experienced populations they have not presented here again.

The company model used the generated iterations (n=1000) of NMA-derived clinical efficacy parameters related to remission and response. These were hard coded into the Excel model, while the other parameters (namely costs, utilities, discontinuation due to AE, surgery and spontaneous remission related probabilities etc.) used distributions to sample the parameter values probablistically with each PSA run. A table containing the list of parameters varied in the PSA with the respective distributions were presented in the CS (Appendix J.2). The conclusion of PSA results were in line with the base case results; however, in the TNFi-experienced subgroup for the comparison of ozanimod versus vedolizumab the PSA ICER was higher than that of the base case. As per the CS, the company noted that this difference was due to smaller base case incremental QALYs with marginal variations resulting in significant variations in the ICER (though still in the SW quadrant). The ERG noted this difference in incremental QALYs between the base case ([redacted]) and PSA ([redacted]); however, did not find any further issues associated with it. Further, the CS Section B.3.8.1 indicated that AE rates were sampled using a log-normal distribution, utilities were sampled using a beta distribution and the costs using a gamma distribution. However, the ERG noted that a (1-Gamma) distribution was used to sample utilities in the model, although the impact on the results were negligible.

The ERG viewed the approach used to derive the samples for parameters from NMA using Convergence Diagnosis and Output Analysis (CODA) software as appropriate given it takes into account the joint posterior distribution of the parameters included. However, ERG considered that the correlation between the parameters has not been represented adequately as described earlier in Section 3.4.2. The ERG also had reproducibility issues with the PSA as the CODA parameters were hard coded in the model and the settings used for Markov chain Monte Carlo (MCMC) simulations to derive those CODA samples were not fully transparent. Furthermore, the fact that tofacitinib has not been included as a relevant comparator renders the CEAC less

useful for decision making as the probability of ozanimod being cost-effective could change with tofacitinib inclusion.

5.2.3. Scenario analyses

The company conducted scenario analyses to explore uncertainty surrounding key model parameters/assumptions. The company's base case inputs and alternative scenario analysis inputs used are outlined in Table below. The ERG considered the range of scenario analyses conducted by the company to be comprehensive; however it should be noted that results do not include comparator PAS discounts. Furthermore, the company did not conduct a scenario analysis whereby tofacitinib is considered as a relevant comparator in both subgroups. As such, results should be interpreted with caution.

Table 55: Base case and scenario analysis parameters/assumptions used by the company

Model parameter	Base case value	Scenario analysis value(s)
Spontaneous remission	0.5%	0%, 1%
Extended induction	Excluded	Included
Dose escalation	30%	0%, 50%
Treatment waning	Excluded	Included- 25% treatment waning after 2 years
Vial sharing	Excluded	Included
Subsequent treatment	Excluded	Included- applied to TNFi-naïve subgroup only (subsequent treatments were vedolizumab and ustekinumab)
Alternative utility values	Woehl et al.[29] and Arseneau et al. [30]	• TRUENORTH[27,28] • Vedolizumab (TA342)[87] • Tofacitinib (TA547)[21]
CvT/BSC (treatment distribution)		Tofacitinib TA547[21]
Proportion of patients receiving vedolizumab SC	50%	0%, 30%

Abbreviations: BSC, best supportive care; CvT, conventional therapy; NICE, National Institute for Health and Care Excellence; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor; TA, technology appraisal

For the company's full list of scenario analyses results, see Section B.3.8.3 in the CS. The ERG noted that results were mostly sensitive to alternative assumptions with respect to extended induction, dose escalation, utility values and proportion of patients receiving SC vedolizumab.

Incremental results and ICERs for these scenario analyses are presented in Table 56 to Table 59.

Table 56: Extended induction

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	XXXX	XXXX	£28,686	0.003
	Infliximab	XXXX	XXXX	£167,024*	0.175
	Golimumab	XXXX	XXXX	£71,023*	0.101
	Vedolizumab	XXXX	XXXX	£52,736*	0.205
Extended induction included	Adalimumab	XXXX	XXXX	£28,686	0.003
	Infliximab	XXXX	XXXX	£95,490*	0.178
	Golimumab	XXXX	XXXX	£53,607*	0.116
	Vedolizumab	XXXX	XXXX	£49,151*	0.250
TNFi-experienced					
Base case	Vedolizumab	XXXX	XXXX	£199,551	0.170
	Ustekinumab	XXXX	XXXX	Ozanimod dominant	0.156
Extended induction included	Vedolizumab	XXXX	XXXX	£81,131	0.234
	Ustekinumab	XXXX	XXXX	Ozanimod dominant	0.184

Abbreviations: ICER, incremental cost-effectiveness ratio; NHB, net health benefit; QALY, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor

Note: *SW quadrant ICER; costs saved per QALY forgone. Threshold used to calculate NHB was £30,000

Table 57: Alternative dose escalation assumption

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024*	0.175
	Golimumab	[redacted]	[redacted]	£71,023*	0.101
	Vedolizumab	[redacted]	[redacted]	£52,736*	0.205
0% dose escalation	Adalimumab	[redacted]	[redacted]	£52,734	-0.047
	Infliximab	[redacted]	[redacted]	£105,530*	0.097
	Golimumab	[redacted]	[redacted]	£32,908*	0.007
	Vedolizumab	[redacted]	[redacted]	£41,492*	0.104
50% dose escalation	Adalimumab	[redacted]	[redacted]	£12,655	0.036
	Infliximab	[redacted]	[redacted]	£208,020*	0.228
	Golimumab	[redacted]	[redacted]	£96,434*	0.163
	Vedolizumab	[redacted]	[redacted]	£60,233*	0.272

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551*	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.156
0% dose escalation	Vedolizumab	[redacted]	[redacted]	£147,551*	0.118
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.134
50% dose escalation	Vedolizumab	[redacted]	[redacted]	£234,217*	0.205
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.171

Abbreviations: ICER, incremental cost-effectiveness ratio; NHB, net health benefit; QALY, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor

Note: *SW quadrant ICER; costs saved per QALY forgone. Threshold used to calculate NHB was £30,000

Table 58: Alternative utility values

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Adalimumab	[redacted]	[redacted]	£28,686	0.003
	Infliximab	[redacted]	[redacted]	£167,024*	0.175
	Golimumab	[redacted]	[redacted]	£71,023*	0.101
	Vedolizumab	[redacted]	[redacted]	£52,736*	0.205
TRUENORTH	Adalimumab	[redacted]	[redacted]	£54,046	-0.026
	Infliximab	[redacted]	[redacted]	£337,782*	0.195
	Golimumab	[redacted]	[redacted]	£143,381*	0.138
	Vedolizumab	[redacted]	[redacted]	£103,454*	0.337
TA342	Adalimumab	[redacted]	[redacted]	£29,933	0.000
	Infliximab	[redacted]	[redacted]	£170,401*	0.176
	Golimumab	[redacted]	[redacted]	£72,272*	0.102
	Vedolizumab	[redacted]	[redacted]	£54,142*	0.212
TA547	Adalimumab	[redacted]	[redacted]	£64,906	-0.032
	Infliximab	[redacted]	[redacted]	£418,880*	0.198
	Golimumab	[redacted]	[redacted]	£175,903*	0.144
	Vedolizumab	[redacted]	[redacted]	£123,157*	0.359
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551*	0.170
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.156
TRUENORTH	Vedolizumab	[redacted]	[redacted]	£440,991*	0.187
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.121
TA342	Vedolizumab	[redacted]	[redacted]	£197,216*	0.170

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.153
TA547	Vedolizumab	[redacted]	[redacted]	£517,373*	0.189
	Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	0.115

Abbreviations: ICER, incremental cost-effectiveness ratio; NHB, net health benefit; QALY, quality-adjusted life years; TA, technology appraisal; TNFi, tumour necrosis factor inhibitor

Note: *SW quadrant ICER; costs saved per QALY forgone. Threshold used to calculate NHB was £30,000

Table 59: Proportion of patients receiving SC Vedolizumab SC

Scenario	Treatment	Inc. costs	Inc. QALYs	ICERs	NHB (QALY)
TNFi-naïve					
Base case	Vedolizumab	[redacted]	[redacted]	£52,736*	0.205
0% patients receive SC		[redacted]	[redacted]	£68,803*	0.330
30% patients receive SC		[redacted]	[redacted]	£59,039*	0.256
TNFi-experienced					
Base case	Vedolizumab	[redacted]	[redacted]	£199,551*	0.170
0% patients receive SC		[redacted]	[redacted]	£1,982,556*	0.231
30% patients receive SC		[redacted]	[redacted]	£338,194*	0.196

Abbreviations: ICER, incremental cost-effectiveness ratio; NHB, net health benefit; QALY, quality-adjusted life years; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor

Note: *SW quadrant ICER; costs saved per QALY forgone. Threshold used to calculate NHB was £30,000

5.3. Model validation and face validity check

The company described their approach to model validation briefly in the CS Section B.3.10.1, which stated that cell-by-cell model verification was performed by an independent modelling team and clinical opinion was sought to ensure face validity of model structure, inputs and the assumptions. However, the company did not provide a comparison of their model outcomes (QALYs) with that of the previous TAs/publications. Therefore, ERG compared the modelled QALYs from current model with that of the of the previous TAs/publications as discussed in section 4.2.6.5.

Besides a few labelling issues, ERG noted a discrepancy between the CS Document B and the model in the distribution used for utilities in the PSA, as discussed in Section 5.2.2, however it did not have any material impact on the results. Further, during clarification (clarification question B14) the ERG indicated that a fully incremental analysis with the associated CE frontier was missing from the model, after which it was added. Otherwise, no serious errors were found in the company's model that impacted the results.

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1. Exploratory and sensitivity analyses undertaken by the ERG

As noted throughout the report, the ERG conducted a number of scenario analyses to explore uncertainty surrounding certain model parameters and assumptions. The scenario analyses are listed below and the results are presented in Section 6.2.

6.1.1. Cost comparison versus tofacitinib

As noted in 4.2.4 , due to the lack of model flexibility, the ERG were unable to include tofacitinib as a comparator into the analysis. However, in order to explore the uncertainty the ERG conducted a cost comparison which compared ozanimod to tofacitinib in both the TNFi-naïve and TNFi-experienced subgroups. This scenario analysis assumed clinical equivalency between treatments in terms of efficacy and only included differences in drug acquisition costs, monitoring costs and adverse event costs over the modelled time horizon (without considering discontinuation from the active treatment). However, please note that the extended induction and the concomitant medications costs were not considered in this analysis.

Though the clinical equivalency assumption is simplistic, in reality this would likely be a pessimistic assumption for tofacitinib given its clinical response and remission in the maintenance phase were better compared to ozanimod as found in the NMA. The ERG is of the opinion that the committee may find this comparison of costs, though only naïve, to be useful. Further, this analysis could be considered a starting point in addressing the uncertainty associated with the exclusion of tofacitinib as a relevant comparator.

Based on this analysis ozanimod resulted in a cost saving of [redacted] and [redacted] in the TNFi-naïve and TNFi-experienced subgroups respectively, where the PAS price was considered for ozanimod and cPAS was not considered for tofacitinib (see Section 6.2 for results). However, the conclusion changed with the consideration of cPAS for tofacitinib resulting in cost savings compared to ozanimod over lifetime horizon of the model (see cPAS Appendix).

6.1.2. Spontaneous remission

Based on clinical input to the ERG, spontaneous remission is likely to occur for approximately 5% to 30% of flare ups, which would result in a higher per year rate than the company's modelled yearly rate of 12%. This scenario analysis used a higher rate of spontaneous remission reflective of 0.75% per model cycle (18% per year), which also closely corresponds to

the mid-point of clinical expert opinion based estimates (see Section 4.2.2 for further discussion). The ERG noted that this is also in line with the observation mentioned in TA 633 that 1% per model cycle is likely to be an overestimate. Based on this analysis, the total costs were found to decrease across all treatments as the patients from the 'Active UC' state were re-distributed between 'Remission' and 'Response No Remission'. The total QALYs increased as the utility value for the response states were higher. See Section 6.2 for results.

6.1.3. Discontinuation due to AEs

The ERG noted that in the maintenance phase ozanimod was associated with the lowest discontinuation rate compared to all other treatments, resulting in a per cycle discontinuation rate of [redacted]. In order to test uncertainty surrounding modelled discontinuation rate for ozanimod, this scenario analysis assumes a higher discontinuation rate during the maintenance phase i.e. ozanimod treatment discontinuation is assumed to be 5% of the induction discontinuation rate. This rate was chosen as the AE discontinuation rate in the maintenance phase was at least 5% that of the induction for all other treatments based on their respective trials. As expected, the total QALYs and costs decreased marginally with higher discontinuation for ozanimod which resulted in minor impact with regard to cost-effectiveness of ozanimod versus the comparators. See Section 6.2 for results.

6.1.4. Ozanimod AE rate in the maintenance phase

In the company's base case, the per cycle AE rate was based on the rates within the CSR. As noted in 3.2.5.1, the ERG noted there to be some uncertainty surrounding the estimation of ozanimod rates and considered these to be somewhat low when compared to AE rates for comparator treatments (particularly in the maintenance phase). Also, the ERG noted that the rate used in the model was not tested as part of sensitivity analysis (albeit the AE cost per cycle was varied). Therefore, in this scenario, the ERG assumed that the maintenance AE rate for ozanimod was 20% higher, to be in line with the modelled rates for comparator treatments. A very minor increase in the total costs of ozanimod was noted which did not have any impact on its cost-effectiveness versus the comparators. See Section 6.2 for results.

6.1.5. Proportion of vedolizumab SC

As noted in 4.2.8.1, the company assumed that 50% of patients would receive SC vedolizumab and 50% would receive IV vedolizumab. Based on clinical input to the ERG, a 50% split is likely to be a reasonable assumption, however it was noted that patients are being steadily phased

onto SC vedolizumab over time and therefore the majority of patients are likely to receive SC vedolizumab after one year. In order to reflect this opinion, in this scenario, the ERG assumed that 80% of patients receive SC vedolizumab in the maintenance phase (patients typically start treatment on SC vedolizumab after the 6-week induction period). Based on this analysis, ozanimod incremental savings reduced from [redacted] to [redacted], due to reduced administration costs associated with SC vedolizumab and the reduction in the proportion of vedolizumab IV patients modelled to receive dose escalation. See Section 6.2 for results. The results become even more sensitive to the SC proportion when cPAS was considered for vedolizumab (see cPAS Appendix).

6.1.6. Treatment regimen costs applied per treatment cycle

In the company's base case, treatment regimen costs were applied per model cycle in the maintenance phase (in line with TA547).[20] The ERG noted at the clarification stage that the company model included the option of modelling treatment costs per treatment cycle as well, and the results were sensitive to this setting (however it was not tested as a scenario in the CS Section B.3.8.3). Subsequently, the company indicated in the clarification response to question B10 that if the regimen costs were applied per treatment cycle the entire cohort would receive the full cost of the treatment upfront even if they discontinue treatment in subsequent model cycles. As there may be some deviation in the dosing schedule in practice, the company indicated that the application of costs per model cycle was preferred in the base case.

Though the ERG agreed with company's choice of modelling treatment costs per model cycle for the base case, the ERG considered it would still be worth exploring the option of modelling the treatment costs per treatment cycle, given its noticeable impact on the results. Through this scenario the ERG noted the sensitivity of ICER to minor change in treatment costs given the difference in the QALYs between the treatments were lower. For instance, for the comparison of ozanimod versus adalimumab, although the difference in the total drug acquisition costs with per treatment cycle approach was only around [redacted], the ICER increased to >£33k (versus £28k in the base case) as the incremental QALYs were lower ([redacted]). See Section 6.2 for the results.

6.2. Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

The ERG made the changes described in Sections 6.1.2 to Section 6.1.6. Each change has been made individually. Please note that the individual impact of revised modelled efficacy

estimates for BSC was not captured in the ERG base case and hence included here. The results of the ERG's exploratory analyses are provided in Table 60 and Table 61 for the TNFi-naïve and TNFi experienced subgroups respectively. The ERG acknowledged that fully incremental results are considered to be appropriate and suitably robust for decision making by NICE. However, due to the company's exclusion of tofacitinib from the analysis, the ERG have only presented pairwise results on the basis that presentation of fully incremental results (without a relevant active comparator) is likely to be misleading.

Table 60: ERG scenario analysis (TNFi-naïve subgroup)

	Section in ERG report	Incremental costs	Incremental QALYs	ICER £/QALY	% Change from company base case
Company base-case					
ozanimod	5.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£28,686	
infliximab		[redacted]	[redacted]	£167,024*	
golimumab		[redacted]	[redacted]	£71,023*	
vedolizumab		[redacted]	[redacted]	£52,736*	
Cost comparison with tofacitinib					
Incremental cost associated with ozanimod	6.1	[redacted]	Not applicable		
Spontaneous remission (0.75% per model cycle)					
ozanimod	6.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£29,830	4%
infliximab		[redacted]	[redacted]	£169,731*	2%
golimumab		[redacted]	[redacted]	£72,123*	2%
vedolizumab		[redacted]	[redacted]	£53,983*	2%
Ozanimod AE discontinuation rate in maintenance phase (5% that of induction)					
ozanimod	6.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£29,790	4%
infliximab		[redacted]	[redacted]	£137,368*	-18%
golimumab		[redacted]	[redacted]	£65,285*	-2%
vedolizumab		[redacted]	[redacted]	£51,677*	-8%
Ozanimod AE rate in the maintenance phase (20% increase)					

	Section in ERG report	Incremental costs	Incremental QALYs	ICER £/QALY	% Change from company base case
ozanimod	6.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£28,750	0%
inflixumab		[redacted]	[redacted]	£166,869*	
golimumab		[redacted]	[redacted]	£70,961*	
vedolizumab		[redacted]	[redacted]	£52,720*	
% patients receiving SC vedolizumab (80% after year 1)					
ozanimod	6.1	-	-	-	-
adalimumab		Not applicable			
infliximab					
golimumab					
vedolizumab		[redacted]	[redacted]	£44,204*	-16%
Treatment regimen costs applied per treatment cycle					
ozanimod	6.1	-	-	-	-
adalimumab		[redacted]	[redacted]	£33,815	18%
infliximab		[redacted]	[redacted]	£188,210*	13%
golimumab		[redacted]	[redacted]	£71,528*	1%
vedolizumab		[redacted]	[redacted]	£53,501*	1%
Revised modelled efficacy estimates for BSC in the post-active treatment phase					
ozanimod	6.3	-	-	-	-
adalimumab		[redacted]	[redacted]	£28,797	0%
infliximab		[redacted]	[redacted]	£167,294*	0%
golimumab		[redacted]	[redacted]	£71,133*	0%
vedolizumab		[redacted]	[redacted]	£52,859*	0%

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

Table 61: ERG scenario analysis (TNFi-experienced subgroup)

	Section in ERG report	Incremental costs	Incremental QALYs	ICER £/QALY	% Change from company base case
Company base-case					
ozanimod	5.1	-	-	-	-
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£33,725)	
vedolizumab		[redacted]	[redacted]	£199,551*	
Cost comparison with tofacitinib					
Incremental cost associated with ozanimod	6.1	[redacted]	Not applicable		
Spontaneous remission (0.75% per model cycle)					
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£34,594)	3%
vedolizumab		[redacted]	[redacted]	£198,146*	-1%
Ozanimod AE discontinuation rate in maintenance phase (5% that of induction)					
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£41,096)	22%
vedolizumab		[redacted]	[redacted]	£160,695*	-19%
Ozanimod AE rate in the maintenance phase (20% increase)					
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£33,689)	0%
vedolizumab		[redacted]	[redacted]	£199,367*	
% patients receiving SC vedolizumab (80% after year 1)					
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£33,725)	0%

	Section in ERG report	Incremental costs	Incremental QALYs	ICER £/QALY	% Change from company base case
vedolizumab		[redacted]	[redacted]	£161,152*	-19%
Treatment regimen costs applied per treatment cycle					
ozanimod	6.1	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£47,464)	41%
vedolizumab		[redacted]	[redacted]	£208,721*	5%
Revised modelled efficacy estimates for BSC in the post-active treatment phase					
ozanimod	6.3	-	-	-	
ustekinumab		[redacted]	[redacted]	Dominated by ozanimod (-£33,354)	-1%
vedolizumab		[redacted]	[redacted]	£200,192*	0%

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

6.3. ERG's preferred assumptions

This section presents the results based on ERG preferred assumptions for the base case. The results below present both the incremental and cumulative impact of ERG preferences.

The ERG's preference would have been to include tofacitinib as a comparator within the economic analysis. However, due to the lack of model flexibility, it was not possible to include tofacitinib in the economic model. As an exploratory analysis, the ERG has conducted a cost comparison versus tofacitinib (see Table 60 and Table 61 for results).

As part of the ERG preferred base case, the ERG considered the following:

- Revised remission and response probability estimates for the treatments and BSC derived from the ERG run of the NMA using the alternative placebo baseline risks (as per 3.4.2.4)
- Revised post-active treatment transition probabilities for BSC which include an alternative means of estimating remission probabilities for BSC based on 'loss of remission' (directly from the sustained remission estimates) as opposed to using the

BSC response rates for the TNFi-experienced population for both populations in the base case.

It is to be noted that due to the use of alternative placebo baseline estimates derived by including only trials which are relevant to decision making, the overall response and remission decreases across all treatments. Due to higher utility associated with the remission, less patients entering that state over the modelled horizon caused reduction in the total QALYs as shown in Table 62. The total costs also decreased owing to a reduction in remission costs which could not be offset by the corresponding increase in active UC state costs.

However, the incremental impact of revised post-active treatment transition probabilities for BSC was different for TNFi-naïve and TNFi-experienced populations. For the TNFi-naïve subgroup, the overall response increased, resulting in marginal total QALY increase while it decreased marginally for the TNFi-experienced subgroup. The increase or decrease in the overall response was driven by the proportional increase or decrease in the 'remission' and 'response no remission' probabilities, which differed between the subgroups. On the other hand, the increase or decrease in total costs was driven by whether the reduction in response health state costs were offset by the corresponding increase in the active UC state costs.

The cumulative effect of these changes in the base case resulted in decreased total costs and QALYs across all treatments for both the subgroups. Please note that the cumulative effect of the ERG base case changes were the same as the incremental impact following revised modelled efficacy estimates for BSC (as shown in Table 62 and Table 63), as there were only two changes as part of the ERG base case.

In the TNFi-naïve subgroup, pairwise deterministic analysis indicated that the ICER for ozanimod compared to adalimumab was £27,794, based on an incremental QALY gain of [redacted] and an incremental cost of [redacted]. Compared to infliximab, golimumab and vedolizumab, ozanimod resulted in SW ICERs i.e., ozanimod was less costly and less effective. Please note that the fully incremental analysis has not been presented here as it would be inaccurate without considering tofacitinib as a relevant comparator.

Probabilistic analysis resulted in similar conclusions with an ICER for ozanimod compared to adalimumab of £27,842. With respect to other comparators, ozanimod was less costly and less effective. Similar to the fully incremental analysis, the CEAC would be inaccurate and not suitable for decision making without considering tofacitinib. Hence, it has not been presented

here. The scatterplots of the cost-effectiveness plane for ozanimod versus each of comparators have been presented in the Appendix B.

Table 62: Summary of ERG's preferred assumptions and ICER (TNFi-naïve subgroup)

Scenario	Total costs	Total QALYs	Incremental cost	Incremental QALYs	ICER: Ozanimod vs. comparators (£/QALY)
Company's base case					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£28,686
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£167,024*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£71,023*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£52,736*
ERG's preferred base case assumptions (applied incrementally over company's base case)					
Re-estimation of baseline placebo risks					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,479
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£169,098*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£82,608*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,298*
Revised modelled efficacy estimates for BSC in the post-active treatment phase					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,794
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£169,791*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£82,863*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,640*
Cumulative impact of ERG preferences (deterministic)					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,794
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£169,791*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£82,863*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,640*
Cumulative impact of ERG preferences (probabilistic)					
ozanimod	[redacted]	[redacted]	-	-	-
adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,842

Scenario	Total costs	Total QALYs	Incremental cost	Incremental QALYs	ICER: Ozanimod vs. comparators (£/QALY)
infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£158,721*
golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£87,452*
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£68,470*

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; vs., versus

Note: * ICER in SW quadrant

In the TNFi-experienced subgroup, pairwise deterministic analysis indicated that ozanimod was considered less costly and less effective compared to vedolizumab, resulting in a SW ICER of £436,080 [redacted]. Compared to ustekinumab, ozanimod was dominant i.e. less costly and more effective. Please note that the fully incremental analysis has not been presented here as it would be inaccurate without considering tofacitinib as a relevant comparator.

In the probabilistic analysis, however, ozanimod was found to be dominant compared to both ustekinumab and vedolizumab. As shown in Table 63, for the comparison against vedolizumab the incremental cost savings reduced to [redacted] and the QALY gain increased to [redacted] resulting in the treatment being dominated by ozanimod. However, this should be interpreted with caution as the ICER was found to be highly sensitive to even marginal changes in the incremental costs and QALYs. Furthermore, there is uncertainty around the proportion of patients treated with vedolizumab receiving the treatment as an SC formulation in clinical practice. The ERG noted that it is likely that considering any cPAS for vedolizumab in conjunction with a higher proportion of SC vedolizumab would alter this conclusion, possibly resulting in a SW ICER.

The scatterplots of the cost-effectiveness plane for ozanimod versus each of comparators have been presented in the Appendix B. Like the fully incremental analysis, the CEAC too would be inaccurate and not suitable for decision making without considering tofacitinib. Hence, it has not been presented here.

Table 63: Summary of ERG's preferred assumptions and ICER (TNFi-experienced subgroup)

Scenario	Total costs	Total QALYs	Incremental cost	Incremental QALYs	ICER: Ozanimod vs. comparators (£/QALY)
Company's base case					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£33,725)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£199,551*
ERG's preferred base case (applied incrementally over company's base case)					
Re-estimation of baseline placebo risks					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£71,524)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£427,683*
Revised modelled efficacy estimates for BSC in the post-active treatment phase					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£70,807)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£436,080*
Cumulative impact of ERG preferences (deterministic)					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£70,807)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£436,080*
Cumulative impact of ERG preferences (probabilistic)					
ozanimod	[redacted]	[redacted]	-	-	-
ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£56,635)
vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated by ozanimod (-£12,926)

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; vs., versus

Note: * ICER in SW quadrant

6.4. Conclusions of the cost-effectiveness section

For the TNFi-naïve subgroup, based on the ERG's preferred results, ozanimod was cost effective compared to adalimumab at a conventional willingness-to-pay threshold of £30,000, resulting in an ICER of £27,794, based on an incremental QALY gain of [redacted] and an incremental cost of [redacted]. Compared to infliximab, golimumab and vedolizumab, ozanimod resulted in ICERs in the SW quadrant i.e., ozanimod was less costly and less effective. For the TNFi-experienced subgroup, ozanimod was considered less costly and less effective compared to vedolizumab, resulting in a SW ICER of £436,080 [redacted]. Compared to ustekinumab, ozanimod was found to be dominant i.e. less costly and more effective.

The ERG noted that a key strength of the company's submission was the use of precedent to inform the majority of model parameters and assumptions. Furthermore, as discussed throughout the ERG report, the company addressed several key concerns raised previously in prior UC appraisals including TA633 [20] and TA547 [21]. As a result, the ERG's preferred base case assumptions were broadly aligned with the company's (with the exception of baseline placebo risk estimates and revised assumptions with respect to modelled efficacy for BSC). As outlined by the ERG's preferred base case analysis, results were not particularly sensitive to these changes (with the exception of the comparison to vedolizumab in the TNFi-experienced subgroup, see Table 63).

However, there were some key limitations with the company's analysis. In addition to uncertainty surrounding the NMA (and modelled clinical effectiveness estimates), the company did not present results comparing ozanimod to tofacitinib. As noted in Section 4.2.4, the ERG considered tofacitinib to be a potentially relevant comparator. The exclusion of this comparison introduces additional uncertainty and means that the incremental cost effectiveness results (both pairwise and fully incremental) should be interpreted with caution. This concern extends to the interpretation of PSA results as well as the CEAC. Overall, the ERG suggest that NICE deliberate on the appropriateness of tofacitinib as a relevant comparator.

7. END OF LIFE

The ERG considered that ozanimod does not meet NICE end of life criteria as the treatment is not indicated for people with a short life expectancy (normally defined as less than 24 months).

References

1. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 5: the baseline natural history model. Medical Decision Making. 2013;33(5):657-70.
2. Pasvol TJ, Horsfall L, Bloom S, Segal AW, Sabin C, Field N, et al. Incidence and prevalence of inflammatory bowel disease in UK primary care: a population-based cohort study. BMJ Open. 2020;10(7):e036584.
3. Roberts-Thomson IC, Bryant RV, Costello SP. Uncovering the cause of ulcerative colitis. JGH Open. 2019;3(4):274-6.
4. Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG clinical guideline: ulcerative colitis in adults. American Journal of Gastroenterology. 2019;114(3):384-413.
5. NHS. Ulcerative colitis, 2019. Available from: https://www.nhs.uk/conditions/ulcerative-colitis/.
6. Bertani L, Tricò D, Zanzi F, Baiano Svizzero G, Coppini F, de Bortoli N, et al. Oral sucrosomial iron is as effective as intravenous ferric carboxy-maltose in treating anemia in patients with ulcerative colitis. Nutrients. 2021;13(2).
7. Jonefjäll B, Simrén M, Lasson A, Öhman L, Strid H. Psychological distress, iron deficiency, active disease and female gender are independent risk factors for fatigue in patients with ulcerative colitis. United European Gastroenterol Journal. 2018;6(1):148-58.
8. 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 & Colitis. 2017;11(6):649-70.
9. RCGP Learning. Inflammatory bowel disease toolkit, 2019. Available from: https://elearning.rcgp.org.uk/mod/book/view.php?id=12536&chapterid=386.
10. Pola S, Patel D, Ramamoorthy S, McLemore E, Fahmy M, Rivera-Nieves J, et al. Strategies for the care of adults hospitalized for active ulcerative colitis. Clinical Gastroenterology and Hepatology. 2012;10(12):1315-25.e4.
11. Adams SM, Bornemann PH. Ulcerative colitis. American Family Physician. 2013;87(10):699-705.
12. Holvoet T, Lobaton T, Hindryckx P. Optimal management of acute severe ulcerative colitis (ASUC): Challenges and solutions. Clinical and Experimental Gastroenterology. 2021;14:71-81.
13. NHS. Ulcerative colitis. Complications, 2019. Available from: https://www.nhs.uk/conditions/ulcerative-colitis/complications/.
14. Crohn's & Colitis UK. Ulcerative colitis, your guide, 2019. Available from: https://www.crohnsandcolitis.org.uk/about-crohns-and-colitis/publications/ulcerative-colitis.
15. National Institute for Health and Care Excellence (NICE). Ulcerative Colitis: management. NICE guideline [NG130], 2019. Available from: https://www.nice.org.uk/guidance/NG130.
16. 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.
17. NHS. Ulcerative colitis. Diagnosis, 2019. Available from: https://www.nhs.uk/conditions/ulcerative-colitis/diagnosis/.
18. National Institute for Health and Care Excellence (NICE). NICE Pathway: Inducing remission in people with ulcerative colitis, 2021. Available from: https://pathways.nice.org.uk/pathways/ulcerative-colitis/inducing-remission-in-people-with-ulcerative-colitis#content=view-info-category%3Aview-about-menu.

19. NHS Nottingham and Nottinghamshire Clinical Commissioning Group. Nottinghamshire commissioning guidance for the treatment of ulcerative colitis (UC), 2020. Available from: https://www.nottsapc.nhs.uk/media/1636/ulcerative_colitis_algorithm.pdf.
20. National Institute for Health and Care Excellence (NICE). Ustekinumab for treating moderately to severely active ulcerative colitis. Technology appraisal guidance [TA633], 2020. Available from: https://www.nice.org.uk/guidance/TA633.
21. National Institute for Health and Care Excellence (NICE). Tofacitinib for moderately to severely active ulcerative colitis. Technology appraisal guidance [TA547], 2018. Available from: https://www.nice.org.uk/guidance/TA547.
22. Honap S, Chee D, Chapman TP, Patel M, Kent AJ, Ray S, et al. Real-world effectiveness of tofacitinib for moderate to severe ulcerative colitis: A multicentre UK experience. Journal of Crohn's & Colitis. 2020;14(10):1385-93.
23. Celgene Corporation. Zeposia®(ozanimod) capsules for oral use. Package insert, 2021: Bristol Myers Squibb. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/209899s000lbl.pdf.
24. Scott FL, Clemons B, Brooks J, Brahmachary E, Powell R, Dedman H, et al. Ozanimod (RPC1063) is a potent sphingosine-1-phosphate receptor-1 (S1P1) and receptor-5 (S1P5) agonist with autoimmune disease-modifying activity. British Journal of Pharmacology. 2016;173(11):1778-92.
25. Choi D, Stewart AP, Bhat S. Ozanimod: A first-in-class sphingosine 1-phosphate receptor modulator for the treatment of ulcerative colitis. Annals of Pharmacotherapy. 2021:10600280211041907.
26. Cohen JA, Comi G, Selmaj KW, Bar-Or A, Arnold DL, Steinman L, et al. Safety and efficacy of ozanimod versus interferon beta-1a in relapsing multiple sclerosis (RADIANCE): a multicentre, randomised, 24-month, phase 3 trial. Lancet Neurology. 2019;18(11):1021-33.
27. Celgene. Phase 3, multicenter, randomized, double-blind, placebo-controlled trial of oral RPC1063 as induction and maintenance therapy for moderate to severe ulcerative colitis - TRUE NORTH. Clinical Study Report, 2020.
28. Sandborn WJ, Feagan BG, D'Haens G, Wolf DC, Jovanovic I, Hanauer SB, et al. Ozanimod as induction and maintenance therapy for ulcerative colitis. New England Journal of Medicine. 2021;385(14):1280-91.
29. Woehl A, Hawthorne AB, McEwan P. The relation between disease activity, quality of life and health utility in patients with ulcerative colitis. Gut. 2008;57(A1-A172):153-4.
30. 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.
31. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.
32. CRD. Systematic reviews: CRD's guidance for undertaking reviews in healthcare, 2009 York, UK: Centre for Reviews and Dissemination, University of York. Available from: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf.
33. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
34. Sandborn WJ, Feagan BG, Wolf DC, D'Haens G, Vermeire S, Hanauer SB, et al. Ozanimod induction and maintenance treatment for ulcerative colitis. New England Journal of Medicine. 2016;374(18):1754-62.
35. King D, Reulen RC, Thomas T, Chandan JS, Thayakaran R, Subramanian A, et al. Changing patterns in the epidemiology and outcomes of inflammatory bowel disease in the United Kingdom: 2000-2018. Alimentary Pharmacology & Therapeutics. 2020;51(10):922-34.

36. Tran JQ, Zhang P, Ghosh A, Liu L, Syto M, Wang X, et al. Single-dose pharmacokinetics of ozanimod and its major active metabolites alone and in combination with gemfibrozil, itraconazole, or rifampin in healthy subjects: A randomized, parallel-group, open-label study. Advances in Therapy. 2020;37(10):4381-95.
37. Van Hout B, Janssen MF, Feng Y-S, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health. 2012;15(5):708-15.
38. Gordon JP, McEwan PC, Maguire A, Sugrue DM, Puelles J. Characterizing unmet medical need and the potential role of new biologic treatment options in patients with ulcerative colitis and Crohn's disease: a systematic review and clinician surveys. European Journal of Gastroenterology & Hepatology. 2015;27(7):804.
39. Hoentjen F, Haarhuis BJT, Drenth JPH, de Jong DJ. Elective switching from infliximab to adalimumab in stable Crohn's disease. Inflammatory Bowel Diseases. 2013;19(4):761-6.
40. Patil SA, Rustgi A, Langenberg P, Cross RK. Comparative effectiveness of anti-TNF agents for Crohn's disease in a tertiary referral IBD practice. Digestive Diseases and Sciences. 2013;58(1):209-15.
41. Van Assche G, Vermeire S, Ballet V, Gabriels F, Noman M, D'Haens G, et al. Switch to adalimumab in patients with Crohn's disease controlled by maintenance infliximab: prospective randomised SWITCH trial. Gut. 2012;61(2):229-34.
42. Singh S, Murad MH, Fumery M, Dulai PS, Sandborn WJ. First- and second-line pharmacotherapies for patients with moderate to severely active ulcerative colitis: An updated network meta-analysis. Clinical Gastroenterology and Hepatology. 2020;18(10):2179-91.e6.
43. 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.
44. 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.
45. Sands BE, Tremaine WJ, Sandborn WJ, Rutgeerts PJ, Hanauer SB, Mayer L, et al. Infliximab in the treatment of severe, steroid-refractory ulcerative colitis: a pilot study. Inflammatory Bowel Diseases. 2001;7(2):83-8.
46. Panaccione R, Ghosh S, Middleton S, Marquez JR, Scott BB, Flint L, et al. Combination therapy with infliximab and azathioprine is superior to monotherapy with either agent in ulcerative colitis. Gastroenterology. 2014;146(2):392-400.e393.
47. 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.
48. Macaluso FS, Maida M, Ventimiglia M, Renna S, Cottone M, Orlando A. Factors affecting clinical and endoscopic outcomes of placebo arm in trials of biologics and small molecule drugs in ulcerative colitis: A meta-analysis. Inflammatory Bowel Diseases. 2018;25(6):987-97.
49. Olfatifar M, Zali MR, Pourhoseingholi MA, Balaii H, Ghavami SB, Ivanchuk M, et al. The emerging epidemic of inflammatory bowel disease in Asia and Iran by 2035: A modeling study. BMC Gastroenterology. 2021;21(1):204.
50. Prideaux L, Kamm MA, De Cruz PP, Chan FK, Ng SC. Inflammatory bowel disease in Asia: a systematic review. Journal of Gastroenterology and Hepatology. 2012;27(8):1266-80.
51. Motoya S, Watanabe K, Ogata H, Kanai T, Matsui T, Suzuki Y, et al. Vedolizumab in Japanese patients with ulcerative colitis: A phase 3, randomized, double-blind, placebo-controlled study. PLoS ONE. 2019;14(2):e0212989.

52. Sands BE, Sandborn WJ, Panaccione R, O'Brien CD, Zhang H, Johanns J, et al. Ustekinumab as induction and maintenance therapy for ulcerative colitis. New England Journal of Medicine. 2019;381(13):1201-14.
53. Sandborn WJ, Ghosh S, Panes J, Vranic I, Su C, Rousell S, et al. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. New England Journal of Medicine. 2012;367(7):616-24.
54. Sandborn WJ, Peyrin-Biroulet L, Zhang J, Chiorean M, Vermeire S, Lee SD, et al. Efficacy and safety of etrasimod in a phase 2 randomized trial of patients with ulcerative colitis. Gastroenterology. 2020;158(3):550-61.
55. Seibold F, Peyrin-Biroulet L, Rubin DT, Feagan B, Oh YS, Tyrrell H, et al. SO2O4O3O1 Efficacy of etrolizumab in patients with moderate to severe UC who had failed TNF antagonist therapy (HICKORY). Swiss Medical Weekly. 2019;149 (Suppl 240):2S.
56. Vermeire S, O'Byrne S, Keir M, Williams M, Lu TT, Mansfield JC, et al. Etrolizumab as induction therapy for ulcerative colitis: a randomised, controlled, phase 2 trial. Lancet. 2014;384(9940):309-18.
57. Rutgeerts P, Feagan BG, Marano CW, Padgett L, Strauss R, Johanns J, et al. Randomised clinical trial: a placebo-controlled study of intravenous golimumab induction therapy for ulcerative colitis. Alimentary Pharmacology & Therapeutics. 2015;42(5):504-14.
58. AbbVie. Study to evaluate the safety and efficacy of two adalimumab dosing regimens in subjects with moderate to severe ulcerative colitis, 2014. Available from: https://clinicaltrials.gov/ct2/show/study/NCT02065622.
59. Probert CS, Hearing SD, Schreiber S, Kuhbacher T, Ghosh S, Arnott ID, et al. Infliximab in moderately severe glucocorticoid resistant ulcerative colitis: A randomised controlled trial. Gut. 2003;52(7):998-1002.
60. Sands BE, Peyrin-Biroulet L, Loftus EV, Jr., Danese S, Colombel JF, Toruner M, et al. Vedolizumab versus adalimumab for moderate-to-severe ulcerative colitis. New England Journal of Medicine. 2019;381(13):1215-26.
61. Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel JF, Sandborn WJ, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. New England Journal of Medicine. 2013;369(8):699-710.
62. Colombel JF, Panes J, D'Haens GR, Schreiber S, Panaccione R, Peyrin-Biroulet L, et al. 945 Higher versus standard adalimumab maintenance regimens in patients with moderately to severely active ulcerative colitis: Results from the SERENE-UC maintenance study. Gastroenterology. 2020;158(6 Suppl 1):S-192.
63. 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.
64. Sandborn WJ, Su C, Sands BE, D'Haens GR, Vermeire S, Schreiber S, et al. Tofacitinib as induction and maintenance therapy for ulcerative colitis. New England Journal of Medicine. 2017;376(18):1723-36.
65. 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.
66. 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.
67. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. International Journal of Epidemiology. 2012;41(3):818-27.

68. Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JP. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. Statistics in Medicine. 2015;34(6):984-98.
69. 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. Report by the Decision Support Unit, 2011 (last updated 2016): ScHARR, University of Sheffield. Available from: https://nicedsu.sites.sheffield.ac.uk/tsds/completed-or-in-progress.
70. Dias S, Ades AE, Welton NJ, Jansen J, Sutton AJ. Network meta-analysis for decision making: Wiley; 2018.
71. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 5: Evidence synthesis in the baseline natural history model. Report by the Decision Support Unit, 2011 (last updated April 2012): ScHARR, University of Sheffield. Available from: https://nicedsu.sites.sheffield.ac.uk/tsds/completed-or-in-progress.
72. 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.
73. Sandborn WJ, Baert F, Danese S, Krznaric Z, Kobayashi T, Yao X, et al. Efficacy and safety of vedolizumab subcutaneous formulation in a randomized trial of patients with ulcerative colitis. Gastroenterology. 2020;158(3):562-72.e12.
74. 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.
75. 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. Journal of Clinical Gastroenterology. 2015;49(7):582-8.
76. Pugliese D, Variola A, Privitera AC, Allocca M, Bossa F, Cappello M, et al. Real-life study (GORE-UC) evaluating the effectiveness of golimumab for the treatment of ulcerative colitis: An interim analysis from Italian group for the study of inflammatory bowel disease (IG-IBD). United European Gastroenterology Journal. 2017;5(5 Suppl 1):A750-A1.
77. Samaan MA, Cunningham G, Tamilarasan AG, Beltran L, Pavlidis P, Ray S, et al. Therapeutic thresholds for golimumab serum concentrations during induction and maintenance therapy in ulcerative colitis: results from the GO-LEVEL study. Alimentary Pharmacology & Therapeutics. 2020;52(2):292-302.
78. Yu J, Park SJ, Kim HW, Lim YJ, Park J, Cha JM, et al. Effectiveness and safety of golimumab in patients with ulcerative colitis: A multicenter, prospective, postmarketing surveillance study. Gut and Liver. 2021:Online first.
79. Strik A, Van De Vrie W, Bloemsaat-Minekus J, Nurmohamed M, Bossuyt P, Bodelier A, et al. Unchanged infliximab serum concentrations after switching from the originator infliximab to the biosimilar CT-P13 in patients with quiescent inflammatory bowel disease: Results from a prospective study (SECURE). Journal of Crohn's & Colitis. 2018;12(Suppl 1):S477-S8.
80. Park SH, Kim Y-H, Lee JH, Kwon HJ, Lee S-H, Park DI, et al. Post-marketing study of biosimilar infliximab (CT-P13) to evaluate its safety and efficacy in Korea. Expert Review of Gastroenterology & Hepatology. 2015;9 Suppl 1:35-44.
81. Coletta M, Paroni M, Alvisi MF, De Luca M, Rulli E, Mazza S, et al. Immunological variables associated with clinical and endoscopic response to vedolizumab in patients with inflammatory bowel diseases. Journal of Crohn's & Colitis. 2020;14(9):1190-201.
82. Danese S, Subramaniam K, Van Zyl J, Adsul S, Lindner D, Roth J, et al. Vedolizumab treatment persistence and safety in a 2-year data analysis of an extended access programme. Alimentary Pharmacology & Therapeutics. 2021;53(2):265-72.

83. Osterman M, Yarur A, Navaneethan U, James A, Lopez C, Rosario M, et al. Characterization of early clinical and pharmacokinetic response profiles of vedolizumab: An interim analysis of enterpret, a phase 4 clinical study. Gastroenterology. 2020;158(3 Suppl):S15.
84. Vermeire S, Lukas M, Magro F, Adsul S, Lindner D, Rosario M, et al. Vedolizumab efficacy, safety, and pharmacokinetics with reduced frequency of dosing from every 4 weeks to every 8 weeks in patients with Crohn's disease or ulcerative colitis. Journal of Crohn's & Colitis. 2020;14(8):1066-73.
85. Dolan P. Modeling valuations for EuroQol health states. Medical Care. 1997;35(11):1095-108.
86. Wu B, Wang Z, Zhang Q. Cost-effectiveness of different strategies for the treatment of moderate-to-severe ulcerative colitis. Inflammatory Bowel Diseases. 2018;24(11):2291-302.
87. National Institute for Health and Care Excellence (NICE). Vedolizumab for treating moderately to severely active ulcerative colitis. Technology appraisal guidance [TA342], 2015. Available from: https://www.nice.org.uk/guidance/ta342.
88. National Institute for Health and Care Excellence (NICE). Infliximab, adalimumab and golimumanb for treating moderately to severely active ulcerative colitis after the failure of conventional therapy. Technology appraisal guidance [TA329]. 2015. Available from: https://www.nice.org.uk/guidance/ta329.
89. 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. Canadian Journal of Gastroenterology and Hepatology. 2016;2016:8723949.
90. 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. Alimentary Pharmacology & Therapeutics. 2015;41(1):87-98.
91. Royal College of Physicians. National clinical audit of inpatient care for adults with ulcerative colitis, UK inflammatory bowel disease (IBD) audit, National Report, 2014. 2014.
92. Segal JP, McLaughlin SD, Faiz OD, Hart AL, Clark SK. Incidence and long-term implications of prepouch ileitis: An observational study. Diseases of the Colon & Rectum. 2018;61(4):472-5.
93. 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.
94. Jess T, Gamborg M, Munkholm P, Sørensen TI. Overall and cause-specific mortality in ulcerative colitis: meta-analysis of population-based inception cohort studies. American Journal of Gastroenterology. 2007;102(3):609-17.
95. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value in Health. 2010;13(5):509-18.
96. Swinburn P, Elwick H, Bean K, Curry A, Patel S, Bodger K. The impact of surgery on health related quality of life in ulcerative colitis. Gut. 2012;61:A237.
97. 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.
98. Stevenson M, Archer R, Tosh J, Simpson E, Everson-Hock E, Stevens J, et al. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying antirheumatic drugs and after the failure of conventional disease-modifying antirheumatic drugs only: systematic review and economic evaluation. Health Technology Assessment. 2016;20(35):1-610.
99. Lindsay JO, Armuzzi A, Gisbert JP, Bokemeyer B, Peyrin-Biroulet L, Nguyen GC, et al. Indicators of suboptimal tumor necrosis factor antagonist therapy in inflammatory bowel disease. Digestive and Liver Disease. 2017;49(10):1086-91.

100. NHS England. National schedule of NHS costs, 2019/2020. Available from: https://www.england.nhs.uk/national-cost-collection.
101. 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 & Therapeutics. 2008;28(10):1230-9.

Appendix A: Additional searches conducted by the ERG

Additional search strategy for phase 4 trials of ozanimod for ulcerative colitis

Ovid MEDLINE (1946 to February 15, 2022)

1 Colitis, Ulcerative/ 37741

2 ((colitis or colorectitis or proctocolitis or procto colitis) adj3 (ulcer* or mucosa* or gravis or idiopathic*)).tw,kf. 46305

3 (((colon or colonic) adj3 ulceration) and chronic*).tw,kf. 48

4 (UC and (ulcer* or colitis*)).tw,kf. 14733

5 1 or 2 or 3 or 4 54678

6 (ozanimod or rpc 1063 or rpc1063 or HSDB 7852 or OZM or Zeposia$2 or UNII-Z80293URPV or Z80293URPV or 1306760-87-1).tw,kf,rn. 153

7 5 and 6 52

8 clinical trial, phase iv/ 2276

9 ("phase 4" or "phase IV").ti,ab. 4739

10 8 or 9 5949

11 7 and 10 0

12 6 and 10 1

Ovid Embase (1974 to February 15, 2022)

1 exp Colitis, Ulcerative/ 81807

2 ((colitis or colorectitis or proctocolitis or procto colitis) adj3 (ulcer* or mucosa* or gravis or idiopathic*)).tw,kf. 73582

3 (((colon or colonic) adj3 ulceration) and chronic*).tw,kf. 94

4 (UC and (ulcer* or colitis*)).tw,kf. 32686

5 1 or 2 or 3 or 4 92875

6 ozanimod/ 504

7 (ozanimod or rpc 1063 or rpc1063 or HSDB 7852 or OZM or Zeposia$2 or UNII-Z80293URPV or Z80293URPV or 1306760-87-1).tw,kf,rn. 574

8 6 or 7 574

9 5 and 8 219

10 phase 4 clinical trial/ 4659
11 ("phase 4" or "phase IV").ti,ab. 7957
12 10 or 11 9728
13 9 and 12 0
14 8 and 12 1

ClinicalTrials.gov (www.clinicaltrials.gov)

Search: Ozanimod (Other terms field). Limited to Phase 4. 1 record – ozanimod for MS
Search: Zeposia (Other terms field). Limited to Phase 4. 1 record – ozanimod for MS
Search: rpc1063 (Other terms field). Limited to Phase 4. 1 record – ozanimod for MS

WHO ICTRP (https://trialsearch.who.int/)

Search: ozanimod (intervention). Recruitment status: ALL. Limited to Phase 4. 0 records
Search: Zeposia (intervention). Recruitment status: ALL. Limited to Phase 4. 0 records
Search: rpc1063 (intervention). Recruitment status: ALL. Limited to Phase 4. 0 records

EU Clinical Trials Register (https://www.clinicaltrialsregister.eu)

Search: ozanimod. Limited to Phase 4. 0 records
Search: zeposia. Limited to Phase 4. 0 records
Search: rpc1063. Limited to Phase 4. 0 records

Additional Ovid MEDLINE and Ovid Embase search strategy for phase 4 trials of comparator treatments for ulcerative colitis

Ovid MEDLINE (1946 to February 16, 2022)

1 Colitis, Ulcerative/ 37738
2 ((colitis or colorectitis or proctocolitis or procto colitis) adj3 (ulcer* or mucosa* or gravis or idiopathic*)).tw,kf. 46303
3 (((colon or colonic) adj3 ulceration) and chronic*).tw,kf. 48
4 (UC and (ulcer* or colitis*)).tw,kf. 14732

5 1 or 2 or 3 or 4 54676

6 clinical trial, phase iv/ 2274

7 ("phase 4" or "phase IV").ti,ab. 4737

8 6 or 7 5947

9 Ustekinumab/ 1437

10 (ustekinumab or "cnto 1275" or cnto1275 or stelara$2 or UNII-FU77B4U5Z0 or FU77B4U5Z0 or 15610-63-0).tw,kf,rn. 2597

11 Infliximab/ 11320

12 (infliximab or CT-P13 or CTP13 or SB2 or "abp 710" or abp710 or avakine$2 or flixabi$2 or "gp 1111" or gp1111 or inflectra$2 or ixifi$2 or "pf 06438179" or "pf 6438179" or pf06438179 or pf6438179 or remicade$2 or remsima$2 or renflexis$2 or revellex$2 or "ta 650" or ta650 or zessly$2).tw,kf,rn. 16945

13 Adalimumab/ 6267

14 (adalimumab or ADA or "abp 501" or abp501 or "abt d2e7" or abtd2e7 or adaly$2 or amgevita$2 or amjevita$2 or "avt 02" or "avt 02" or "bat 1406" or bat1406 or "bax 2923" or bax2923 or "bax 923" or bax923 or "bi 695501" or bi695501 or "chs 1420" or chs1420 or "ct p17" or ctp17 or cyltezo$2 or "da 3113" or da3113 or "dmb 3113" or dmb3113 or exemptia$2 or "fkb 327" or fkb327 or fyzoclad$2 or "gp 2017" or gp2017 or hadlima$2 or halimato$2 or hefiya$2 or "hlx 03" or hlx03 or hulio$2 or humira$2 or hyrimoz$2 or "ibi 303" or ibi303 or imraldi$2 or kromeya$2 or lu 200134 or lu200134 or "m 923" or m923 or mabura$2 or (monoclonal adj3 antibod$ adj3 D2E7) or "msb 11022" or msb11022 or "ons 3010" or ons3010 or "pf 06410293" or "pf 6410293" or pf06410293 or pf6410293 or raheara$2 or "sb 5" or sb5 or solymbic$2 or trudexa$2 or "zrc 3197" or zrc3197 or FYS6T7F842 or 331731-18-1 or 1446410-95-2).tw,kf,rn. 20606

15 (vedolizumab or entyvio$2 or "ldp 02" or ldp02 or "mln 0002" or mln0002 or "mln 02" or mln02 or "ldp 02" or UNII-9RV78Q2002 or 9RV78Q2002 or 943609-66-3).tw,kf,rn. 1430

16 (tofacitinib or "cp 690 550" or "cp 690550" or "cp690 550" or cp690550 or HSDB 8311 or xeljanz$2 or UNII-87LA6FU830 or 87LA6FU830 or 477600-75-2 or 540737-29-9).tw,kf,rn. 2180

17 (golimumab or CNTO-148 or CNTO148 or Simponi$2 or UNII-91X1KLU43E or 91X1KLU43E or 476181-74-5).tw,kf,rn. 1470

18 (filgotinib or "g 146034" or "g 146034 101" or g146034 or "g146034 101" or "glpg 0634" or glpg0634 or "gs 6034" or gs6034 or Jyseleca$2).tw,kf,rn. 195

19 (etrasimod or APD334 or UNII-6WH8495MMH or 6WH8495MMH or 1206123-37-6).tw,kf,rn. 18

20 (etrolizumab or pro 145223 or pro145223 or rhumab beta7 or UNII-I2A72G2V3J or I2A72G2V3J or 1044758-60-2).tw,kf,rn. 87

21 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 37270

22 5 and 8 and 21 13

Ovid Embase (1974 to February 16, 2022)

1 exp Colitis, Ulcerative/ 81818

2 ((colitis or colorectitis or proctocolitis or procto colitis) adj3 (ulcer* or mucosa* or gravis or idiopathic*)).tw,kf. 73598

3 (((colon or colonic) adj3 ulceration) and chronic*).tw,kf. 94

4 (UC and (ulcer* or colitis*)).tw,kf. 32694

5 1 or 2 or 3 or 4 92891

6 phase 4 clinical trial/ 4661

7 ("phase 4" or "phase IV").ti,ab. 7959

8 6 or 7 9731

9 Ustekinumab/ 9542

10 (ustekinumab or "cnto 1275" or cnto1275 or stelara$2 or UNII-FU77B4U5Z0 or FU77B4U5Z0 or 15610-63-0).tw,kf,rn. 9821

11 Infliximab/ 56559

12 (infliximab or CT-P13 or CTP13 or SB2 or "abp 710" or abp710 or avakine$2 or flixabi$2 or "gp 1111" or gp1111 or inflectra$2 or ixifi$2 or "pf 06438179" or "pf 6438179" or pf06438179 or pf6438179 or remicade$2 or remsima$2 or renflexis$2 or revellex$2 or "ta 650" or ta650 or zessly$2).tw,kf,rn. 58355

13 Adalimumab/ 39262

14 (adalimumab or ADA or "abp 501" or abp501 or "abt d2e7" or abtd2e7 or adaly$2 or amgevita$2 or amjevita$2 or "avt 02" or "avt 02" or "bat 1406" or bat1406 or "bax 2923" or bax2923 or "bax 923" or bax923 or "bi 695501" or bi695501 or "chs 1420" or chs1420 or "ct p17" or ctp17 or cyltezo$2 or "da 3113" or da3113 or "dmb 3113" or dmb3113 or exemptia$2 or "fkb 327" or fkb327 or fyzoclad$2 or "gp 2017" or gp2017 or hadlima$2 or halimato$2 or hefiya$2 or "hlx 03" or hlx03 or hulio$2 or humira$2 or hyrimoz$2 or "ibi 303" or ibi303 or imraldi$2 or kromeya$2 or lu 200134 or lu200134 or "m 923" or m923 or mabura$2 or

(monoclonal adj3 antibod$ adj3 D2E7) or "msb 11022" or msb11022 or "ons 3010" or ons3010 or "pf 06410293" or "pf 6410293" or pf06410293 or pf6410293 or raheara$2 or "sb 5" or sb5 or solymbic$2 or trudexa$2 or "zrc 3197" or zrc3197 or FYS6T7F842 or 331731-18-1 or 1446410-95-2).tw,kf,rn. 56135

15 vedolizumab/ 5526

16 (vedolizumab or entyvio$2 or "ldp 02" or ldp02 or "mln 0002" or mln0002 or "mln 02" or mln02 or "ldp 02" or UNII-9RV78Q2002 or 9RV78Q2002 or 943609-66-3).tw,kf,rn. 5799

17 tofacitinib/ 6597

18 (tofacitinib or "cp 690 550" or "cp 690550" or "cp690 550" or cp690550 or HSDB 8311 or xeljanz$2 or UNII-87LA6FU830 or 87LA6FU830 or 477600-75-2 or 540737-29-9).tw,kf,rn. 7079

19 golimumab/ 8467

20 (golimumab or CNTO-148 or CNTO148 or Simponi$2 or UNII-91X1KLU43E or 91X1KLU43E or 476181-74-5).tw,kf,rn. 8667

21 (filgotinib or "g 146034" or "g 146034 101" or g146034 or "g146034 101" or "glpg 0634" or glpg0634 or "gs 6034" or gs6034 or Jyseleca$2).tw,kf,rn. 780

22 (etrasimod or APD334 or UNII-6WH8495MMH or 6WH8495MMH or 1206123-37-6).tw,kf,rn. 103

23 (etrolizumab or pro 145223 or pro145223 or rhumab beta7 or UNII-I2A72G2V3J or I2A72G2V3J or 1044758-60-2).tw,kf,rn. 365

24 filgotinib/ 727

25 etrasimod/ 100

26 etrolizumab/ 347

27 or/9-26 101011

28 5 and 8 and 27 50

Additional search of Ovid MEDLINE to identify HRQoL literature not identified by company searches

Ovid MEDLINE(R) <1946 to February Week 3 2022>

1 Inflammatory Bowel Diseases/ or exp Colitis, Ulcerative/ 61154

2 (inflammatory bowel disease or ulcerative Colitis).ti,ab. 65527

3 1 or 2 78192

4 (hrql or hrqol or patient reported outcome$ or satisfaction or preference or disability adjusted life or daly$ or activities of daily living or adl).ab,ti. 292822

5 ((health adj3 (utility$ or status)) or (utilit$ adj3 (valu$ or measur$ or health or life or estimate$ or elicit$ or disease or score$ or weight)) or (disutility$ and health) or (disutility$ and scor$) or (disutility$ and valu$) or standard gamble or time trade off or time tradeoff or tto or rosser or willingness to pay or visual analog scale or visual analogue scale or discrete choice experiment or qwb or 15d or health utilities index or hui or hui1 or hui2 or hui3).ab,ti. 135634

6 (sf36 or sf 36 or sf6 or sf 6 or short form 6 or sf6d or sf 6d or short form 6d or eq 5d or eq5d or euroqol or euro qol or health status or hye or hyes or rosser index or quality of wellbeing or qwb or CUCQ or (Crohn$ adj1 Ulcerative Colitis Questionnaire) or RFIPC or Rating Form of Inflammatory Bowel Disease Patient Concerns or IBDQ or IBDQ-32 or Inflammatory Bowel Disease Questionnaire or SIBDQ or Short Inflammatory Bowel Disease Questionnaire or (health$ adj year$ adj equivalent$)).ti,ab. 87449

7 3 and (4 or 5 or 6) 1729

8 exp Longitudinal Studies/ or (longitudinal study or retrospective study or prospective study or cohort$ or follow up or cross-sectional study or cross sectional study or followup study or observational study or registry or registries or real world or cross sectional).ti,ab. or exp Retrospective studies/ or exp Prospective studies/ or exp Cohort Studies/ or exp Cross-Sectional Study/ or exp Cohort Studies/ or exp Observational Study/ 3423002

9 7 and 8 867

10 (Ephemera or "Introductory Journal Article" or News or "Newspaper Article" or Editorial or Comment or Overall or Letter or Short Survey or Tombstone or Books).pt. or in vitro Techniques/ or in vitro study/ or (commentary or editorial or comment or letter or mice or rat or mouse or animal or murine).ti. 3583247

11 (case report or case study or case series or woman or man or child or adolescent or female or male or boy or girl or infant).ti. 771190

12 case reports/ or case study/ or case report$.jw. 2094080

13 ((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) not ((child$ or juvenile or babies or infant$ or adolescent$ or pediatric$ or paediatric$) and adults)).ti. 1193387

14 review.pt. not (((systematic or meta) and analy$) or ((indirect or mixed) and treatment comparison)).ti,ab. 2583450

15 (exp animal/ or nonhuman/) not exp human/ 4960851

16 or/10-15 12513723

17 9 not 16 774

18 limit 17 to yr="2010-current" 559

19 Quality-Adjusted Life Years/ 14384

20 (quality adjusted or adjusted life year$).ti,ab,kf. 17490

21 (qaly$ or qald$ or qale$ or qtime$).ti,ab,kf. 10972

22 (illness state$1 or health state$1).ti,ab,kf. 6519

23 (hui or hui1 or hui2 or hui3).ti,ab,kf. 1496

24 (multiattribute$ or multi attribute$).ti,ab,kf. 864

25 (utility adj3 (score$1 or valu$ or health$ or cost$ or measur$ or disease$ or mean or gain or gains or index$)).ti,ab,kf. 15133

26 utilities.ti,ab,kf. 6921

27 (eq-5d or eq5d or eq-5 or eq5 or euro qual or euroqual or euro qual5d or euroqual5d or euro qol or euroqol or euro qol5d or euroqol5d or euro quol or euroquol or euro quol5d or euroquol5d or eur qol or eurqol or eur qol5d or eur qol5d or eur?qul or eur?qul5d or euro$ quality of life or european qol).ti,ab,kf. 12212

28 (euro$ adj3 (5 d or 5d or 5 dimension$ or 5dimension$ or 5 domain$ or 5domain$)).ti,ab,kf. 4228

29 (sf36$ or sf 36$ or sf thirtysix or sf thirty six).ti,ab,kf. 21683

30 (time trade off$1 or time tradeoff$1 or tto or timetradeoff$1).ti,ab,kf. 1859

31 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 71426

32 3 and 31 429

33 32 not 18 310

34 limit 33 to yr="2010-current" 188

Appendix B: Scatterplots from probabilistic sensitivity analysis for ERG base case

TNFi-naïve population

For the ERG base case, scatter plots showing the incremental costs and QALYs for ozanimod versus the relevant comparators considered in the TNFi-naïve population across all PSA iterations (n=1000) are presented in Figure 12 to Figure 15.

Figure 12: Cost-effectiveness plane for ozanimod versus adalimumab

Figure 13: Cost-effectiveness plane for ozanimod versus infliximab

Figure 14: Cost-effectiveness plane for ozanimod versus vedolizumab

Figure 15: Cost-effectiveness plane for ozanimod versus golimumab

TNFi-experienced population

For the ERG base case, scatter plots showing the incremental costs and QALYs for ozanimod versus the relevant comparators considered in the TNFi-experienced population across all PSA iterations (n=1000) are presented in Figure 16 and Figure 17.

Figure 16: Cost-effectiveness plane for ozanimod versus vedolizumab

Abbreviations: QALYs, quality-adjusted life years; WTP, willingness-to-pay

Figure 17: Cost-effectiveness plane for ozanimod versus ustekinumab

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

ERG report – factual accuracy check and confidential information check

Ozanimod for treating moderate to severe ulcerative colitis [ID3841]

The ERG response to the issues raised by the company during the factual accuracy check (FAC) is provided in the tables below.

Section 1: Factual inaccuracies

Executive summary

Issue 1 Continuation on conventional therapy

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 18 Section 1.2: "Patients discontinuing treatment received conventional therapy and entered the 'Active' ulcerative colitis (UC) health state accruing costs and QALYs associated with this health state." Page 127 Section 4.2.2.2: "Finally, it should be noted that in the post active model component, patients were assumed to receive conventional therapy" Page 135 Section 4.2.6.3: "In the company model (post active treatment phase), the modelled cohort progress to the 'Active UC' health state where some may continue to receive conventional treatment and still continue to experience 'Remission' or 'Response No Remission' since UC is a relapse-remitting disease."	Please amend as follows: "Patients discontinuing treatment received **best supportive care, comprising components of conventional therapy,** and entered the 'Active' ulcerative colitis (UC) health state accruing costs and QALYs associated with this health state." "Finally, it should be noted that in the post active model component, patients were assumed to receive **best supportive care, comprising components of conventional therapy**" "In the company model (post active treatment phase), the modelled cohort progress to the 'Active UC' health state where some may continue to receive **best supportive care, comprising components of conventional therapy,** and still continue to experience 'Remission' or 'Response No Remission' since UC is a relapse-remitting disease."	Patients entering the "Active UC" state following discontinuation may receive certain components of CvT, which is termed BSC in order to avoid confusion. The distinction was considered important by clinical experts as CvT is typically viewed as a first-line treatment prior to advanced therapies, whereas BSC only comprises certain components of CvT following failure of advanced treatments.	The ERG notes the importance of the distinction between CvT and BSC, comprising certain components of CvT, and agrees with this improvement to avoid confusion. Therefore, the sections of text on p.18, p.128 and p.135 have been amended according to the company's proposed changes.

Issue 2 Change from company base case

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG response</th></tr>
<tr>
<td>Pages 27 and 28 Section 1.7; Table 5.</td>
<td>Please amend as follows:

<table>
<tr><th colspan="6">Ozanimod AE discontinuation rate in maintenance phase (5% that of induction)</th></tr>
<tr><td colspan="2"></td><td>Incremental costs</td><td>Incremental QALYs</td><td>ICER £/QALY</td><td>% Change from company base case</td></tr>
<tr><td>ozanimod</td><td rowspan="5">Error! Reference source not found.</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>adalimumab</td><td>[redacted]</td><td>[redacted]</td><td>£29,790</td><td>4%</td></tr>
<tr><td>infliximab</td><td>[redacted]</td><td>[redacted]</td><td>£137,368*</td><td>-18%</td></tr>
<tr><td>golimumab</td><td>[redacted]</td><td>[redacted]</td><td>£65,285*</td><td>-8%</td></tr>
<tr><td>vedolizumab</td><td>[redacted]</td><td>[redacted]</td><td>£51,677*</td><td>-2%</td></tr>
</table>
</td>
<td>The cells in the column presenting “% Change from company base case” for golimumab and vedolizumab were incorrectly reported.</td>
<td>The ERG thanks the company for pointing out this inconsistency and agrees that the “% Change from company base case” values were swapped for golimumab and vedolizumab. Values have been corrected in Section 1.7, Table 5, p.27-28, of the ERG report.</td>
</tr>
</table>

Introduction and background

Issue 3 Description of biological treatments

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 32 Section 2.1: “The remaining biological treatments are vedolizumab, ustekinumab and tofacitinib.”	Please amend as follows: “The remaining treatments are biologics, vedolizumab and ustekinumab, and a small molecule drug, tofacitinib.	Tofacitinib is not a biological treatment, but rather a small molecule drug.	The ERG agrees with the company’s justification that tofacitinib is not a biological treatment, but a small molecule drug. The text on

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
			p.32 has been amended as proposed.

Issue 4 Safety warnings for tofacitinib

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 34 Section 2.2.1: "The CS refers to safety warnings regarding tofacitinib from the US Food and Drug Administration (FDA), though clinical advice to the ERG mentioned the more conservative approach to the safety of tofacitinib in the US."	Please amend as follows: "The CS refers to safety warnings regarding tofacitinib from the US Food and Drug Administration (FDA) **and the European Medicines Agency (EMA)**, though clinical advice to the ERG mentioned the more conservative approach to the safety of tofacitinib in the US. **Monitoring of safety warning regarding tofacitinib by the Medicines and Healthcare products Regulatory Agency (MHRA) were also referenced by the company in its response to ERG clarification question B.9.**"	The ERG report implies that the CS only refers to safety warnings issued by the FDA, when the CS/clarification response also refer to safety warnings issued by the EMA and MHRA. These safety warnings should be noted in the ERG report to accurately reflect the full company justification for tofacitinib not representing a relevant comparator for ozanimod.	The ERG notes the company's justification that the ERG report does not fully capture safety warnings issued for tofacitinib. In the interest of reporting the full company justification for excluding tofacitinib from the CS, the ERG agrees with the amendment in principle. The proposed text has been included, but with some changes (in **bold**) to provide the correct context for the ERG qualifier. The ERG report now states the following on p.34-35: "The CS refers to safety warnings regarding tofacitinib from the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA). Monitoring of safety warning**s** regarding tofacitinib by the Medicines and Healthcare products Regulatory Agency (MHRA)

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
			were also referenced by the company in its response to ERG clarification question B.9. **Clinical advice to the ERG mentioned the more conservative approach to the safety of tofacitinib in the US and, notably,** [that use of tofacitinib is increasing in the UK, driven largely by patients' preference for an oral treatment and its fast-acting nature, and estimated the use of tofacitinib to be approximately 5% in the first line and 25% in the second line in the Royal Devon and Exeter NHS Foundation Trust.]" The ERG requests it be noted that text in square brackets was originally included in the ERG report, and does not represent a further addition to this point.

Issue 5 Distinction of treatment lines by TNFi experience

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 37 Section 2.3; Table 7: While the ERG agreed with the company's decision to stratify its analyses by subpopulations	While the ERG agreed with the company's decision to stratify its analyses by subpopulations related to treatment experience, it **considered the stratification**	The CS makes no claim that TNFi experience provides an absolute distinction of first- and second-line treatments. Rather, the company	The ERG notes the company's request to include specific reference to inconsistency with the NICE

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
related to treatment experience, it did not agree that TNFi experienced provides an absolute distinction between the first and second line following CvT.	**to be inconsistent with the NICE scope in that TNFi experience does not provide an absolute distinction between the first and second line following CvT**.	has aimed to stratify the population to reflect patient groups who would and would not be suitable for treatment with a TNFi, which is largely, but not exclusively, dependent on prior exposure to TNFis. Accordingly, the CS states that TNFi exposure forms the basis for clinical decision-making, with treatment options differing in two distinct sub-populations. Whilst it is acknowledged that this differs from the stratification specified in the NICE scope (i.e. patients who have or have not been previously treated with one or more biologic), in the company's view, TNFi experience is a more relevant way of stratifying patients in terms of its impact on clinical decision-making. It is also more consistent with the subgroup data available from clinical trials that feed into the NMA (the vast majority of trials stratified patients according to TNFi experience). We therefore ask the ERG to include specific reference to inconsistency with the NICE scope to avoid misrepresenting the company's intention.	scope to avoid misrepresenting the company's intention. It agrees that this is a reasonable request to improve the representation of the company's approach, and have made the amendment to Table 7, p.38, as suggested.

Clinical effectiveness

Issue 6 TOUCHSTONE study enrolment

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 46 Section 3.2.1: "The key supporting trial, TOUCHSTONE,[34] is a multicentre, phase 2 dose-finding study with an 8-week induction and 24-week maintenance phase. The study enrolled 300 participants who were randomised to receive 0.5 mg ozanimod, 1 mg ozanimod or placebo."	Please amend The key supporting trial, TOUCHSTONE,[34] is a multicentre, phase 2 dose-finding study with an 8-week induction and 24-week maintenance phase. The study enrolled **199** participants who were randomised to receive 0.5 mg ozanimod, 1 mg ozanimod or placebo.	Sandborn *et al.* (2016)[1] reports that TOUSTONE screened 347 patients, with 199 randomised to either ozanimod 0.5 mg, ozanimod 1 mg or placebo, with 197 receiving at least one dose of study drug.	The ERG thanks the company for pointing out this inaccuracy. The text on p.47 has been amended as suggested, and further references to the number of participants in this study were revised and found to be reported correctly in Table 9 (p.48).

Issue 7 TOUCHSTONE eligibility criteria

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 47 Section 3.2.1 and Table 10 on page 49 Section 3.2.2.1: Eligibility criteria of the TOUCHSTONE trial are stated as "adults aged 18 to 73"	"Adults aged 18 to **75**"	Sandborn *et al.* (2016)[1] reports TOUCHSTONE to have an eligibility criterion of adults aged 18 to 75.	The ERG notes the reference to an eligibility criterion of adults aged 18 to 75 reported in Sandborn *et al*. (2016)[1], but further notes that the trial registry (NCT01647516) stipulates this criterion for ages 18 to 73. Given the higher rigour of a peer-reviewed publication relative to a trial registry entry, the ERG accepts this amendment and has made the change to Table 10 (p.50), and additionally to Table 9 (p.48). No further references to 73

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
			years were found in the text of Section 3.2.1.

Issue 8 TRUENORTH eligibility criteria

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 48 Section 3.2.2.1: Potential participants were identified through endoscopically confirmed UC of moderate to severe activity, defined by a Mayo score of 6 to 12.	Potential participants were identified through endoscopically confirmed UC of moderate to severe activity, defined by a **four-component** Mayo score of 6 to 12.	The four-component Mayo score was used to screen potential trial participants. We suggest including this additional context.	This is not a factual inaccuracy. The ERG notes that the four-component Mayo score at screening is not specified in the trial registry (NCT02435992) or peer-reviewed publication (Sandborn *et al.* (2021)[2]). This change has not been made.

Issue 9 Reporting of baseline characteristics in TRUENORTH

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG response</th></tr>
<tr>
<td>Pages 51 and 52 Section 3.2.2.2; Table 11: Concomitant medication and Previous medication baseline characteristics in TRUENORTH are reported as “NR” in table 11 of the ERG report.</td>
<td>Please include these baseline characteristics for TRUENORTH, as reported in the Sandborn et al. (2021) paper, e.g. :
<table>
<tr><th></th><th>Ozanimod (Cohort 1)</th><th>Placebo</th><th>Ozanimod (Cohort 2)</th></tr>
<tr><td>Concomitant medication</td><td>Glucocorticoid – 27.7%
Aminosalicylate – 87.2%</td><td>Glucocorticoid – 32.4%
Aminosalicylate – 84.3%</td><td>Glucocorticoid – 33.8%
Aminosalicylate – 85.8%</td></tr>
<tr><td>Previous medication</td><td>TNFi – 30.3%</td><td>TNFi – 30.1%</td><td>TNFi – 43.4%</td></tr>
</table>
</td>
<td>NR is an inaccurate, since baseline characteristics relating to concomitant and previous medication use are reported in the Sandborn et al. (2021)[2] publication and the CSR.</td>
<td>The ERG agrees with the company’s justification that NR in the relevant table cells is inaccurate, and that concomitant and previous medication use are reported in the trial publication and CSR. It has made the amendments to Table 11 (p.52-53) as suggested.</td>
</tr>
</table>

Issue 10 Mean age of diagnosis SD in ozanimod 0.5 mg group in TOUCHSTONE trial

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 52 Section 3.2.2.2; Table 11: Mean age of diagnosis SD in ozanimod 0.5 mg group in TOUCHSTONE trial is reported as 12.1.	Please amend the SD of mean age of diagnosis in ozanimod 0.5 mg group in TOUCHSTONE trial as **11.3**.	Typographical error.	The ERG notes the inconsistency raised by the company between the SD reported in the ERG report and the value reported in Sandborn *et al*. (2016)[1]. However, it further notes that the reporting inaccuracy stems from Appendix L.1.3, Table 73, of the CS, where this value is incorrectly reported. Given the higher relative rigour of the peer-reviewed publication, this amendment has been made to Table 11 (p.53).

Issue 11 Minor misreporting of p values

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 64 Section 3.2.5.1: "The company also reported a greater reduction in the three-component Mayo score in patients treated with ozanimod compared to those treated with placebo during maintenance at the 52-week time point (LS mean (SE) change from baseline [redacted] for ozanimod and [redacted] for placebo, [redacted].	Please amend as follows: Page 64 Section 3.2.5.1: "The company also reported a greater reduction in the three-component Mayo score in patients treated with ozanimod compared to those treated with placebo during maintenance at the 52-week time point (LS mean (SE) change from baseline [redacted] for ozanimod and [redacted] for placebo, [redacted]. In TOUCHSTONE, a significantly greater reduction in the three-component Mayo score	Typographical errors.	The ERG notes the typographical errors pointed out by the company. However, it notes that the values **p<0.020, p<0.0042** and **p<0.0004** are reported in the CS (Document B, p. 69; Appendix L.1.3, p.302; Appendix L.3.2, p.304; respectively). Despite this, the ERG accepts that these typographical inaccuracies

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
In TOUCHSTONE, a significantly greater reduction in the three-component Mayo score was reported in the 1 mg ozanimod group when compared to placebo following induction up to week 10 (mean (SD) change from baseline -3.4 (2.79) for 1 mg ozanimod and -2.0 (2.52) for placebo, p<0.0042). A significantly greater reduction was also observed in the 1 mg ozanimod group when compared to placebo after maintenance at the 32-week time point (mean (SD) change from baseline -3.4 (2.93) for 1 mg ozanimod and -1.6 (2.72) for placebo, p<0.0004)." Page 65 Section 3.2.5.1: "Though the ERG noted a slight discrepancy in the reporting of the p-value presented in the CS appendices (Appendix L.3.2., p.302 and Table 80, p.303) in the maintenance phase at week 32, a significantly greater proportion (by either value) of those in the 1 mg ozanimod arm also achieved clinical remission than in the placebo arm (21% vs. 6%, p=0.02)."	was reported in the 1 mg ozanimod group when compared to placebo following induction up to week 10 (mean (SD) change from baseline -3.4 (2.79) for 1 mg ozanimod and -2.0 (2.52) for placebo, **p=0.0042**). A significantly greater reduction was also observed in the 1 mg ozanimod group when compared to placebo after maintenance at the 32-week time point (mean (SD) change from baseline -3.4 (2.93) for 1 mg ozanimod and -1.6 (2.72) for placebo, **p=0.0004**)." Page 65 Section 3.2.5.1: "Though the ERG noted a slight discrepancy in the reporting of the p-value presented in the CS appendices (Appendix L.3.2., p.302 and Table 80, p.303) in the maintenance phase at week 32, a significantly greater proportion (by either value) of those in the 1 mg ozanimod arm also achieved clinical remission than in the placebo arm (21% vs. 6%, **p=0.01**)."		are due to the company's errors and have changed these values on p.65, as suggested, in the interest of generating an accurate record of the outcomes measured in these trials. The ERG notes the company's stated preference for **p=0.01** in the last case, and accepts this suggestion as it had already pointed out the inconsistency in reporting of the p-value in its report. Therefore, in the interest of generating an accurate record of the outcomes measured, it has amended the text on p.66 accordingly.

Issue 12 Definition of clinical remission in TRUNEORTH trial

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 64 Section 3.2.5.1: "Achievement of clinical remission with the four-component Mayo score (as defined in Section **Error! Reference source not found.**) was the primary endpoint during both the induction and maintenance phases of the TRUENORTH[27,28] study."	Please amend as follows: "Achievement of clinical remission with the **three**-component Mayo score (as defined in Section **Error! Reference source not found.**) was the primary endpoint during both the induction and maintenance phases of the TRUENORTH[27,28] study."	The Sandborn *et al.* (2021)[2] publication states: "The primary end point for both periods was the percentage of patients with clinical remission, as assessed with the three-component Mayo score."	The ERG agrees with the company's justification that the primary end point was the three-component Mayo score, as reported in Sandborn *et al*. (2021)[2]. This change has been made on p.65, as suggested.

Issue 13 Endoscopic improvement during induction period in placebo arm in TRUENORTH trial

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 66 Section 3.2.5.1: "Endoscopic improvement was significantly greater in the ozanimod arm than in the placebo arm at week 10 of the induction phase (27.3% vs. 12.0%, p<0.0001; OR (95% Wald CI) 2.88 (1.80 to 4.60)) and week 52 of the maintenance phase (45.7% vs. 26.4%, p<0.0001; OR (95% Wald CI) 2.48 (1.65 to 3.72))."	Please amend as follows: "Endoscopic improvement was significantly greater in the ozanimod arm than in the placebo arm at week 10 of the induction phase (27.3% vs. **11.6**%, p<0.001; OR (95% Wald CI) 2.88 (1.80 to 4.60)) and week 52 of the maintenance phase (45.7% vs. 26.4%, **p<0.001**; OR (95% Wald CI) 2.48 (1.65 to 3.72))."	Endoscopic improvement during induction period in placebo arm in TRUENORTH trial was 11.6% rather than 12.0%. The p value was also misreported.	The ERG notes the errors pointed out by the company. However, these values were not misreported by the ERG. The exact values reported in the ERG report, i.e. 27.3% vs. **12.0%**, p<0.0001; 45.7% vs. 26.4%, **p<0.0001** are reported in the CS (Document B, p.35, p.62 and p.67). Despite this, the ERG accepts that these instances of misreporting are due to the company's errors and have changed these values on p.67, as suggested, in the interest of generating an accurate record of the outcomes measured in these trials.

Issue 14 MCID in PCS score in induction period of TRUENORTH trial

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 69 Section 3.2.5.1: "In addition, a significantly greater proportion of patients treated with ozanimod compared to placebo achieved MCID for this score ([redacted] vs. [redacted], [redacted])."	Please amend as follows: "In addition, a significantly greater proportion of patients treated with ozanimod compared to placebo achieved MCID for this score ([redacted] vs. [redacted], [redacted])."	The p-value for this reported outcomes was [redacted], rather than [redacted] as reported in the ERG report.	The ERG thanks the company for pointing out this misreporting. The text on p.70 has been amended as suggested.

Issue 15 Justification for stratification by TNFi experience in company NMA

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 73 Section 3.3.2.3: "The company justified this approach by indicating that TNFis are used almost exclusively in the first line; clinical advice to the ERG presented a more complex situation in UK clinical practice, as described in Section 2.2.1."	Please amend as follows: "The company justified this approach by indicating the following: • **TNFis are used almost exclusively in the first line** • **With the exception of UNIFI, across all trials included in the NMA, including TRUENORTH, subgroups were stratified by TNFi experience rather than biologic experience. This terminology is therefore a more accurate classification of the subgroups in which efficacy results are available** • **This approach is in line with previous NMAs in UC (TA547)"**	The ERG's representation of the company's justification for separate subgroups by TNFi experience in its NMA is lacking the full context given in the CS.	The ERG notes the company's position that its justification for separate subgroups by TNFi experience in its NMA is not represented in its entirety in the ERG report. The ERG accepts the amendment and has changed the text on p.74-75, but with the additional changes (in **bold**) shown to include the original ERG qualifier: "The company justified this approach by indicating the following: • TNFis are used almost exclusively in the first line. • With the exception of UNIFI, across all trials

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
			included in the NMA, including TRUENORTH, subgroups were stratified by TNFi experience rather than biologic experience. This terminology is therefore a more accurate classification of the subgroups in which efficacy results are available. • This approach is in line with **a** previous **NMA** in UC (TA547)." **With respect to the first of these points, clinical advice to the ERG presented a more complex situation in UK clinical practice, as described in Section 2.2.1.**

Issue 16 Heterogeneity due to subgrouping by TNFi experience

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 74 Section 3.3.2.3: "As a result, the ERG is of the opinion that subgrouping by prior TNFi experience may have introduced heterogeneity by misclassification and also limits generalisability of the results to the NICE scope."	Please include the full context for the subgrouping of the NMA by TNFi experience rather than biologic experience. Page 74 Section 3.3.2.3: "As a result, the ERG is of the opinion that **subgrouping by prior TNFi experience may limit generalisability of the results to the**	It is inaccurate to describe subgrouping by prior TNFi experience a misclassification, since this stratification is in line with the vast majority of the trials included in the NMA. The ERG notes the Motoya and TOUCHSTONE trials as including patients with prior biologic	The ERG notes the company's justification as to why subgrouping by prior TNFi experience is not a misclassification as the majority of trials included in the NMA stratified their reporting based on prior TNFi experience. The ERG accepts

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 76 Section 3.3.3: "As discussed in Section 3.3.2, the ERG considered the stratification of the company's analyses by prior TNFi experience to be a departure from NICE scope which limits the generalisability, both to populations which are naïve to and have experience of biologics."	**NICE scope, but is in line with the method of stratification used by the majority of the trials included in the NMA**." Page 76 Section 3.3.3: "As discussed in Section 3.3.2, the ERG considered the stratification of the company's analyses by prior TNFi experience to be a departure from NICE scope which **may** limit the generalisability, both to populations which are naïve to and have experience of biologics, **but is in line with the method of stratification used by the majority of the trials included in the NMA**."	experience. However, both these trials, and all other trials included in the NMA with the exception of UNIFI, stratified enrolment by prior TNFi experience, not experience of any biologic. Patients in the TNFi-experienced subgroup therefore should all have received a prior TNFi by definition. Stratification by TNFi-experience in the NMA was therefore necessary for consistency with the available data from the trials informing the NMA. Reporting these data to reflect populations which are naïve to or have experience of biologics would be a misclassification of the available data. Heterogeneity results from the fact that trials may have differed in the proportions of TNFi-experienced patients who *also* received additional biologic therapies, an issue which is addressed separately in the CS (Document B, page 90).	the suggested amendments and have made these changes on p.75 of the report.

Issue 17 Detail of ordinal response-remission NMA

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 75 Section 3.3.2.5: "In cases where studies only reported clinical remission, but	"In cases where studies only reported clinical remission, but not clinical response, the company reported leveraging an ordinal	Further details of the ordinal response-remission NMA are given	The ERG notes the company's justification for the suggested amendment and agrees that

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
not clinical response, the company reported leveraging an ordinal response-remission NMA to retain studies – this approach was not described in greater detail in the CS or its appendices."	response-remission NMA to retain studies – this approach **is described in further detail in Section B.2.8.4 of the CS**."	on page 96 Section B.2.8.4 of the CS.	the approach, which allowed for the inclusion of studies which did not report on every outcome, is described in sufficient detail in Section B.2.8.4 of the CS. The ERG therefore accepts the proposed amendment and made these changes to p.76.

Issue 18 Inclusion of tofacitinib as a comparator treatment in the company NMA eligibility criteria

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 76 Section 3.3.3: "The selection criteria used by the company are described in the CS appendices, with specific selection criteria presented in Appendix D.2.2. (Table 7, p.50-51). The ERG considered these criteria to be broadly appropriate, and noted specifically the inclusion of tofacitinib as a comparator treatment."	Please amend as follows: "The selection criteria used by the company are described in the CS appendices, with specific selection criteria presented in Appendix D.2.2. (Table 7, p.50-51). The ERG considered these criteria to be broadly appropriate, and noted specifically the inclusion of tofacitinib as a comparator treatment, **resulting from the company's decision to include all treatments specified in the NICE final scope in the NMA.**"	The CS reports that all treatments specified in the NICE final scope were included in the NMA. This context should be reported here.	The ERG thanks the company for pointing out this additional context and agrees with its inclusion. The amendment has been made in the text (p.77) of the ERG report.

Issue 19 Outcomes reported by TNFi experience in trial maintenance periods

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 77 Section 3.3.4: "A total of 12 trials reported at least one outcome in at least one of the subgroups related to TNFi experience."	Please amend as follows: "**For the maintenance phase of treatments**, a total of 12 trials reported at least one outcome in at least one of the subgroups related to TNFi experience."	As per the first sentence in this paragraph this sentence should specify which treatment phase is being considered.	The ERG notes and agrees with the company's position that the treatment phase under consideration should be specified. This amendment has

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
			been made to p.78 in the interest of greater clarity.

Issue 20 Specification of summary statistics for baseline characteristics

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 92 Section 3.4.2.4: Table 20 reports summary statistics for Age, CRP (mg/L), Years since diagnosis, and Mayo score without denoting whether they are mean or median values.	Please amend as follows: Change "Age, years" to "Age, years (mean)". Change "CRP (mg/L)" to "CRP (mg/L) (mean)" and footnote the value for OCTAVE 2 as being a median. Change "Years since diagnosis" to "Years since diagnosis (mean)" and footnote the value for OCTAVE 2 as being a median. Change "Mayo score" to "Mayo score (mean)"	Please provide description of summary statistics.	The ERG notes and agrees with the company's position that the summary statistics for baseline characteristics of ERG-selected trials for alternative baseline placebo risk should be specified. These amendments have been made to Table 20 (p.93) in the interest of accurate and comprehensive reporting.

Issue 21 Specification or use of baseline data for Induction/Experienced

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 93 Section 3.4.2.4: The "Induction/experienced" row of Table 21 is intended to display data for OCTAVE 2, however the data values for Response or remission of 29/124 (0.23) and for Remission of 1/124 (0.008) appear to be those from OCTAVE 1+2, not OCTAVE 2.	There are two options for amendment and the optimal choice is subjective and may depend on clinical advice that the ERG had originally sought. Given "response or remission" data are not reported for OCTAVE 2 alone, please clearly state that values were from OCTAVE 1+2 and state this as a limitation due to being a departure from the selected OCTAVE 2 trial. Alternatively, consider altering the choice of trial to instead be the combination of OCTAVE 1+2 and then subsequently update baseline characteristics in Table 20 to be those from OCTAVE 1+2, as available.	The values presented appear to be misaligned with the study chosen to best align with the UK population and this needs to be either stated as a limitation or changed for accuracy.	The ERG thanks the company for pointing out this inconsistency and accepts this point. It has made changes to acknowledge the misaligned values as a limitation by amending the ERG report as follows: • A sentence has been added to the text on p.92: 'The company highlighted that remission or response

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
			data for the ERG's selected baseline trial for TNFi-experienced participants during induction (OCTAVE 2) were only available when pooled with results from the OCTAVE 1 trial. The baseline values for OCTAVE 1 are therefore also supplied in Table 20, where it can be seen that there is generally good correspondence, but that compared with OCTAVE 2 the percentage of males is about 13% higher, and the percentage with TNFi exposure about 5% less. The need to pool the ERG's selected trial (OCTAVE 2) with a similar trial (OCTAVE 1) is a limitation of the ERG's exploratory analysis.' • A column of baseline values for OCTAVE 1 has been added to Table 20 (p.93) • The entry in Table 21 (p.94) for 'Trial supplying baseline risk' has been changed from 'OCTAVE 2' to 'OCTAVE 1 + 2'

Issue 22 Three-component Mayo score sensitivity analysis

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 109 Section 3.4.4.5: "Sensitivity analyses reported by the company comprised an assumption of not pooling doses of the same active treatment if it had the same method of administration, the exclusion of trials with a treat-though design, using the three-component Mayo score instead of the four-component Mayo score, the exclusion of trials conducted exclusively in Asian populations, and the inclusion of TOUCHSTONE[34] in the TNFi-naïve analysis."	Please amend as follows: "Sensitivity analyses reported by the company comprised an assumption of not pooling doses of the same active treatment if it had the same method of administration, the exclusion of trials with a treat-though design, using the three-component Mayo score instead of the four-component Mayo score **in the TRUENORTH trial**, the exclusion of trials conducted exclusively in Asian populations, and the inclusion of TOUCHSTONE[34] in the TNFi-naïve analysis."	Data from the TRUENORTH sensitivity analysis using the 4-component Mayo score was used in the base case NMAs. The sensitivity analysis conducted to explore the influence of using the 3-component was *only* conducted using the three-component data from the TRUENORTH trial.	The ERG notes the company's point that the sensitivity analysis exploring the influence of using data from three- vs. four-component Mayo scores was only conducted using three-component Mayo score data from TRUENORTH. The ERG agrees with this point and made this amendment to p.110 in the interest of further clarity.

Cost-effectiveness analysis

Issue 23 Transition to the 'Active UC' health state

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 126 Section 4.2.2.2: "Patients that discontinued active treatment due to AEs or loss of response entered the post active treatment component of the model and were assumed to initially enter the 'Active UC' health state."	Please amend as follows: "Patients that discontinued active treatment due to AEs, loss of response **or failure to achieve response**, entered the post active treatment component of the model and were assumed to initially enter the 'Active UC' health state."	In line with Section B.3.3.7 of the CS the company's model allows patients to enter the post active treatment phase of the model due to three reasons: 1. Discontinuation due to AEs 2. Loss of response 3. Failure to achieve response	The ERG considers the company's amendment to be reasonable. Section 4.2.2.2 (p.127) in the report has been updated to reflect the company's wording.

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
		Section 4.2.2.2 of the ERG report only captures the first two reasons, excluding patients who transition to the 'Active UC' health state due to failure to achieve response during induction.	

Issue 24 Spontaneous remission calculations

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 127, Section 4.2.2.2: "The ERG noted that the modelled spontanous remission rate of 0.5% per model cycle (approximately 13% per year), was not based on clinical data, but was an arbitrary value chosen by the company to align with NICE committee preferences in TA633." Page 160, Section 6.1.2: "spontaneous remission is likely to occur for approximately 5% to 30% of flare ups, which would result in a higher per year rate than the company's modelled yearly rate of approximately 13%. This scenario analysis used a higher rate of spontaneous remission reflective of 0.75% per model cycle, approximately 20% per year"	Please amend values to 12% and 18% (rounded to nearest integer) annually respectively for each of 0.5% and 0.75% per model cycle.	Values for annual probabilities have been calculated by simply multiplying the probability per model cycle by 52/2 (number of model cycles in a year) which is an inaccurate way of converting probabilities between different time frames. Instead values should be converted to a rate per cycle, then an annual rate, and subsequently back to an annual probability.	The ERG would like to clarify that the values provided per year were only approximate estimates (which were intended to be indicative in nature) and not the exact calculated values (and hence "approximately" had been used). Nevertheless, ERG has made the amendment as follows: Page 128, Section 4.2.2.2: "The ERG noted that the modelled spontanous remission rate of 0.5% per model cycle (**12%** per year), was not based on clinical data, but was an arbitrary value chosen by the company to align with NICE committee preferences in TA633."

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
			Pages 161-162, Section 6.1.2: "…spontaneous remission is likely to occur for approximately 5% to 30% of flare ups, which would result in a higher per year rate than the company's modelled yearly rate of **12**%. This scenario analysis used a higher rate of spontaneous remission reflective of 0.75% per model cycle (**18%** per year), **which also closely corresponds to the mid-point of clinical expert opinion based estimates**".

Issue 25 Analysis of uncertainty relating to spontaneous remission

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 127 Section 4.2.2.2: "In order to explore uncertainty, the ERG has conducted a scenario analysis which used a higher rate of spontaneous remission compared to the company's base case estimate (reflective of 0.75% per model cycle)."	Please amend as follows: "**In order to explore uncertainty, the company conducted scenario analyses in which 0% and 1% rates of spontaneous response were tested.** The ERG has **additionally** conducted a scenario analysis which used a higher rate of spontaneous **response** compared to the company's base case estimate (reflective of 0.75% per model cycle)."	The ERG report has omitted discussion of the scenario analyses relating to spontaneous remission conducted by the company, which should be included for completeness.	The ERG acknowledges that the company did provide scenario analyses which tested uncertainty surrounding spontaneous remission, and the ERG agrees that commentary should reflect this for completeness. Text on p.128 has been updated to reflect the company's wording.

Issue 26 Subsequent treatments TNFi-experienced population

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 129 Section 4.2.2: "Within the current appraisal for ozanimod, the company provided limited scenario analysis which allowed for subsequent treatment usage in the TNFi-naïve subgroup (this was not conducted for the TNF-experienced subgroup as the company stated that there was a lack of available data to inform efficacy)."	Please amend as follows: "Within the current appraisal for ozanimod, the company provided limited scenario analysis which allowed for subsequent treatment usage in the TNFi-naïve subgroup. This was not conducted for the TNF-experienced subgroup as the company stated that there was a lack of available data to inform efficacy **and clinical consultation indicated that treatment decisions after failure on multiple biologics are patient-dependent and highly variable**."	As specified in Section B.3.3.5 of the CS (p.157) the company's decision to exclude subsequent treatments in the third-line setting was also informed by clinical expert opinion. Clinical experts consulted stated that subsequent treatments after failure on multiple biologics are not routine clinical practice in the UK, are patient-dependent and highly variable. The full justification for limiting subsequent treatments to 2nd-line for the TNFi-naïve subgroup should be reported.	The ERG notes the company's point and considers this to be reasonable. For completeness, the ERG have amended text on p.130, to reflect variability in subsequent treatment use within the TNFi-experienced subgroup. This was done broadly according to the company's suggestion, by stating the following: "Within the current appraisal for ozanimod, the company provided limited scenario analysis which allowed for subsequent treatment usage in the TNFi-naïve subgroup (this was not conducted for the TNF-experienced subgroup as the company stated that there was a lack of available data to inform efficacy **and clinical opinion to the company noted that treatments provided after failure on multiple biologics were likely to be patient dependent and variable**)."

Issue 27 Extended induction

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 129 Section 4.2.2: "The ERG noted that extended induction was not considered as part of the company's base case analysis on the basis that it is not standard clinical practice in the UK for all treatments."	Please amend as follows: "The ERG noted that extended induction was not considered as part of the company's base case analysis on the basis that it is not standard clinical practice in the UK for all treatments **and the limitations associated with using trial data to inform the patient distributions into the health states. Such limitations were in line with those in recent appraisal TA633 and are described in detail in Section B.3.3.3 (p.154) of the CS.**"	As specified in Section B.3.3.3 of the CS (p.154) it was not possible for the company to perform an NMA (including ozanimod) which assessed response after extended induction, as TRUENORTH did not include an extended induction period (in line with the SmPC for ozanimod). As such, direct clinical trial data were used to inform patient distribution into the 'Remission', 'Response No Remission' and 'Active UC' health states for the scenario analysis where extended induction was selected. This approach is in line with TA633, where probabilities of response and remission at the end of extended induction were derived directly from trial data, using results for individual treatment arms. However, use of within-trial data results in 'breaking of trial randomisation'. As such the company's decision to exclude extended induction in the base case resulted from a combination of feedback from clinical experts and the limitations of the available data to inform response and remission rates after extended induction. The full justification should be reported.	This is not a factual inaccuracy. However, the ERG consider that it is reasonable to include some of the company's additional justification. The text has been amended broadly according to the company's suggestion, by stating the following on p.130 of the report: "The ERG noted that extended induction was not considered as part of the company's base case analysis on the basis that it is not standard clinical practice in the UK for all treatments **and further noted limitations associated with using trial data to inform the patient distributions into the health states (see p.154 of the CS)**."

Issue 28 Exclusion of TNFis as a relevant comparator in the TNFi-experienced population

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 130 Section 4.2.4: "In the TNFi-experienced subgroup, comparators were ustekinumab and vedolizumab. The company stated that TNFis were not appropriate comparators for this subgroup, as TNFi switching no longer occurs in clinical practice."	Please amend as follows: "In the TNFi-experienced subgroup, comparators were ustekinumab and vedolizumab. The company stated that TNFis were not appropriate comparators for this subgroup, as TNFi switching no longer routine clinical practice, and thus a second TNFi is only clinically relevant in a small subgroup of TNFi-experienced patients"	The company submission (CS) does not state that TNFi switching no longer occurs in clinical practice, rather it is states that switching is *no longer routine clinical practice.* The company states this is due to advancements in therapeutic drug monitoring and increased availability of drugs with different modes of action.	The ERG considers the company's point to be partially reasonable i.e., the CS does in fact state that switching is no longer routine clinical practice. Text on p.132 of the ERG report has been updated to reflect this. The ERG does not consider the final sentence proposed by the company "…and thus a second TNFi is only clinically relevant in a small subgroup of TNFi-experienced patients" a necessary addition.

Issue 29 Time-horizon length

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 132, Section B.2.5: "The company used a lifetime horizon (58 years) in the base case analysis and justified this on the basis that a lifetime horizon has been used in previous UC appraisals including ustekinumab TA633[20] and toficitinib TA547."	Please amend to: "The company used a lifetime horizon (**58 and 60 years in the TNFi-naïve and TNFi-experienced populations, respectively**) in the base case analysis and justified this on the basis that a lifetime horizon has been used in previous UC appraisals	The mean age of the model populations differed; 40 and 42 years in the TNFi-naïve and TNFi-experienced populations, respectively. As the model had a maximum age of 100 years, the lifetime horizon had a length of 58 and 60 years in the TNFi-naïve and TNFi-experienced population, respectively.	The ERG considers the company's amendment to be reasonable. Text on p.133 has been updated to reflect the company's amendment.

Issue 30 Justification of exclusion of half-cycle correction

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 133 Section 4.2.5: "Finally, the ERG noted that the model did not incorporate a half-cycle correction. The company justified this on the basis that the model uses a short two-week cycle length."	Please amend as follows: ""Finally, the ERG noted that the model did not incorporate a half-cycle correction. The company justified this on the basis that the model uses a short two-week cycle length, **and the fact that a half-cycle correction was not applied in TA547 despite the submitted model having an eight-week cycle length.**"	Further justification based on precedence in previous appraisals for the omission of a half-cycle correction was provided by the company in answer to clarification question B.10, which has not been reported by the ERG in its report. The full justification should be reported.	This is not a factual inaccuracy. However, the ERG considers the additional rationale provided by the company in B.10 of the clarification document a reasonable addition to include in the report. Text on p.134 has been updated to reflect this.

Issue 31 Assumption of BSC post-active transitions

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 134, Section 4.2.6.3: "the ERG noted that the loss of response and loss of response (no remission) estimates were noticeably different between the non-biologic failure and biologic failure subgroups in TA633[20] (Table 43 and Table 44 TA633 committee papers), in contrast to the same pooled estimate used for both the subgroups in the current company's base case" Page 186, Section 6.3: "Revised post-active treatment transition probabilities for BSC which include an alternative means of estimating remission probabilities for BSC based on 'loss of	Please amend to: "the ERG noted that the loss of response and loss of response (no remission) estimates were noticeably different between the non-biologic failure and biologic failure subgroups in TA633[20] (Table 43 and Table 44 TA633 committee papers), **and that the company had used the data for the TNFi-experienced group in both populations, given patients receiving BSC in the model (regardless of the population selected) do so in the post-active treatment setting, and thus have failed at least one active treatment by definition.**" "(…) as opposed to **using the BSC response rates for the TNFi-experienced population for both populations in the base case.**"	The ERG may have misinterpreted the efficacy data informing the transitions for BSC. Data for the TNFi-experienced population were used to inform BSC transitions for both the TNFi-naïve and TNFi-experienced populations. "Pooled" refers to these data being derived from the placebo arms of all trials included in the NMA. TNFi-experienced data were considered more appropriate because all patients receiving BSC in the model (regardless of the population selected) do so in the post-active treatment setting, and thus have failed at least one active treatment by definition.	The ERG found the company's description in the CS to be lacking in places, thus making interpretation of the exact approach very challenging. The explanation here is much clearer, and the ERG considers that the report should be updated to reflect the clarity provided by the company. Text in the ERG report has been updated on p.137 and p.167-168.

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
remission' (directly from the sustained remission estimates) and different BSC response rates for the TNF-naïve and TNF-experienced populations, as opposed to an overall pooled estimate in the company's base case (as per 4.2.6.3)."			

Issue 32 Rounding error (QALY for golimumab in the TNFi-naïve population)

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 138 Section 4.2.6; Table 42	Please amend as follows: Study name (time horizon) / QALYs: **TNFi-naïve** / **TNFi-experienced** **Ozanimod company model (lifetime)**: Oza: [redacted] / Oza: [redacted] Ada: [redacted] / Ved: [redacted] Inf: [redacted] / Ust: [redacted] Ved: [redacted] Gol: [redacted]	In Table 68 of the CS the QALY for golimumab is reported as '[redacted]', this would result in a QALY of '[redacted]' when restricted to 2 decimal places.	The ERG notes that the value provided for golimumab total QALYs is as per the company submitted model ([redacted] without rounding) and therefore did not deem the amendment to be necessary.

Issue 33 Exclusion of conference abstracts from database searches

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 146 Section 4.2.8.1: "The ERG noted that the company did not use a recognised filter for HRQoL studies and restricted the	Please amend to: "The ERG noted that the company did not use a recognised filter for HRQoL studies and restricted the bibliographic database searches to exclude conference abstracts, **as these**	As clarified in response to the ERG clarification B.2, conference abstracts were excluded from the database searches as these had	The ERG believes the company is referring to Section 4.2.7.1 (previously p.142). This is not a factual inaccuracy, however for completeness the

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
bibliographic database searches to exclude conference abstracts."	**were separately hand searched and excluded to avoid double counting.**"	already been hand searched previously.	additional rationale provided by the company has been added to text on p.141 of the ERG report.

Issue 34 Acquisition cost of adalimumab

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 146, Section 4.2.8.1: "However, based on cross validation of the company's medicine acquisition costs with those reported in the BNF and TA633,[20] the ERG noted that the company's cost for adalimumab (£633 for 40 mg/0.8 mL) represented the solution for injection pre filled syringes."	Please amend to: "However, based on cross validation of the company's medicine acquisition costs with those reported in the BNF and TA633,[20] the ERG noted that the company's cost for adalimumab (**£633.60** for 40 mg/0.8 mL) represented the solution for injection pre filled syringes."	As specified in Table 57 of the CS the pack cost for adalimumab was £633.60.	This is a very minor point; however, the ERG agrees that the pence can be added (see p.147 of the ERG report).

Issue 35 Resource use cost for care without colectomy

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 150, Section 4.2.8.5; Table 49.	Please amend to: **Resource item** / **Unit cost** **Outpatient** Consultant visit: £183.43 Blood test: £1.81 **Inpatient** Emergency endoscopy: £814.46 Elective endoscopy: £330.51	As specified in Table 64 of the CS the cost of care without colectomy is £2,301.47.	The ERG notes the very minor typographical error. The correction has been made to Table 49 (p.151).

Description of problem	Description of proposed amendment		Justification for amendment	ERG response
	Care without colectomy	**£2,301.47**		
	Stoma care (post-colectomy)	£541.75		

Issue 36 Administration company (Pairwise) base case results: TNFi-experienced (probabilistic)

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 152, Section 5.1.2; Table 52.	Please amend to: *Company probabilistic base case* **Intervention** / **Total costs** / **LYG** / **Total QALYs** / **Inc. costs** / **Inc. LYG** / **Inc. QALYs** / **Cost/QALY gained** Ozanimod / [redacted] / NR / [redacted] / - / / - / - Vedolizumab / [redacted] / / [redacted] / [redacted] / - / [redacted] / £1,324,054* Ustekinumab / [redacted] / / [redacted] / [redacted] / - / [redacted] / Ozanimod dominant	As the total costs and QALYs were not reported in the CS, it is assumed that the ERG re-ran the PSA to obtain the results presented here. However, ICERs presented here are aligned with the CS. The company has updated the values with data extracted from the original PSA reported in the CS (Table 72) for consistency.	As the company has noted, the results provided in the report were based on the ERG re-run of PSA and therefore the ERG did not deem the company's amendment to be necessary. However, in the interest of completeness, a sentence has been added to Section 5.1.2 (p.153) to clarify that the company's probabilistic results presented for the TNFi-experienced population were based on the ERG's re-run of the PSA, as these were not initially provided by the company.

Issue 37 Fully incremental results

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 152 Section 5.1.3:	Please amend to:	In line with the the cost-effectiveness frontier, the interpretation of the fully	The ERG considers the company's point to be reasonable. The amendment

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
"Golimumab and infliximab were extendedly dominated and dominated respectively by adalimumab."	"**Infliximab was dominated by golimumab, and golimumab was extendedly dominated by vedolizumab.**"	incremental analyses should be amended to note that infliximab was dominated by golimumab (which was associated with lower total costs but higher total QALYs) not adalimumab. Similarly, golimumab was extendedly dominated by vedolizumab.	has been made in Section 5.1.3 (p.153-154) as suggested.

Issue 38 Comparators PAS discounts

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 153 Section 5.2.1: "Overall, the ERG considered the company's OWSA to be useful in deteriming the sensitivity of model parameters to variation, however the results were be of limited use for decision making/interpretation as most parameters were varied by an abitrary percentage, and cPAS results were not included for comparator treatments.	Please amend as follows: "(…) and cPAS results were not included for comparator treatments, **as these were not available to the company**."	The company does not have information relating to the comparators' agreed PAS discounts, and as such could not run any analyses which included these.	This is not a factual inaccuracy. The ERG does not consider that the amendment is required.

Issue 39 Cost comparison versus tofacitinib, clinical efficacy

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 160, Section 6.1.1: "This scenario analysis assumed clinical equivalency between treatments in terms of efficacy and only included differences in drug acquisition costs, monitoring	Please amend to: "(…) This scenario analysis assumed clinical equivalency between treatments in terms of efficacy, **with patients spending the whole modelled time horizon on each active**	The way that this cost comparison has been conducted calculates the costs per year on each treatment and then multiplies by the model time horizon minus the baseline starting age. Whilst this technically	The ERG notes that this is not strictly a factual inaccuracy as the assumptions mentioned were not exhaustive, but only indicative of a few key assumptions. Nevertheless,

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
costs and adverse event costs over the modelled time horizon."	**treatment for simplicity**, and only included differences in drug acquisition costs, monitoring costs and adverse event costs over the modelled time horizon."	assumes equal efficacy in that all patients remain on active treatment for their whole lifetime, a simplifying assumption is made that no patients discontinue treatment, which should be reported here.	the ERG considers the suggestion to be reasonable and it has been incorporated as follows: Page 161, Section 6.1.1: "This scenario analysis assumed clinical equivalency between treatments in terms of efficacy and only included differences in drug acquisition costs, monitoring costs and adverse event costs over the modelled time horizon **(without considering discontinuation from the active treatment)**."

Issue 40 Tofacitinib cost-savings

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 160 Section 6.1.1: "Based on this analysis ozanimod resulted in a cost saving of [redacted] and [redacted] in the TNFi-naïve and TNFi-experienced subgroups respectively (...)"	Please amend as follows: "Based on this analysis ozanimod resulted in a cost saving of [redacted] and [redacted] in the TNFi-naïve and TNFi-experienced subgroups respectively (…)"	Currency symbols have been omitted in the reporting of the cost-savings in the ERG's analysis.	The ERG considers this to be a very minor point. The currency symbols have been added to the text on p.161.

Issue 41 Sensitivity of AE incidence

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 161, Section 6.1.4: "Also, the ERG noted that the rate used in the model was not tested as part of sensitivity analysis, which introduced further uncertainty."	Please amend to: "Also, the ERG noted that the **incidence of AEs in the model was not explicitly varied, however, the AE cost per cycle was varied as part of sensitivity analysis. This implicitly tests the sensitivity of AE incidence as AE costs per cycle are calculated as the AE incidence per cycle multiplied by the unit cost of managing an AE.**"	While the incidence of AEs per cycle are not explicitly varied in the model, the cost per cycle for AEs is varied. As AE costs per cycle are calculated as AE incidence multiplied by the unit cost of an AE, then varying the cost by 20% implicitly tests the sensitivity of the model to AE incidence.	The ERG does not consider this a factual inaccuracy. However, the ERG has amended text on p.162 to reflect that the AE cost per cycle was varied.

Issue 42 Increased proportion of SC vedolizumab

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 162, Section 6.1.5: "Based on clinical input to the ERG, a 50% split is likely to be a reasonable assumption, however it was noted that patients are being steadily phased onto SC vedolizumab over time and therefore the majority of patients are likely to receive SC vedolizumab after one year. In order to reflect this opinion, in this scenario, the ERG assumed that 80% of patients receive SC vedolizumab after year 1."	Please amend to: "(…) In order to reflect this opinion, in this scenario, the ERG assumed that 80% of patients receive SC vedolizumab in the maintenance phase. Although, this is likely an overestimate as it is unlikely that 80% patients will start treatment on SC vedolizumab after the 6 week induction, and the discounted mToT calculated for vedolizumab in the model is only 1.59 years."	This scenario has been implemented such that all vedolizumab usage in maintenance is 80% SC. Induction for vedolizumab is 6 weeks therefore this percentage will be applied for the majority of vedolizumab treatment in all years, not just applied after year 1. Based on the clinical feedback mentioned, this is likely an overestimate.	Though not strictly a factual inaccuracy, the ERG considers the company's suggestion to be reasonable and text in the report has been amended as follows: Page 163, Section 6.1.5: "In order to reflect this opinion, in this scenario, the ERG assumed that 80% of patients receive SC vedolizumab **in the maintenance phase (patients typically start treatment on SC vedolizumab after the 6-week induction period).**"

Issue 43 Costs savings associated with SC vedolizumab

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 162 Section 6.1.5: "Based on this analysis, ozanimod incremental savings reduced from [redacted] to [redacted], due to reduced administration costs associated with SC vedolizumab."	Please amend as follows: "Based on this analysis, ozanimod incremental savings reduced from [redacted] to [redacted], due to reduced administration costs associated with SC vedolizumab, **as well as the reduction in the proportion of patients modelled to receive dose escalation (since this is not applied to patients receiving SC vedolizumab)**."	By increasing the share of patients receiving SC vedolizumab, the number of patients receiving dose escalated IV vedolizumab will be reduced. As such, the cost savings associated with increased share of SC vedolizumab are not only due to reduced administration costs.	The ERG considers the company's suggestion to be reasonable and amended text in its report as follows: Page 163, Section 6.1.5: "Based on this analysis, ozanimod incremental savings reduced from [redacted] to [redacted], due to reduced administration costs associated with SC vedolizumab **and the reduction in the proportion of vedolizumab IV patients modelled to receive dose escalation**."

Issue 44 Implementation of the loss of remission values used for post-active transitions is incorrect in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Model tab Efficacy Active Treatment, cells L55-L58 and T55. Values have been converted using the time periods for incorrect treatments.	Please amend to: In cell L55, the formula for golimumab Loss of Remission =1-('Sensitivity Analysis Filter'!P49)^(model_cycle/(I42-I18)). This should be updated to =1-('Sensitivity Analysis Filter'!P49)^(model_cycle/(**I44**-I18)).	In cell L55, the formula for golimumab Loss of Remission =1-('Sensitivity Analysis Filter'!P49)^(model_cycle/(I42-I18)), where I42 corresponds to the 52 week follow up of vedolizumab (IV), not the 60 week follow-up of golimumab (cell I44). This only has an effect on the ERGs base case due to the use of the Loss of Remission values for post-active BSC which has subsequently also been calculated slightly	The ERG thanks the company for pointing this out and acknowledges that this is an error. The updated version of the ERG model has subsequently been corrected to reflect this. The ERG apologises for this error.

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
	All formuale should be updated in this way, most importantly, the value for TNFi-naïve BSC Loss of Remission (cell L58) should be updated from "=1-('Sensitivity Analysis Filter'!P52)^(model_cycle/(l45-l21))" to "=1-('Sensitivity Analysis Filter'!P52)^(model_cycle/(**l47**-l21))"	incorrectly. Please note that, despite this error, the results for the TNFi-experienced population are unaffected.	

Issue 45 Change in ERG base case results due to model error

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 164, section 6.2, table 60. Based on the above model innacuracy, the ERG base case results for the TNFi-naïve population require amendment.	Please amend to: **Revised modelled efficacy estimates for BSC in the post-active treatment phase** ozanimod 6.3 - - - - adalimumab [redacted] [redacted] £28,797 0% infliximab [redacted] [redacted] £167,294* 0% golimumab [redacted] [redacted] £71,133* 0% vedolizumab [redacted] [redacted] £52,859* 0%	The post-active treatment BSC transitions for the TNFi-naïve population have been calculated incorrectly. This subsequently affects all results for the TNFi-naïve population.	Following the model error fix in the updated ERG model (as mentioned above in Issue 44), the ERG has amended the results reported in Table 3 (p.25), Table 5 (p.27-28), Table 60 (p.164-165) and Table 62 (p.169-170) of the ERG report.
Page 168, section 6.3, table 62. Based on the above innacuracy, the ERG base case results for the TNFi-naïve population require amendment.	Please amend to: **Revised modelled efficacy estimates for BSC in the post-active treatment phase** ozanimod [redacted] [redacted] - - -	The post-active treatment BSC transitions for the TNFi-naïve population have been calculated incorrectly. This	The ERG have updated these results based on the fix above (see Table 62, p.169-170, in the ERG report). Note that for the revised modelled efficacy estimates for

	adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,794	subsequently affects all results for the TNFi-naïve population.	BSC in the post-active treatment phase, the ERG identified the incremental cost of vedolizumab to be [redacted] (as opposed to the previous value of [redacted]).
	infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£169,791*		
	golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£82,863*		
	vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,640*		
	Cumulative impact of ERG preferences (deterministic)							
	ozanimod	[redacted]	[redacted]	-	-	-		
	adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,794		
	infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£169,791*		
	golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£82,863*		
	vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£56,640*		
	Cumulative impact of ERG preferences (probabilistic)							
	ozanimod	[redacted]	[redacted]	-	-	-		
	adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	£27,794		
	infliximab	[redacted]	[redacted]	[redacted]	[redacted]	£158,636*		
	golimumab	[redacted]	[redacted]	[redacted]	[redacted]	£87,723*		
	vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£68,499*		
Page 167, section 6.3. “In the TNFi-naïve subgroup, pairwise deterministic analysis indicated that the ICER for ozanimod compared to adalimumab was £27,851.” “Probabilistic analysis resulted in similar conclusions with an ICER for ozanimod compared to adalimumab of £27,897”	Please amend to: “In the TNFi-naïve subgroup, pairwise deterministic analysis indicated that the ICER for ozanimod compared to adalimumab was **£27,794**.” “Probabilistic analysis resulted in similar conclusions with an ICER for ozanimod compared to adalimumab of **£27,794**”						Subsequent reporting of model results in the text should be updated to reflect the corrected cells in the model. Note the incremental costs or QALYs are not affected at the	Based on the ERG’s probabilistic analysis, the ICER for ozanimod compared to adalimumab was £27,842 (not £27,794, as estimated by the company). The ERG has implemented this edit in

		reported level of significance.	the report. See p.168 of the ERG report.
Page 170, section 6.4. “For the TNFi-naïve subgroup, based on the ERG’s preferred results, ozanimod was cost effective compared to adalimumab at a conventional willingness-to-pay threshold of £30,000, resulting in an ICER of £27,851”	Please amend to: “(…) resulting in an ICER of **£27,794**”	Subsequent reporting of model results in the text should be updated to reflect the corrected cells in the model. Note the incremental costs or QALYs are not affected at the reported level of significance.	As above.
Appendix B. Scatterplots for the TNFi-naïve population ERG base case.	Please amend to: 	As the ERG base case values for BSC post-active transitions were implemented incorrectly the PSA was rerun to generate these scatter plots. Please note they are also CIC.	The figures have been updated in Appendix B of the ERG report (p.188-193) following the model fix.

Section 2: Confidentiality highlighting amendments

Location of incorrect marking	Description of incorrect marking	Amended marking	ERG response
Page 51 Section 3.2.2; Table 11	The ranges are not reported in the Sandborn *et al.* (2021) publication, which only reports interquartile ranges for these characteristics. The ranges are only reported in the CSR and should therefore be highlighted as AIC.	Please amend to: **Characteristic** / **TRUENORTH** Median C-reactive protein (mg/L) (range): 4.0 ([redacted]) / 5.0 ([redacted]) / 5.0 ([redacted]) Median faecal calprotectin (µg/g) (range): 1079.48 ([redacted]) / 1349.79 ([redacted]) / 1259.85 ([redacted])	The ERG notes and agrees with the request to amend the marking for ranges for C-reactive protein and faecal calprotectin in TRUENORTH, as these were only reported in the CSR. These changes to marking have been made in Table 11 (p.52).

<table>
<tr><th>Location of incorrect marking</th><th>Description of incorrect marking</th><th>Amended marking</th><th>ERG response</th></tr>
<tr><td>Pages 95–102, Figures 1–8,
Pages 104–109, Tables 23–30,
Pages 113–114, Tables 31 and 32,
Pages 116–117, Figure 9 and Table 33</td><td>NMA results have not yet been published, so should be marked AIC.</td><td>Please add AIC highlighting to all results presented in Tables 23–33, and Figures 1–9.</td><td>The ERG thanks the company for pointing out that the NMA results have not yet been published and are therefore AIC. These changes have been made.</td></tr>
<tr><td>Page 130 Section 4.2.3; Table 38</td><td>These data should be marked as AIC as they have not yet been published.</td><td>Please amend to:<table><tr><td rowspan="2">Characteristic</td><td colspan="2">Population</td></tr><tr><td>TNFi-naïve</td><td>TNFi-experienced</td></tr><tr><td>Mean weight, kg</td><td>[redacted]</td><td>[redacted]</td></tr><tr><td>Proportion of female, %</td><td>[redacted]</td><td>[redacted]</td></tr><tr><td>Mean age, years</td><td>[redacted]</td><td>[redacted]</td></tr></table></td><td>The ERG notes the company’s point that these data have not yet been published and are consequently AIC. These changes to the marking have been made on p.96-103, p.105-110, p.114-115 and p.117-118.</td></tr>
<tr><td>Page 138 Section 4.2.6; Table 42</td><td>These data should be marked as CIC in line with Table 68 of the CS. The QALY for vedolizumab in the TNFi-naïve population has also been amended to 9.81 in line with Table 68 of the CS (see Issue 12).</td><td>Please amend to:<table><tr><td rowspan="2">Study name (time horizon)</td><td colspan="2">QALYs</td></tr><tr><td>TNFi-naïve</td><td>TNFi-experienced</td></tr><tr><td>Ozanimod company model (lifetime)</td><td>Oza: [redacted]</td><td>Oza: [redacted]</td></tr><tr><td></td><td>Ada: [redacted]</td><td>Ved: [redacted]</td></tr><tr><td></td><td>Inf: [redacted]</td><td>Ust: [redacted]</td></tr><tr><td></td><td>Ved: [redacted]</td><td></td></tr><tr><td></td><td>Gol: [redacted]</td><td></td></tr></table></td><td>The ERG thanks the company for pointing out this inconsistency in CIC marking when compared to Table 68 of the CS. These changes to the marking have been made in Table 42 (p.139).

The ERG does not agree with the proposed amendment of the QALY for golimumab to [redacted], as the exact value is [redacted] (see Issue 12).</td></tr>
<tr><td>Page 154, Section 5.2.2</td><td>The incremental QALYs are marked as CIC on p.181 of the CS as they provide</td><td>Please amend to:</td><td>The ERG thanks the company for pointing out this inconsistency in CIC marking when compared to</td></tr>
</table>

Location of incorrect marking	Description of incorrect marking	Amended marking	ERG response
	information on the confidential PAS price for ozanimod. These values should therefore be marked as CIC in the ERG report.	"The ERG noted this difference in incremental QALYs between the base case ([redacted]) and PSA ([redacted]); however, did not find any further issues associated with it."	p.181 of the CS. These changes to the marking have been made on p.155.
Page 162, Section 6.1.6	The incremental costs and QALYs are as they provide information on the confidential PAS price for ozanimod. These values should therefore be marked as CIC in the ERG report.	Please amend to: For instance, for the comparison of ozanimod versus adalimumab, although the difference in the total drug acquisition costs with per treatment cycle approach was only around [redacted], the ICER increased to >£33k (versus £28k in the base case) as the incremental QALYs were lower ([redacted]). See Section **Error! Reference source not found.** for the results.	The ERG thanks the company for pointing out this omission in CIC marking, given that these values provide information on the confidential PAS price for ozanimod. These changes to the marking have been made on p.163.

References

1. Sandborn, WJ; Feagan, BG; Wolf, DC *et al.* Ozanimod Induction and Maintenance Treatment for Ulcerative Colitis. New England Journal of Medicine. 2016;374(18):1754-1762.
2. Sandborn, WJ; Feagan, BG; D'Haens, G *et al.* Ozanimod as Induction and Maintenance Therapy for Ulcerative Colitis. New England Journal of Medicine. 2021;385(14):1280-1291.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Ozanimod for treating moderately to severely active ulcerative colitis [ID3841]

As a stakeholder you have been invited to comment on the evidence review group (ERG) report for this appraisal.

Your comments and feedback on the key issues below are really valued. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the Company involved in this appraisal, please complete the 'Summary of changes to the Company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

NICE National Institute for Health and Care Excellence

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by **5pm** on **18 May 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	Seyavash Najle-Rahim
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Celgene (a Bristol Myers Squibb Company)
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	N/A

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the ERG report.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Tofacitinib was excluded as a comparator in TNFi-naïve and -experienced subgroups	Yes (New analyses; Appendix 4)	The Company maintain that the reservation of tofacitinib to later treatment lines in UK clinical practice due to concerns regarding its safety profile negate it as a relevant comparator to ozanimod. However, for completeness, tofacitinib has been included in both populations as part of a revised economic model for the Committee's consideration. The ERG expressed concerns about the exclusion of tofacitinib as a comparator in both the tumour necrosis factor-alpha inhibitor (TNFi) naïve and TNFi-experienced populations in the Company model, citing clinical opinion that the use of tofacitinib in clinical practice was increasing. As noted in the Company submission (Section B.1.1), tofacitinib was not viewed as a relevant comparator as clinical consultation received as part of the appraisal noted that whilst tofacitinib may be effective for some patients, concerns regarding its safety profile mean it is not routinely used in UK clinical practice, and when used is typically reserved for later treatment lines.[1] This is in line with the opinion of clinicians consulted in TA633, and was re-affirmed in additional clinical consultation conducted by the Company as part of this response. As noted in the Company's response to the ERG's clarification question B.9, there has since been no downgrading of the European Medicines Agency (EMA) warnings and restrictions regarding the use of tofacitinib;[2] rather, a safety a

		review of JAK inhibitors has recently been commissioned by the EMA, which may lead to further restrictions on the use of tofacitinib in patients with moderate to severe ulcerative colitis (UC).[3] In the USA, the Food and Drugs Administration (FDA) has also issued warnings associated with tofacitinib, noting increased risk of serious heart-related events, cancer, blood clots, and death.[4] In reference to these safety warnings the FDA has restricted approved use of tofacitinib in UC to only certain patients who are not treated effectively or who experience severe side effects with TNFis. Clinical feedback sought as part of this response acknowledged that whilst safety concerns surrounding the use of tofacitinib do affect the US market differently, these concerns are also relevant to European practices. Safety is a critical factor in the choice of treatment and is highly prioritised in clinical decision making, which may account for the use of tofacitinib later in the treatment pathway. Clinical feedback sought by the ERG estimated that 5% of TNFi-naïve patients received tofacitinib first line in the Royal Devon and Exeter NHS Foundation Trust. Clinician feedback sought as part of this response agreed that tofacitinib may be used as a first-line treatment option in a minority of TNFi-naïve patients. However, it was noted that the Royal Devon and Exeter NHS Foundation Trust consulted by the ERG is a specialised tertiary referral centre, and as such is likely to represent one end of the clinical spectrum and may not be generalisable to UK clinical practice as a whole. In particular, clinical feedback received as part of the technical engagement process indicated that the proposed estimate of 25% of TNFi-experienced patients receiving tofacitinib was likely an overestimate compared with what might be expected in broader UK practice, potentially reflecting the greater expertise and resource available to specialised tertiary centres. Clinical feedback also noted that 5% of TNFi-naïve patients receiving tofacitinib likely represented a maximum estimate for the use of tofacitinib in this patient group across the UK.

		In light of the above arguments, the Company maintain that tofacitinib should not be considered a relevant comparator in the TNFi-naïve or TNFi-experienced populations. However, given the likely heterogeneity in the use of tofacitinib in UK clinical practice, for completeness and to reduce uncertainty in the committee's decision making, tofacitinib has been included in the cost-effectiveness model in both populations. The results of these analyses when the updated PAS price for ozanimod has been applied are summarised in Table 3 below. Ozanimod was found to be cost-effective compared to tofacitinib in both populations at a willingness-to-pay threshold of £30,000. Full details of the model inputs utilised in these analyses, as well as the impact on the sensitivity and scenario analyses are presented in Appendix 4.
Key issue 2: Baseline risks for placebo anchors in the NMAs taken from the same trials those used for relative risk	Yes (New analyses; Appendix 1)	To address the ERG's concerns, the Company has updated the sources informing baseline risk for placebo anchors in the network meta-analyses (NMAs) to better reflect the UK patient population. The ERG highlighted concerns over the methods used to derive baseline risk for placebo anchors used in the NMAs in the Company submission. The Company acknowledges the limitations associated with the calculation of baseline risk using all trials included in the NMA, and agrees that it may be more appropriate to select individual sources that better reflect the decision problem and UK clinical practice, in line with the recommendations in TSD5. In the Company submission, the baseline risk for placebo anchors included in the NMA (the probability of being in non-response and non-remission under placebo) were calculated from the same set of trials used to calculate the relative treatment effect. This approach was utilised in order to maximise the use of available data, given that baseline characteristics vary across trials and thus selection of an individual trial that is most representative of the UK patient population across all characteristics is challenging. This approach was intended to average trial variability and provide an unbiased method of

		selecting baseline risk, which the Company maintain represents a valid approach. However, it is acknowledged that this method is not strictly consistent with TSD5, which recommends separate modelling of relative treatment effects and baseline effects in most cases.[5] It was further noted that this approach may have limitations due to the heterogeneity in baseline characteristics across the placebo arms of included trials, meaning some trials were likely to have been included which may be less reflective of the UK patient population than other included trials. The Company agrees that conducting a systematic review to identify studies which are highly generalisable to the UK context may represent an optimal approach. However, given the high number of hits expected from a systematic search for additional evidence sources such as real-world evidence (RWE) and observational studies in UC, conducing a proper, protocol-driven SLR was not feasible within the technical engagement timeframe. Instead, a targeted search of the literature was conducted to identify potential real-world evidence sources which may be more generalisable to UK clinical practice. A review of UK national registry reports[6, 7] as well as recent guidelines on the management of ulcerative colitis[8, 9] and their reference lists revealed a scarcity of demographic data, and no data that could inform baseline risk. Clinical consultation conducted by the Company as part of the technical engagement process did not highlight any new sources of baseline risk as being more generalisable than those identified by the ERG. In addition, it was noted that any recent observational or RWE studies in patients with moderately to severely active UC would be unlikely to provide evidence for placebo response required for the NMA, given patients in real-world practice receive active treatments. The Company therefore considers the ERG's approach of utilising placebo arm values from individual trials included in the NMAs that were more generalisable to the UK context to be a suitable method for deriving baseline risk for placebo anchors, given the time available. As acknowledged by the ERG, limitations do remain with this approach.

		Whilst individual trials may better reflect certain UK patient characteristics, they may not be reflective of UK patients in all respects, and thus their selection could be considered arbitrary. Despite these limitations, the method proposed by the ERG allows for an estimation of baseline risk which may be more generalisable to the UK population than the Company's original base case, and more closely aligns with the recommendations in TSD5. The Company has therefore amended their method for deriving baseline risk for placebo anchors to align with the changes proposed by the ERG. The trials used to derive baseline risk in the NMA were: • PURSUIT SC for the TNFi-naïve subgroup in the induction period NMA • OCTAVE 1 + 2 pooled[1] for the TNFi-experienced subgroup in the induction period NMA • PURSUIT M for the TNFi-naïve subgroup in the maintenance period NMA • GEMINI1 for the TNFi-experienced subgroup in the maintenance period NMA The methodology and results of this analysis are presented in the Appendix 1, and their impact on cost-effectiveness results in isolation are presented in Table 3. Overall, with the exception of the comparison with infliximab in the TNFi-naïve population, this change to the Company base case improves the cost-effectiveness of ozanimod in all comparisons. This approach to modelling baseline risk has been incorporated into the Company's revised base case, the full results of which are presented in Appendix 4.

[1]Remission or response data for the ERG's selected baseline trial for TNFi-experienced participants during induction (OCTAVE 2) were only available when pooled with results from the OCTAVE 1 trial.

Key issue 3: A random effects model may be more appropriate for use in the maintenance phase NMAs	Yes (New analyses; **Error! Not a valid result for table.**)	The Company's base case NMA has been revised, with all analyses conducted using random effects models using informative priors, in line with the ERG's preferred approach. The ERG raised concerns over the use of fixed effect models in the Company NMAs for both the TNFi-naïve and TNFi-experienced populations. In the original base case analysis, fixed effects models were selected in both the TNFi-naïve and TNFi-experienced populations for the maintenance period NMAs, as well as the induction period NMA for the TNFi-experienced population. A fixed effects approach was utilised owing to the fact that the random effects model did not converge and a reasonable statistical fit was obtained from the fixed effects model (Document, B Section B.2.8.5). In the induction period in the TNFi-naïve population NMA, a random effects model was used, as this was associated with greater statistical fit. The Company acknowledges that due to the presence of heterogeneity between included studies (Document B, Section B.2.8.3) random effects models were likely to have better clinical validity if an appropriate prior distribution could be identified. As recommended by the ERG, key publications by Turner *et al.* (2012) and Turner et al. (2015) were consulted.[10, 11] The Turner *et al.* (2015) publication was preferred as it represented a more recent extension of the work done by Turner *et al.* (2012). The methodology and results of the analysis exploring random effects models with an informative prior distribution are presented in Appendix 2, and the impact of this change on the cost-effectiveness results in isolation is presented in Table 3. As outlined in **Error! Not a valid result for table.**, the change to efficacy estimates when using a random effects NMA as opposed to a fixed effects were small, with similar point estimates for all comparisons, in both populations, and across both the induction and maintenance phases observed. However, the use of random effects models led to widened credible intervals, resulting in fewer statistically significant results. Of note, tofacitinib and vedolizumab no longer showed statistically significant differences in

		response or remission in the maintenance phase NMA for the TNFi-naïve population, compared with ozanimod. Use of random effects models to generate clinical efficacy estimates for all treatments had limited impact on the cost-effectiveness results (see Appendix 4). Appendix 3 presents the results of an analysis where the source of placebo baseline risk was revised as detailed in the response to key issue 2, and where random effects models with an informative prior distribution were applied as detailed in this response to key issue 3. The clinical efficacy estimates derived from this analysis were incorporated into the revised base case for the Company's cost-effectiveness analysis, the results of which are summarised in Table 3 and presented in full in Appendix 4. With the exception of infliximab in the TNFi-naïve population, as with the change resulting from key issue 2, this approach improves the cost-effectiveness for ozanimod compared to all comparators.
Key issue 4: Modelled efficacy estimates for BSC in the post-active treatment phase	Yes (New analyses; Appendix 4)	The Company maintain that TNFi-experienced data are more appropriate to inform transition probabilities for best supportive care (BSC) when the TNFi-naïve subpopulation is selected, given patients transitioning to active treatment in the model have failed at least one treatment. The Company however agree with the ERG that it is more suitable to calculate transition probabilities directly from the sustained remission estimates via 'loss of remission'. **Population-specific transition probabilities for BSC** The ERG raised concerns over the Company's estimation of modelled transition probabilities for BSC in the post active treatment phase of the model. In the Company model, the modelled cohort may progress to the 'Active UC' health state upon failure of active treatment, where some may continue to receive BSC comprising of

		components of conventional therapy (CvT). In line with the relapsing-remitting nature of the disease, patients in this health state may experience 'Remission', or 'Response no Remission'. In the Company model, transition probabilities among the 'Active UC', 'Remission' and 'Response No Remission' health states for BSC in both the TNFi-naïve and TNFi-experienced populations were informed by data from the TNFi-experienced NMAs. This was owing to the fact that patients receiving BSC in the model (regardless of the population selected) do so in the post-active treatment setting, and thus have failed at least one active treatment by definition. The Company therefore maintain that TNFi-experienced data are more appropriate to inform transition probabilities for BSC when the TNFi-naïve subpopulation is selected. The ERG considered it more appropriate to directly calculate the transition probabilities from available subgroup data, so loss of response and loss of response (no remission) were based on TNFi-naïve and TNFi-experienced estimates, as appropriate. Transition probabilities for BSC in the ERG's revised base case therefore differed between the TNFi-naïve and TNFi-experienced subpopulations. In order to reduce uncertainty, a scenario analysis has been explored in which the ERG's preferred approach for BSC transitions has been incorporated, but have not been included in the revised Company base case. The results of this scenario analysis are presented in Appendix 4. **Remission transition probabilities** The ERG raised concerns over the Company's estimation of response rates in the post active treatment phase of the model. The Company used data for loss of overall response (including remission) to inform remission transition probabilities for BSC. The ERG considered that the remission probabilities for BSC could have been calculated through 'loss of remission' (calculated directly from sustained remission estimates). It is acknowledged that this approach may

		be more suitable as it aligns with the approach taken for active treatments in the model. This amendment has therefore been incorporated into the revised Company base case. Full details of the inputs utilised to inform these transitions in the Active UC health state are provided in Appendix 4, and their impact on cost-effectiveness results in isolation is presented in Table 3. The implementation of BSC transition probabilities derived from "loss of remission" resulted in very marginal changes to the base case cost-effectiveness results, with all changes in incremental cost-effectiveness ratios (ICER) within approximately 1% for all comparators.
Key issue 5: There is uncertainty surrounding the handling of subsequent treatments/treatment sequencing within the model	Yes (New analyses; Appendix 5)	The Company has updated the cost-effectiveness model to allow greater flexibility in exploring treatment sequences. The ERG noted that there was uncertainty surrounding the handling of subsequent treatments within the Company model. The Company have therefore updated the economic model to allow for the consideration of various treatment sequencing options, but maintain the most relevant treatment sequences were those explored in the scenario analyses presented in the Company submission. Given the availability of multiple treatment options in clinical practice for patients with moderately to severely active UC, scenario analyses were explored in the TNFi-naïve population in the Company submission (Document B, Section B.3.8.3) where both vedolizumab and ustekinumab were modelled as subsequent treatments following receipt of either ozanimod or TNFis first line. This approach is similar to that taken in TA633.[1] Ustekinumab was explored as a subsequent treatment for the small subset of TNFi-naïve patients who receive vedolizumab first line. The subsequent treatments explored for ozanimod and relevant comparators were informed by clinical expert

		opinion and represent the most common subsequent treatments received by patients in UK clinical practice. In order to reduce clinical uncertainty in the economic evaluation, increased flexibility has been added to the Company model to allow for modelling of additional subsequent treatments in both populations. Subsequent treatment options include any treatments where data were available from the TNFi-experienced NMAs, in line with approach taken by the ERG in TA547:[12] adalimumab, vedolizumab, ustekinumab and tofacitinib. For completeness, the results of scenario analyses exploring all subsequent treatment options are provided in Appendix 5 below. Rows depicting treatment sequences rarely used in UK clinical practice have been marked in grey. These scenario analyses were found to have limited impact on the base case cost-effectiveness results, showing the results to be robust to clinical uncertainty regarding the use of treatment sequences in clinical practice. No efficacy data are available to inform efficacy of biologic therapies specifically in the third-line or later, and clinical consultation indicated that treatment decisions after failure on multiple biologics are patient-dependent and highly variable. Clinical expert feedback sought as part of this response suggested that modelling treatment sequences beyond the second line setting would not be informative for decision making given the number of assumptions required to model such sequences. However, for completeness, the model has also been updated with functionality to permit subsequent treatments to be explored when the TNFi-experienced population is selected. The results of these additional scenario analyses are provided in Appendix 5 below. In all scenarios explored, in the absence of relevant data specifically in the third-line setting, efficacy data derived from the TNFi-experienced NMAs were applied for the selected third line treatment. These analyses should therefore be interpreted with caution. In line with the results in the TNFi-naïve population, these scenario analyses were found to have limited impact on the cost-effectiveness results, showing the Company's revised base case to be robust to

		clinical uncertainty regarding subsequent treatments in the TNFi-experienced population.
Key issue 6: The PSA provided by the Company was not considered helpful for decision making	Yes (New analyses; Appendix 4)	The Company has updated its economic model to include tofacitinib (see response to key issue 1) and has amended the NMA to align with the approaches preferred by the ERG (see responses to key issue 2 and 3). The ERG voiced concerns relating to the applicability of the probabilistic sensitivity analyses (PSA) provided by the Company. In particular, the ERG raised concerns over the omission of tofacitinib as a comparator in the Company model and the methods used to generate baseline risk estimates for the NMA. The omission of tofacitinib from the cost-effectiveness analysis, including the PSA, has been addressed by including tofacitinib in both populations as part of a revised Company model, in order to reduce uncertainty regarding the exclusion of tofacitinib as a comparator. As stated in Key issue 1, the Company maintains its position that tofacitinib does not represent a relevant comparator in either TNFi-naïve or TNFi-experienced populations due to concerns regarding its safety profile, typically restricting its use to later treatment lines in UK clinical practice. The source of baseline risk in the Company NMAs has also been updated in response to Key issue 2 to align with the ERGs preferred approach, with revised efficacy parameters incorporated into the economic model accordingly. The results of this revised cost effectiveness analysis and PSA are presented in Appendix 4.

NICE National Institute for Health and Care Excellence

Summary of changes to the Company's cost-effectiveness estimate(s)

The revised economic model includes a new patient access scheme (PAS) for ozanimod, representing a discount to the list price of [redacted]. This discount has been included in all analyses explored to address the key issues in Table 3 below, with percentage changes in incremental cost-effectiveness ratios (ICERs) presented relative to the company base case before technical engagement but including the revised discount (Table 3, row 2). Positive changes to ICERs in the SW quadrant, denoted by the footnote [a], and those in which the intervention are dominant, represent a more cost-effective ICER for ozanimod.

Table 3: Changes to the Company's cost-effectiveness estimate

Key issue(s) in the ERG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the Company's base-case incremental cost-effectiveness ratio (ICER)	
			TNFi-naïve	TNFi-experienced
Company base case before technical engagement (original PAS discount of [redacted]%)	**Incremental QALYs** ***TNFi-naïve*** • Adalimumab: [redacted] • Infliximab: [redacted] • Golimumab: [redacted] • Vedolizumab: [redacted] ***TNFi-experienced*** • Vedolizumab: [redacted] • Ustekinumab: [redacted]	**Incremental costs** ***TNFi-naïve*** • Adalimumab: [redacted] • Infliximab: [redacted] • Golimumab: [redacted] • Vedolizumab: [redacted] ***TNFi-experienced*** • Vedolizumab: [redacted] • Ustekinumab: [redacted]	**Adalimumab** • ICER (£/QALYs) = £28,686 • NHB (QALYS) = 0.003 **Infliximab** • ICER (£/QALYs) = £167,024[a] • NHB (QALYS) = 0.175 **Golimumab** • ICER (£/QALYs) = £71,023[a] • NHB (QALYS) = 0.100 **Vedolizumab** • ICER (£/QALY) = £52,736[a]	**Vedolizumab:** • ICER (£/QALYs) = £199,551[a] • NHB (QALYS) = 0.170 **Ustekinumab** • ICER (£/QALYs) = Ozanimod dominant (-£33,725) • NHB (QALYS) = 0.156

			• NHB (QALYS) = 0.205	
Company base case before technical engagement (revised PAS discount of [redacted]%)	**Incremental QALYs** ***TNFi-naïve*** • Adalimumab: [redacted] • Infliximab: [redacted] • Golimumab: [redacted] • Vedolizumab: [redacted] ***TNFi-experienced*** • Vedolizumab: [redacted] • Ustekinumab: [redacted]	**Incremental costs** ***TNFi-naïve*** • Adalimumab: [redacted] • Infliximab: [redacted] • Golimumab: [redacted] • Vedolizumab: [redacted] ***TNFi-experienced*** • Vedolizumab: [redacted] • Ustekinumab: [redacted]	**Adalimumab** • ICER (£/QALYs) = Ozanimod dominant (-£53,637) • NHB (QALYS) = 0.173 **Infliximab** • ICER (£/QALYs) = £300,112[a] • NHB (QALYS) = 0.346 **Golimumab** • ICER (£/QALYs) = £140,576[a] • NHB (QALYS) = 0.271 **Vedolizumab** • ICER (£/QALYs) = £71,650[a] • NHB (QALYS) = 0.375	**Vedolizumab:** • ICER (£/QALYs) = £351,376[a] • NHB (QALYS) = 0.322 **Ustekinumab** • ICER (£/QALYs) = Ozanimod dominant (-£95,897) • NHB (QALYS) = 0.308
Key issue 1: Tofacitinib was excluded as a comparator in TNFi-naïve and -experienced subgroups	Tofacitinib was not included as a comparator in either the TNFi-naïve or TNFi-experienced population as the safety concerns associated with tofacitinib mean its use is reserved for later treatment lines.	For completeness, the Company has included tofacitinib in the cost-effectiveness model in both populations, however the Company maintains their stance that the exclusion of tofacitinib as a relevant comparator in the TNFi-naïve and TNFi-experienced populations is	**Tofacitinib** • Incremental costs= [redacted] • Incremental QALY = [redacted] • ICER (£/QALY) = £45,201[a] • NHB (QALYS) = 0.100	**Tofacitinib** • Incremental costs= [redacted] • Incremental QALY = [redacted] • ICER (£/QALY) = £88,643[a] • NHB (QALYS) = 0.122

		appropriate and therefore that these analyses are not considered relevant to UK clinical practice. The inclusion of tofacitinib does not form part of the Company's revised base case, but is instead presented for completeness for the Committee's consideration. The results for ozanimod compared to tofacitinib have been obtained using the revised company base case following technical engagement (see Appendix 4), including the revised PAS for ozanimod, representing a discount to the list price of [redacted]%.		
Key issue 2: Baseline risks for placebo anchors in the NMAs taken from the same trials those used for relative risk	Baseline risk for placebo anchors included in the NMA were calculated from the same set of trials used to calculate the relative treatment effect in order to maximise the use of available data.	The Company agrees that the ERG's proposed method for deriving baseline risk may be more suitable, in the absence of more appropriate data sources. The Company has amended their base case to reflect the changes proposed by the ERG. Baseline risk for placebo anchors were obtained from individual trails included in the NMAs that were more considered generalisable to the UK context.	**Adalimumab** • ICER (£/QALYs) = Ozanimod dominant (-£56,761) • NHB (QALYS) = 0.143 • Change from original base case ICER = +5.8% **Infliximab** • ICER (£/QALYs) = £269,815[a] • NHB (QALYS) = 0.298 • Change from original base	**Vedolizumab:** • ICER (£/QALYs) = £510,264[a] • NHB (QALYS) = 0.266 • Change from original base case ICER = +45.2% **Ustekinumab** • ICER (£/QALYs) = Ozanimod dominant (-£140,604) • NHB (QALYS) = 0.273 • Change from original base case ICER = +46.6%

			case ICER = -10.1% **Golimumab** • ICER (£/QALYs) = £155,027[a] • NHB (QALYS) = 0.234 • Change from original base case ICER = +10.3% **Vedolizumab** • ICER (£/QALYs) = £76,957[a] • NHB (QALYS) = 0.313 • Change from original base case ICER = +7.4%	
Key issue 3: A random effects model may be more appropriate for use in the maintenance phase NMAs	Fixed effects models were used in both the TNFi-naïve and TNFi-experienced populations for the maintenance period NMAs. A fixed effects approach was utilised owing to the fact that the random effects model did not converge and a reasonable statistical fit was obtained from the fixed effects model (Document, B Section B.2.8.5).	The Company acknowledges that due to the presence of heterogeneity between included studies, random effects models may have better clinical validity. The Company has therefore explored revised NMAs using random effects models with an appropriate prior distribution.	**Adalimumab** • ICER (£/QALYs)= Ozanimod dominant (-£58,276) • NHB (QALYS) = 0.171 • Change from original base case ICER = +8.6% **Infliximab** • ICER (£/QALYs) = £306,675[a] • NHB (QALYS) = 0.344 • Change from original base case ICER = +2.2%	**Vedolizumab:** • ICER (£/QALYs) = £447,631[a] • NHB (QALYS) = 0.322 • Change from original base case ICER = +27.4% **Ustekinumab** • ICER (£/QALYs) = Ozanimod dominant (-£97,312) • NHB (QALYS) = 0.309 • Change from original base case ICER = +1.5%

			Golimumab • ICER (£/QALYs) = £110,930[a] • NHB (QALYS) = 0.265 • Change from original base case ICER = -21.1% **Vedolizumab** • ICER (£/QALYs) = £72,145[a] • NHB (QALYS) = 0.373 • Change from original base case ICER = +0.7%	
Key issue 4: Modelled efficacy estimates for BSC in the post-active treatment phase	The Company used data for loss of overall response (including remission) to inform remission transition probabilities for BSC.	The Company acknowledges the ERG's proposed approach of calculating remission probabilities directly from sustained remission estimates may be more suitable as it aligns with the approach taken for active treatments in the model. This amendment has therefore been incorporated into the economic model.	**Adalimumab** • ICER (£/QALYs) = Ozanimod dominant (-£52,954) • NHB (QALYS) = 0.174 • Change from original base case ICER = -1.3% **Infliximab** • ICER (£/QALYs) = £298,010[a] • NHB (QALYS) = 0.345 • Change from original base case ICER = -0.7% **Golimumab**	**Vedolizumab:** • ICER (£/QALYs) = £352,490[a] • NHB (QALYS) = 0.323 • Change from original base case ICER = +0.3% **Ustekinumab** • ICER (£/QALYs) = Ozanimod dominant (-£94,699) • NHB (QALYS) = 0.310 • Change from original base case ICER = -1.2%

			• ICER (£/QALYs) = £139,654[a] • NHB (QALYS) = 0.270 • Change from original base case ICER = -0.7% **Vedolizumab** • ICER (£/QALYs) = £70,930[a] • NHB (QALYS) = 0.372 • Change from original base case ICER = -1.0%	
Company's base case following technical engagement (revised PAS discount of [redacted]% applied)	**Incremental QALYs** ***TNFi-naïve*** • Adalimumab: [redacted] • Infliximab: [redacted] • Golimumab: [redacted] • Vedolizumab: [redacted] ***TNFi-experienced*** • Vedolizumab: [redacted] • Ustekinumab: [redacted]	**Incremental costs** ***TNFi-naïve*** • Adalimumab: [redacted] • Infliximab: [redacted] • Golimumab: [redacted] • Vedolizumab: [redacted] ***TNFi-experienced*** • Vedolizumab: [redacted] • Ustekinumab: [redacted]	**Adalimumab** • ICER (£/QALYs) = Ozanimod dominant (-£59,307) • NHB (QALYS) = 0.145 • Change from original base case ICER = +10.6% **Infliximab** • ICER (£/QALYs) = £273,773[a] • NHB (QALYS) = 0.300 • Change from original base case ICER = -8.8% **Golimumab** • ICER (£/QALYs) = £121,137[a]	**Vedolizumab:** • ICER (£/QALYs) = £786,412[a] • NHB (QALYS) = 0.266 • Change from original base case ICER = +123.8% **Ustekinumab** • ICER (£/QALYs) = Ozanimod dominant (-£141,946) • NHB (QALYS) = 0.274 • Change from original base case ICER = +48.0%

			• NHB (QALYS) = 0.232 • Change from original base case ICER = -13.8% **Vedolizumab** • ICER (£/QALY) = £76,103[a] • NHB (QALYS) = 0.314 • Change from original base case ICER = +6.2%	

[a]SW quadrant ICER; costs saved per QALY forgone. NHB calculated at a willingness-to-pay threshold of £30,000.
Abbreviations: ICER: incremental cost-effectiveness ratio; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Sensitivity analyses around revised base case

The following additional scenario analysis were explored in the Company model to reduce uncertainty relating to the key issues raised by the ERG:

- Increased flexibility in subsequent treatment options in the TNFi-naïve population, including subsequent treatment with adalimumab, vedolizumab, ustekinumab and tofacitinib (key issue 5; Appendix 5)
- All scenario analyses explored in the original Company submission have been rerun using revised base case inputs (Appendix 4)
- In addition, the following exploratory scenario analyses suggest by the ERG have been explored: 0.75% spontaneous remission rate per model cycle, 80% of patients receiving SC vedolizumab, higher maintenance period rates of discontinuation due to AEs and AE incidence for ozanimod, costs applied per treatment cycle (Appendix 4)
- The deterministic sensitivity analysis (DSA) and PSA have been rerun in line with the original Company submission, using the Company's revised base case, which includes the ERG's preferred assumptions (including the addition of tofacitinib to the cost-effectiveness analysis; Appendix 4)

Ozanimod remained cost-effective across the vast majority of scenarios explored in the TNFi-naïve population, and all scenarios explored in the TNFi-experienced population.

References

1. National Institute for Health and Care Excellence (NICE). TA633: Ustekinumab for treating moderately to severely active ulcerative colitis. Available at: https://www.nice.org.uk/guidance/ta633/documents/committee-papers-3 [Last accessed: 13/05/2022].
2. European Medicines Agency. PRAC recommendations on signals. https://www.ema.europa.eu/en/documents/prac-recommendation/prac-recommendations-signals-adopted-7-10-june-2021-prac-meeting_en.pdf [Accessed 21 February 2022].
3. European Medicines Agency. EMA starts safety review of Janus kinase inhibitors for inflammatory disorders. Available at https://www.ema.europa.eu/en/news/ema-starts-safety-review-janus-kinase-inhibitors-inflammatory-disorders. [Accessed 21 February 2022].
4. Food and Drug Administration. FDA Drug Safety Communication. Available at https://www.fda.gov/media/151936/download. [Accessed 16 February 2022].
5. Dias S, Welton NJ, Sutton AJ, et al. NICE DSU Technical Support Document 5: Evidence synthesis in the baseline natural history model. Report by the Decision Support Unit, 2011 (last updated April 2012): ScHARR, University of Sheffield.
6. IBD Registry. IBD Biological Therapies Audit Annual Report, 2021. Available at: https://ibdregistry.org.uk/wp-content/uploads/2021/11/2021-07-Annual-Biologics-Report-V01.pdf [Last accessed: 13/05/2022].
7. Royal College of Physicians. National clinical audit of biological therapies: UK inflammatory bowel disease (IBD) audit, 2016. Available at: https://www.rcplondon.ac.uk/projects/outputs/national-clinical-audit-biological-therapies-annual-report-2016 [Last accessed: 13/05/2022].
8. Lamb CA, Kennedy NA, Raine T, et al. British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. Gut 2019;68:s1-s106.
9. National Institute for Health and Care Excellence (NICE). NG130: Ulcerative colitis: management, 2019. Available at: https://www.nice.org.uk/guidance/ng130 [Last accessed: 13/05/2022].
10. Turner RM, Davey J, Clarke MJ, et al. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol 2012;41:818-27.
11. Turner RM, Jackson D, Wei Y, et al. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. Stat Med 2015;34:984-98.
12. National Institute for Health and Care Excellence (NICE). TA547: Tofacitinib for moderately to severely active ulcerative colitis. Available at: https://www.nice.org.uk/guidance/ta547 [Last accessed: 13/05/2022].
13. Dias S, Welton NJ, Sutton AJ, et al. NICE DSU Technical Support Document 5: Evidence synthesis in the baseline natural history model. 2011.
14. Physicians RCo. National clinical audit of biological therapies: UK inflammatory bowel disease (IBD) audit, 2016.
15. King D, Reulen RC, Thomas T, et al. Changing patterns in the epidemiology and outcomes of inflammatory bowel disease in the United Kingdom: 2000-2018. Aliment Pharmacol Ther 2020;51:922-934.

Appendix 1: NMA exploring alternative sources of baseline risk

Sources of baseline risk

Although in agreement with the ERG's statement that the ideal approach to inform baseline risk estimates would be through a systematic literature review (SLR), the Company ascertained that conducting a proper, protocol-driven SLR would not be feasible within the timeframe of the response to the appraisal due to its broad scope and lack of restrictions on study design.[13] Instead, a targeted search was conducted to identify potential real-world data sources to inform baseline risk. A review of UK national registry reports[6, 7] as well as recent guidelines on the management of ulcerative colitis[8, 9] and their reference lists revealed a scarcity of demographic data. Only one of the identified sources, the National Clinical Audit of Biologic Therapies (2016) annual report, provided detailed demographic information (i.e., gender, median age at diagnosis, median age at treatment initiation, median time from diagnosis to treatment and disease distribution) specific to adult UC patients (n = 903), however this study did not provide data on clinical remission or clinical response rates for patients receiving BSC without advanced/biologic treatment, and therefore could not be used to inform baseline risk in the NMAs.[14] Furthermore, expert opinion consulted as part of the technical engagement process noted that any recent observational or RWE studies in patients with moderately to severely active UC would be unlikely to provide evidence for placebo response required for the NMA, given patients in real-world practice receive active treatments.

The Company therefore considers the ERG's approach of utilising placebo arm values from individual trials included in the NMAs that were more generalisable to the UK context to be a suitable method for deriving baseline risk for placebo anchors, given the time available. As acknowledged by the ERG, limitations do remain with this approach. Whilst individual trials may better reflect certain UK patient characteristics, they may not be reflective of UK patients in all respects, and thus their selection could be considered arbitrary. Despite these limitations, the method proposed by the ERG allows for an estimation of baseline risk which may be more generalisable to the UK population than the Company's original base case, and more closely aligns with the recommendations in TSD5.

In alignment with the ERG approach, the cross-sectional, retrospective UK cohort involving 37,793 subjects with UC as presented by King et al. (2020) was selected to support inspection of alignment of NMA trials to the UK population.[15] For each NMA, trial-level baseline risk (probability of no response, no remission in the placebo arm) and key baseline characteristics for each included trial were evaluated for comparability with the UK population. Baseline characteristics evaluated included mean age, % male, mean duration of UC, severity of disease (mean Mayo score, % left-sided, % extensive), % prior TNFi exposure, % prior TNFi failure, % patients on concomitant corticosteroids, and mean C-reactive protein. Among these key baseline characteristics, the UK cohort as reported by King *et al.* 2020 was limited to median age (which is not exactly comparable to means reported for trials) and % male, reducing the scope of comparability.[15] Considering these limitations, a clinical expert was consulted to review the studies, their baseline characteristics, and trial-level baseline risks. The clinical expert did not

highlight any new sources of baseline risk as being more generalisable than those identified by the ERG. The studies selected by the NICE ERG represented the studies that were closest in distribution of available baseline characteristics (% male) to the retrospective UK cohort,[15] were assessed to have low risk of bias, and were not trials of exclusively Asian populations. The Company has therefore amended their method for deriving baseline risk for placebo anchors to align with the changes proposed by the ERG. The trials used to derive baseline risk in the NMA are listed here, and key baseline characteristics are presented in Table 4:

- PURSUIT-SC for treatment-naïve patients at induction
- PURSUIT M for treatment-naïve patients during maintenance
- OCTAVE 1 + 2 for treatment-experienced patients at induction
- GEMINI 1 for treatment-experienced patients during maintenance

Table 4: Key baseline characteristics for the trials selected to estimate baseline risk for NMAs presented in this Appendix.

Baseline Characteristic[a]	TNFi-naïve (induction phase): PURSUIT-SC	TNFi-naïve (maintenance phase): PURSUIT M	TNFi-experienced (induction phase): OCTAVE 1 + 2	TNFi-experienced (maintenance phase): GEMINI 1
N	331	156	*234*	126
Age, years (mean)	39.0	40.2	*41.1*	40.3
Sex, male (%)	52.9	48.1	*56.4*	54.8
Duration of UC, years (mean)	6.0	6.9	NR	7.8
Mayo score (mean)	8.3	8.3	*9.0*	8.4
Left-sided (%)	57.0	NR	*32.6*	42.1
Extensive (%)	43.0	NR	*52.4*	13.5
Prior TNFi exposure (%)	NA	NA	*55.5*	37
Prior TNFi failure (%)	NA	NA	*53.0*	30.2
Concomitant corticosteroid use (%)	*42.9*	*56.4*	*48.3*	*57*
CRP, mg/L (mean)	10.7	9.6	NR	NR
Endoscopic reading	Local	Local	Central	Local
No response, no remission in population of interest[b] (%)	69.5	68.8	76.6	84.2

Abbreviations: CRP, C-reactive protein; NA, not applicable; NR, not reported; TNFi, tumour necrosis factor inhibitor; UC, ulcerative colitis
Notes: [a] Baseline characteristics for TNFi-naïve and TNFi-experienced patient population were not provided separately; characteristics are presented for the overall patient population.
[b] Percentage of patients with no response, no remission relates to TNFi-naïve or TNFi-experienced population as specified in the column header for each population.
Italicized values are based on calculations or assumptions using data reported in the studies.

NMA Methodology

Each of the originally submitted base case NMAs had estimated the baseline risk model parameter by taking the unweighted average of the probabilities, on the standard normal distribution scale, of no response, no remission among trials with placebo arms within the NMA data. In a revised analysis, for each NMA, the baseline risk model parameter has been updated to estimate only from the single study listed described above, mimicking that performed by the ERG and presented within the ERG Report. Precisely, model parameter 'A' now draws directly from trial baseline parameter 'mu[s]', where 's' is the study index corresponding to the study selected to represent baseline risk. Results from this revised analysis are now presented within this Appendix.

NMA Results

Results for the revised TNFi-naïve (induction, maintenance) and TNFi-experienced (induction) NMAs replicated those presented by the ERG within a margin of error attributable to chance (i.e., random sampling). Replication of the ERG's TNFi-experienced (maintenance) results was achieved only when VISIBLE 1 was selected for the baseline risk instead of GEMINI 1 – indicating a possible programming error. The Company inquired but did not receive correspondence from the ERG regarding the source of this discrepancy and therefore this has been amended in the revised Company NMA to correctly select GEMINI 1 to estimate the baseline risk. Therefore, the results presented below for the TNFi-experienced (maintenance) NMA marginally differ (by ~1% to 3%) in absolute probability to those reported by the ERG, varying across response category and treatment (see Table 29).

TNFi-naïve population (induction)

Table 5: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-naïve participants during the induction phase (baseline risk)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Infliximab	Pooled	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Golimumab	200/100 mg SC	[redacted]	[redacted]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

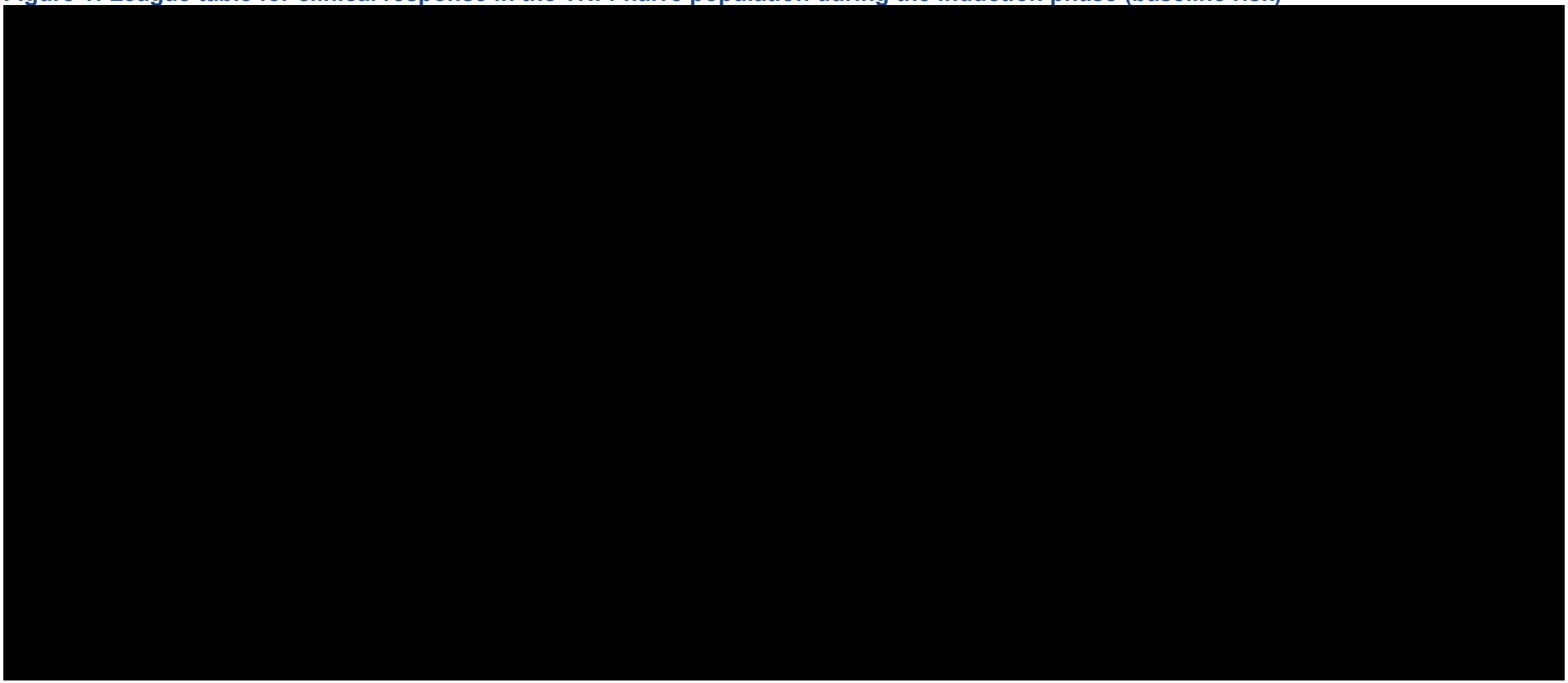
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a]random effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 6: NMA outcomes for comparators versus placebo in TNFi-naïve participants during the induction phase (baseline risk)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Infliximab	Pooled	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Golimumab	200/100 mg SC	[redacted]	[redacted]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]

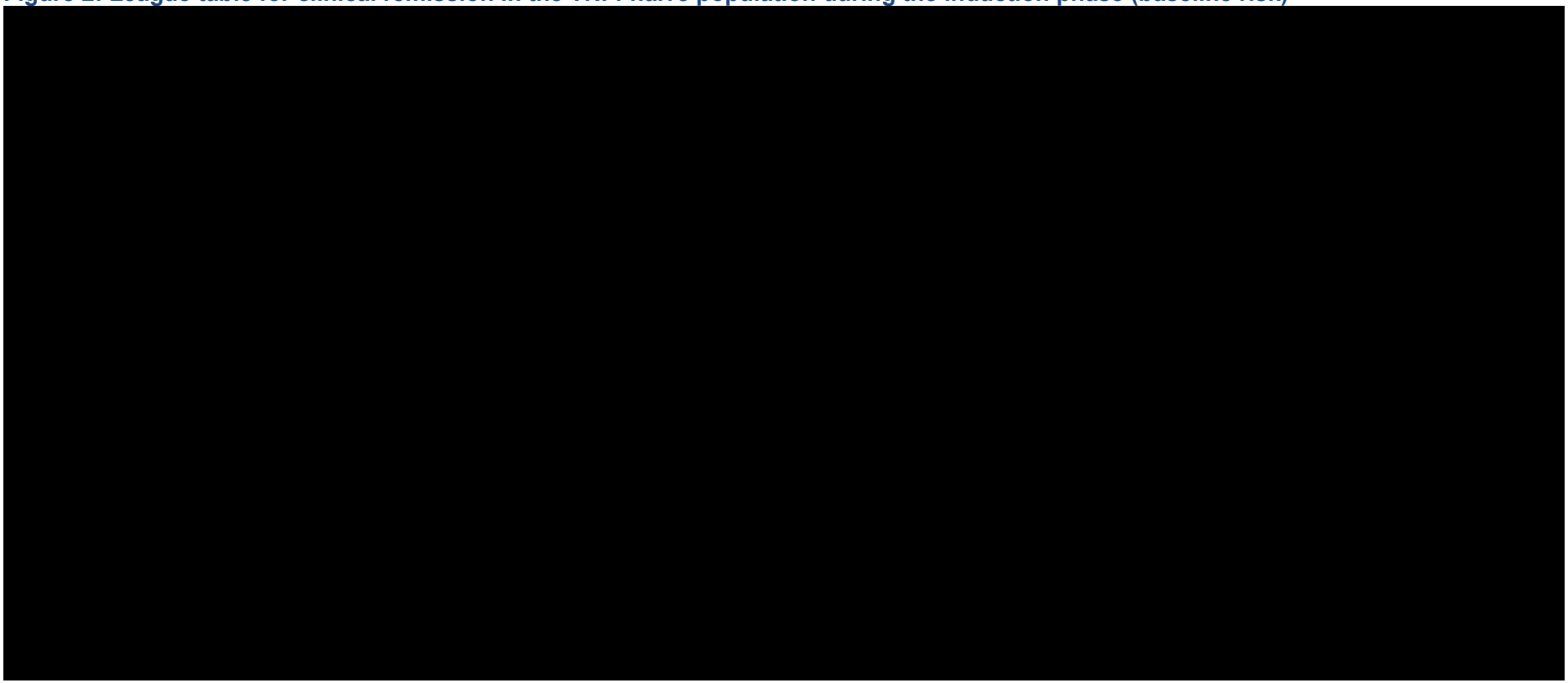
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a]random effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour the comparator; grey cells favour placebo.

Figure 1: League table for clinical response in the TNFi-naïve population during the induction phase (baseline risk)



Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

Figure 2: League table for clinical remission in the TNFi-naïve population during the induction phase (baseline risk)



Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

TNFi-naïve population (maintenance)

Table 7: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-naïve participants during the maintenance phase (baseline risk)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Infliximab	Pooled	[redacted]	[redacted]
Golimumab	Pooled	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

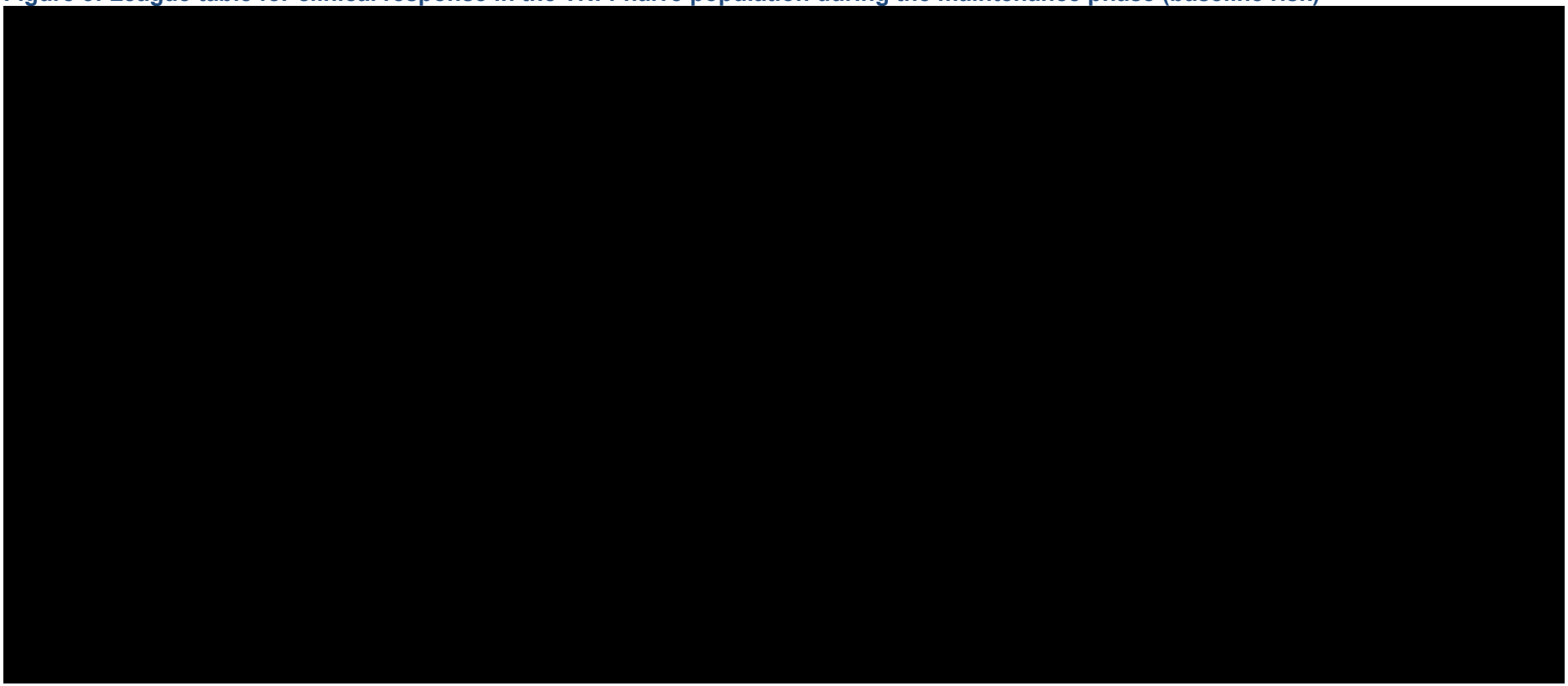
Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effect NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 8: NMA outcomes for comparators versus placebo in TNFi-naïve participants during the maintenance phase (baseline risk)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Infliximab	Pooled	[redacted]	[redacted]
Golimumab	Pooled	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]

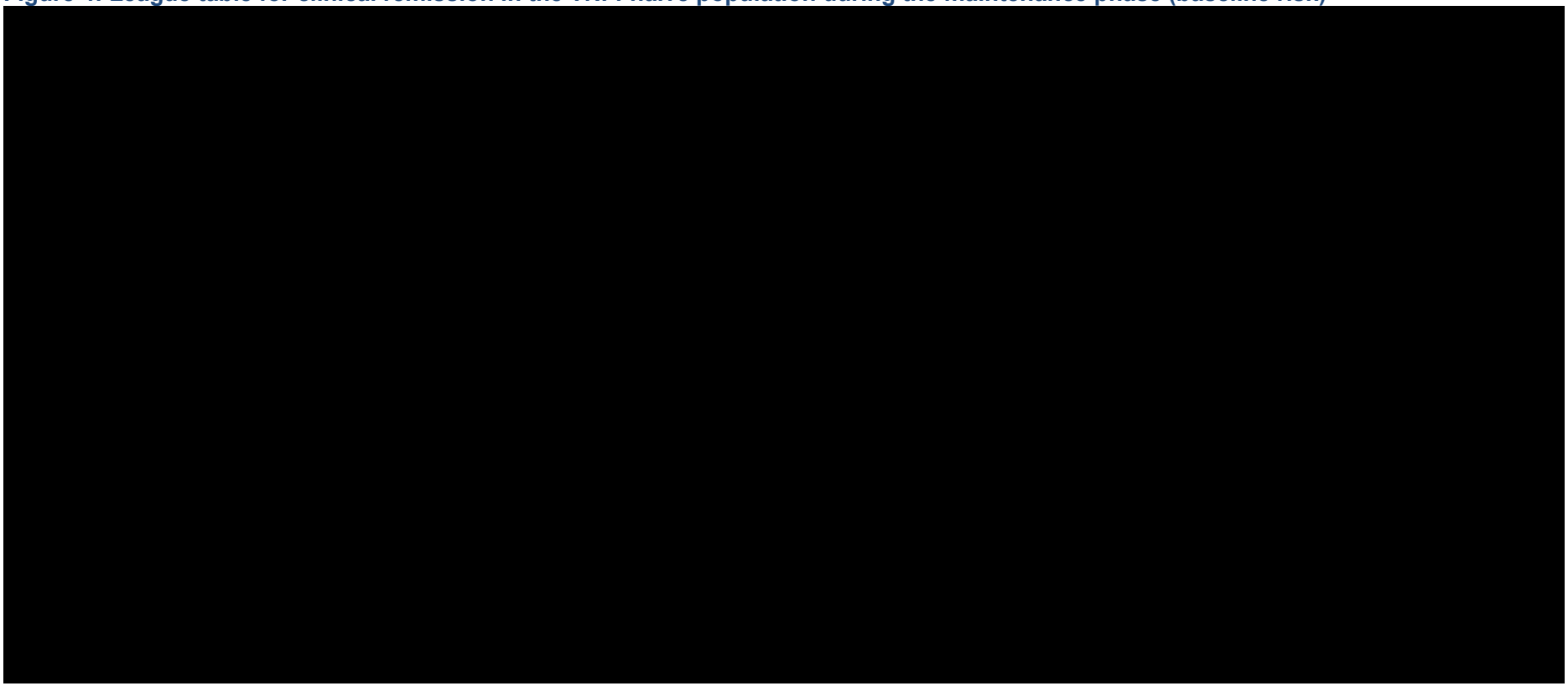
Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effect NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour the comparator; grey cells favour placebo.

Figure 3: League table for clinical response in the TNFi-naïve population during the maintenance phase (baseline risk)



Fixed effect NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab pooled; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

Figure 4: League table for clinical remission in the TNFi-naïve population during the maintenance phase (baseline risk)



Fixed effect NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab pooled; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

TNFi-experienced population (induction)

Table 9: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-experienced participants during the induction phase (baseline risk)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

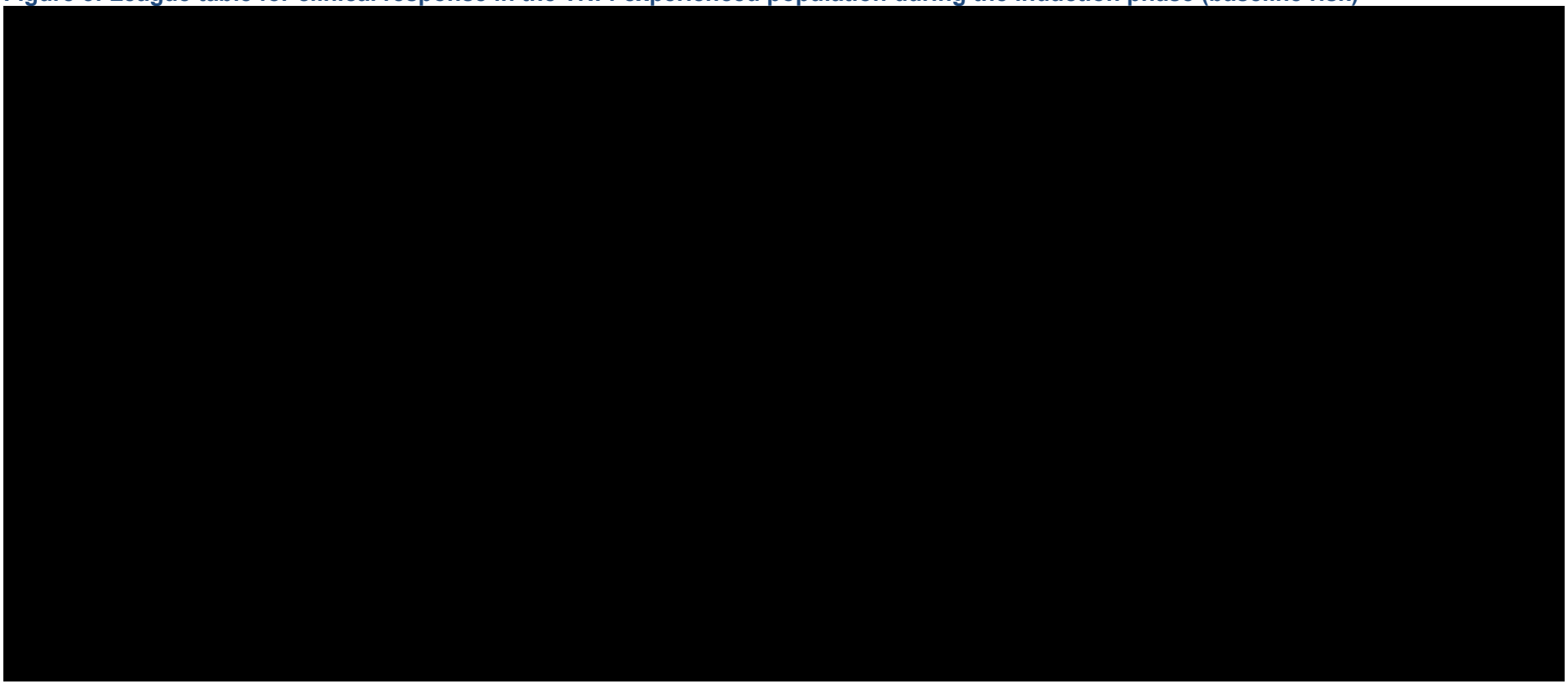
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effect NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 10: NMA outcomes for comparators versus placebo in TNFi-experienced participants during the induction phase (baseline risk)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]

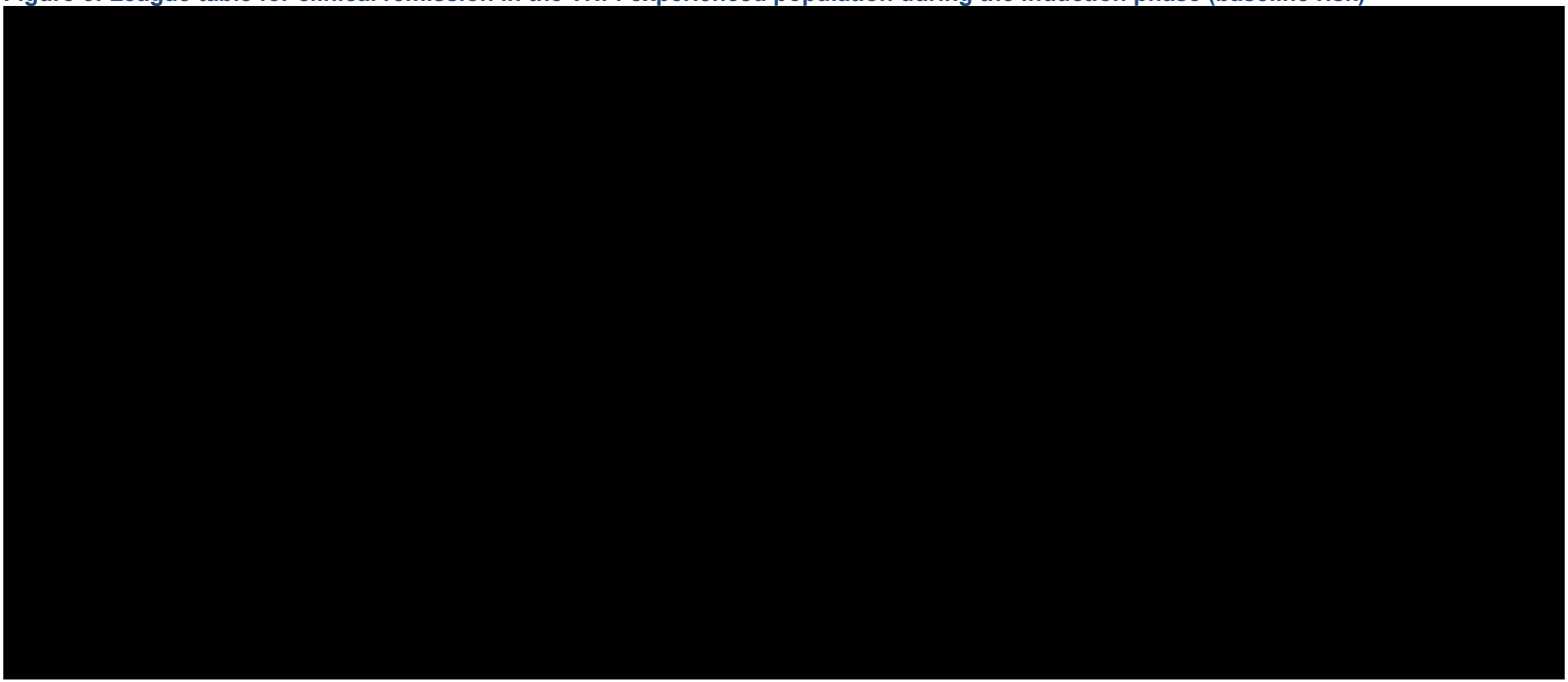
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effect NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour the comparator; grey cells favour placebo.

Figure 5: League table for clinical response in the TNFi-experienced population during the induction phase (baseline risk)



Fixed effect NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

Figure 6: League table for clinical remission in the TNFi-experienced population during the induction phase (baseline risk)



Fixed effect NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

TNFi-experienced population (maintenance)

Table 11: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-experienced participants during the maintenance phase (baseline risk)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Ustekinumab	Pooled	[redacted]	[redacted]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

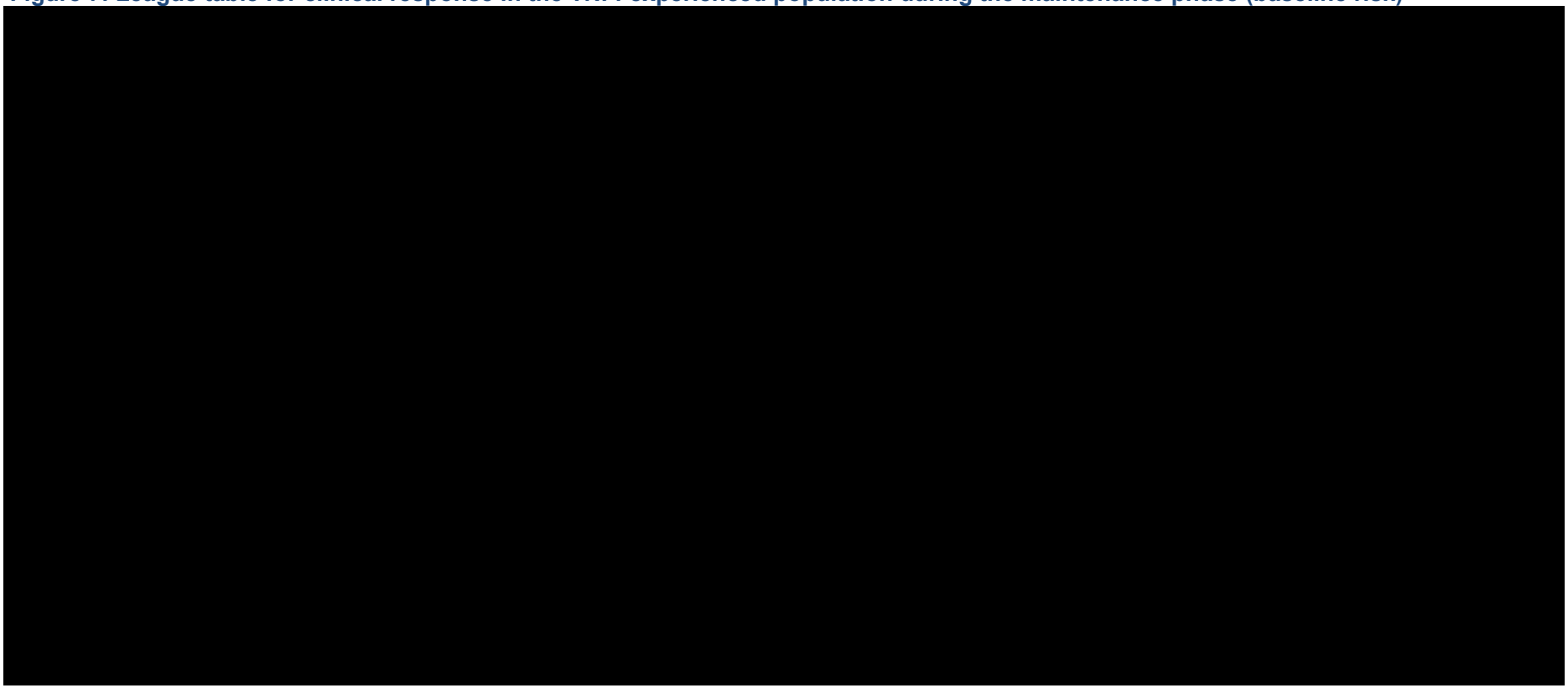
Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effect NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 12: NMA outcomes for comparators versus placebo in TNFi- experienced participants during the maintenance phase (baseline risk)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Ustekinumab	Pooled	[redacted]	[redacted]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]

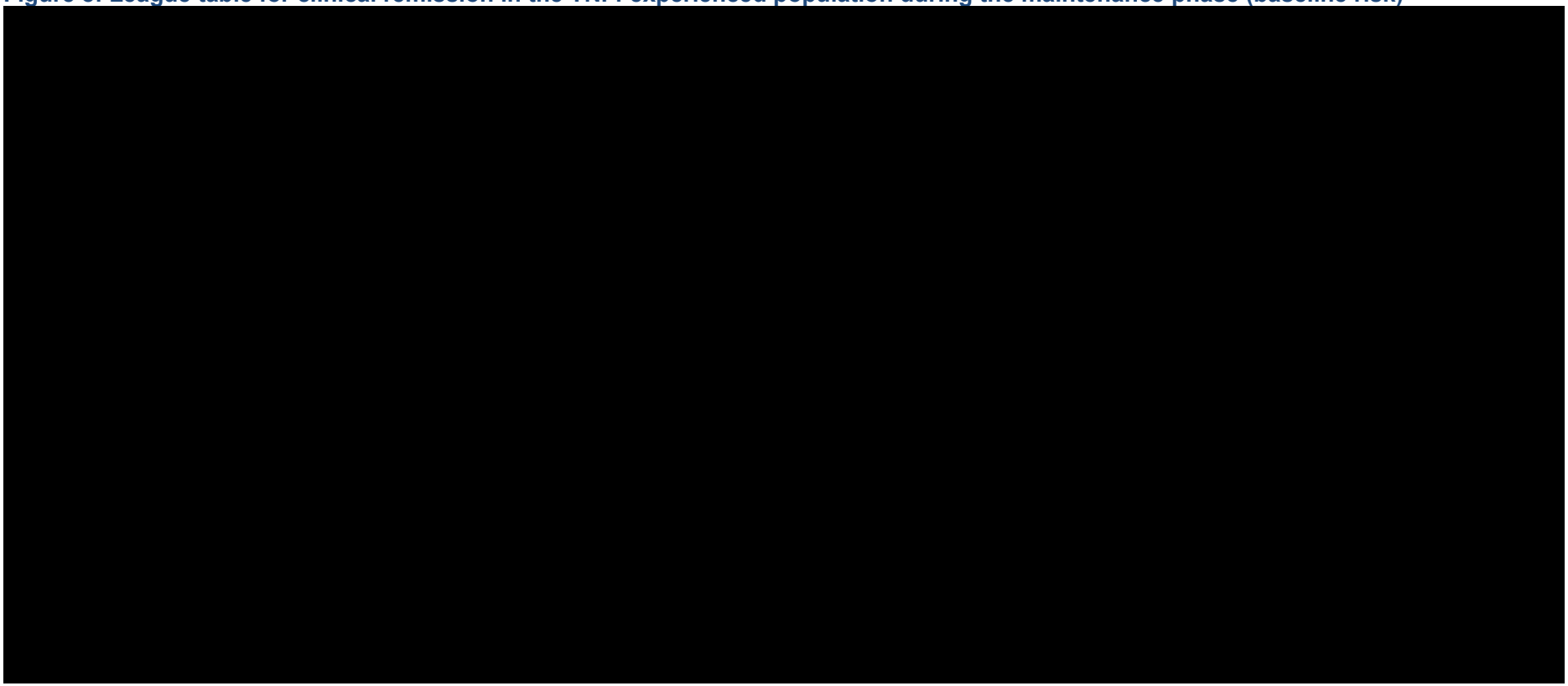
Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effect NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour the comparator; grey cells favour placebo.

Figure 7: League table for clinical response in the TNFi-experienced population during the maintenance phase (baseline risk)



Fixed effect NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

Figure 8: League table for clinical remission in the TNFi-experienced population during the maintenance phase (baseline risk)



Fixed effect NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

Appendix 2: NMA exploring random effects models with informative prior distributions from the literature

Sources of informative prior distributions

In agreement with the ERG's suggestions, suitable informative priors were sought for the between-study heterogeneity parameter to potentially overcome estimation and model convergence challenges experienced for random effects NMAs of TNFi-naïve at maintenance phase and TNFi-experienced at induction or maintenance data. As recommended by the ERG, key publications by Turner *et al.* (2012) and Turner *et al.* (2015) were consulted.[10, 11] The Turner *et al.* (2015) publication was preferred as it represented a more recent extension of the work done by Turner *et al.* (2012). Briefly, the authors construct and provide predictive distributions of the between-study heterogeneity parameter using thousands of binary outcome meta-analyses from the Cochrane Database of Systematic Reviews (CDSR). The author's work was assessed to be generally appropriate to support the selection of an informative prior distribution for the between-study heterogeneity parameter within the context of modelling clinical response and clinical remission outcomes in the current submission. The most appropriate prior distribution was selected to be a log-normal (LN) distribution of $LN(\mu = -2.70, \sigma^2 = 1.52^2)$ from the outcome type "Subjective outcomes (various)" and intervention comparison type "Pharmacological vs. Placebo/control" presented in Table IV of Turner *et al.* (2015), on the basis of clinical response and clinical remission being underpinned by the composite Mayo score and the networks almost exclusively involving placebo-controlled trials.[11]

NMA Methodology

Each of the originally submitted base case NMAs had used a vague, or uninformative prior distribution for the between-study heterogeneity parameter of random effects models, as follows (using JAGS syntax, where $dunif(a, b)$ represents a uniform distribution with minimum a and maximum b):

sd ~ dunif(0, 2)

In a revised analysis, for each NMA, the following informative prior has been specified for the between-study heterogeneity parameter of random effects models, in alignment with that selected from Turner et al. (2015) (using JAGS syntax, where $dlnorm(\mu, 1/\sigma^2)$ represents a LN distribution with mean μ and precision $1/\sigma^2$):[11]

sd ~ dlnorm(-2.70, 0.43283)

As a result of applying this informative prior distribution, random effects models for all base case NMAs converged, including those for TNFi-naïve at maintenance phase and TNFi-experienced at induction or maintenance data, which failed previously. Results for the random effects models from this revised analysis are now presented within this Appendix.

NMA Results

Table 13: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-naïve participants during the induction phase (informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Infliximab	Pooled	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Golimumab	200/100 mg SC	[redacted]	[redacted]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

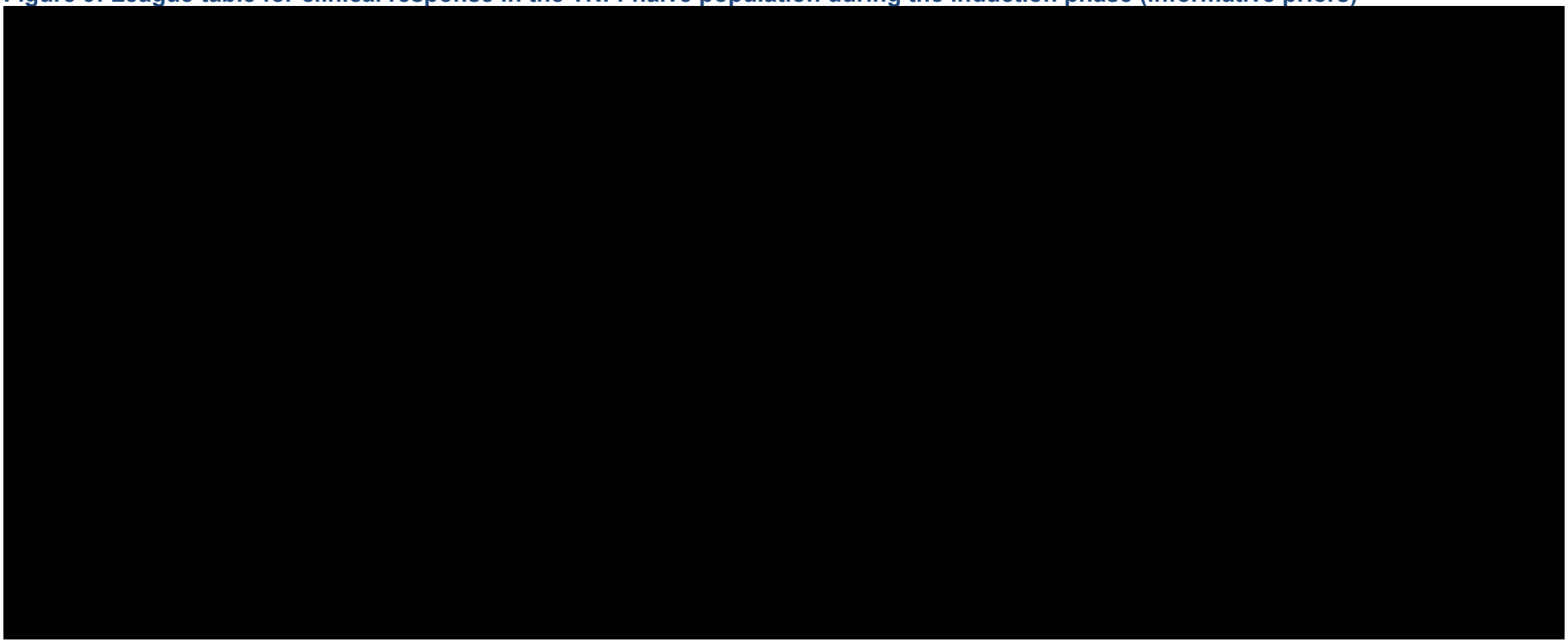
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 14: NMA outcomes for comparators versus placebo in TNFi-naïve participants during the induction phase (informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Infliximab	Pooled	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Golimumab	200/100 mg SC	[redacted]	[redacted]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]

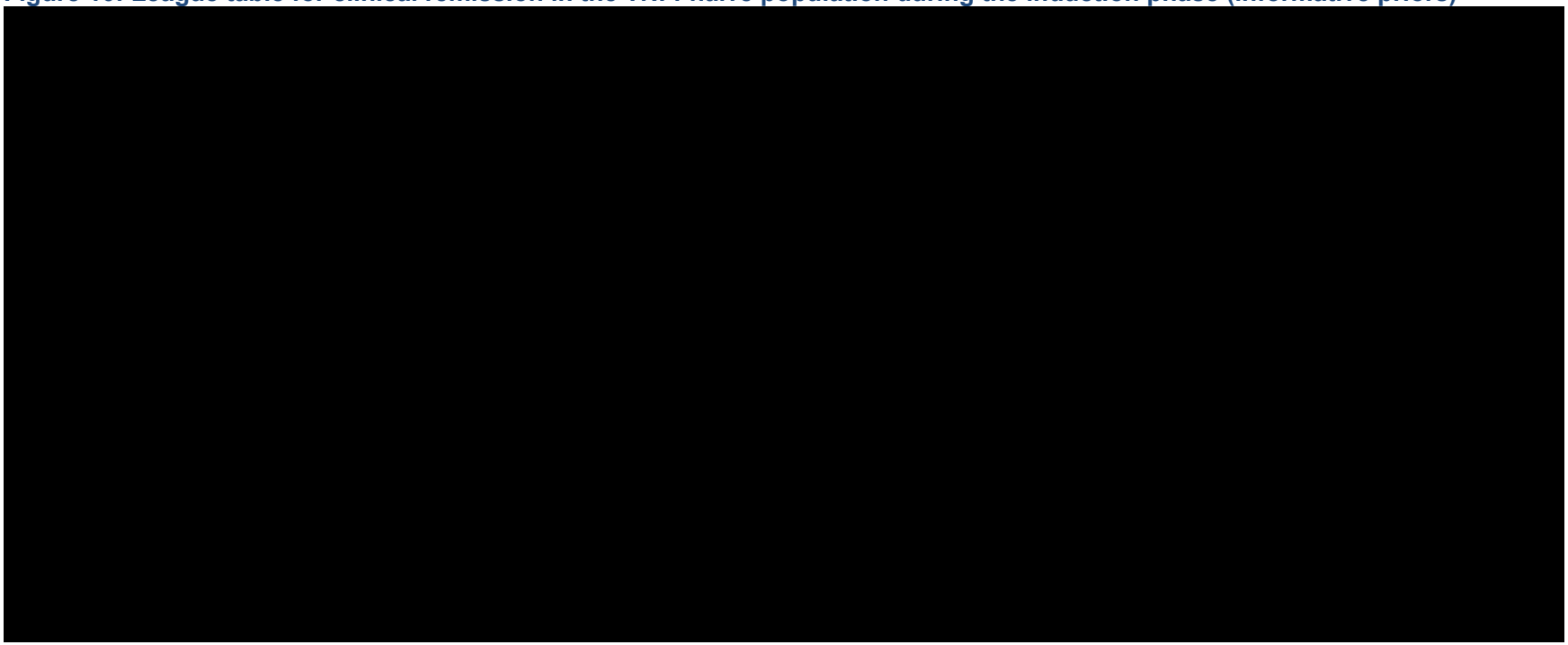
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour treatment, grey cells favour placebo.

Figure 9: League table for clinical response in the TNFi-naïve population during the induction phase (informative priors)



Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

Figure 10: League table for clinical remission in the TNFi-naïve population during the induction phase (informative priors)



Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

TNFi-naïve population (maintenance)

Table 15: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-naïve participants during the maintenance phase (informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Infliximab	Pooled	[redacted]	[redacted]
Golimumab	Pooled	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

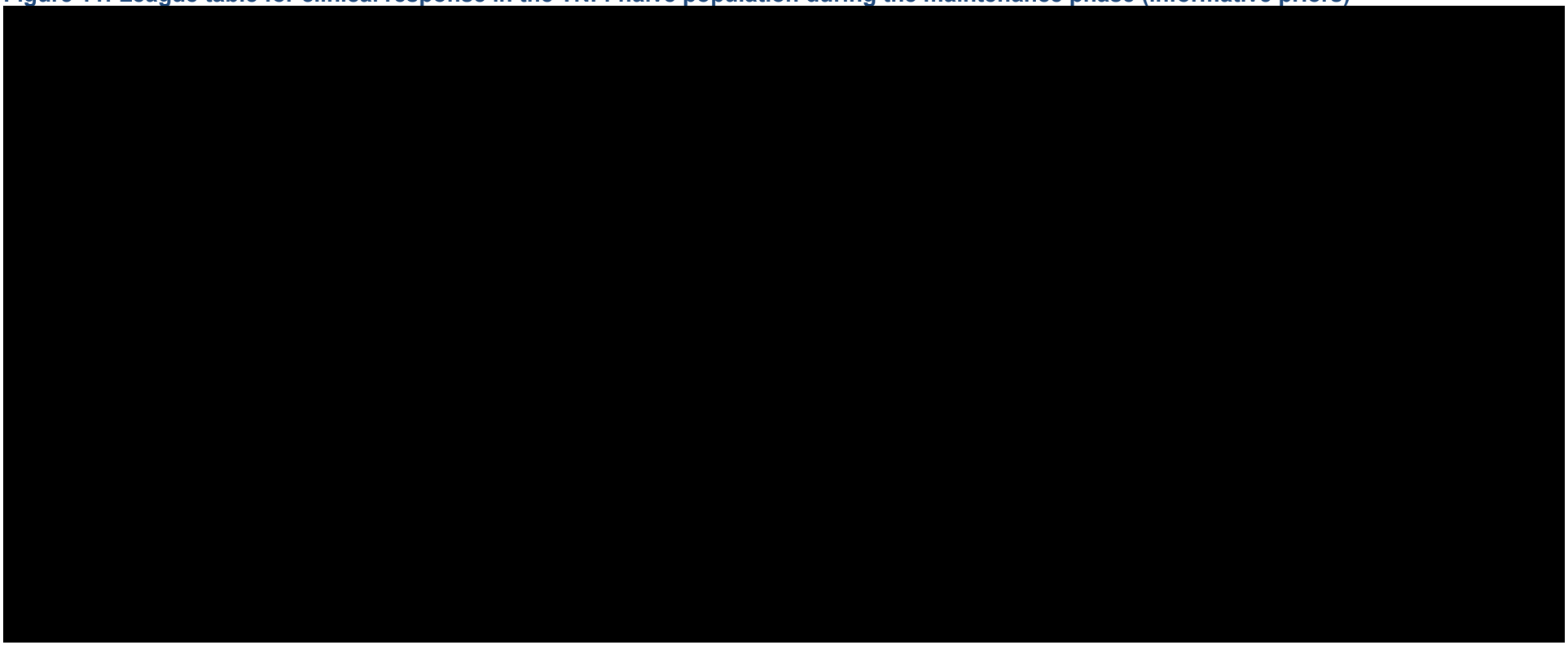
Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 16: NMA outcomes for comparators versus placebo in TNFi-naïve participants during the maintenance phase (informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Infliximab	Pooled	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Golimumab	Pooled	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]

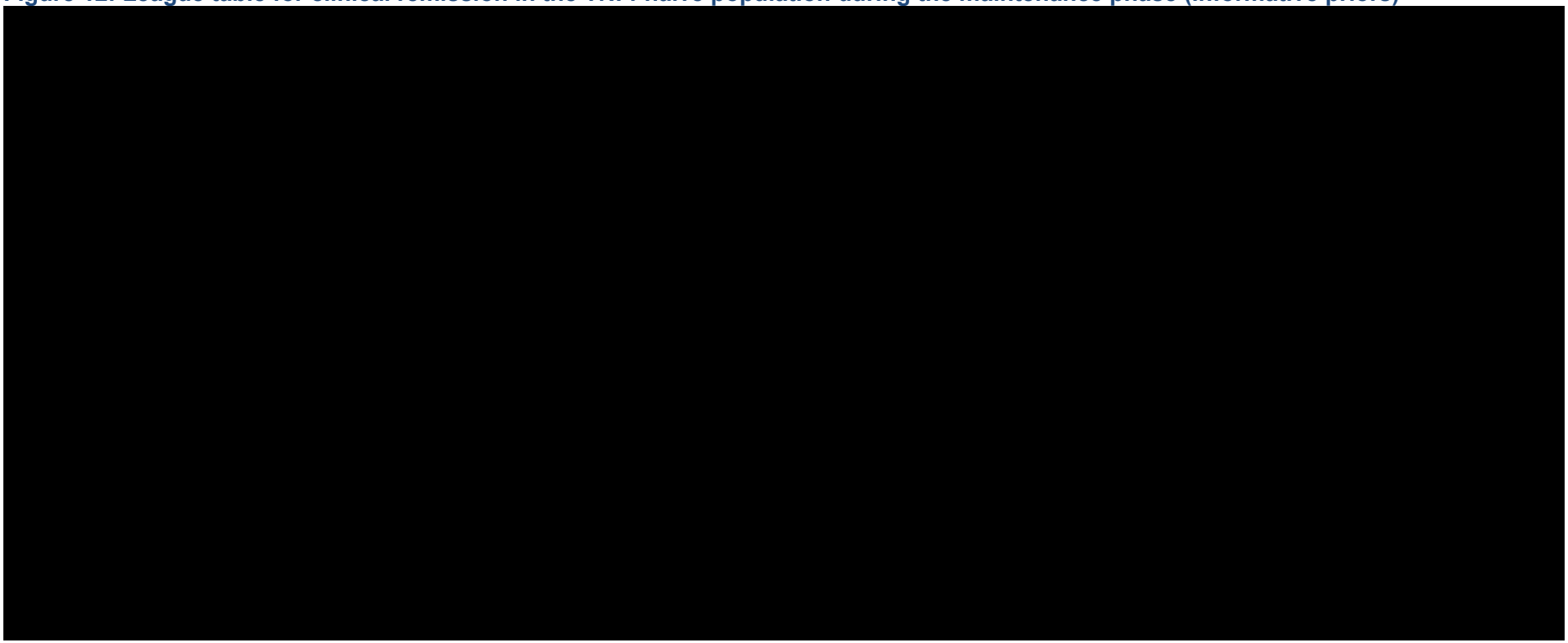
Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour treatment, grey cells favour placebo.

Figure 11: League table for clinical response in the TNFi-naïve population during the maintenance phase (informative priors)



Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab pooled; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

Figure 12: League table for clinical remission in the TNFi-naïve population during the maintenance phase (informative priors)



Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab pooled; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

TNFi-experienced population (induction)

Table 17: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-experienced participants during the induction phase (informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

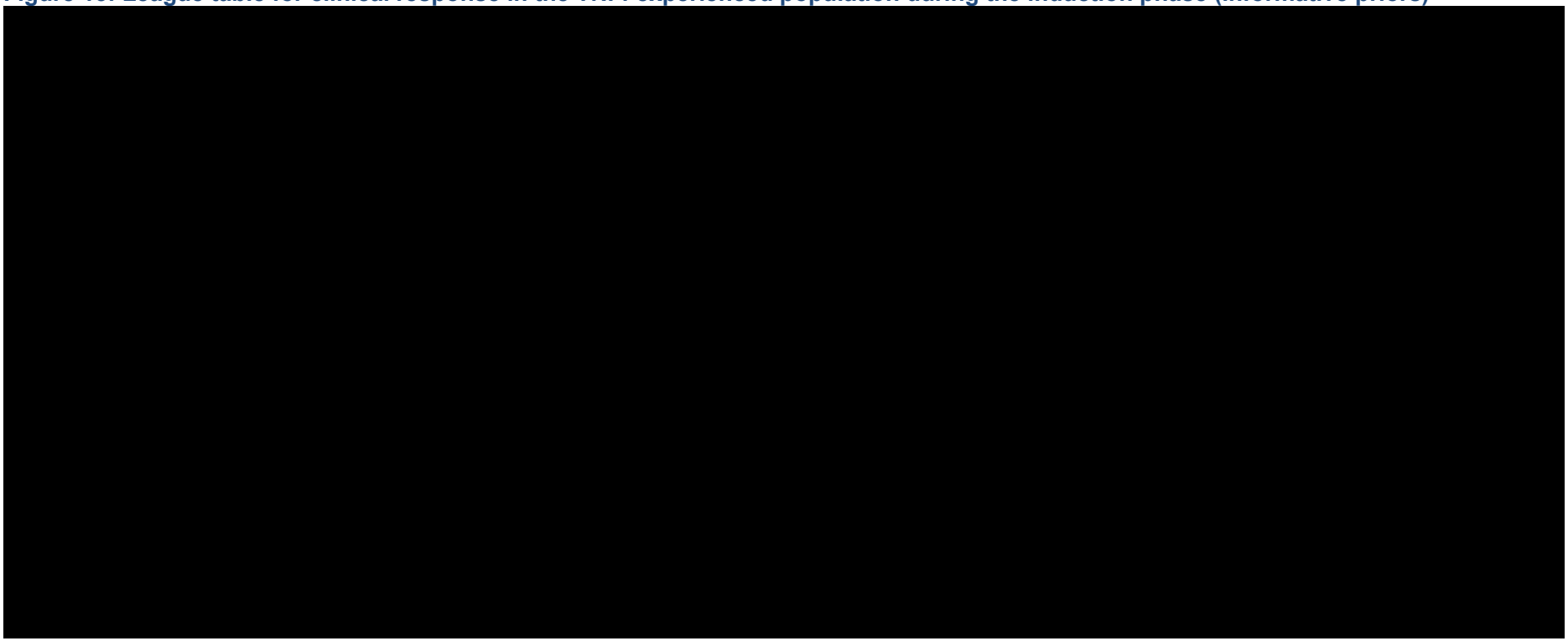
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 18: NMA outcomes for comparators versus placebo in TNFi-experienced participants during the induction phase (informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]

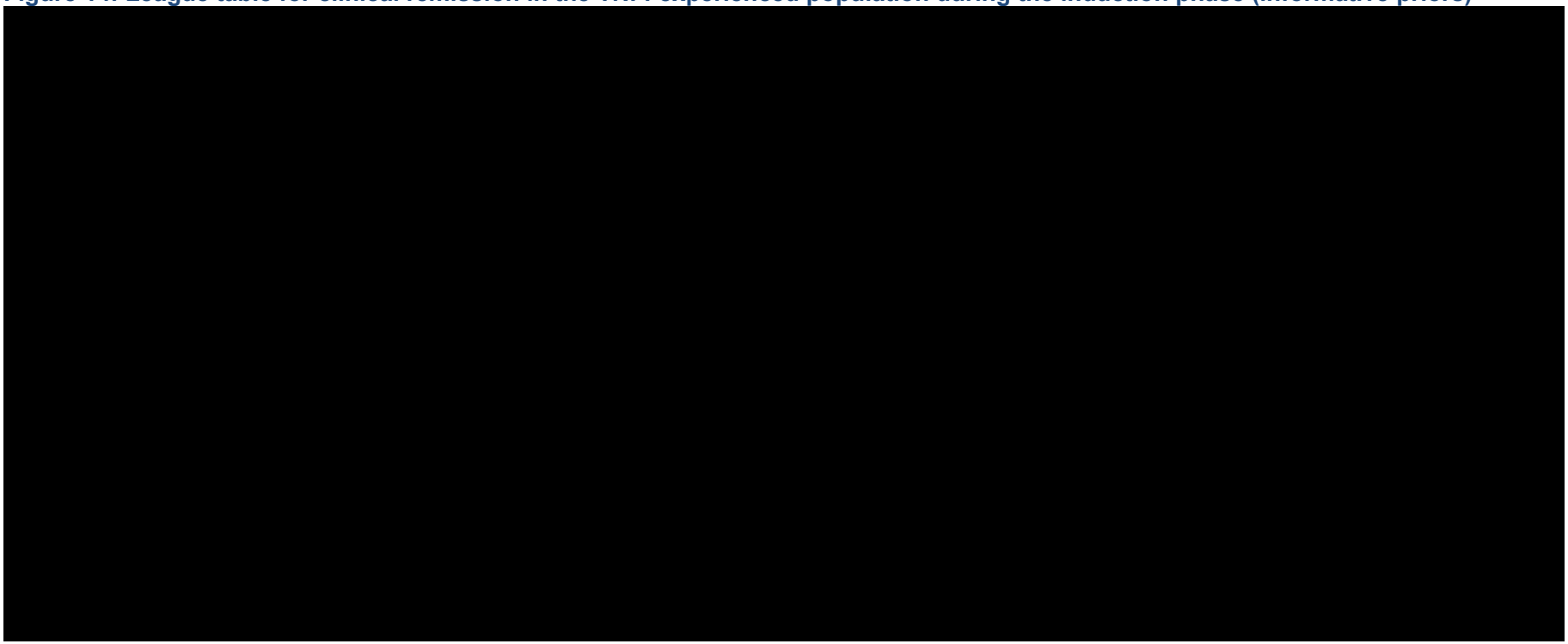
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour treatment, grey cells favour placebo.

Figure 13: League table for clinical response in the TNFi-experienced population during the induction phase (informative priors)



Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

Figure 14: League table for clinical remission in the TNFi-experienced population during the induction phase (informative priors)



Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

TNFi-experienced population (maintenance)

Table 19: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-experienced participants during the maintenance phase (informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Ustekinumab	Pooled	[redacted]	[redacted]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

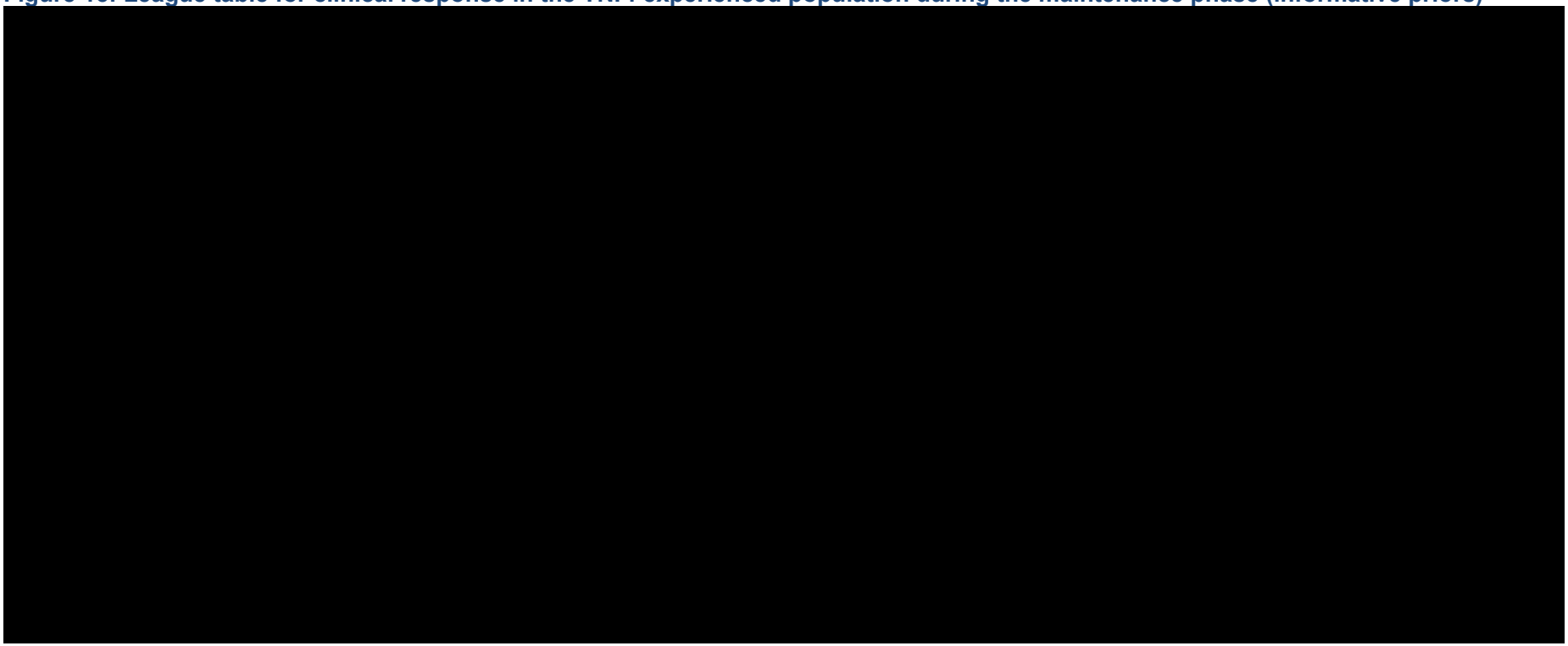
Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 20: NMA outcomes for comparators versus placebo in TNFi- experienced participants during the maintenance phase (informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Ustekinumab	Pooled	[redacted]	[redacted]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] fixed effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour treatment, grey cells favour placebo.

Figure 15: League table for clinical response in the TNFi-experienced population during the maintenance phase (informative priors)



Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

Figure 16: League table for clinical remission in the TNFi-experienced population during the maintenance phase (informative priors)

Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

Appendix 3: NMA exploring random effects models with alternative baseline placebo risk

A revised analysis was conducted by implementing the combination of baseline risk parameters selected per Appendix 1 and informative priors with random effects models used per Appendix 2. This analysis now represents the revised base case NMAs, and the results are presented within this section. The absolute probabilities of being in clinical response and clinical remission for each treatment in these revised base case NMAs are presented in Table 29, alongside those from the originally submitted base case NMAs and those produced by the ERG.

NMA Results

TNFi-naïve population (induction)

Table 21: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-naïve participants during the induction phase (baseline risk and informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Infliximab	Pooled	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Golimumab	200/100 mg SC	[redacted]	[redacted]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

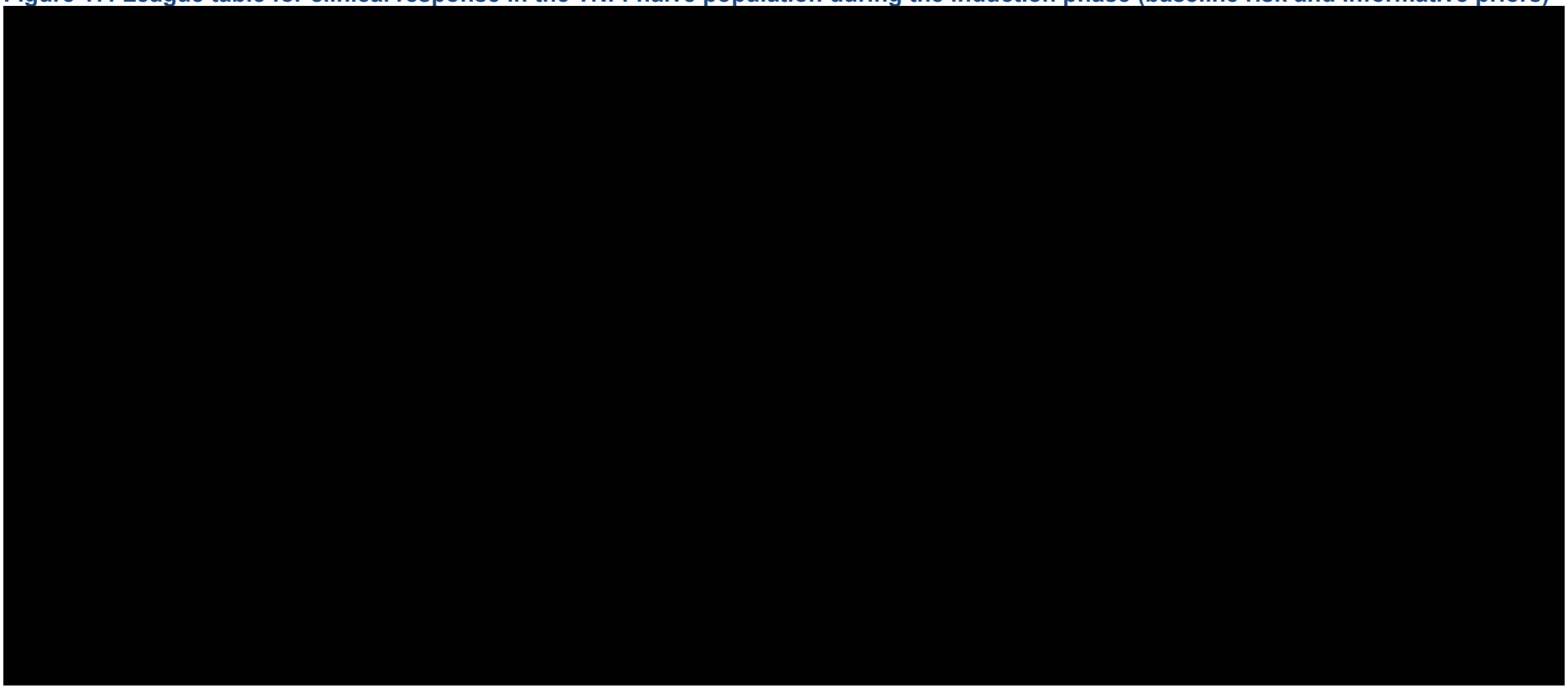
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 22: NMA outcomes for comparators versus placebo in TNFi-naïve participants during the induction phase (baseline risk and informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Infliximab	Pooled	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Golimumab	200/100 mg SC	[redacted]	[redacted]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]

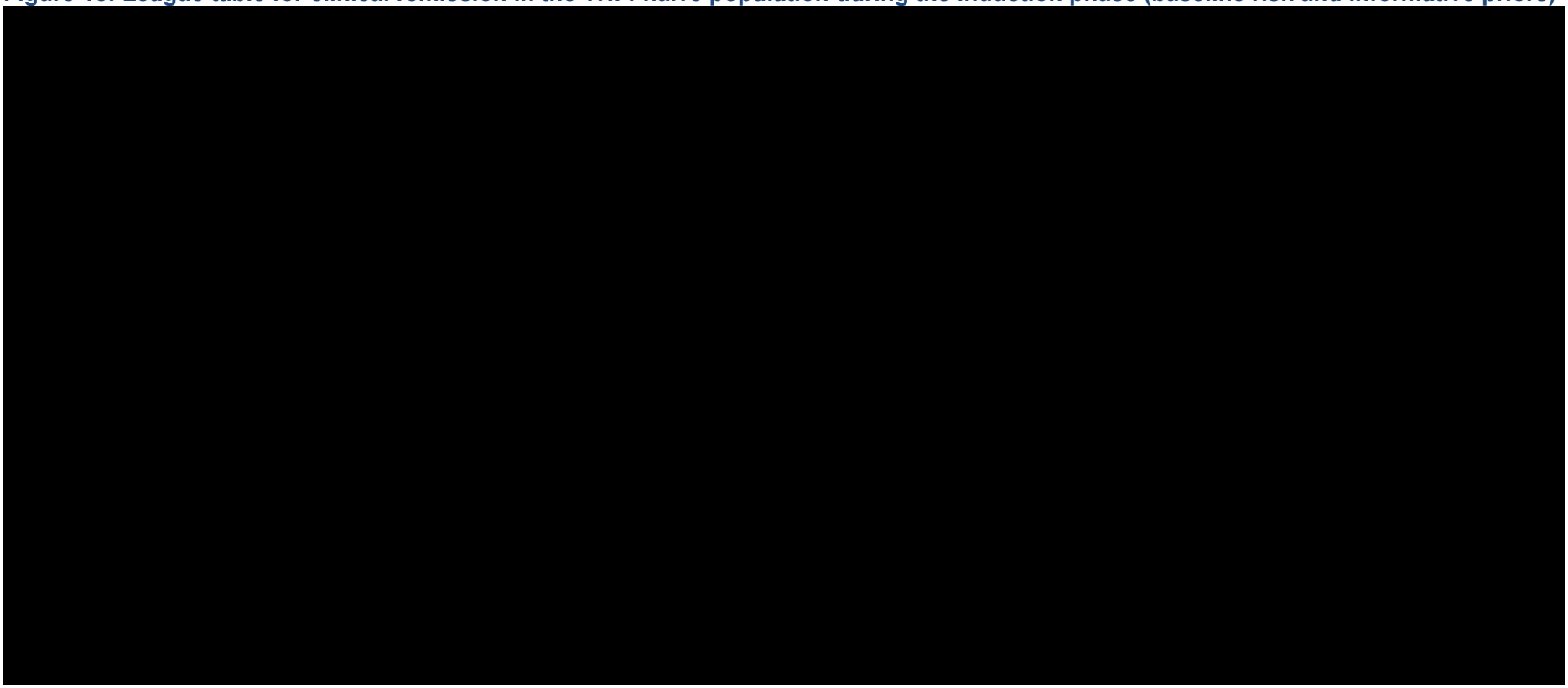
Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour the comparator, grey cells favour placebo.

Figure 17: League table for clinical response in the TNFi-naïve population during the induction phase (baseline risk and informative priors)



Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

Figure 18: League table for clinical remission in the TNFi-naïve population during the induction phase (baseline risk and informative priors)



Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

TNFi-naïve population (maintenance)

Table 23: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-naïve participants during the maintenance phase (baseline risk and informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Infliximab	Pooled	[redacted]	[redacted]
Golimumab	Pooled	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 24: NMA outcomes for comparators versus placebo in TNFi-naïve participants during the maintenance phase (baseline risk and informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Infliximab	Pooled	[redacted]	[redacted]
Golimumab	Pooled	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour comparator, grey cells favour placebo.

Figure 19: League table for clinical response in the TNFi-naïve population during the maintenance phase (baseline risk and informative priors)

Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab pooled; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

Figure 20: League table for clinical remission in the TNFi-naïve population during the maintenance phase (baseline risk and informative priors)

Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; GOL, golimumab pooled; IFX, infliximab pooled; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

TNFi-experienced population (induction)

Table 25: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-experienced participants during the induction phase (baseline risk and informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 26: NMA outcomes for comparators versus placebo in TNFi-experienced participants during the induction phase (baseline risk and informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Adalimumab	160/80/40 mg Q2W	[redacted]	[redacted]
Vedolizumab	300 mg IV	[redacted]	[redacted]
Ustekinumab	Pooled	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Tofacitinib	10 mg BID	[redacted]	[redacted]

Abbreviations: BID, twice a day; CrI, credible interval; IV, intravenous; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour comparator, grey cells favour placebo.

Figure 21: League table for clinical response in the TNFi-experienced population during the induction phase (baseline risk and informative priors)

Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

Figure 22: League table for clinical remission in the TNFi-experienced population during the induction phase (baseline risk and informative priors)

Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab pooled; VEDO, vedolizumab.

TNFi-experienced population (maintenance)

Table 27: NMA outcomes for ozanimod 1 mg QD versus comparators in TNFi-experienced participants during the maintenance phase (baseline risk and informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Ustekinumab	Pooled	[redacted]	[redacted]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]
Placebo	-	[redacted]	[redacted]

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Comparators serve as the reference group; darker coloured cells represent statistically significant differences: blue cells favour ozanimod, grey cells favour the comparator.

Table 28: NMA outcomes for comparators versus placebo in TNFi- experienced participants during the maintenance phase (baseline risk and informative priors)

Comparator	Dose	Clinical response; OR (95% CrI) [a]	Clinical remission; OR (95% CrI) [a]
Ustekinumab	Pooled	[redacted]	[redacted]
Adalimumab	40 mg Q2W	[redacted]	[redacted]
Ozanimod	1 mg QD	[redacted]	[redacted]
Vedolizumab	Pooled	[redacted]	[redacted]
Tofacitinib	Pooled	[redacted]	[redacted]
Vedolizumab	108 mg Q2W SC	[redacted]	[redacted]

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; Q2W, every two weeks; QD, once a day; SC, subcutaneous; TNFi, tumour necrosis factor inhibitor.
Notes: [a] random effects NMA. Placebo serves as the reference group; darker coloured cells represent statistically significant differences: blue cells favour comparator, grey cells favour placebo.

Figure 23: League table for clinical response in the TNFi-experienced population during the maintenance phase (baseline risk and informative priors)

Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

Figure 24: League table for clinical remission in the TNFi-experienced population during the maintenance phase (baseline risk and informative priors)

Random effects NMA. Treatments are ordered from top-left to bottom-right by decreasing rank. For each pairwise comparison, the row treatment serves as the reference group. Pairwise comparisons from the best fitting NMA model are shown in terms of odds ratios and 95% credible interval. An odds ratio > 1 favours the treatment in each column. Odds ratios with credible intervals that do not span unity are shown with a pink background to indicate significance.
Abbreviations: ADA, adalimumab; OZA, ozanimod; PBO, placebo; SC, subcutaneous; TOF, tofacitinib pooled; UST, ustekinumab pooled; VEDO, vedolizumab pooled.

Table 29: Numerical results of original Company base case NMAs and ERG-replicated results of Company base case NMAs compared with results of revised Company base case NMAs

Treatment	Setting (TNFi experience/ treatment phase)	Response			Remission			No response		
		Original Company base-case	ERG	Revised Company base-case	Original Company base-case	ERG	Revised Company base-case	Original Company base-case	ERG	Revised Company base-case
PBO	Naïve/induction	[redacted]								
PBO	Experienced/induction	[redacted]								
PBO	Naïve/maintenance	[redacted]								
PBO	Experienced/maintenance	[redacted]								
OZA	Naïve/induction	[redacted]								
OZA	Experienced/induction	[redacted]								
OZA	Naïve/maintenance	[redacted]								
OZA	Experienced/maintenance	[redacted]								
ADA	Naïve/induction	[redacted]								
ADA	Experienced/induction	[redacted]								
ADA	Naïve/maintenance	[redacted]								
ADA	Experienced/maintenance	[redacted]								
GOL	Naïve/induction	[redacted]								
GOL	Naïve/maintenance	[redacted]								
IFX	Naïve/induction	[redacted]								
IFX	Naïve/maintenance	[redacted]								
TOF	Naïve/induction	[redacted]								
TOF	Experienced/induction	[redacted]								
TOF	Naïve/maintenance	[redacted]								
TOF	Experienced/maintenance	[redacted]								
UST	Naïve/induction	[redacted]								

UST	Experienced/induction	[redacted]								
UST	Naïve/maintenance	[redacted]								
UST	Experienced/maintenance	[redacted]								
VEDO	Naïve/induction	[redacted]								
VEDO	Experienced/induction	[redacted]								
VEDO	Naïve/maintenance	[redacted]								
VEDO	Experienced/maintenance	[redacted]								
VEDO 108	Naïve/maintenance	[redacted]								
VEDO 108	Experienced/maintenance	[redacted]								

[a]NICE ERG confirmed the value of 0.179 for TOF Naïve/maintenance in their report was a transcription error; a copy of the value for Response category. A value of 0.32 was provided in the corrected ERG report.
Abbreviations: ADA, adalimumab; ERG, evidence review group; GOL, golimumab; IFX, infliximab; OZA, ozanimod; PBO, placebo; TOF, tofacitinib; UST, ustekinumab; VEDO, vedolizumab; VEDO 108, vedolizumab 108 mg Q2W SC.

Appendix 4: Updated Company base case following technical engagement

To address the key issues raised by the ERG, the following changes have been implemented to yield the Company's revised base case:

- All transition probabilities are now derived from NMAs with random effects with informative priors, and baseline risk informed by selected trials suggested by the ERG (see Appendix 3)
- Transitions probabilities in the post-active treatment phase are now modelled using efficacy estimates for BSC calculated from 'loss of remission', estimated from sustained remission data
- A revised PAS discount of [redacted]% has been included

The updated Company base case following technical engagement is presented in Table 30 and Table 31 for the TNFi-naïve and TNFi-experienced populations, respectively.

Revised Company base case results

Revised base case pair-wise cost effectiveness results, including a revised PAS discount of [redacted]%, are presented for the TNFi-naïve and TNFi-experienced populations, in Table 30 and Table 31, respectively. For completeness, tofacitinib has been included in the cost-effectiveness model in both populations, however the Company maintains their stance that the exclusion of tofacitinib as a relevant comparator in the TNFi-naïve and TNFi-experienced populations is appropriate and therefore that these analyses are not considered relevant to UK clinical practice. The inclusion of tofacitinib does not form part of the Company's revised base case, but is instead presented for completeness for the Committee's consideration

Ozanimod was found to be cost-effective against all comparators in both populations at a willingness-to-pay threshold of £30,000. Of note, ozanimod was cost-effective against tofacitinib in both populations. The revised base case produced improved cost-effectiveness estimates for ozanimod for all comparisons (i.e. increased SW quadrant ICERs or made ozanimod more dominant), with the exceptions of adalimumab and infliximab in the TNFi-naïve population, which showed marginally decreased SW quadrant ICERs.

The base case fully incremental analysis in both the TNFi-naïve and TNFi-experienced populations showed ozanimod to be the most cost-effective treatment option, with fully incremental ICERs of £45,201 and £88,643 saved per QALY forgone compared to tofacitinib.

Table 30: Revised base case pair-wise cost-effectiveness results – TNFi-naïve population

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER for ozanimod versus comparator (£/QALY)	NHB (at a WTP threshold of £30,000)
Ozanimod	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Ozanimod dominant	0.145
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£273,773[a]	0.300
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£121,137[a]	0.232
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£76,103[a]	0.314
Tofacitinib[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	*£45,201[a]*	*0.100*

[a]SW quadrant ICER; costs saved per QALY forgone.
[b]Not considered a relevant comparator to ozanimod in UK clinical practice.
Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years.

Table 31: Revised base case pair-wise cost-effectiveness results – TNFi-experienced population

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER for ozanimod versus comparator (£/QALY)	NHB (at a WTP threshold of £30,000)
Ozanimod	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£786,412[a]	0.266
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Ozanimod dominant	0.274
Tofacitinib[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	*£88,643[a]*	*0.122*

[a]SW quadrant ICER; costs saved per QALY forgone.
[b]Not considered a relevant comparator to ozanimod in UK clinical practice.
Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; NHB: net health benefit; QALYs: quality-adjusted life years.

Table 32: Revised base case fully incremental cost-effectiveness results – TNFi-naïve population

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Ozanimod	[redacted]	[redacted]	-	-	[redacted]
Adalimumab/biosimilar	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib[b]	[redacted]	[redacted]	[redacted]	[redacted]	*£45,201*
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Infliximab/biosimilar	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£959,385

[a]Not considered a relevant comparator to ozanimod in UK clinical practice.
Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Table 33: Revised base case fully incremental cost-effectiveness results – TNFi-experienced population

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Ozanimod	[redacted]	[redacted]	-	-	-
Tofacitinib[a]	[redacted]	[redacted]	[redacted]	[redacted]	*£88,643*
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated

[a]Not considered a relevant comparator to ozanimod in UK clinical practice.
Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years.

Probabilistic sensitivity analyses

The probabilistic sensitivity analyses were run in line with the methodology outlined in the Company submission. Given the model structure is limited to 5 model engines, and the fact that golimumab is rarely used in clinical practice, golimumab was excluded to permit inclusion of tofacitinib.

Table 34: Probabilistic results (TNFi-naïve population)

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Ozanimod	-	-	-	[redacted]
Adalimumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]
Infliximab	[redacted]	[redacted]	£267,599[b]	[redacted]
Golimumab[c]	-	-	-	-
Vedolizumab	[redacted]	[redacted]	£87,482[b]	[redacted]
Tofacitinib[d]	[redacted]	[redacted]	*£45,526[b]*	[redacted]

[a]The probability of ozanimod being cost-effective versus the comparator technology at a cost-effectiveness threshold of £30,000/QALY.
[b]SW quadrant ICER; costs saved per QALY forgone
[c]Given that the model structure is limited to 5 model engines, and the fact that golimumab is rarely used in clinical practice, golimumab has been excluded to permit inclusion of tofacitinib.
[d]Not considered a relevant comparator to ozanimod in UK clinical practice.
Abbreviations: ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year gained; TNFi: tumour necrosis factor-alpha inhibitor.

Table 35: Probabilistic results (TNFi-experienced population)

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Ozanimod	-	-	-	[redacted]
Vedolizumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]
Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]
Tofacitinib[c]	[redacted]	[redacted]	*£92,592[b]*	[redacted]

[a]The probability of ozanimod being cost-effective versus the comparator technology at a cost-effectiveness threshold of £30,000/QALY.
[b]SW quadrant ICER; costs saved per QALY forgone.
[c]Not considered a relevant comparator to ozanimod in UK clinical practice.
Abbreviations: ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year gained

TNFi-naïve

Figure 25: Cost-effectiveness plane for ozanimod versus adalimumab (TNFi-naïve population)

Abbreviations: QALYs: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; WTP: willingness-to-pay threshold.

Figure 26: Cost-effectiveness plane for ozanimod versus infliximab (TNFi-naïve population)

Abbreviations: QALYs: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness-to-pay threshold.

Figure 27: Cost-effectiveness plane for ozanimod versus vedolizumab (TNFi-naïve population)[a]

Abbreviations: QALYs: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; WTP: willingness-to-pay threshold.

Figure 28: Cost-effectiveness plane for ozanimod versus tofacitinib[a] (TNFi-naïve population)

[a]Not considered a relevant comparator to UK clinical practice.
Abbreviations: QALYs: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness-to-pay threshold.

Figure 29: Cost-effectiveness acceptability curve for ozanimod versus adalimumab, infliximab, vedolizumab and tofacitinib (TNFi-naïve population)

Abbreviations: QALYs: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness-to-pay threshold.

TNFi-experienced

Figure 30: Cost-effectiveness plane for ozanimod versus vedolizumab (TNFi-experienced population)

Abbreviations: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness to pay.

Figure 31: Cost-effectiveness plane for ozanimod versus ustekinumab (TNFi-experienced population)

Abbreviations: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness to pay.

Figure 32: Cost-effectiveness plane for ozanimod versus tofacitinib (TNFi-experienced population)[a]

[a]Not considered a relevant comparator in UK clinical practice.
Abbreviations: QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor; vs: versus; WTP: willingness to pay.

Figure 33: Cost-effectiveness acceptability curve for ozanimod versus ustekinumab, vedolizumab, tofacitinib (TNFi-experienced population)

Abbreviations: TNFi: tumour necrosis factor-alpha inhibitor.

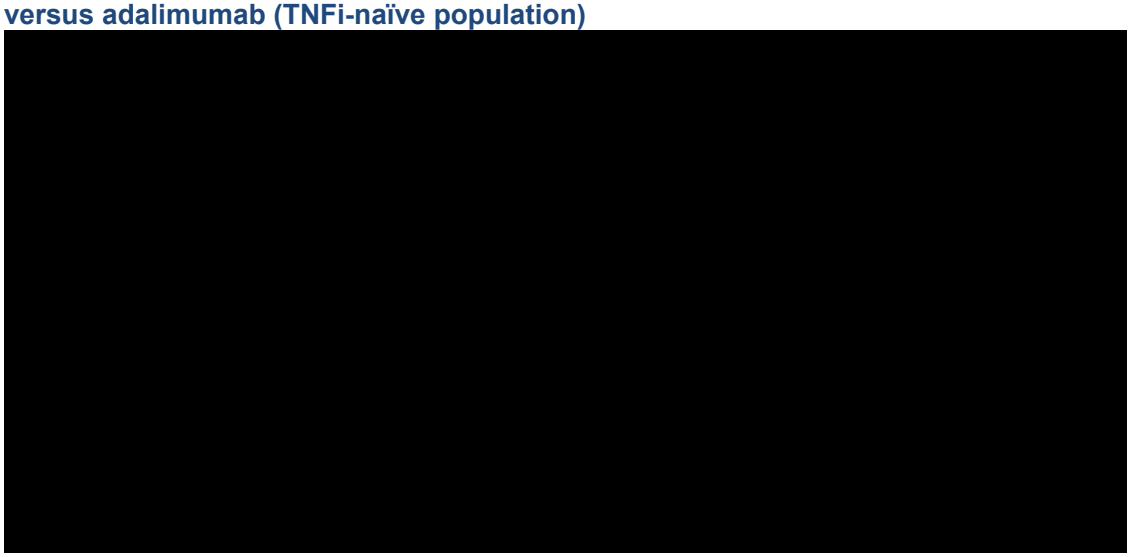
Deterministic sensitivity analysis

Deterministic sensitivity analyses (DSA) were conducted by varying the input for each parameter in the model by ±20% of their mean value, whilst keeping all other inputs the same. For certain parameters where standard errors of the mean were available, the lower and upper limits were defined by the 95% CI around the mean. As some of the comparisons resulted in a South-West (SW) quadrant ICER, to improve the readability of the results, the NHB of ozanimod at a WTP threshold of £30,000 compared to the comparators has been presented. NHB calculated at the upper and lower bounds for the 10 most influential parameters are shown graphically in tornado plots in Figure 34–Figure 38 for the TNFi-naïve population and Figure 39–Figure 41 for the TNFi-experienced population, respectively.

In line with the Company's original base case analysis, the parameters with the greatest impact on the NHB in both the TNFi-naïve and TNFi-experienced population were those related to the proportion of patients achieving sustained clinical response and remission at maintenance.

TNFi-naïve

Figure 34 Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus adalimumab (TNFi-naïve population)



Abbreviations: DSA: deterministic sensitivity analysis; IFX: infliximab; OZA: ozanimod; RNR: response no remission; TNFi: tumour necrosis factor alpha inhibitor.

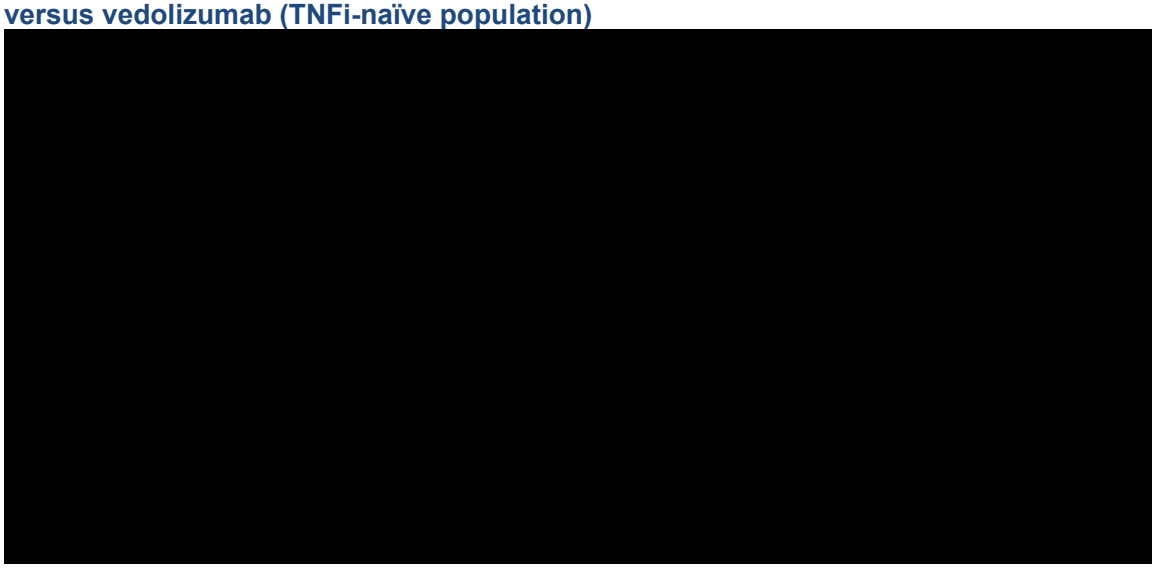
Figure 35 Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus infliximab (TNFi-naïve population)

Abbreviations: ADA: adalimumab; BSC: best supportive care; DSA: deterministic sensitivity analysis; OZA: ozanimod; RNR: response no remission; TNFi: tumour necrosis factor alpha inhibitor.

Figure 36 Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus golimumab (TNFi-naïve population)

Abbreviations: ADA: adalimumab; BSC: best supportive care; DSA: deterministic sensitivity analysis; OZA: ozanimod; RNR: response no remission; TNFi: tumour necrosis factor alpha inhibitor.

Figure 37 Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus vedolizumab (TNFi-naïve population)



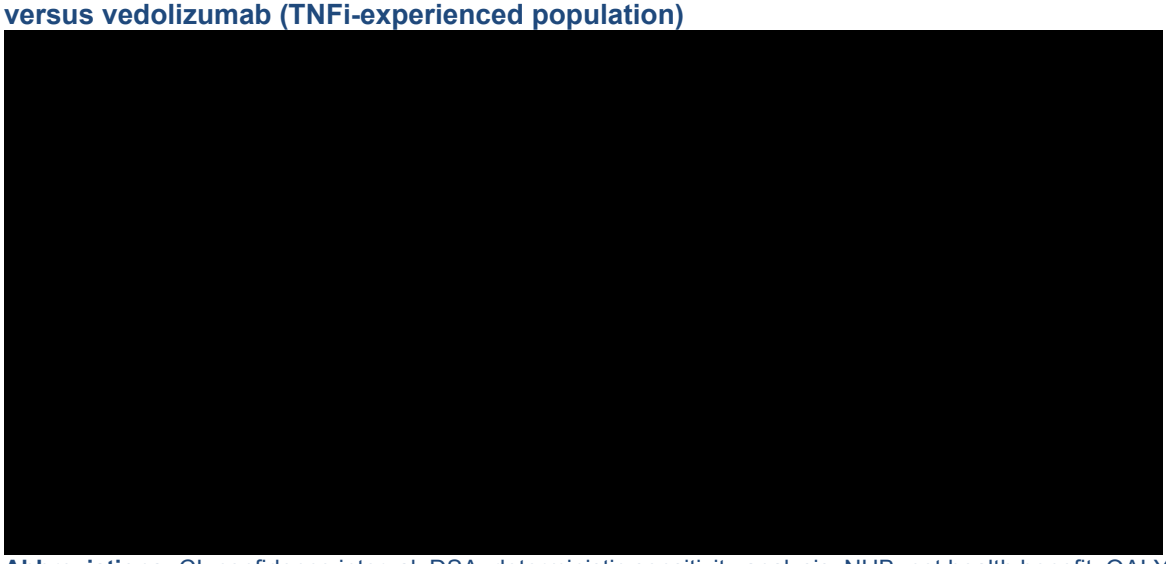
Abbreviations: DSA: deterministic sensitivity analysis; OZA: ozanimod; RNR: response no remission; TNFi: tumour necrosis factor alpha inhibitor; VDZ: vedolizumab.

Figure 38 Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus tofacitinib (TNFi-naïve population)[a]

[a]Not considered a relevant comparator to UK clinical practice.
Abbreviations: DSA: deterministic sensitivity analysis; OZA: ozanimod; RNR: response no remission; TNFi: tumour necrosis factor alpha inhibitor; TOF: tofacitinib.

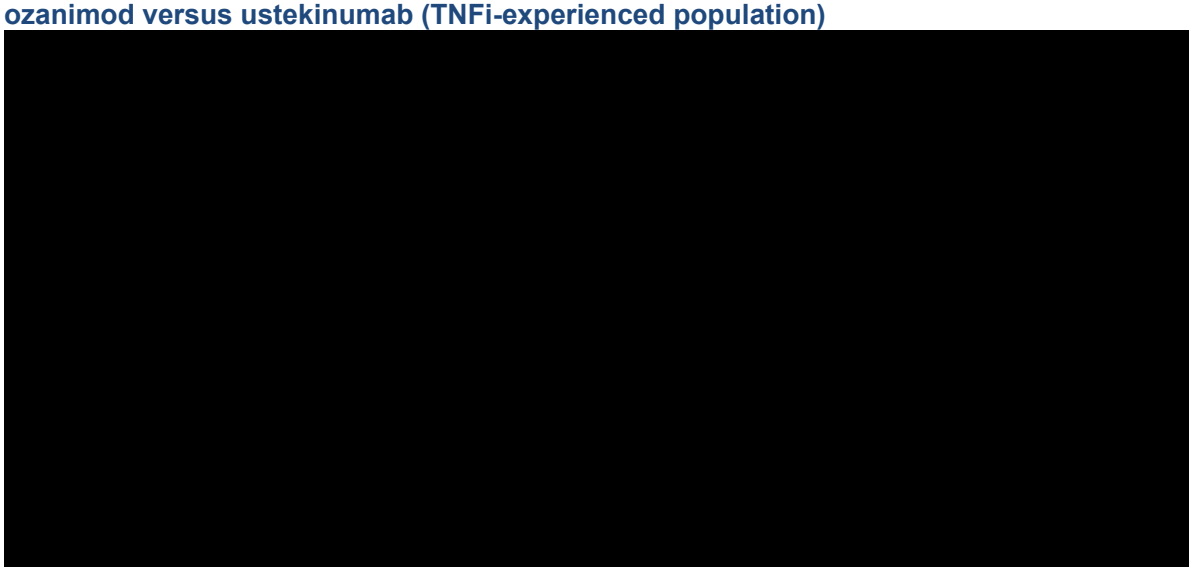
TNFi-experienced

Figure 39: Tornado diagram for the drivers of NHB – top ten parameters for ozanimod versus vedolizumab (TNFi-experienced population)



Abbreviations: CI: confidence interval; DSA: deterministic sensitivity analysis; NHB: net health benefit; QALY: quality-adjusted life year; TNFi: tumour necrosis factor alpha inhibitor.

Figure 40: Tornado diagram for the driver of NHB results – top ten parameters for ozanimod versus ustekinumab (TNFi-experienced population)



Abbreviations: CI: confidence interval; DSA: deterministic sensitivity analysis; IV: intravenous; QALY: quality-adjusted life year; NHB: net health benefit; OZA: ozanimod; SC: subcutaneous; TNFi: tumour necrosis factor alpha inhibitor; VDZ: vedolizumab.

Figure 41: Tornado diagram for the driver of NHB results – top ten parameters for ozanimod versus tofacitinib (TNFi-experienced population)[a]

[a]Not considered a relevant comparator to UK clinical practice.
Abbreviations: CI: confidence interval; DSA: deterministic sensitivity analysis; IV: intravenous; QALY: quality-adjusted life year; NHB: net health benefit; OZA: ozanimod; SC: subcutaneous; TNFi: tumour necrosis factor alpha inhibitor; VDZ: vedolizumab.

Scenario analyses

All scenario analyses explored in the original Company submission have been rerun using revised base case inputs, as well as the additional exploratory scenarios conducted by the ERG. Ozanimod remained cost-effective across the vast majority of scenarios explored in the TNFi-naïve population, and all scenarios explored in the TNFi-experienced population.

Table 36: Results from scenario analyses for ozanimod versus comparators (TNFi-naïve population)

#	Description	Adalimumab			Infliximab			Golimumab			Vedolizumab			*Tofacitinib*[a]		
		Inc. costs	Inc. QALYs	NHB (QALY)	Inc. costs	Inc. QALYs	NHB (QALY)	Inc. costs	Inc. QALYs	NHB (QALY)	Inc. costs	Inc. QALYs	NHB (QALY)	*Inc. costs*	*Inc. QALYs*	*NHB (QALY)*
Scenario analyses explored in original Company submission																
1	1% Spontaneous response/remission	[redacted]	[redacted]	0.142	[redacted]	[redacted]	0.301	[redacted]	[redacted]	-0.075	[redacted]	[redacted]	0.321	[redacted]	[redacted]	*0.107*
2	0% Spontaneous response/remission	[redacted]	[redacted]	0.147	[redacted]	[redacted]	0.299	[redacted]	[redacted]	-0.078	[redacted]	[redacted]	0.307	[redacted]	[redacted]	*0.092*
3	Include extended induction	[redacted]	[redacted]	0.143	[redacted]	[redacted]	0.300	[redacted]	[redacted]	-0.156	[redacted]	[redacted]	0.317	[redacted]	[redacted]	*0.104*
4	0% dose escalation in maintenance phase	[redacted]	[redacted]	0.145	[redacted]	[redacted]	0.301	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.361	[redacted]	[redacted]	*0.115*
5	50% dose escalation in maintenance phase	[redacted]	[redacted]	0.106	[redacted]	[redacted]	0.237	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.236	[redacted]	[redacted]	*0.005*
6	25% treatment waning after 2 years	[redacted]	[redacted]	0.170	[redacted]	[redacted]	0.342	[redacted]	[redacted]	-0.068	[redacted]	[redacted]	0.366	[redacted]	[redacted]	*0.163*
7	Include vial sharing	[redacted]	[redacted]	0.141	[redacted]	[redacted]	0.295	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.307	[redacted]	[redacted]	*0.104*
8	TRUENORTH utilities	[redacted]	[redacted]	0.145	[redacted]	[redacted]	0.253	[redacted]	[redacted]	-0.038	[redacted]	[redacted]	0.314	[redacted]	[redacted]	*0.100*
9	TA342 utilities	[redacted]	[redacted]	0.122	[redacted]	[redacted]	0.318	[redacted]	[redacted]	-0.031	[redacted]	[redacted]	0.414	[redacted]	[redacted]	*0.196*
10	TA547 utilities	[redacted]	[redacted]	0.118	[redacted]	[redacted]	0.322	[redacted]	[redacted]	-0.075	[redacted]	[redacted]	0.431	[redacted]	[redacted]	*0.213*
11	TA547 treatment distribution of	[redacted]	[redacted]	0.142	[redacted]	[redacted]	0.300	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.319	[redacted]	[redacted]	*0.105*

	concomitant therapy and BSC															
12	30% VDZ patients receive SC	[redacted]	[redacted]	0.139	[redacted]	[redacted]	0.294	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.308	[redacted]	[redacted]	*0.094*
13	0% VDZ patients receive SC	[redacted]	[redacted]	0.145	[redacted]	[redacted]	0.300	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.353	[redacted]	[redacted]	*0.100*
Exploratory scenarios conducted by the ERG																
14	0.75% spontaneous response rate	[redacted]	[redacted]	0.145	[redacted]	[redacted]	0.300	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.411	[redacted]	[redacted]	*0.100*
15	Higher proportion of VDZ SC (80%)	[redacted]	[redacted]	0.145	[redacted]	[redacted]	0.300	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.254	[redacted]	[redacted]	*0.100*
16	Larger AE discontinuation in maintenance for ozanimod	[redacted]	[redacted]	0.140	[redacted]	[redacted]	0.295	[redacted]	[redacted]	-0.083	[redacted]	[redacted]	0.310	[redacted]	[redacted]	*0.096*
17	Larger AE incidence in maintenance for ozanimod	[redacted]	[redacted]	0.144	[redacted]	[redacted]	0.300	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.314	[redacted]	[redacted]	*0.100*
18	Population-specific BSC transition probabilities	[redacted]	[redacted]	0.143	[redacted]	[redacted]	0.300	[redacted]	[redacted]	-0.076	[redacted]	[redacted]	0.317	[redacted]	[redacted]	*0.103*
19	Per treatment cycle costs	[redacted]	[redacted]	0.138	[redacted]	[redacted]	0.333	[redacted]	[redacted]	0.238	[redacted]	[redacted]	0.326	[redacted]	[redacted]	*0.100*

Footnotes: [a]Not considered a relevant comparator in UK clinical practice.
Abbreviations: BSC: best supportive care; NHB: net health benefit; QALYs: quality-adjusted life years; SC: subcutaneous; TA: technology assessment; VDZ: vedolizumab.

Table 37: Results from scenario analyses for ozanimod versus comparators (TNFi-experienced population)

#	Description	Vedolizumab			Ustekinumab			*Tofacitinib*[a]		
		Inc. costs	Inc. QALYs	NHB (QALY)	Inc. costs	Inc. QALYs	NHB (QALY)	*Inc. costs*	*Inc. QALYs*	*NHB (QALY)*
Scenario analyses in original Company submission										
1	1% Spontaneous response/remission	[redacted]	[redacted]	0.264	[redacted]	[redacted]	0.271	[redacted]	[redacted]	*0.124*
2	0% Spontaneous response/remission	[redacted]	[redacted]	0.267	[redacted]	[redacted]	0.276	[redacted]	[redacted]	*0.120*
3	Include extended induction	[redacted]	[redacted]	0.329	[redacted]	[redacted]	0.303	[redacted]	[redacted]	*0.155*
4	0% dose escalation in maintenance phase	[redacted]	[redacted]	0.229	[redacted]	[redacted]	0.256	[redacted]	[redacted]	*0.065*
5	50% dose escalation in maintenance phase	[redacted]	[redacted]	0.290	[redacted]	[redacted]	0.285	[redacted]	[redacted]	*0.161*
6	25% treatment waning after 2 years	[redacted]	[redacted]	0.261	[redacted]	[redacted]	0.271	[redacted]	[redacted]	*0.122*
7	Include vial sharing	[redacted]	[redacted]	0.265	[redacted]	[redacted]	0.266	[redacted]	[redacted]	*0.122*
8	TRUENORTH utilities	[redacted]	[redacted]	0.272	[redacted]	[redacted]	0.251	[redacted]	[redacted]	*0.153*
9	TA342 utilities	[redacted]	[redacted]	0.273	[redacted]	[redacted]	0.247	[redacted]	[redacted]	*0.159*
10	TA547 utilities	[redacted]	[redacted]	0.265	[redacted]	[redacted]	0.271	[redacted]	[redacted]	*0.124*
11	TA547 treatment distribution of concomitant therapy and BSC	[redacted]	[redacted]	0.260	[redacted]	[redacted]	0.268	[redacted]	[redacted]	*0.117*
12	30% VDZ patients receive SC	[redacted]	[redacted]	0.283	[redacted]	[redacted]	0.274	[redacted]	[redacted]	*0.122*
13	0% VDZ patients receive SC	[redacted]	[redacted]	0.308	[redacted]	[redacted]	0.274	[redacted]	[redacted]	*0.122*
Exploratory scenarios conducted by the ERG										
14	0.75% spontaneous response rate	[redacted]	[redacted]	0.265	[redacted]	[redacted]	0.272	[redacted]	[redacted]	*0.123*
15	Higher proportion of VDZ SC (80%)	[redacted]	[redacted]	0.236	[redacted]	[redacted]	0.274	[redacted]	[redacted]	*0.122*
16	Larger AE discontinuation in maintenance for ozanimod	[redacted]	[redacted]	0.262	[redacted]	[redacted]	0.270	[redacted]	[redacted]	*0.119*
17	Larger AE incidence in maintenance for ozanimod	[redacted]	[redacted]	0.265	[redacted]	[redacted]	0.273	[redacted]	[redacted]	*0.122*
18	Population-specific BSC transition probabilities	[redacted]	[redacted]	0.265	[redacted]	[redacted]	0.274	[redacted]	[redacted]	*0.122*
19	Per treatment cycle costs	[redacted]	[redacted]	0.280	[redacted]	[redacted]	0.313	[redacted]	[redacted]	*0.122*

Footnotes: [a]Not considered a relevant comparator in UK clinical practice.
Abbreviations: BSC: best supportive care; NHB: net health benefit; QALYs: quality-adjusted life years; SC: subcutaneous; TA: technology assessment; VDZ: vedolizumab.

Appendix 5: Cost-effectiveness model scenarios exploring treatment sequences

As noted in the Company's response to Key issue 5, in order to reduce clinical uncertainty in the economic evaluation, increased flexibility has been added to the Company model to allow for modelling of additional subsequent treatments in both populations. Subsequent treatment options include any treatments where data were available from the TNFi-experienced NMAs, in line with approach taken by the ERG in TA547: adalimumab, vedolizumab, ustekinumab and tofacitinib. For completeness, the results of scenario analyses exploring all subsequent treatment options are presented in Table 38 and Table 39 for the TNFi-naïve and TNFi-experienced populations, respectively. No efficacy data are available to inform efficacy of biologic therapies specifically in the third-line or later. In the absence of relevant data specifically in the third-line setting, efficacy data derived from the TNFi-experienced NMAs were applied for the selected third line treatment. Results presented in the TNFi-experienced population should therefore be interpreted with considerable caution.

Overall, these scenario analyses show comparable results to the revised base case, with only marginal changes to NHB observed when including different subsequent treatment options, demonstrating the base case to be robust to uncertainty regarding the use of treatment sequences in clinical practice.

Table 38: Treatment sequencing scenarios in the TNFi-naïve population

Scenario	First treatment	Second treatment	Incr. Costs	Incr. QALYs	ICER	NHB
Revised Company base case	Ozanimod	-	-	-	-	-
	Adalimumab	-	[redacted]	[redacted]	-£59,307	0.145
	Infliximab	-	[redacted]	[redacted]	£273,773[a]	0.300
	Vedolizumab	-	[redacted]	[redacted]	£76,103[a]	0.314
	Tofacitinib	-	[redacted]	[redacted]	*£45,201[a]*	*0.100*
Subsequent treatment: Vedolizumab	Ozanimod	Vedolizumab	-	-	-	-
	Adalimumab	Vedolizumab	[redacted]	[redacted]	-£61,735	0.145
	Infliximab	Vedolizumab	[redacted]	[redacted]	£275,813[a]	0.299
	Tofacitinib	*Vedolizumab*	[redacted]	[redacted]	*£44,904[a]*	*0.097*
Subsequent treatment: Ustekinumab	Ozanimod	Ustekinumab	-	-	-	-
	Adalimumab	Ustekinumab	[redacted]	[redacted]	-£61,050	0.145
	Infliximab	Ustekinumab	[redacted]	[redacted]	£275,082[a]	0.299
	Vedolizumab	Ustekinumab	[redacted]	[redacted]	£75,989[a]	0.311
	Tofacitinib	*Ustekinumab*	[redacted]	[redacted]	*£44,794[a]*	*0.097*
Subsequent treatment: Tofacitinib	*Ozanimod*	*Tofacitinib*	-	-	-	-
	Adalimumab	*Tofacitinib*	[redacted]	[redacted]	*-£61,891*	*0.145*
	Infliximab	*Tofacitinib*	[redacted]	[redacted]	*£276,749[a]*	*0.300*
	Vedolizumab	*Tofacitinib*	[redacted]	[redacted]	*£76,697[a]*	*0.313*
Subsequent treatment: Adalimumab	*Ozanimod*	*Adalimumab*	-	-	-	-
	Infliximab	*Adalimumab*	[redacted]	[redacted]	*£274,544[a]*	*0.300*
	Vedolizumab	*Adalimumab*	[redacted]	[redacted]	*£76,123[a]*	*0.313*
	Tofacitinib	*Adalimumab*	[redacted]	[redacted]	*£45,066[a]*	*0.099*

[a]SW quadrant ICER; costs saved per QALY forgone
Note: Rows marled in grey represent treatment sequences rarely used in UK clinical practice, according to clinical expert feedback.
Abbreviations: ICER: incremental cost-effectiveness ratio; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

Table 39: Treatment sequencing scenarios in the TNFi-experienced population

Scenario	First treatment	Second treatment	Incr. Costs	Incr. QALYs	ICER	NHB
Revised Company base case	Ozanimod	-	-	-	-	-
	Vedolizumab	-	[redacted]	[redacted]	£786,412[a]	0.266
	Ustekinumab	-	[redacted]	[redacted]	-£141,946	0.274
	Tofacitinib	-	[redacted]	[redacted]	*£88,643[a]*	*0.123*
Subsequent treatment: Vedolizumab	Ozanimod	Vedolizumab	-	-	-	-
	Ustekinumab	Vedolizumab	[redacted]	[redacted]	-£146,079	0.275
	Tofacitinib	*Vedolizumab*	[redacted]	[redacted]	*£88,875[a]*	*0.122*
Subsequent treatment: Ustekinumab	Ozanimod	Ustekinumab	-	-	-	-
	Vedolizumab	Ustekinumab	[redacted]	[redacted]	£771,104[a]	0.265
	Tofacitinib	*Ustekinumab*	[redacted]	[redacted]	*£88,618[a]*	*0.121*
Subsequent treatment: Tofacitinib	Ozanimod	Tofacitinib	-	-	-	-
	Vedolizumab	Tofacitinib	[redacted]	[redacted]	£754,627[a]	0.265
	Ustekinumab	Tofacitinib	[redacted]	[redacted]	-£146,699	0.274

[a]SW quadrant ICER; costs saved per QALY forgone
Notes: Rows italicised and greyed out represent treatment sequences rarely used in UK clinical practice. Rows highlighted in blue represent the most plausible treatment sequences in current UK clinical practice.
Abbreviations: ICER: incremental cost-effectiveness ratio; NHB: net health benefit; QALYs: quality-adjusted life years; TNFi: tumour necrosis factor-alpha inhibitor.

NICE National Institute for Health and Care Excellence

Patient expert statement

[ID3841] Ozanimod for treating moderate to severe ulcerative colitis

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	**Nancy Greig**
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	
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.)	**x**☐ yes
7. How did you gather the information included in your statement? (please tick all that apply)	**x**☐ 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 UC in 2007 and had a subtotal colectomy with formation of an ileostomy for severe UC in 2011 aged 34 I am now relatively well although I will require further surgery to create a permanent ileostomy and remove the rectum in the near future. I continue to have inflammation in the rectal stump. Prior to my surgery and since before diagnosis, I had recurrent flare-ups with explosive diarrhoea, blood and pus in the stools, pain and urgency. I also had abdominal pain, fatigue, joint pain and mouth ulcers. During my worst flare-ups I lost significant amounts of weight and would need to go to the toilet over 25

	times a day. This made it extremely difficult to continue working all day in an office. I had to commute 35 miles and was very anxious about taking the train to work. Sometimes I would sleep for less than 3 hours a night because of the pain and toilet visits, go to the toilet 8-10 times before leaving the house and then avoid eating all day so there was nothing left in my bowel and I could avoid incontinence. Even walking from the station to my office made me feel weak and experience heart palpitations. Following diagnosis I was continuously on various doses of 5-ASAs and during flare-ups, I was treated with steroids. On three occasions I was hospitalised and given IV-steroids. Every time I took steroids I suffered insomnia, anxiety and depression and had to take antidepressants. Each time I took a course I became more resistant to them and the severity of my UC increased. Anti-TNF therapy was not available for UC at the time (2007-2011). My consultant tried both Azathioprine and 6-Mercaptopurine, but both of these caused intolerable nausea and vomiting and did not improve my symptoms. I would have been keen to try other options at this stage if they had been available. After 10 days of IV steroids in December 2011 an emergency subtotal colectomy was performed. Over the next few years I suffered a number of complications. In 2014 I was readmitted with a complete small bowel obstruction caused by adhesions. Following surgery to divide the adhesions and 'unstick' my womb and fallopian tubes, I then developed a pelvic abscess and sepsis. I was informed that due to my surgeries my fallopian tubes were probably immobile. I eventually managed to become pregnant after 2 rounds of IVF and our son was born in 2016 when I was 38. My partner has been the main person caring for me throughout most of my illness and any subsequent complications. When I was suffering the worst effects of UC it was difficult for me to be able to leave the house and for us to have a normal social life. This was very difficult for my partner as I was severely ill for much of the first 6 years of our relationship. We were not able to try for a baby until after my first surgery and then discovered that we had to embark on the difficult journey of fertility treatment. The effects of UC and surgery have put a strain on our

	relationship at times. I still need to have a proctectomy and I am worried about the impact of further surgery on my family, particularly my son who is now 6 years old.
9. What do patients or carers think of current treatments and care available on the NHS?	Patients and carers generally feel that there should be a wider range of treatment options available on the NHS and there are significant side effects with the treatments that exist. Although Anti-TNF therapies and tofacitinib are now available for UC, these are not always effective at achieving long term remission. In many cases people either cannot tolerate side effects or become resistant to particular drugs so it can take a long time to find an appropriate treatment regime or this may not be possible at all with the current treatment options.
10. Is there an unmet need for patients with this condition?	In my opinion there is still significant unmet need in terms of a range of treatments that spare patients the worst effects of steroids, keep their condition in remission and allow them to delay or avoid surgery, for example to start a family or complete higher education. Patients with moderate to severe UC frequently run out of treatment options and are left with either an elective or emergency colectomy as their only option.
Advantages of the technology	
11. What do patients or carers think are the advantages of the technology?	As an oral therapy, patients not have to visit hospital for infusions so the benefits for the patients and carers are clear in terms of being able to enjoy a normal family life. There is also the added convenience for family and patients in terms of fewer hospital visits, lower travel costs, time off work and less burden on carers.

Disadvantages of the technology	
12. What do patients or carers think are the disadvantages of the technology?	I am not aware of any disadvantages of the technology.
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.	I agree with the statements made in the Crohn's & Colitis UK submission about this.
Equality	
14. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	I agree with the statements made in the Crohn's & Colitis UK submission about this.

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:

- I have personal experience of severe UC, which had a significant detrimental impact on my life and my opportunities as a young adult. This is an all consuming and disabling condition.
- Though I am now relatively well after a colectomy, the surgery led to a number of complications which have had a far reaching impact on my life and continue to do so.
- Surgery is not a ‘cure’ for people with moderate to severe UC, nor do I believe it is a less costly option than medical management when the cost of further surgery, readmissions/ complications, infertility treatment and a lifetime of ostomy products are considered.
- There is still significant unmet need for people with moderate to severe UC so Ozanimod would provide another option, particularly for those for whom other treatments have failed or those who wish to avoid or delay surgery.

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

Technical engagement response form

Ozanimod for treating moderately to severely active ulcerative colitis [ID3841]

As a stakeholder you have been invited to comment on the evidence review group (ERG) report for this appraisal.

Your comments and feedback on the key issues below are really valued. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the ERG report that are likely to be discussed by the committee. The key issues in the ERG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the ERG report.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by **5pm** on **18 May 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Janssen
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the ERG report.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Tofacitinib was excluded as a comparator in TNFi-naïve and -experienced subgroups	None	No comment
Key issue 2: Baseline risks for placebo anchors in the NMAs taken from the same trials those used for relative risk	None	No comment
Key issue 3: A RE model may be more appropriate for use in the maintenance phase NMAs	None	No comment
Key issue 4: Modelled efficacy estimates for BSC in the post-active treatment phase	None	No comment
Key issue 5: There is uncertainty surrounding the handling of subsequent treatments/treatment sequencing within the model	None	No comment

Key issue 6: The PSA provided by the company was not considered helpful for decision making	None	No comment

Additional issues

All: Please use the table below to respond to additional issues in the ERG report that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this appraisal (for example, at the clarification stage).

Table 3 Additional issues from the ERG report

Issue from the ERG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Ustekinumab was excluded as a comparator in TNFi-naïve patients	ERG report Section 2.2.1. Current treatments for ulcerative colitis (page 34) ERG report table 7: Summary of the decision problem, comparators (page 40)	None	Janssen would like to bring attention to ERG that ustekinumab is routinely commissioned for biological naïve patients unsuitable to TNFi treatment, in-line with TA633 (described below). Therefore, ustekinumab should have been considered a relevant comparator in both subgroups of patients, TNFi-experienced and TNFi-Naïve, as delineated in the Final Scope of this appraisal and following TA633 recommendation. TA633 recommends ustekinumab as an option for treating moderately to severely active ulcerative colitis in adults when conventional therapy or a biological agent cannot be tolerated, or the disease has responded inadequately or lost response to treatment, only if: • a tumour necrosis factor-alpha inhibitor has failed (that is the disease has responded inadequately or has lost response to treatment) or • **a tumour necrosis factor-alpha inhibitor cannot be tolerated or is not suitable, and** • the company provides ustekinumab at the same price or lower than that agreed with the Commercials Medicines Unit.

NICE National Institute for Health and Care Excellence

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the ERG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the ERG report	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the ERG report	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the ERG report	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
[PLEASE DESCRIBE HERE]

Ozanimod for treating moderately to severely active ulcerative colitis [ID3841]

ERG Review of Company's Response to Technical Engagement Response

---

Produced by

Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Authors

Amanda Brand[1]
Brian O'Toole[1]
Madhusubramanian Muthukumar[1]
Justin Matthews[1]
Fraizer Kiff[1]
Naomi Shaw[1]
Edward CF Wilson[1]
Jonathan Digby-Bell[2]
Louise Crathorne[1]
G.J. Melendez-Torres[1]

[1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter

[2] Consultant Gastroenterologist, Royal Devon & Exeter Hospital

Correspondence to

Amanda Brand
3.09 South Cloisters, St Luke's Campus, Heavitree Road, Exeter, EX1 2LU; a.brand2@exeter.ac.uk

1. INTRODUCTION

This document provides the Evidence Review Group's (ERG's) critique of the company's response to the technical engagement (TE) report produced by the National Institute for Health and Care Excellence (NICE) for the appraisal of ozanimod (ID3841).

In response to technical engagement, the company have sought clinical consultation, presented a series of new analyses, and have updated their economic model to incorporate new clinical efficacy inputs as well as a revised patient access scheme (PAS) price for ozanimod. The company responded only to key issues raised by the ERG; no additional key issues were raised by the company.

The ERG has reviewed the additional evidence presented by the company to address key uncertainties raised in the ERG report. A response to each of the issues, including those raised by the company, is presented in the sections below.

The ERG response is structured as follows:

- Section 2: ERG response to the company's submission at technical engagement
- Section 0: ERG response to updates in the company's base case
- Section 3.3: ERG response to stakeholder comments received during technical engagement.

In addition, this response is accompanied by an appendix containing the results of the company's economic model after confidential patient access scheme (cPAS) discounts have been applied for comparators to ozanimod. Please note that the results in this document therefore only contain the PAS discount agreed for ozanimod.

2. ERG RESPONSE TO COMPANY'S SUBMISSION AT TECHNICAL ENGAGEMENT

This section contains the ERG's response to the company's submission at TE.

2.1. Key issue 1: Tofacitinib was excluded as a comparator in TNFi-naïve and -experienced subgroups

The ERG thanks the company for providing a revised economic model that includes tofacitinib as a comparator both in the TNFi-naïve and -experienced populations as clinical opinion to the ERG indicated that tofacitinib is used in both these populations in UK practice. The ERG notes clinical feedback to the company indicating that the Royal Devon and Exeter NHS Foundation Trust, where both clinicians advising the ERG work, is a specialised tertiary referral centre and likely represent one end of the clinical spectrum. The ERG agrees with the company that this is likely the case as clinical advice to the ERG acknowledged that tofacitinib may be used more at this centre than most others.

Clinical advice to the ERG further indicated that the use of tofacitinib is increasing over time, with one clinician indicating an increase even in the past two years. While the ERG notes the company's reiteration of safety concerns and highlighting of clinical opinion consulted in TA633[1] indicating that tofacitinib is not routinely used in UK practice, it considers the inclusion of tofacitinib to best represent the evolving treatment landscape for the condition. Furthermore, the ERG notes that, despite aforementioned clinical opinion in TA633[1] assessing ustekinumab, this appraisal still considered tofacitinib as a comparator in both the NMA and cost-effectiveness model.

As a result, the ERG agrees with the company that the inclusion of tofacitinib reduces the uncertainty in the committee's decision-making within the context of the heterogeneous and evolving use of tofacitinib throughout the UK. The ERG notes, however, the company's reiteration that tofacitinib is not considered a relevant comparator to ozanimod in UK clinical practice.

Overall, the ERG considers the company's approach to incorporating tofacitinib into the economic model as reasonable; however, due to time constraints, a cell-by-cell validation in MS-Excel model was not possible.

2.2. Key issue 2: Baseline risks for placebo anchors in the NMAs taken from the same trials those used for relative risk

The NMA in the company submission used a baseline risk (i.e. non-response under no treatment or placebo) formed by averaging over the placebo arms of the identical set of selected trials for each setting, that is, the same sources of information. The ERG report pointed out that TSD5[2] recommends separate modelling and sources of evidence for relative treatment effects and baseline effects. The ERG notes the company has accepted the principle of this amendment to their approach during TE.

In line with the guidance in TSD5[2], and in order to develop its own base case, the ERG had preferred a restriction of the full set of placebo arms for the baseline risk in each setting, in order to represent UK clinical practice more closely. The ERG carried out an exploratory analysis in which, for practical reasons, only one trial was adopted for baseline risk in each NMA setting. The ERG notes that the company has adopted exactly the same set of trials (one for each of the four settings) during TE as those used by the ERG in its exploratory analysis. The company also identified and corrected an error in the ERG exploratory analysis (incorrect trial used from that selected for TNFi-experienced at maintenance).

The ERG had contextualised its own rapid exploratory analysis as indefinite, and noted in the ERG report that a more systematic approach was needed:

"The ERG acknowledges the limitations of using these trials for placebo baseline risk, given the unsystematic selection of these based on limited information related to demographics, settings and methodological quality. This approach was selected due to time constraints within the appraisal and should therefore be seen as an attempt at improving the generalisability of results, vis-à-vis that of an unweighted average of all placebo arms, albeit with its own uncertainty."

As a result, the ERG indicated in its report that studies informing baseline placebo risk (which need not necessarily be RCTs) should undergo a separate searching/filtering process oriented to the baseline placebo setting in the target population. The ERG recommended a protocol-driven systematic search as per TSD5[2] to identify sources 'closest to UK clinical practice', with both real-world evidence (RWE) or observational evidence as well as trials being eligible. As stated previously, the company agreed in principle but stated that it was not feasible within the timeframe of TE to carry out a systematic review that included non-RCT sources due to an anticipated high number of records from the suggested search yield.

Instead, the company reports:

'…a targeted search of the literature was conducted to identify potential real-world evidence sources which may be more generalisable to UK clinical practice. A review of UK national registry reports as well as recent guidelines on the management of ulcerative colitis and their reference lists revealed a scarcity of demographic data, and no data that could inform baseline risk.'

The company found few RWE sources and stated that a consultation with clinicians did not reveal any further sources. The details of this search process were not supplied, and the ERG could not specifically determine whether the search would have fully captured reports from UK electronic health records. The company argued that the few RWE sources that were found were not suitable for the purpose of baseline risk estimation.

The ERG acknowledges the difficulty of comprehensive searches within the timeframe of TE, but would reiterate that a full systematic review (including searches of bibliographic databases, plus web searches of grey literature sources and registry reports) may have provided additional baseline-appropriate studies beyond the trials selected in the company submission. However, in the absence of a comprehensive systematic review, the ERG would have preferred the company to provide:

(a) A formal scheme or protocol for selecting evidence for the baseline risk, perhaps building on the ERG's brief outline, namely to find evidence 'selecting sources closest to UK clinical practice, with most appropriate choices for factors identified as important determinants of outcome, i.e. concomitant steroid use, duration of disease, prior TNFi experience, endoscopic central reading and severity of disease defined by a modified Mayo score of 9 or 10 (see ERG report, Section 3.4.2.4).

(b) A more detailed and comprehensive assessment of the existing set of trials following this scheme. This would have included as a first step the exclusion of trials with clearly inappropriate characteristics in the UK setting (in particular excluding studies on Asian populations, which also appeared to differ markedly in the level of extensive disease), and then a thorough systematic examination of characteristics of remaining trials in relation to the UK target, resulting most likely in a subset of suitable trial placebo arms.

Therefore, while preferable to the catch-all approach in the original company submission, the ERG believes that the single-trial analysis presented by the company during TE is suboptimal. The ERG had indicated that its exploratory analysis made no guarantee that the

most suitable selections per setting had been found. More importantly, if there is a subset of trials suitable for baseline estimation the ERG believes that the pooling of this information would have provided a more robust estimation of baseline placebo risk.

Nevertheless, the ERG believes the selected single trials are at least broadly appropriate, as the baseline characteristics of participants in these trials have been validated as broadly representative of UK clinical practice by consulting expert opinion and through examining literature, both by the ERG and the company. The clinical expert to the company was consulted 'to review the studies, their baseline characteristics, and trial-level baseline risks'; it is implied in the company's TE response that the clinician was satisfied. In addition, the company assessed comparability of gender and age demographics to those reported for a UK cohort by King et al. (2020)[3] included in the ERG report, though the company noted the limited scope of comparison. As an aid to further assessment, the ERG has included Table 1 which summarises some relevant characteristics. The company also provided a table of baseline characteristics (Table 4 of the company's TE response, which itself extends information supplied in Table 20 of the ERG report).

As a result of this updated baseline selection, combined with the inclusion of an informative prior in three settings (see Section 2.3), company NMA results are updated. Modelled probabilities of non-response are increased, and response or remission are decreased, over the original company submission (Table 29 of the company's TE response).

Table 1. Information on sites and inclusion/exclusion criteria for selected trials

Setting	Selected trial	Sites	Inclusion Criteria	Exclusion Criteria
TNFi-naïve/ induction	PURSUIT-SC	217 sites in Eastern Europe (400 patients), North America (278 patients), Asia Pacific and South Africa (204 patients), and Western Europe and Israel (183 patients).[4]	Mayo Score 6-12 and endoscopic subscore ≥2 Inadequate response to, or had failed to tolerate, ≥1 of following conventional therapies: oral ASAs, oral CSs, AZA, and/or MP; or CS dependent	Prior use within 12 months of biologic TNFi agent(s) natalizumab or other agents targeting the a-4 integrin, B-cell depleting agents (rituximab), or T-cell depleting agents (alemtuzumab, visilizumab) Prior use of oral CSs at a dose >40 mg prednisone or its equivalent per day Treatment with CSP, TAC, sirolimus, or MMF within 8 weeks
TNFi-naïve /maintenance	PURSUIT M	251 centers.[5]	Mayo Score 6-12 and endoscopic subscore ≥2 Inadequate response to, or had failed to tolerate, ≥1 of following conventional therapies: oral ASAs, oral CSs, AZA, and/or MP; or CS dependent	Prior use within 12 months of biologic TNFi agent(s) natalizumab or other agents targeting the a-4 integrin, B-cell depleting agents (rituximab), or T-cell depleting agents (alemtuzumab, visilizumab) Prior use of oral CSs at a dose >40 mg prednisone or its equivalent per day Receipt of CSP, TAC, sirolimus, or MMF within 8 weeks Patients receiving 5-ASAs, IMMs, CSs at baseline of the PURSUIT-IV or PURSUIT-SC studies had to have maintained doses throughout induction
TNFi-experienced /induction	OCTAVE 1, OCTAVE 2	OCTAVE 1 conducted at 144 sites worldwide; OCTAVE 2, at 169 sites.[6]	Aged ≥18 years UC diagnosis for at least 4 months Mayo Score 6-12 and endoscopic subscore ≥2 Inadequate response to, or had failed to tolerate, ≥1 of following therapies: oral or intravenous glucocorticoids, AZA, MP, IFX or ADA	Crohn's disease, fulminant colitis, toxic megacolon, or indeterminate, microscopic, ischemic, or infectious colitis Treatment with TNFi therapy or interferon therapy within 8 weeks Treatment with CSP, MMF/mycophenolic acid, or TAC within 4 weeks Treatment with intravenous CSs, rectally administered CS or 5-aminosalicylic acid within 2 weeks

Setting	Selected trial	Sites	Inclusion Criteria	Exclusion Criteria
TNFi-experienced /maintenance	GEMINI 1	conducted at 211 medical centers in 34 countries.[7]	Aged 18 to 80 years Mayo score of 6-12 with a sigmoidoscopy subscore of ≥2 Loss of response to, inadequate response to or intolerance to ≥ 1 of: IMMs, TNFi or CSs	Toxic megacolon, abdominal abscess, symptomatic colonic stricture, stoma, a history of colectomy Treatment with TNFi agents within 60 days Treatment with CSP, thalidomide, or investigational drugs within 30 days Prior treatment with VDZ, natalizumab, efalizumab, or rituximab

Source: inclusion and exclusion criteria from Doc B Appendices, Table 11; trial site information from publications
Abbreviations: ADA, adalimumab; ASA, aminosalicylic acid; AZA, azathioprine; CS, corticosteroid; CSP, cyclosporine; IFX, infliximab; IMM, immunomodulator; MP, mercaptopurine; MMF, mycophenolate mofetil; TAC, tacrolimus; TNFi, tumour necrosis factor inhibitor; N/A, not applicable; UC, ulcerative colitis; VDZ, vedolizumab

2.3. Key issue 3: A random effects model may be more appropriate for use in the maintenance phase NMAs

The ERG thanks the company for implementing the recommendation to conduct random effects (RE) models, including an informative prior for the between-trial variation, in its network meta-analyses (NMA) for all four settings (both induction and maintenance phases for TNFi-naïve and -experienced populations). As outlined in the ERG report, it is the position of the ERG that RE models are preferable due to the presence of clinical and methodological heterogeneity. This heterogeneity was acknowledged by the company, but non-convergence or estimation problems were encountered with RE models using uninformative priors in the company submission in three of the four NMA settings (the maintenance phase of both TNFi-naïve and -experienced populations as well as the induction phase of the TNFi-experienced population). This resulted in the company submission using fixed effect (FE) models for the three settings where RE models were not feasible in its NMA.

As part of its TE response, the company used a generic informative prior for placebo-controlled pharmacological studies with subjective outcomes, as supplied by Turner et al. (2015)[8], which the ERG believes is an appropriate choice. The company have reported that by doing so RE models were now feasible, as they were found to be converging in all settings. The company report these results when applied to the base case in Appendix 2 of its TE response. As anticipated, the resulting point estimates of efficacy from the NMA were similar to the original FE models, but credible intervals are wider. The ERG believes that this step has improved the NMA estimates since the original FE modelling approach made a highly unrealistic assumption of no clinical or methodological heterogeneity.

2.4. Key issue 4: Modelled efficacy estimates for BSC in the post-active treatment phase

There are two aspects to this key issue: the first is around the company's use of TNFi-experienced data to inform transition probabilities for the TNFI-naïve population in the BSC arm (Section 2.4.1.1); and the second around estimating response rates in the post active treatment phase of the model i.e. the use of overall response data (including remission) to inform remission transition probabilities for BSC (Section 2.4.1.2).

2.4.1.1. The company's use of TNFi-experienced data to inform transition probabilities for the TNFI-naïve population in the BSC arm

As noted in the ERG report, the ERG preferred the use of subgroup-specific data to inform transition probabilities in the BSC arm, i.e. loss of response and loss of response (no

remission) calculated based on TNFi-naïve and TNFi-experienced estimates. In their response the company maintained that data from the TNFi-experienced group were more appropriate to inform BSC transition probabilities for the TNFi-naïve population, on the basis that patients in the active treatment phase of the model have already failed at least one treatment. The company therefore did not use the subgroup-specific data preferred by the ERG, as part of their revised base case outlined in Section 0.

The ERG maintains that the use of data from the TNFi-experienced group to estimate BSC transitions in both arms ignores the availability of data from the TNFi-naïve population and is therefore inappropriate. Furthermore, the use of subgroup-specific data to inform BSC transitions is consistent with the approach taken TA633[1]. The company acknowledged uncertainty surrounding their approach, therefore as part of their response, a scenario analysis was provided which used the ERG's preferred approach for BSC transitions (see Appendix 4 of the TE response form). See Section 3.1 for further commentary.

2.4.1.2. Estimating response rates in the post active treatment phase of the model i.e. the use of overall response data (including remission) to inform remission transition probabilities for BSC

In their response the company agreed with the ERG's preference that remission probabilities for BSC are more appropriately estimated through 'loss of remission', calculated directly from sustained remission estimates. The company has therefore incorporated this approach into their revised base case, as reported in Section 0.

2.5. Key issue 5: There is uncertainty surrounding the handling of subsequent treatments/treatment sequencing within the model

As part of the ERG's initial critique, it was noted that the model was not flexible enough to capture treatment switching and that some uncertainty remained surrounding the impact of alternative treatment sequencing assumptions on the base case ICER. The ERG noted that the company has now provided model flexibility to allow for the selection of different subsequent treatments, which is helpful. However, the model remains unable to capture the impact of treatment switching or the 'step-up' and 'step-down' assumptions, which are likely to occur in clinical practice, though ERG notes that the impact of these assumptions would likely be less influential than other drivers in terms of the cost-effectiveness.

2.6. Key issue 6: The PSA provided by the Company was not considered helpful for decision making

In its report, the ERG raised concerns surrounding the appropriateness of the company's PSA, on the basis that tofacitinib was not included as a comparator. As part of this response, the company has included tofacitinib as a comparator in both the TNFi-naïve and TNFi-experienced subgroups, as per the ERG's preference. The ERG noted that the company has submitted a revised model as part of their TE response. The ERG noted that in the company's model the TNF-naïve subgroup only allowed for five comparators, i.e. in order to include tofacitinib, the company had to exclude golimumab; similarly, to include golimumab as a comparator tofacitinib had to be excluded. As a result, the cost-effectiveness acceptability curve (CEAC) is not entirely correct; however, it is sufficiently accurate to support decision-making as the probability of golimumab being cost-effective is negligible (close to 0%).

It should also be noted that the company's approach to estimating baseline placebo risk in the NMA was a further element of PSA uncertainty based on the ERG's initial critique of the company submission. As outlined in Section 2.2, the company has attempted to address this uncertainty by adopting the approach used by the ERG during its exploratory analysis, i.e. selecting a single trial per setting that is closer to UK clinical practice to inform placebo risk. Although the ERG notes the improvement from the original approach, it does not consider this amendment to have resolved uncertainty as comprehensively as the approach that was recommended by the ERG.

3. ERG RESPONSE TO CHANGES TO THE COMPANY'S COST-EFFECTIVENESS ESTIMATE(S)

3.1. List of changes made by the company

As noted previously, the company made several changes to their base case analysis in order to address the issues highlighted by the ERG in the report (see Table 2 below for the complete list of changes). Overall, the ERG considered the changes made by the company to be mostly acceptable. However, the ERG did not agree with the company's decision to estimate BSC transition probabilities in the TNFi-naïve subgroup using clinical efficacy data from the TNFi-experienced subgroup (see Section 2.4) and did not fully agree with the company's approach to estimating baseline placebo risk in the NMA (see Section 2.2). The use of BSC transitions (based on subgroup-specific data), has been included as part of the ERG's preferred base case (see Section 3.3 and the cPAS Appendix).

Table 2: List of changes made by the company at Technical Engagement

	Company's original base case	Company's revised base case	Aligns with ERG preference
Tofacitinib as a comparator	Not included	Included	Yes
Baseline risk for placebo anchors	Included pooled baseline placebo risk from all studies included in the NMA	Included baseline placebo risk from one generalisable included trial per NMA setting, as per ERG scenario	Yes, partially
Model type used for maintenance NMA	Fixed effects	Random effects with informative prior	Yes
Method of estimating remission transition probabilities for BSC	Estimated based on loss of overall response (including remission)	Estimated based on 'loss of remission', calculated directly from sustained remission estimates	Yes
BSC transition probabilities for the TNFi-naïve population	Use of TNFi-experienced data	Use of TNFi-experienced data	No
Increased PAS for ozanimod	[redacted]	[redacted]	N/A

Abbreviations: BSC, best supportive care; ERG, Evidence Review Group; N/A, not applicable; NMA, network meta-analysis; PAS, patient access scheme; TNFi, tumour necrosis factor inhibitors

3.2. Company's revised base case results

The company's revised base case results are presented below. Results are presented as follows:

- Pair-wise results are presented in **Table *3*** and **Table *4***.
- Fully incremental results are presented in Table 5 and **Table *6***.
- Probabilistic results are presented in **Table *7*** and **Table *8***.

These results include the revised ozanimod PAS of [redacted]. It should be noted that these results do not include the relevant confidential pricing information provided by NICE and use the company-provided prices for concomitant medications; they therefore do not reflect accurate treatment costs. Please see the appendix to this document, which contains results incorporating those discounts.

Table 3: Company's revised base case results (pairwise): TNFi-naïve subgroup

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER for ozanimod versus comparator (£/QALY)	NHB (at a WTP threshold of £30,000)
Ozanimod	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Ozanimod dominant	0.145
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£273,773[a]	0.300
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£121,137[a]	0.232
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£76,103[a]	0.314
Tofacitinib[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	*£45,201*[a]	*0.100*

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NHB, net health benefit; QALYs, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor; WTP, willingness-to-pay

Note:
[a]SW quadrant ICER; costs saved per QALY forgone
[b]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 4: Company's revised base case results (pairwise): TNFi-experienced subgroup

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER for ozanimod versus comparator (£/QALY)	NHB (at a WTP threshold of £30,000)
Ozanimod	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£786,412[a]	0.266
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Ozanimod dominant	0.274
Tofacitinib[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	*£88,643*[a]	*0.122*

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NHB, net health benefit; QALYs, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor; WTP, willingness-to-pay

Note:
[a]SW quadrant ICER; costs saved per QALY forgone
[b]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 5: Company's revised base case results (fully incremental): TNFi-naïve subgroup

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Ozanimod	[redacted]	[redacted]	-	-	-
Adalimumab/biosimilar	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib[a]	[redacted]	[redacted]	[redacted]	[redacted]	*£45,201*
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Infliximab/biosimilar	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£959,385

Abbreviations: ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor
Note:
[a]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 6: Company's revised base case results (fully incremental): TNFi-experienced subgroup

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Ozanimod	[redacted]	[redacted]	-	-	-
Tofacitinib[a]	[redacted]	[redacted]	[redacted]	[redacted]	*£88,643*
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated

Abbreviations: ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor
Note:
[a]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 7: Company's revised probabilistic results: TNFi-naïve subgroup

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Ozanimod	-	-	-	[redacted]
Adalimumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]
Infliximab	[redacted]	[redacted]	£267,599[b]	[redacted]
Golimumab[c]	-	-	-	-
Vedolizumab	[redacted]	[redacted]	£87,482[b]	[redacted]
Tofacitinib[d]	[redacted]	[redacted]	*£45,526*[b]	[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year gained; TNFi, tumour necrosis factor inhibitor
Notes:
[a]The probability of ozanimod being cost-effective versus the comparator technology at a cost-effectiveness threshold of £30,000/QALY.
[b]SW quadrant ICER; costs saved per QALY forgone
[c]Given that the model structure is limited to 5 model engines, and the fact that golimumab is rarely used in clinical practice, golimumab has been excluded to permit inclusion of tofacitinib.
[d]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 8: Company's revised probabilistic results: TNFi-experienced subgroup

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Ozanimod	-	-	-	[redacted]
Vedolizumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]
Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]
Tofacitinib[c]	[redacted]	[redacted]	*£92,592*[b]	[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year gained; TNFi, tumour necrosis factor inhibitor
Notes:
[a]The probability of ozanimod being cost-effective versus the comparator technology at a cost-effectiveness threshold of £30,000/QALY
[b]SW quadrant ICER; costs saved per QALY forgone
[c]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

3.3. ERG-preferred revised base case results (excluding NICE-provided cPAS and CMU prices)

As noted in Section 3.1, the company's revised changes were considered mostly reasonable by the ERG, with the exception of the estimation of BSC transition probabilities in the TNFi-naïve subgroup. The ERG-preferred revised base case therefore broadly followed the company's revised base case, however, subgroup-specific data were used to inform transition probabilities in the BSC arm. See Please note that, since the TNFi-experienced population already includes subgroup-specific data, the results are the same as that of the company's revised base case.

Table *9* below for full set of ERG-preferred base case revisions – as for the company's revised base case, it should be noted that these results do not include the relevant confidential pricing information provided by NICE and use the company-provided prices for concomitant medications.

ERG-preferred revised base case results are presented as follows;

- Pair-wise results are presented in Table 10 and Table 11
- Fully incremental results are presented in Table 12 and Table 13.
- Probabilistic results are presented in Table 14 and Table 15

Please note that, since the TNFi-experienced population already includes subgroup-specific data, the results are the same as that of the company's revised base case.

Table 9: ERG-preferred revised base case assumptions

	ERG preferred base case assumptions
Tofacitinib as a comparator	included
Baseline risk for placebo anchors	Included baseline placebo risk from one generalisable included trial per NMA setting, as per ERG scenario
Model type used for maintenance NMA	Random effects with informative prior
Method of estimating remission transition probabilities for BSC	Estimated based on 'loss of remission', calculated directly from sustained remission estimates
BSC transition probabilities for the TNFi-naïve population	Use of subgroup-specific data
Increased PAS for ozanimod	XXXX

Abbreviations: BSC, best supportive care; ERG, Evidence Review Group; NMA, network meta-analysis; PAS, patient access scheme; TNFi, tumour necrosis factor inhibitor

Table 10: ERG revised base case results (pairwise): TNFi-naïve subgroup

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER for ozanimod versus comparator (£/QALY)
Ozanimod	[redacted]	[redacted]	[redacted]	-	-	-	-
Adalimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Ozanimod dominant
Infliximab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£276,197[a]
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£122,194[a]
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£77,055[a]
Tofacitinib[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	*£45,852*[a]

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor
Note:
[a]SW quadrant ICER; costs saved per QALY forgone
[b]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 11: ERG revised base case results (pairwise): TNFi-experienced subgroup

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER for ozanimod versus comparator (£/QALY)
Ozanimod	[redacted]	[redacted]	[redacted]	-	-	-	-
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£786,412[a]
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Ozanimod dominant
Tofacitinib[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	*£88,643*[a]
Adalimumab/biosimilar[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	*-£12,593*[a]

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor
Note:
[a]SW quadrant ICER; costs saved per QALY forgone
[b]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 12: ERG revised base case results (fully incremental): TNFi-naïve subgroup

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Ozanimod	[redacted]	[redacted]	-	-	-
Adalimumab/biosimilar	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib[a]	[redacted]	[redacted]	[redacted]	[redacted]	*£45,852*
Golimumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Infliximab/biosimilar	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	£939,694

Abbreviations: ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor
Note:
[a]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 13: ERG revised base case results (fully incremental): TNFi-experienced subgroup

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER incremental (£/QALY)
Ozanimod	[redacted]	[redacted]	-	-	-
Adalimumab/biosimilar[a]	[redacted]	[redacted]	[redacted]	[redacted]	*Dominated*
Tofacitinib[a]	[redacted]	[redacted]	[redacted]	[redacted]	*£88,643*
Ustekinumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
Vedolizumab	[redacted]	[redacted]	[redacted]	[redacted]	Dominated

Abbreviations: ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life years; TNFi, tumour necrosis factor inhibitor
Note:
[a]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 14: ERG revised probabilistic results: TNFi-naïve subgroup

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Ozanimod	-	-	-	[redacted]
Adalimumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]
Infliximab	[redacted]	[redacted]	£272,548[b]	[redacted]
Golimumab[c]	-	-	-	-
Vedolizumab	[redacted]	[redacted]	£88,748[b]	[redacted]
Tofacitinib[d]	[redacted]	[redacted]	*£46,243[b]*	[redacted]

Abbreviations: QALY, quality-adjusted life year gained; TNFi, tumour necrosis factor inhibitor
Notes:
[a]The probability of ozanimod being cost-effective versus the comparator technology at a cost-effectiveness threshold of £30,000/QALY.
[b]SW quadrant ICER; costs saved per QALY forgone
[c]Given that the model structure is limited to 5 model engines, and the fact that golimumab is rarely used in clinical practice, golimumab has been excluded to permit inclusion of tofacitinib.
[d]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

Table 15: ERG revised probabilistic results: TNFi-experienced subgroup

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Ozanimod	-	-	-	[redacted]
Vedolizumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]
Ustekinumab	[redacted]	[redacted]	Ozanimod dominant	[redacted]
Tofacitinib[c]	[redacted]	[redacted]	*£92,592[b]*	[redacted]
Adalimumab/biosimilar[c]	[redacted]	[redacted]	Ozanimod dominant	[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year gained; TNFi, tumour necrosis factor inhibitor
Notes:
[a]The probability of ozanimod being cost-effective versus the comparator technology at a cost-effectiveness threshold of £30,000/QALY
[b]SW quadrant ICER; costs saved per QALY forgone
[c]Not considered a relevant comparator to ozanimod in UK clinical practice by the company

4. ERG RESPONSE TO ISSUES RAISED BY STAKEHOLDERS

Responses to technical engagement were received by the following stakeholder:

- Representative for companies who manufacture a comparator product to ozanimod (Janssen: manufacturer of ustekinumab (STELARA®)).

The stakeholder provided no comment on the key issues raised by the ERG, but raised one additional issue. The ERG has provided specific feedback to the issue raised by the stakeholder:

1. *Ustekinumab was excluded as a comparator in TNFi-naïve patients*

Details of additional issue: Referencing the ERG report, Sections 2.2.1 (p.34) and Table 7 (p.40): 'Janssen would like to bring attention to ERG that ustekinumab is routinely commissioned for biological naïve patients unsuitable to TNFi treatment, in-line with TA633 (described below).

Therefore, ustekinumab should have been considered a relevant comparator in both subgroups of patients, TNFi-experienced and TNFi-Naïve, as delineated in the Final Scope of this appraisal and following TA633 recommendation.

TA633 recommends ustekinumab as an option for treating moderately to severely active ulcerative colitis in adults when conventional therapy or a biological agent cannot be tolerated, or the disease has responded inadequately or lost response to treatment, only if:

- a tumour necrosis factor -alpha inhibitor has failed (that is the disease has responded inadequately or has lost response to treatment) or
- **a tumour necrosis factor -alpha inhibitor cannot be tolerated or is not suitable, and**
- the company provides ustekinumab at the same price or lower than that agreed with the Commercials Medicines Unit.'

ERG response: The ERG thanks Janssen for highlighting this issue. Clinical opinion on the appropriateness of this exclusion was sought by the ERG, and experts agreed that the omission of ustekinumab for TNFi-naïve patients was aligned with routine clinical practice in the UK.

5. REFERENCES

1. National Institute for Health and Care Excellence (NICE). Ustekinumab for treating moderately to severely active ulcerative colitis. Technology appraisal guidance [TA633], 2020. Available from: https://www.nice.org.uk/guidance/TA633.
2. Dias, S., Welton, N.J., Sutton, A.J., Ades, A.E. NICE DSU Technical Support Document 5: Evidence synthesis in the baseline natural history model. 2011; last updated April 2012; available from http://www.nicedsu.org.uk.
3. King D, Reulen RC, Thomas T, Chandan JS, Thayakaran R, Subramanian A, et al. Changing patterns in the epidemiology and outcomes of inflammatory bowel disease in the United Kingdom: 2000-2018. Alimentary Pharmacology & Therapeutics. 2020;51(10):922-34.
4. 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.
5. 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.
6. Sandborn WJ, Su C, Sands BE, D'Haens GR, Vermeire S, Schreiber S, et al. Tofacitinib as induction and maintenance therapy for ulcerative colitis. New England Journal of Medicine. 2017;376(18):1723-36.
7. Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel JF, Sandborn WJ, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. New England Journal of Medicine. 2013;369(8):699-710.
8. Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JP. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. Statistics in Medicine. 2015;34(6):984-98.