EAG STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

# Etrasimod for treating moderately to severely active ulcerative colitis [ID5091]

Confidential until published

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135815

Completed 26th October 2023

CONTAINS [redacted] DATA



UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

**Title:** Etrasimod for treating moderately to severely active ulcerative colitis [ID5091]

**Produced by:** Liverpool Reviews & Implementation Group (LR*i*G)

**Authors:** Janette Greenhalgh, Senior Research Fellow (Clinical Effectiveness), LR*i*G, University of Liverpool

Sam Bryning, Research Associate (Health Economic Modeller), LR*i*G, University of Liverpool

James Mahon, Director, Coldingham Analytical Services, Berwickshire

Marty Chaplin, Research Associate (Medical Statistician), LR*i*G, University of Liverpool

Sophie Beale, Director, HARE Research, North Yorkshire

Angela Boland, Director, LR*i*G, University of Liverpool

Yenal Dundar, Research Fellow (Clinical Effectiveness), LR*i*G, University of Liverpool

Joanne McEntee, Senior Medicines Advice Pharmacist, North West Medicines Information Centre, Liverpool

Keith Bodger, Reader, University of Liverpool and Honorary Consultant Gastroenterologist, Liverpool University Hospital NHS Foundation Trust

**Correspondence to:** Janette Greenhalgh, Senior Research Fellow, Liverpool Reviews and Implementation Group, University of Liverpool, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

**Date completed:** 26th October 2023

**Source of funding:** This report was commissioned by the NIHR Evidence Synthesis Programme as project number 135815

**Acknowledgements:** The authors would like to thank Chris Probert, Professor of Gastroenterology, University of Liverpool who provided feedback on a draft version of the report.



**Rider on responsibility for report:** The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

**Declared competing interests of the authors**: Professor Probert has received funding from Janssen for attending a symposium, speaker fees from Dr Falk Pharma and hospitality from Takeda Pharmaceuticals.

**This report should be referenced as follows:** Greenhalgh J, Bryning S, Mahon J, Chaplin M, Beale S, Boland A, Dundar Y, McEntee J, Bodger K. Etrasimod for treating moderately to severely active ulcerative colitis [ID5091]: A Single Technology Appraisal. LRiG, University of Liverpool, 2023

**Contributions of authors**

| | |
|---|---|
| Janette Greenhalgh | Project lead, critical appraisal of the clinical evidence and supervised the final report |
| Sam Bryning | Critical appraisal of the economic model |
| James Mahon | Critical appraisal of the economic model |
| Marty Chaplin | Critical appraisal of the statistical evidence |
| Sophie Beale | Critical appraisal of the clinical and economic evidence, editorial input |
| Angela Boland | Critical appraisal of the clinical and economic evidence, editorial input |
| Yenal Dundar | Critical appraisal of the company search strategies |
| Joanne McEntee | Critical appraisal of the company submission |
| Keith Bodger | Clinical advice and critical appraisal of the clinical evidence |

## Table of contents

# LIST OF TABLES



# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AE | Adverse event |
| CrI | Credible interval |
| DIC | Deviance Information Criterion |
| EAG | External Assessment Group |
| ECG | Electrocardiogram |
| EMA | European Medicines Agency |
| HRQoL | Health-related quality of life |
| IBDQ | Inflammatory bowel disease questionnaire |
| JAKi | Janus kinase inhibitor |
| MHRA | Medicines and Healthcare products Regulatory Agency |
| MMS | Modified Mayo score |
| NHS | National Health Service |
| NICE | National Institute for Health and Care Excellence |
| NMA | Network meta-analysis |
| OLE | Open-label extension |
| PAS | Patient Access Scheme |
| PEAS | Primary efficacy analysis set |
| PRO | Patient-reported outcome |
| RCT | Randomised controlled trial |
| S1P | Sphingosine-1-phosphate |
| SAE | Serious adverse event |
| SLR | Systematic literature review |
| SmPC | Summary of Product Characteristics |
| SUCRA | Surface Under the Cumulative RAnking curve |
| TEAEs | Treatment-emergent adverse events |
| TNFi | Tumour necrosis factor alpha inhibitor |
| UME | Unrelated mean effects |
| WPAI-UC | Work Productivity and Activity Impairment Questionnaire – Ulcerative Colitis |

# 1 EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the External Assessment Group (EAG) as being potentially important for decision making. It also includes the EAG's cost comparison results.

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the results of the cost comparison analysis. Where appropriate, Sections 1.3 to 1.6 explain the key issues in more detail.

All issues identified represent the EAG's view, not the opinion of NICE.

## *1.1 Overview of the EAG's key issues*

Table A Summary of key issues

| Issue | Summary of issue | Report sections |
|---|---|---|
| Issue 1 | Lack of direct evidence for the comparison of etrasimod versus relevant comparators | Section 2.4 |
| Issue 2 | The company's NMA results do not demonstrate conclusively that etrasimod provides similar or greater health benefits compared to any other drug in the biologic-experienced setting | Section 2.4.6 |
| Issue 3 | For patients in the biologic-experienced setting, it is not clear that a cost comparison approach is the appropriate method of economic evaluation for the comparison of etrasimod versus all other drugs | Section 3.7.8 |
| Issue 4 | Impact of subsequent treatments on cost comparison results is unknown | Section 4.6.4 |

There are no major differences between the company and the EAG's cost comparison analysis results. The EAG has only implemented two minor corrections in the company model and has not proposed any model revisions.

## *1.2 Overview of key model outcomes*

NICE technology appraisals usually compare how much a new technology improves length (overall survival) and quality of life in a QALY. As the company has carried out a cost comparison analysis, the technology is not modelled to affect QALYs. The company has assumed that the results of the five network meta-analyses (NMAs) presented in the company submission demonstrate that etrasimod is at least as efficacious and as safe as the comparators in the cost comparison analysis. The EAG considers that, for biologic-naïve patients, this is a robust conclusion to draw from the results of the company's NMAs of

etrasimod versus adalimumab (induction and maintenance phases). For biologic-experienced patients, company NMA results are mixed; however, no other drug is statistically significantly better than etrasimod (induction and maintenance phases).

The company's base case analysis comprised drug acquisition and administration costs, pre-initiation electrocardiogram (for etrasimod only) and concomitant treatment costs over a 5-year period. The company/EAG cost comparison analysis results are driven by the drug acquisition costs and whether there is a confidential discount in place.

## *1.3 The decision problem: summary of the EAG's key issues*

None

## *1.4 The clinical effectiveness evidence: summary of the EAG's key issues*

Issue 1 Lack of direct evidence for the comparison of etrasimod versus the relevant comparators

| | |
|---|---|
| **Report section** | Section 2.4 |
| **Description of issue and why the EAG has identified it as important** | The company has provided clinical effectiveness evidence from two RCTs, namely the ELEVATE UC 12 and ELEVATE UC 52 trials. Trial results demonstrate the clinical effectiveness of etrasimod versus placebo. There is no direct effectiveness evidence for the comparison of etrasimod versus any of the relevant comparators listed in the final scope issued by NICE, i.e., adalimumab, infliximab, filgotinib, golimumab, ozanimod, tofacitinib, ustekinumab and vedolizumab<br>The company has carried out NMAs to generate indirect clinical effectiveness evidence for the comparison of etrasimod versus relevant comparators[a] |
| **What alternative approach has the EAG suggested?** | None |
| **What is the expected effect on the cost effectiveness estimates?** | See Issue 2 and Issue 3 |
| **What additional evidence or analyses might help to resolve this key issue?** | See Issue 2 and Issue 3 |

[a]Due to a lack of clinical effectiveness data, infliximab and golimumab were not included in the biologic-experienced networks
NMA=network meta-analysis

Issue 2 The company's NMA results do not demonstrate conclusively that etrasimod provides similar or greater health benefits compared to any other drug in the biologic-experienced setting

| | |
|---|---|
| **Report section** | Section 2.4.6 |
| **Description of issue and why the EAG has identified it as important** | The company's NMA results do not demonstrate **conclusively** that etrasimod provides similar or greater health benefits compared to any other drug in the biologic-experienced setting |
| **What alternative approach has the EAG suggested?** | None |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown. If there is not enough evidence to demonstrate similarity, then a cost utility analysis is required |
| **What additional evidence or analyses might help to resolve this key issue?** | Seek clinical advice about whether it is appropriate to assume that, compared to other drugs, etrasimod provides similar or greater health benefits in the biologic-experienced setting |

NMA=network meta-analysis

## *1.5 The cost effectiveness evidence: summary of the EAG's key issues*

Issue 3 For patients in the biologic-experienced setting, it is not clear that a cost comparison approach is the appropriate method of economic evaluation for the comparison of etrasimod versus all other drugs

| | |
|---|---|
| **Report section** | Section 2.4.6 |
| **Description of issue and why the EAG has identified it as important** | Lack of clinical effectiveness evidence to demonstrate **conclusively** that etrasimod provides similar or greater health benefits compared to other drugs in the biologic-experienced setting means it is not clear that a cost comparison approach is the appropriate method of economic evaluation for the comparison of etrasimod versus all other drugs |
| **What alternative approach has the EAG suggested?** | None |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown. If there is not enough evidence to demonstrate similarity, then a cost utility analysis is required |
| **What additional evidence or analyses might help to resolve this key issue?** | Seek clinical advice about whether it is appropriate to assume that, compared to other drugs, etrasimod provides similar or greater health benefits in the biologic-experienced setting. If this assumption is not appropriate then a cost utility analysis is required |

Issue 4 Impact of subsequent treatments on cost comparison results is unknown

| | |
|---|---|
| **Report section** | Section 3.7.8 |
| **Description of issue and why the EAG has identified it as important** | The company base case analysis cost comparison results are only valid for patients who stay on a single treatment for 5 years. Clinical advice to the EAG is that some patients switch treatments during this time interval. It is not possible to make a reliable assumption about second or subsequent treatment(s) for patients in either the biologic-naïve or biologic-experienced setting. As the costs of the drugs available to treat moderately to severely active UC differ, subsequent treatment costs are difficult to capture in an economic model (cost comparison analysis or cost utility analysis) |
| **What alternative approach has the EAG suggested?** | None |
| **What is the expected effect on the cost effectiveness estimates?** | Unknown |
| **What additional evidence or analyses might help to resolve this key issue?** | Seek clinical consensus on current treatment sequencing patterns |

UC=ulcerative colitis

## *1.6 Other key issues: summary of the EAG's view*

None

## *1.7 Summary of EAG's cost comparison analysis results*

Table B Summary of company and EAG cost comparison results

| Treatment | Total 5-year cost per patient | |
|---|---|---|
| | **Company results** | **EAG results** |
| Etrasimod | £55,215 ([redacted]) | £55,215 ([redacted]) |
| Adalimumab | £43,308 | £42,991 |
| Filgotinib | £52,901 | £52,901 |
| Golimumab | £52,040 | £51,659 |
| Infliximab (IV then SC) | £52,527 | £53,754 |
| Infliximab (IV only) | £52,129 | £57,035 |
| Ozanimod | £89,460 | £89,460 |
| Tofacitinib | £46,753 | £46,753 |
| Upadacitinib | £55,464 | £55,464 |
| Ustekinumab | £54,348 | £53,070 |
| Vedolizumab (IV then SC) | £70,506 | £70,301 |
| Vedolizumab (IV only) | £70,408 | £75,314 |

IV=intravenous; PAS=Patient Access Scheme; SC=subcutaneous

# 2 INTRODUCTION AND BACKGROUND

## 2.1 Introduction

This appraisal focuses on the use of etrasimod (Velsipity®) to treat patients 16 years of age and older with moderately to severely active ulcerative colitis (UC) who have had an inadequate response, lost response, or were intolerant to either conventional therapy (biologic-naïve patients), or advanced immunomodulators, i.e., biologic agents or small molecules (biologic-experienced patients). The company has presented evidence for biologic-naïve and biologic-experienced patients for both the induction phase and the maintenance phase of treatments (i.e., four groups).

In the final scope[1] issued by the National Institute for Health and Care Excellence (NICE), the technology (etrasimod) was selected to be appraised as a cost comparison analysis.

In this External Assessment Group (EAG) report, the term 'company submission' (CS) refers to the company's document B, which is the company's full evidence submission.

## 2.2 Etrasimod

Information provided in this section has been extracted from CS, Table 1 and CS, Table 2.

Etrasimod is a sphingosine 1-phosphate receptor modulator that binds to S1P receptors 1, 4 and 5 (S1P1,4,5) and is a balanced G-protein and beta-arrestin agonist at S1P1. Etrasimod partially and reversibly blocks the capacity of lymphocytes to egress from lymphoid organs, reducing the number of lymphocytes in peripheral blood thereby lowering the number of activated lymphocytes in the tissue.

Etrasimod is formulated as a once-daily, orally administered tablet (2mg) that can be taken with or without food. The dosage does not change between induction and maintenance phases of treatment.

The company anticipates that a positive Committee for Medicinal Products for Human Use (CHMP) opinion will be issued on [redacted]. The company plans to submit an application to the Medicines and Healthcare products Regulatory Agency (MHRA) for regulatory approval on [redacted]. The company anticipates that the European Medicines Agency (EMA) will issue marketing authorisation on [redacted] and that the MHRA will issue marketing authorisation on [redacted], with etrasimod becoming available in the UK on [redacted].

The anticipated MHRA marketing authorisation submitted indication is for patients 16 years of age and older with moderately to severely active UC who have had an inadequate response, lost response, or were intolerant to either conventional therapy, or advanced immunomodulators (biologic agents or small molecules).

Etrasimod is available to the NHS at a confidential discounted Patient Access Scheme (PAS) price.

## *2.3 Company's overview of current service provision*

Clinical guidelines for the management of UC include NICE Guidelines (NG30[2]) and British Society of Gastroenterology consensus guidelines 2019.[3] For patients who are the focus of this appraisal, i.e., patients with moderately to severely active UC, when conventional therapy or a biologic agent cannot be tolerated or the disease has responded inadequately or lost response to treatment, current guidelines recommend a biologic or an oral advanced small molecule (non-biologic) therapy.[2,3] The drugs listed in Table 1 are currently recommended by NICE as treatments for NHS patients with moderately to severely active UC.

Table 1 Comparator treatments

| Drug class | Drug | Year | NICE recommendation |
|---|---|---|---|
| TNFi | Adalimumab, infliximab, golimumab TA329[4] | 2015 | An option for treating moderately to severely active UC in adults whose disease has responded inadequately to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who cannot tolerate, or have medical contraindications for, such therapies. |
| Integrin inhibitor | Vedolizumab TA342[5] | 2015 | An option for treating moderately to severely active UC in adults. |
| IL12/23 inhibitor | Ustekinumab TA633[6] | 2020 | An option for treating moderately to severely active UC 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: (i) a TNFi has failed or (ii) a TNFi cannot be tolerated or is not suitable. |
| JAKi | Tofacitinib TA547[7] | 2018 | An option for treating moderately to severely active UC in adults when conventional therapy or a biological agent cannot be tolerated, or the disease has responded inadequately or lost response to treatment. |
| | Filgotinib TA792[8] | 2022 | An option for treating moderately to severely active UC in adults when conventional or biological treatment cannot be tolerated, or the disease has responded inadequately or lost response to treatment. |
| | Upadacitinib TA856[9] | 2023 | An option for treating moderately to severely active UC in adults when conventional therapy or a biological agent cannot be tolerated, or if the condition has not responded well enough or has stopped responding to these treatments. |
| S1P | Ozanimod TA828[10] | 2022 | An option for treating moderately to severely active UC in adults when conventional or biological treatments cannot be tolerated or are not working well enough. |

JAKi=Janus kinase inhibitors; S1P=sphingosine-1-phosphate; TNFi=tumour necrosis factor alpha inhibitors; UC=ulcerative colitis

The company's interpretation of the clinical care pathway for NHS patients with moderately to severely active UC and the proposed placement of etrasimod within this pathway are shown in Figure 1. Clinical advice to the EAG is that Figure 1 is a reasonable reflection of NHS clinical practice for patients with UC.

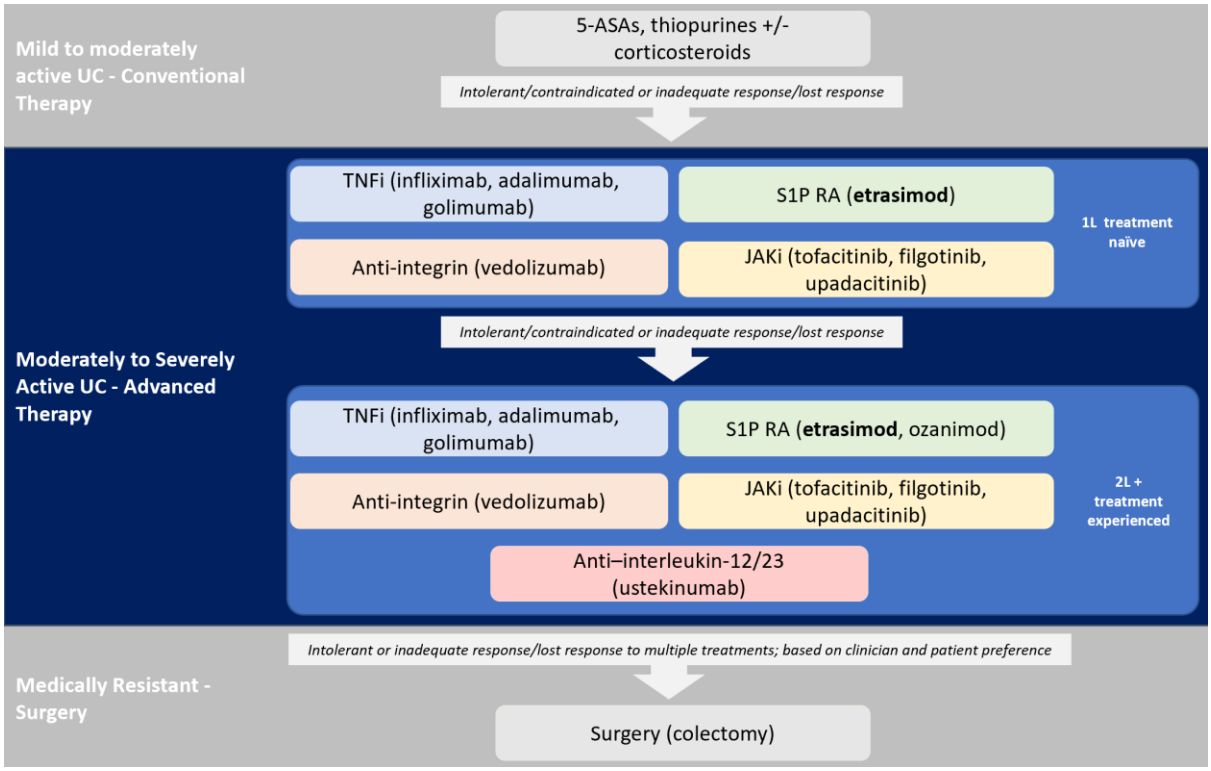


Figure 1 Company representation of the clinical pathway for patients with UC

1L/2L=first/second line; ASA=5-aminosalicylate acid; JAKi=Janus kinase inhibitors; S1P=sphingosine-1-phosphate; RA=receptor antagonist; TNFi=tumour necrosis factor alpha inhibitors; UC=ulcerative colitis
Source: CS, Figure 1

Clinical advice to the EAG is that patients with moderately to severely active UC are typically managed using a sequential treatment approach. The choice of treatment depends on several factors including patient preference, patient contraindications, safety, drug speed of onset, patient antibody responses to prior biologics, any side effects resulting from previous biologics, and cost (CS, p15). Treatment goals extend beyond the alleviation of symptoms to include outcomes such as maintaining a steroid-free remission, mucosal healing, preventing surgery and hospitalisation, and improving patient quality of life.[11]

Clinical advice to the EAG is that:

- in the first instance, most patients who are eligible for treatment with a biologic agent usually receive a TNF-alpha inhibitor (TNFi), such as adalimumab or infliximab (both are available as biosimilars)

- golimumab (a TNFi) is more expensive than adalimumab and infliximab and is therefore used infrequently in NHS clinical practice as a first-line TNFi
- vedolizumab (an integrin inhibitor) may be selected as a first-line biologic agent for patients where there is concern about using a TNFi (i.e., for patients with prior heart failure or increased risk of infections); however, clinical response with vedolizumab is slow compared with TNFi therapies
- ustekinumab (IL 12/23 inhibitor) can be used as a first-line biologic for patients who have contraindications to TNFi therapies

### 2.3.1 Number of patients eligible for treatment with etrasimod

The company provided estimates of the number of patients who would be eligible for treatment with etrasimod in the Budget Impact Model (BIM). Clinical advice to the EAG is that the value used by the company to estimate the proportion of patients with UC who have moderately to severely active disease (52%[10]) may be higher than the proportion of patients with moderately to severely active UC seen in NHS clinical practice.

Table 2 Company estimate of the number of patients with moderately to severely active ulcerative colitis eligible for treatment with etrasimod in Year 1 (prevalent population)

| Population | Proportion | Year 1 (2023) | Source |
|---|---|---|---|
| Total population | - | 61,615,234 | ONS 2021[12] |
| Prevalence of UC in adults | 0.441% | 271,433 | NICE TA856[9] |
| Proportion of UC patients who are adults | 90% | 244,289 | Based on 75% to 80%, 21+ years |
| Proportion of adult UC patients who have moderately to severely active UC | 52% | **127,031** | NICE TA828[10] |
| Proportion with moderately to severely active UC with inadequate response, loss of response or intolerant to CT | 20% | **25,406** | NICE TA828[10] |

CT=conventional therapy; ONS=Office for National Statistics; UC=ulcerative colitis
Source: company BIA model

The company estimates that etrasimod will have a 10% share of the market in Year 1, with this proportion rising to 18% in Year 5 (company BIM).

## *2.4 Critique of company's definition of decision problem*

A summary of the final scope issued by NICE, the decision problem addressed by the company and EAG comments are presented in Table 3. Each parameter is discussed in more detail in the text following Table 3.

Table 3 Summary of decision problem

| | Final scope issued by NICE | Decision problem addressed in the company submission | Rationale if different from the final NICE scope – EAG summary* | EAG comment |
|---|---|---|---|---|
| Population | People with moderately to severely active ulcerative colitis when conventional therapy or a biological agent cannot be tolerated, or the disease has responded inadequately or lost response to treatment. | Patients 16 years of age and older with moderately to severely active UC who have had an inadequate response, lost response, or were intolerant to either conventional therapy, or advanced immunomodulators (biologic agents or small molecules). | The main body of the submission focuses on the advanced treatment naïve population, i.e., naïve to biologics or JAKi. For completeness, advanced treatment experienced analyses are also presented | - |
| Intervention | Etrasimod (Velsipity®) | Etrasimod (Velsipity®) | - | - |
| Comparator(s) | At least 1 of the following treatments, according to NICE guidance:<br>• ozanimod<br>• JAKi (tofacitinib, filgotinib and upadacitinib)<br>• TNFi (infliximab, adalimumab and golimumab)<br>• ustekinumab<br>• vedolizumab<br>• mirikizumab (subject to NICE evaluation) | • TNFi-alpha inhibitors (adalimumab, golimumab and infliximab)<br>• vedolizumab<br>• JAK inhibitors (tofacitinib, filgotinib and upadacitinib) | The target population for etrasimod is patients for whom conventional therapy is inadequately effective, not tolerated or contraindicated. Etrasimod is compared to adalimumab, infliximab and vedolizumab. | Clinical advice to the EAG is that the exclusion of CT and mirikizumab as comparators is reasonable. All relevant comparators have been assessed by the company via network meta-analyses; etrasimod was compared with adalimumab, golimumab, infliximab, vedolizumab, tofacitinib, filgotinib, upadacitinib, ozanimod and ustekinumab. |
| Outcomes | The outcome measures to be considered include:<br>• mortality<br>• disease activity<br>• rates of and duration of response, relapse, and remission<br>• rates of hospitalisation<br>• rates of surgical intervention<br>• endoscopic healing | As per final scope:<br>• measures of disease activity (e.g., rates and duration of response, relapse, and remission<br>• rates of hospitalisation<br>• corticosteroid-free remission<br>• EIHR<br>• HRQoL<br>• rates of surgical intervention<br>• endoscopic improvement | The company has made some assumptions regarding outcome terminology (e.g., around endoscopic healing/normalisation/remission and endoscopic improvement). | Endoscopic remission combined with histological improvement was not captured in the ELEVATE clinical trials.<br><br>The company NMA outcomes are relevant and NMA results can be used to inform treatment decisions. |

| | | | | |
|---|---|---|---|---|
| | • Endoscopic remission combined with histological improvement<br>• corticosteroid-free remission<br>• achieving mucosal healing<br>• AEs<br>• HRQoL | • endoscopic normalisation | | |
| Economic analysis | This technology has been selected to be appraised as a cost comparison.<br>The time horizon should be sufficient to reflect any differences in costs between the technologies being compared.<br>Costs will be considered from an NHS and Personal Social Services perspective.<br>The availability of any commercial arrangements for the intervention and comparator technologies will be taken into account. | Drug acquisition, pre-initiation testing, and administration costs are considered from the NHS and Personal Social Services perspective – all other costs are considered equal across available treatment options.<br>A time horizon of 5 years was selected to reflect differences in initiation costs.<br>The model considers the cost of all available etrasimod comparators.<br>[redacted] | | The company has presented cost comparison analysis results over a 5-year time period.<br>Lack of conclusive clinical effectiveness evidence to demonstrate that etrasimod provides similar or greater health benefits to other drugs in the biologic-experienced setting means that it is not clear if a cost comparison analysis approach is appropriate. |
| Subgroups to be considered | - | Subgroup (or additional analyses, given it is not a subgroup of the naïve population) data for etrasimod is presented among the biologic/JAKi experienced population. | Previous TAs have reported evidence by similar subgroups, therefore for transparency and completeness they have been included in this submission. | The company has provided results of subgroup analyses of trial outcomes according to prior biologic or JAKi therapy exposure (CS, Appendix G) i.e., biologic-naïve or biologic-experienced. |

AEs=adverse events; CT=conventional therapy; EIHR=endoscopic improvement-histologic remission; HRQoL=health-related quality of life; JAKi=Janus kinase inhibitor; TA=technology appraisal; TNFi=tumour necrosis factor alpha inhibitor; UC=ulcerative colitis
*Full details are available in CS, Table 1

### 2.4.1 Source of direct clinical effectiveness data

The company has presented direct clinical effectiveness evidence for the comparison of etrasimod (2mg) versus placebo from two trials designed to assess the efficacy and safety of etrasimod in patients with moderately to severely active UC:

- ELEVATE UC 12[13] (NCT03996369) trial
- ELEVATE UC 52[13] (NCT03945188) trial

Both trials were phase III, randomised, double-blind, placebo-controlled studies. Both trials included a 12-week induction phase and ELEVATE UC 52 also included a 40-week maintenance period. Patients in the two ELEVATE UC trials were eligible to enter an open-label extension (OLE) study: ELEVATE UC OLE study (NCT03950232[14]).

### 2.4.2 Population

The population specified in the final scope issued by NICE is patients with moderately to severely active UC when conventional therapy or a biological agent cannot be tolerated, or the disease has responded inadequately or lost response to treatment. In line with the anticipated licensed indication, the company has presented evidence for people aged 16 years and older and has also included advanced small molecule therapies (non-biologics).

### 2.4.3 Intervention

The intervention is etrasimod (2mg per day). See Section 2.2 for details of the marketing authorisation.

### 2.4.4 Comparators

As the two key etrasimod RCTs were placebo-controlled, the company indirectly compared treatment with etrasimod versus active comparators (n=9) and presented clinical effectiveness results for four different groups:

- biologic-naïve induction phase patients
- biologic-naïve maintenance phase patients
- biologic-experienced induction phase patients
- biologic-experienced maintenance phase patients

The EAG highlights that the terms ‘biologic-naïve’ and ‘biologic-experienced’ encompass patients who are ‘JAKi-naïve’ or ‘JAKi-experienced’ respectively.

For each of the four groups, the company has presented NMA results for the comparison of etrasimod versus adalimumab, infliximab and vedolizumab in the main body of the CS; the NMA results for the comparison of etrasimod versus filgotinib, golimumab, ozanimod, tofacitinib, upadacitinib and ustekinumab are presented in an appendix (CS, Appendix F).

### 2.4.5 Outcomes

The outcomes included in the company's direct and indirect analyses were relevant to this appraisal.

### 2.4.6 Economic analysis

In line with the final scope issued by NICE, the company has carried out a cost comparison analysis; drug costs were assessed over 5 years.

**Appropriateness of a cost comparison analysis**

The EAG considers that, for biologic-naïve patients, the company has conclusively shown that treatment with etrasimod is likely to provide similar or greater health benefits at similar or lower costs compared to adalimumab; NMA results showed that etrasimod was statistically significantly superior to adalimumab. However, as the company NMA results did not show that etrasimod was statistically significantly superior to infliximab or vedolizumab, there is no conclusive evidence of similarity versus these treatments (or versus any other treatments in the network).

The EAG considers that, for biologic-experienced patients, the company has not conclusively shown that treatment with etrasimod is likely to provide similar or greater health benefits at similar or lower cost when compared to adalimumab, infliximab or vedolizumab as there were no statistically significant differences versus any of the comparator treatments.

By carrying out a cost comparison analysis (etrasimod versus nine comparator drugs), the company has implicitly assumed that etrasimod is likely to provide similar or greater health benefits compared to these nine comparator treatments. However, versus most comparator drugs, there is no statistically significant NMA evidence that etrasimod provides similar or greater health benefits.

### 2.4.7 Subgroups

The final scope issued by NICE does not stipulate any subgroup analyses. However, the company has provided results of subgroup analyses of trial outcomes according to prior biologic or JAKi therapy exposure (CS, Appendix G) i.e., biologic-naïve or biologic-experienced. The company cautions that the ELEVATE UC trials were not powered to detect statistically significant treatment effects within subgroups defined by prior biologic exposure status.

### 2.4.8 Other considerations

The company has generated cost comparison analysis results using the Patient Access Scheme (PAS) price for etrasimod and list prices for all other drugs. However, all comparators, are available to the NHS at confidential discounted prices (adalimumab, infliximab and ustekinumab have Commercial Medicines Unit [CMU] prices; filgotinib, golimumab, ozanimod, tofacitinib, upadacitinib and vedolizumab have PAS prices).

Cost comparison analysis results (etrasimod versus all comparators) using all confidential prices are available in a confidential appendix.

# 3 CLINICAL EFFECTIVENESS

## *3.1 Critique of the methods of review(s)*

Full details of the methods used by the company to identify and select clinically relevant evidence to demonstrate the effectiveness of etrasimod are presented in the CS (CS, Appendix F). An assessment of the extent to which the systematic literature review (SLR) was conducted in accordance with the LR*i*G in-house systematic review checklist is presented in Table 4. The EAG conducted its own searches and did not identify any additional trials that provided information on the clinical effectiveness of etrasimod. The EAG considers that the company's review was conducted to a good standard.

Table 4 EAG appraisal of the company's systematic review methods

| Review process | EAG response | Note |
|---|---|---|
| Was the review question clearly defined in terms of population, interventions, comparators, outcomes, and study designs? | Yes | CS, Appendix F.1.1, Table 26 |
| Were appropriate sources searched? | Yes | CS, Appendix F.1.1.1 |
| Was the timespan of the searches appropriate? | Yes | CS, Appendix F.1.1 |
| Were appropriate search terms used? | Yes | CS, Appendix F.1.1.2, Table 27 |
| Were the eligibility criteria appropriate to the decision problem? | Yes | CS, Appendix F.1.1.3, Table 30 |
| Was study selection applied by two or more reviewers independently? | Yes | CS, Appendix F.1.1.3 |
| Was data extracted by two or more reviewers independently? | Yes | CS, Appendix F.1.1.4 |
| Were appropriate criteria used to assess the risk of bias and/or quality of the primary studies? | Yes | CS. Table 9<br>CS, Appendix F.1.1.4 and Appendix F.1.2.4, Table 35 |
| Was the quality assessment conducted by two or more reviewers independently? | Yes | CS, Appendix F.1.1.4 |
| Were attempts to synthesise evidence appropriate? | Yes | NMAs were conducted to allow a comparison of etrasimod with appropriate comparators. The EAG summary and critique of the company's approach are presented in Section 3.7 |

NMA=network meta-analysis
Source: LR*i*G in-house checklist

## *3.2 EAG summary and critique of clinical effectiveness evidence*

### 3.2.1 Trials included in the company systematic review

Two phase III randomised controlled trials (RCTs) that provide clinical effectiveness evidence for etrasimod versus placebo were identified by the company: the ELEVATE UC 12 (NCT03996369) trial and the ELEVATE UC 52 (NCT03945188) trial.

The company also identified a phase II RCT, the OASIS trial (NCT02447302).[15] The OASIS trial compares treatment with two different doses of etrasimod (1mg or 2mg) versus placebo over a period of 12 weeks. OASIS trial results are included in the company's safety NMA (CS, p25).

Patients in the ELEVATE UC 12 and the ELEVATE UC 52 trials were eligible to take part in an open-label extension study, ELEVATE UC OLE (NCT03950232[14]). Patients in the OASIS study were eligible to take part in the OASIS OLE study (NCT02536404[16]). Results from the OLE studies[14,16] were (appropriately) not used to inform the NMAs or the cost comparison analyses; the OLE studies were open-label studies with no control arm.

To compare the clinical effectiveness of treatment with etrasimod versus the comparator treatments listed in the final scope issued by NICE, the company conducted NMAs. The NMAs were conducted for patients with moderately to severely active UC who had not received previous treatment with biologic or JAKi therapies (biologic-naïve) or had received previous treatment with biologic or JAKi therapies (biologic-experienced). The EAG critique and discussion of the company's NMAs is presented in Section 3.7 of this EAG report. Details of the comparator trials included in the company NMAs are available in the CS (CS, Appendix F, Section F.1.2).

### 3.2.2 Etrasimod trials

**Trial characteristics**

A summary of the design and methodologies of the ELEVATE UC 12 and UC 52 trials is presented in the CS (CS, Figure 2 and Table 5). Both trials recruited patients with moderately to severely active UC (defined as modified Mayo score [MMS] of 4 to 9, including an endoscopic subscore ≥2 and a rectal bleeding score ≥1). Patients were randomised (in a 2:1 ratio) to either treatment with etrasimod (2mg) or to placebo. Randomisation was stratified according to previous treatment (biologic or JAKi), baseline use of glucocorticoids and baseline disease activity. The primary endpoint in the trials was the proportion of patients achieving clinical remission defined as a composite of stool frequency subscore=0 (or stool frequency subscore=1 with a ≥1-point decrease from baseline), rectal bleeding subscore=0,

and endoscopic subscore of ≤1 by independent, centrally read assessment (without friability). Patients who did not take part in the ELEVATE OLE study were followed up for 4 weeks.

The ELEVATE UC 12 trial treatment period was 12 weeks and the primary outcome was the proportion of patients achieving clinical remission at 12 weeks. Patients were recruited to the trial from 407 treatment centres across 39 countries. Overall, 238 patients were randomised to receive etrasimod and 116 to receive placebo. Three patients were treated in the UK.

The ELEVATE UC 52 trial treatment period was 52 weeks and the co-primary outcomes were the proportion of patients achieving clinical remission at 12 weeks and at 52 weeks. At the 12-week assessment, patients whose disease activity had shown no improvement, or had worsened compared with baseline, could discontinue treatment and enrol in the ELEVATE UC OLE study (subject to specific criteria being met). Patients whose disease activity had shown no improvement or had worsened during the 40-Week Treatment Period, or who completed all study procedures at Week 52, had the option to enter OLE Study. Patients were recruited to the ELEVATE UC 52 trial from 315 treatment centres across 37 countries. Overall, 289 patients were randomised to receive etrasimod and 144 to receive placebo. One patient was treated in a UK centre.

Clinical advice to the EAG is that the inclusion and exclusion criteria for the ELEVATE UC 12 and the ELEVATE UC 52 trials are reasonable and the results are as generalisable to NHS patients as results from previous trials of UC treatments.

**Patient baseline characteristics**

The baseline characteristics of patients recruited to the ELEVATE UC 12 and the ELEVATE UC 52 trials are presented in the CS (CS, Table 6). Clinical advice to the EAG is that the patients in the trials are comparable to patients recruited to similar trials in this disease area and are representative of patients treated in the NHS. Clinical advice to the EAG is that the ELEVATE trials do not include patients who are hospitalised with acute severe active UC.

### 3.2.3 Quality assessment of the etrasimod trials

The company conducted a quality assessment of the ELEVATE UC 12 and ELEVATE UC 52 trials using the minimum criteria recommended by NICE[17] (CS, Table 9). The EAG agrees with the company's assessments and considers both trials are of good methodological quality.

### 3.2.4 Statistical approaches used to analyse data

In addition to the information provided in the CS, information relevant to the statistical approaches taken by the company to analyse trial data has been extracted from the CSRs,[18,19] the trial statistical analysis plans (TSAP[20]) and the trial protocols.[21,22] The EAG considers that

the approaches adopted by the company were appropriate. See Appendix 6.1 for details.

## *3.3 Efficacy results from the etrasimod trials*

The key efficacy outcome results reported in the CS from the ELEVATE trials are derived from the primary efficacy analysis set (PEAS) and are reported here. The company defines the PEAS population (CS, p35) as patients with an MMS of 5 to 9 who received ≥1 dose of the study drug or placebo. The trial inclusion criteria allowed for the recruitment of patients with an MMS of 4. However, to meet regulatory body requirements,[13] the company has limited the analysis population to patients with an MMS of 5 to 9. In the ELEVATE UC trials, 44 patients had an MMS of 4. The EAG is satisfied that the PEAS population was clearly defined and pre-specified in the TSAP for each of the ELEVATE UC trials.

The EAG has not presented results for the biologic-naïve or biologic experienced patients as trials were not powered to test the statistical significance of subgroup analyses due to the limited numbers in the subgroups (CS, p44). For all subgroup analyses investigated etrasimod showed higher efficacy than placebo. Detailed results are available in CS, Appendix G.

**Summary of patient disposition**

Table 5 shows that few patients in the ELEVATE UC 12 trial discontinued treatment (etrasimod=10.1%, placebo=9.5%). In the ELEVATE UC 52 trial, the rate of discontinuation was lower in the etrasimod arm than in the placebo arm (42.6% versus 68.1%). The main reason for discontinuing treatment in both trials was worsening of disease. The full list of reasons for treatment discontinuation is presented in the CS (CS, Table 8).

Table 5 Summary of patient disposition in the ELEVATE trials

| | ELEVATE UC 12 | | ELEVATE UC 52 | |
|---|---|---|---|---|
| | Etrasimod | Placebo | Etrasimod | Placebo |
| Number of patients randomised | 238 | 116 | 289 | 144 |
| Patients completing treatment | 214 (89.9%) | 105 (90.5%) | 166 (57.4%) | 46 (31.9%) |
| Total discontinuations | 24 (10.1%) | 11 (9.5%) | 123 (42.6%) | 98 (68.1%) |

Source: adapted from CS, Table 8

**Key efficacy results from the ELEVATE UC 12 and ELEVATE UC 52 trials (Week 12)**

The primary outcome of the ELEVATE UC 12 trial was the proportion of patients who achieved clinical remission at Week 12. One of the two co-primary outcomes of the ELEVATE UC 52 trial was the proportion of patients who achieved clinical remission at Week 12. The results for the primary endpoint at Week 12 and for four key secondary endpoints (endoscopic improvement, symptomatic remission, endoscopic improvement-histologic remission and clinical response) for the PEAS population are presented in the CS (CS, Figure 4).

A summary of the outcomes is presented in Table 6. In both ELEVATE trials, statistically significantly more patients treated with etrasimod achieved clinical remission at Week 12 compared with placebo (difference versus placebo in the ELEVATE UC 12 trial was 9.7%; difference versus placebo in the ELEVATE UC 52 trial was 19.8%).

For all four key secondary outcomes, treatment with etrasimod was more effective than placebo (Table 6). Results for the outcomes of clinical remission, endoscopic improvement, endoscopic improvement-clinical remission were all statistically significant. The outcome of clinical response was not included in the company's multiple testing procedure.

Table 6 Primary outcome and key secondary results at Week 12 (ELEVATE UC 12 and ELEVATE UC 52 trials)

| | ELEVATE UC 12 | | ELEVATE UC 52 | |
|---|---|---|---|---|
| | Week 12 | | Week 12 | |
| **Outcome** | **Etrasimod** | **Placebo** | **Etrasimod** | **Placebo** |
| | **N=238** | **N=116** | **N=274** | **N=135** |
| Clinical remission | 24.8% | 15.2% | 27.0% | 7.4% |
| Clinical remission<br>Difference vs placebo | 9.7%<br>(95% CI=1.1 to 18.2)<br>p=0.026 | | 19.8%<br>(95% CI=12.9 to 26.6)<br>p<0.0001 | |
| Endoscopic improvement<br>Difference vs placebo | 12.1%<br>(95% CI=3.0 to 21.2)<br>p=0.0092 | | 21.2%<br>(95% CI=13.0 to 29.3)<br>p<0.0001 | |
| Symptomatic remission<br>Difference vs placebo | 17.5%<br>(95% CI=6.8 to 28.2)<br>p=0.0013 | | 24.6%<br>(95% CI=15.5 to 33.6)<br>p<0.0001 | |
| Endoscopic improvement-histological remission<br>Difference vs placebo | 7.4%<br>(95% CI=0.5 to 14.4)<br>p=0.036 | | 16.9%<br>(95% CI=10.8 to 23.0)<br>p<0.0001 | |
| Clinical response<br>Difference vs placebo | 21.2%<br>(95% CI=10.2 to 32.3)<br>Nominal p=0.0002[a] | | 28.3%<br>(95% CI=18.5 to 38.0)<br>Nominal p<0.0001[a] | |

CI=confidence interval; vs=versus
[a]Hypothesis testing for clinical response was not adjusted for in the company's multiple testing procedure, so the p-value is nominal only
Source: Adapted from CS, Figure 4

**Hospitalisations during the ELEVATE UC 12 trial**

More patients treated with etrasimod (compared with placebo) were admitted to hospital due to UC (1.4% versus 0%). The company highlights (CS, p41) that the small numbers of hospitalised patients do not allow statistically meaningful conclusions to be drawn. None of the patients in the trial had disease-related surgery (CS, p41).

**Key efficacy results from the ELEVATE UC 52 trial at Week 52**

The results for the primary endpoint at Week 52 and for other key secondary outcomes for the PEAS population are presented in the CS (CS, Figure 4). A summary of the outcomes is

presented in Table 7. Statistically significantly more patients treated with etrasimod achieved clinical remission at Week 52 compared with placebo (difference versus placebo was 25.4%). For all key secondary outcomes, treatment with etrasimod was more effective than placebo (Table 7).

Table 7 Primary and key secondary results at Week 52 from the ELEVATE UC 52 trial

| **Clinical** | **Etrasimod** | **Placebo** |
|---|---|---|
| | **N=274** | **N=135** |
| Clinical remission | 32.1% | 18.5% |
| Clinical remission<br>Difference vs placebo | 25.4% (95% CI=18.4 to 32.4)<br>p<0.001 | |
| Endoscopic improvement<br>Difference vs placebo | 26.7% (95% CI=19.0 to 34.4)<br>p<0.0001 | |
| Symptomatic remission<br>Difference vs placebo | 24.9% (95% CI=16.2 to 33.6)<br>p<0.0001 | |
| Endoscopic improvement-histological remission<br>Difference vs placebo | 18.4% (95% CI=11.4 to 25.4)<br>p<0.0001 | |
| Clinical response<br>Difference vs placebo | 24.9% (95% CI=15.8 to 34.1)<br>Nominal p<0.0001[a] | |
| Sustained clinical remission<br>Difference vs placebo | 15.8% (95% CI=10.7 to 21.0)<br>p<0.0001 | |
| Corticosteroid-free clinical remission<br>Difference vs placebo | 25.4% (95% CI=18.4 to 32.4)<br>p<0.0001 | |
| 4-week corticosteroid-free remission among patients with baseline corticosteroid use<br>Difference vs placebo | 23.1% (95% CI=10.2 to 35.9)<br>Nominal p=0.0004[a] | |
| 12-week corticosteroid-free remission among patients with baseline corticosteroid use<br>Difference vs placebo | 23.1% (95% CI=10.2 to 35.9)<br>Nominal p=0.0004[a] | |

CI=confidence interval; vs=versus
[a]Hypothesis testing for this outcome was not accounted for in the company's multiple testing procedure, so the p-value is nominal only
Source:Adapted from CS, Figure 4

Results for other secondary outcomes from the ELEVATE UC 52 trial are reported in the CS (CS, Appendix F) and all show a clinical benefit for etrasimod compared with placebo.

**Hospitalisations during the ELEVATE UC 52 trial**

[redacted] (CS, p43).

## *3.4 Health-related quality of life in the ELEVATE UC 12 and ELEVATE UC 52 trial*

The HRQoL measures used in the etrasimod trials were the Inflammatory Bowel Disease Questionnaire (IBD-Q), the Short Form 36 questionnaire (SF-36), the Short Form 6D (SF-6D) and the Work Productivity and Activity Impairment Questionnaire – Ulcerative Colitis (WPAI-UC). The results of the patient reported outcomes (PROs) from the ELEVATE UC 12 trial and the ELEVATE UC 52 trial at Week 12 are presented in the CS (Table 10). Results of the PROs from the ELEVATE UC 52 trial at Week 52 are presented in the CS (Table 11).

At Week 12, in the ELEVATE UC 12 trial, patients treated with etrasimod reported greater improvements in HRQoL than patients in the placebo arm across most measures. The exceptions were the SF-36 physical component summary and the WPAI-UC work time missed due to absenteeism. In the ELEVATE UC 52 trial, patients treated with etrasimod reported greater improvements in HRQoL compared to patients in the placebo arm across all measures.

At Week 52, patients in the ELEVATE UC 52 trial treated with etrasimod reported greater improvements in HRQoL than patients in the placebo arm for all IBDQ components, for the SF-36 mental component summary and the SF-6D utility index. There were no statistically significant differences between the treatment arms for the SF-36 physical component summary or for any of the WPAI-UC components.

**Results from the OASIS trial**

As noted in Section 3.2.1, the company identified the OASIS trial;[15] a phase II study of etrasimod versus placebo. In the OASIS trial,[15] 50 patients were randomised to receive etrasimod (2mg) and 54 patients were randomised to the placebo arm. The company reports (CS, p43) that the results of the OASIS trial[15] were consistent with the results reported in the ELEVATE trials, i.e., more patients treated with etrasimod achieved clinical remission (33.0% versus 8.1%) and clinical response (50.6% versus 32.5%). OASIS trial[15] clinical remission and clinical response data were not included in the company NMAs as data stratified by prior biologic use were not available; however, OASIS trial[15] safety data were included in the safety NMA (proportion of patients with serious infections).

## 3.5 Subgroup analyses from the ELEVATE UC 12 and UC 52 trials

The final scope issued by NICE does not stipulate any subgroup analyses. However, the company has provided results of subgroup analyses of trial outcomes according to prior biologic or JAKi therapy (CS, Table 12). For the biologic-naïve population, treatment with etrasimod was statistically significantly more effective compared with placebo for all outcomes and all timepoints. For the biologic-experienced population, Week 12 and Week 52 were mixed. The company has used a p-value of <0.05 as a marker of statistical significance and highlights (CS, p44) that p-values should be treated with caution as the ELEVATE trials were not powered to detect statistically significant treatment effects within subgroups defined by prior biologic exposure status.

Additional subgroup analyses (CS, Appendix G) of trial outcomes were conducted for the biologic-naïve and biologic experienced groups based on baseline corticosteroid use (yes/no) and baseline disease activity (MMS 4 to 6 or MMS 7 to 9).

## 3.6 Adverse events

The AEs experienced by patients in the ELEVATE trials are summarised in the CS (CS, Table 18). Specific AEs are reported in the CS, Appendix H (Table 65 and Table 66).

In the ELEVATE UC 12 trial, the company highlights (CS, p58):

- the proportion of patients who reported at least one TEAE was similar in the etrasimod and placebo arms (47.1% versus 46.6%). Most TEAEs were not considered related to the study treatment
- most TEAEs were mild or moderate (Grade 1 or 2) in severity. Grade 3 TEAEs were reported in 7 (2.9%) versus 2 (1.7%) patients in the etrasimod and placebo arms, respectively.
- there was one Grade 4 TEAE in the etrasimod arm (coronary artery disease) and none in the placebo arm
- headache, anaemia, and colitis ulcerative were reported with a >2% difference in the proportion of patients between the etrasimod and placebo arms
- no TEAEs with a fatal outcome were reported during the study.

In the ELEVATE UC 52 trial, the company highlights (CS, p60):

- the proportion of patients who reported at least one TEAE was higher in the etrasimod arm than the placebo arm (71.3% versus 56.3%, exposure adjusted incidence rate: 2.04 versus 1.83, respectively)
- most TEAEs were mild or moderate (Grade 1 or 2) in severity. Grade 3 TEAEs were low and balanced in both treatment arms (etrasimod: 20 [6.9%] patients; placebo: 10 [6.9%] patients)
- there were two Grade 4 TEAEs, one in the etrasimod arm (lymphopenia) and one in the placebo arm (alanine aminotransferase increased)

- the most frequently reported TEAEs were anaemia, headache, colitis ulcerative and coronavirus. Headache and dizziness were reported by >3% more patients in the etrasimod arm than in the placebo arm. Overall, the percentage of patients with TEAEs of colitis ulcerative or abdominal pain was low, and colitis ulcerative TEAEs were lower in etrasimod-treated patients than in patients treated with placebo
- no TEAEs with a fatal outcome were reported during the study.

As noted by the company (CS, Table 2), the SmPC[23] for etrasimod stipulates that all patients should be assessed (using an electrocardiogram) for pre-existing cardiac abnormalities prior to starting treatment and that patients with pre-existing cardiac conditions should be monitored after their first dose.

Beyond the potential impact on patients with pre-existing cardiac conditions, clinical advice to the EAG is that there were no specific or unusual safety concerns or signals in the data presented by the company. Longer-term studies and post-marketing surveillance data would be needed to establish true safety.

## *3.7 Critique of the indirect evidence*

In the absence of head-to-head evidence comparing the clinical effectiveness of etrasimod with the relevant comparators, the company conducted NMAs. The company conducted NMAs for the following outcomes:

- clinical response
- clinical remission
- serious infections

The company performed separate NMAs to assess clinical response and clinical remission for two populations i.e., biologic-naïve, biologic-experienced populations; the NMA for serious infections only includes overall patient population data (Table 8).

Table 8 Main network meta-analyses carried out by the company*

| Population | Induction phase data (Duration: 6-14 weeks) | Maintenance phase data (Duration: 42-54 weeks) |
|---|---|---|
| Biologic-naive | Clinical remission<br>Clinical response | Clinical remission<br>Clinical response |
| Biologic-experienced | Clinical remission<br>Clinical response | Clinical remission<br>Clinical response |
| Overall population | Serious infections | - |

Source: CS, p45

A narrative summary of data for the safety endpoints of SAE and discontinuation due to AE across the studies included in the NMAs is presented in the CS (CS, Appendix F.1.2.10, Table 40 and Table 41).

### 3.7.1 Selection of trials for inclusion in the network meta-analyses

As discussed in Section 3.2.1, the company carried out a global SLR to identify relevant RCTs reporting the efficacy and safety of etrasimod and other relevant comparators for patients with moderately to severely active UC. However, the scope of the company's SLR was broader than the scope required for the NMAs and so the company applied additional selection criteria to identify trials for inclusion in the NMAs. Trials from the SLR were excluded if:

- the trial compared treatments that were out of scope (mirikizumab, risankizumab, and guselkumab)
- the treatment comparison in the trial is not relevant for evidence synthesis (e.g., a comparison between a treatment of interest and a treatment not of interest)
- the trial did not report one of the outcomes of interest (clinical response [induction/maintenance] or clinical remission [induction/maintenance] as measured by the Mayo score, serious infections).

Furthermore, the company only included trials in the NMAs that assessed the efficacy of EMA-licensed doses of therapies specified in the scope. For therapies with a licence that allows for dose increases during the maintenance phase, the company included trials that assessed either the recommended dose or the higher dose. Different doses and/or dosing regimens were treated as unique comparators.

The company states (CS, p46) that, "For RCTs to be eligible for inclusion in the NMA of efficacy outcomes, they were required to report on clinical response and/or clinical remission at the end of an induction (6 to 8 weeks) or maintenance (approximately 1 year) time point". However, the EAG highlights that no trials were excluded based on the induction or maintenance phases not matching these time-points. Indeed, several trials were included that reported induction periods longer than 8 weeks, and several trials were included that reported maintenance phases of less than 1 year (CS, Appendix F, Table 34). For the safety endpoint NMA, trials were required to report on the incidence of serious infections at the end of the induction phase.

Several trials identified in the company's SLR did not meet the inclusion criteria for the NMAs. The company provided reasons for the exclusion of these trials in the CS (CS, Appendix F). The EAG considers that the exclusion of these trials was reasonable.

### 3.7.2 Trials included in the company NMAs

After application of inclusion/exclusion criteria, 31 original trials (116 records) were eligible for inclusion in the company NMAs; a summary of the key characteristics of these 31 trials was included in the CS (CS, Appendix F, Table 34).

A full reference list of the 31 identified trials is presented in the CS (CS, Appendix F, Table 34). Three of the identified trials were not included in any of the company NMAs, the TOUCHSTONE trial,[24] the Sandborn 2012 trial[25] and the LIBERTY-UC trial.[26] These trials do not provide data for subgroups based on prior biologic exposure (CS, p105), or report safety data at the end of an induction period. The remaining 28 trials provide efficacy and safety data for the following treatments:

- adalimumab (6 trials)[27-31]
- etrasimod (3 trials)[13,15]
- filgotinib (1 trial)[32]
- golimumab (3 trials)[33-35]
- infliximab (5 trials)[36-39]
- ozanimod (1 trial)[40]
- tofacitinib (3 trials)[41]
- upadacitinib (2 trials)[42,43]
- ustekinumab (1 trial)[44]
- vedolizumab (4 trials)[31,45-47]

The information presented in Table 9 shows the numbers of RCTs included in the company NMAs, as described in the main body of the CS. The company SLR identified more biologic-naïve population RCT data than biologic-experienced population RCT data, and more induction phase RCT data than maintenance phase RCT data.

Table 9 Number of trials included in the company network meta-analyses

| Population | Induction phase data (duration: 6 to 14 weeks) | | Maintenance phase data (duration: 42 to 54 weeks) | |
|---|---|---|---|---|
| Biologic-naive | Clinical remission/clinical response (n=23) | Adalimumab (n=6)[27-31]<br>Etrasimod (n=2)[13,15]<br>Filgotinib (n=1)[32]<br>Golimumab (n=1)[34]<br>Infliximab (n=5)[36-39]<br>Ozanimod (1)[40]<br>Tofacitinib (n=2)[41]<br>Upadacitinib (n=2)[42,43]<br>Ustekinumab (n=1)[44]<br>Vedolizumab (n=3)[31,45-47] | Clinical remission/clinical response (n=13) | Adalimumab (n=1)[29]<br>Etrasimod (n=1)[13]<br>Filgotinib (n=1)[32]<br>Golimumab (n=2)[33,35]<br>Infliximab (n=1)[39]<br>Ozanimod (n=1)[40]<br>Tofacitinib (n=1)[41]<br>Upadacitinib (n=1)[42,43]<br>Ustekinumab (n=1)[44]<br>Vedolizumab (n=3)[45-47] |
| Biologic-experienced | Clinical remission/clinical response (n=13) | Adalimumab (n=2)[29,31]<br>Etrasimod (n=2) [13,15]<br>Filgotinib (n=1)[32]<br>Ozanimod (n=1)[40]<br>Tofacitinib (n=2)[41]<br>Upadacitinib (n=2)[42,43]<br>Ustekinumab (n=1)[44]<br>Vedolizumab (n=3)[45-47] | Clinical remission/clinical response (n=10) | Adalimumab (n=1)[29]<br>Etrasimod (n=1)[13]<br>Filgotinib (n=1)[32]<br>Ozanimod (n=1)[40]<br>Tofacitinib (n=1)[41]<br>Upadacitinib (n=1)[42,43]<br>Ustekinumab (n=1)[44]<br>Vedolizumab (n=3)[45-47] |
| Overall | Serious infections (n=17) | Adalimumab (n=4)[27,28,30]<br>Etrasimod (n=2)[13,15]<br>Filgotinib (n=1)[32]<br>Golimumab (n=1)[34]<br>Infliximab (n=1)[37]<br>Ozanimod (n=1)[40]<br>Tofacitinib (n=2)[41]<br>Upadacitinib (n=2)[42,43]<br>Ustekinumab (n=1)[44]<br>Vedolizumab (n=2)[46,48] | - | - |

Source: adapted from CS, Table 13, Figure 6, Figure 7 and Figure 8

**Trial characteristics: all included trials**

Key characteristics of the designs of the trials eligible for inclusion in the NMAs are provided in the CS (CS, Table 13 and Appendix F, Table 34). Key patient baseline characteristics are also provided in the CS (CS, Table 36).

The company notes (CS, p102) that most of the RCTs were placebo controlled, except the VARSITY[31] trial (adalimumab versus vedolizumab). It is also noted in the CS that most trials were double blinded, although the VISIBLE 1,[47] PURSUIT-J,[33] TRUE NORTH,[40] GEMINI 1[45] and Motoya 2019[46] trials included an open-label cohort or an open-label induction period.

**Characteristics of trials included in the induction phase NMAs**

The induction phase trials ranged in duration from 6 weeks[35] to 14 weeks.[31] Eleven[27,29,30,34,36-39] trials enrolled biologic-naïve patients only, while the remaining trials enrolled a mixed patient cohort of biologic-naïve and biologic-experienced patients. A comparison of the baseline patient and disease characteristics across each of the arms of the trials included in the induction phase NMAs showed that patients were of a comparable age (mean age ranged from 34.3[36] to 44.5 years[43,49]); however, disease duration (mean 4.4[36] to 10.4[32] years), the proportion of patients with extensive colitis or pan-colitis (9%[28] to 80.8%[37]), and the levels of use of concurrent corticosteroid use varied (13.4%[46] to 77.5%[13]).

**Characteristics of trials included in the maintenance phase NMAs**

The maintenance trials ranged in duration from 42 weeks[40] to 54 weeks.[29,33,34,39] Four[29,33,34,39] of the trials enrolled biologic-naïve patients only. Three[13,29,39] of the trials used a treat-through study design, with the remaining 10[32,33,35,40-47] trials re-randomising patients who entered the maintenance phase. Two other treat-through trials (Suzuki 2014[30] and VARSITY[31]) were eligible for inclusion in the maintenance phase NMAs, however, due to the limitations of the trial data (Section 3.7.4), the company was unable to include data from the Suzuki 2014[30] and VARSITY[31] trials in the maintenance NMAs.

A comparison of the baseline patient and disease characteristics across each of the arms of the trials included in the maintenance phase NMAs, showed that the mean ages of patients were comparable (mean age ranged from 38.1[47] to 44 years[46]); however, there was variation between trials in disease duration (mean 5.9[13,42] to 8.9 years[32]), the proportion of patients with extensive colitis or pan-colitis (6.6%[47] to 58.7%[40]), and levels of concurrent corticosteroid use varied (13.4%[46] to 77.5%[13]).

## 3.7.3 Quality assessment of the trials included in the NMAs

The company quality assessed the trials included in the NMAs using the minimum criteria recommended by NICE.[17] The company quality assessments and EAG comments are

presented in Appendix 6.2. The EAG notes that, in trials where mixed populations were enrolled, patient characteristics were often only reported for the overall population; the EAG therefore considers the assessment of baseline patient comparability is challenging. Overall, the EAG considers that the quality of the trials included in the NMAs was acceptable.

### 3.7.4 Methodological approach to network meta-analyses

A summary of the EAG checks of the company's methodological approach to conducting the NMAs is provided in Appendix 6.3. Overall, the EAG considers that the company's methodological approach was appropriate. Key features of the NMA methodology are outlined in this section.

**Subgroup analysis by prior biologic exposure**

The company performed separate NMAs for biologic-naïve and biologic-experienced patients. Prior biologic exposure was described using different terminology across the included trials. The company assumed the terms 'TNFi-exposure', 'biologic exposure' and 'biologic or JAKi exposure' were interchangeable. If subgroup data based on prior biologic exposure were unavailable, the company used subgroup data based on prior biologic failure. Data for patients who experienced biologic-failure were included in the NMAs for patients with prior exposure to biologic therapy, and data for patients who did not experience biologic-failure were included in the NMAs for patients without prior exposure to biologic therapy. The EAG considers that the different definitions of biologic-exposure status could introduce heterogeneity into the networks of evidence. Trials that did not report subgroup data were excluded from the subgroup analyses.

The NMA for serious infections was conducted using overall trial population data as most included trials did not report this outcome by prior biologic exposure status.

**Treat-through trials versus randomised responder trials**

Of the 31 trials that were eligible for inclusion in the NMA, 15 assessed outcomes at the end of a maintenance phase. These trials were either treat-through trials[13,29-31,39] or randomised responder trials.[32,33,35,40-47] In the treat-through trials, patients were randomised at baseline and outcomes were measured at the end of an induction phase and at the end of a maintenance phase. In the randomised responder trials patients who achieved clinical response during an induction phase (randomised or single-arm) were then randomised to either placebo or to the maintenance dose of the intervention. Outcomes were then measured for these induction-phase responders at the end of the maintenance phase.

The company highlights, and the EAG agrees, that simply combining the reported maintenance phase outcomes from these different types of trial would be inappropriate as it would violate the similarity and homogeneity assumptions necessary for network meta-analysis. Specifically, the patient populations allowed to enter the maintenance phases are incomparable; patients in the randomised responder trials had to have had a response during the induction phase, whereas patients in the treat-through trials may not have had a response during the induction phase. Furthermore, some patients receiving placebo in the maintenance phase of the randomised responder trials would have received active treatment during the induction phase, whereas patients receiving placebo in the maintenance phase of the treat-through trials would have also received placebo during the induction phase.

To account for the differences between the two trial designs, the company converted the outcomes of the treat-through trials to mimic the outcomes of the randomised responder trials. For the ELEVATE 52 trial, the company was able to isolate maintenance phase outcome data for the subset of patients who had responded to treatment during the induction phase as the company had access to individual patient data (IPD) for this trial. For two other treat-through trials (ACT 1[39] and ULTRA 2[29]), the company assumed that the number of responders at the end of induction in each treatment arm could be used as a proxy for the total number of patients who entered the maintenance phase for each treatment arm (if the study had used the randomised responder design). For the induction phase responders, the company established how many of these patients also responded during the maintenance phase by using the number of patients who achieved sustained clinical response. The company would not have been able to use the number of patients who achieved response during the maintenance phase as this would have included some patients who did not respond during the induction phase. For two trials (Suzuki 2014[30] and VARSITY[31]), data were insufficient to apply the adjustments and so these trials were excluded from the maintenance phase analysis.

The EAG considers that the company's approach to accounting for differences between the two trial types was appropriate. However, the EAG highlights that the company's method of adjustment does not account for the fact that the placebo arms of trials included in the company maintenance NMAs are often fundamentally different; some of the placebo arm patients had received and responded to placebo induction (effectively 'skipping' the induction phase), whereas other placebo arm patients had received and responded to active treatment induction. The EAG is unaware of a solution that would account for these differences in placebo arm patients during the maintenance phase.

### 3.7.5 Results of the network meta-analyses: clinical response and clinical remission

The networks of evidence for the analyses of clinical response and clinical remission are provided in Figure 6, Figure 7, Figure 9 and Figure 10 of the CS. A summary of the results from the company's NMAs for clinical response and clinical remission are provided in Table 10. The EAG has not presented results for each comparator versus placebo, or the probabilities of achieving response and remission for each treatment, or surface under cumulative ranking curve (SUCRA) values; these results are available in the CS (CS, Appendix F, Table 48 and Table 49). The EAG has only presented results for comparator doses that are used in NHS clinical practice (and the company economic model).

Table 10 Summary of the company's NMA results: clinical response and clinical remission

| Comparator | Induction phase Etrasimod vs comparator Risk ratio, median (95% CrI) | | Maintenance phase Etrasimod vs comparator Risk ratio, median (95% CrI) | |
|---|---|---|---|---|
| | Clinical response | Clinical remission | Clinical response | Clinical remission |
| **Biologic-naïve subgroup; fixed-effects model[a]** | | | | |
| PBO | [redacted] | [redacted] | [redacted] | [redacted] |
| OZN 1mg | [redacted] | [redacted] | [redacted] | [redacted] |
| TOF 10mg induction, 5mg maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| FIL 200mg | [redacted] | [redacted] | [redacted] | [redacted] |
| UPA 45mg induction, 15mg maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| ADA 160/80/40mg [b] induction, 40mg maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| GOL 200/100mg [c] induction, 50mg maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| IFX 5mg/kg | [redacted] | [redacted] | [redacted] | [redacted] |
| VDZ 300mg induction, 300mg Q8W maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| VDZ 300mg induction, 108mg Q2W maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| UST 90mg Q12W | [redacted] | [redacted] | [redacted] | [redacted] |
| UST 6mg/kg | [redacted] | [redacted] | [redacted] | [redacted] |
| **Biologic-experienced subgroup; random-effects model for induction phase,[d] fixed-effects model for maintenance phase[a]** | | | | |
| PBO | [redacted] | [redacted] | [redacted] | [redacted] |
| OZN 1mg | [redacted] | [redacted] | [redacted] | [redacted] |
| TOF 10mg induction, 5mg maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| FIL 200mg | [redacted] | [redacted] | [redacted] | [redacted] |

| Comparator | Induction phase Etrasimod vs comparator Risk ratio, median (95% CrI) | | Maintenance phase Etrasimod vs comparator Risk ratio, median (95% CrI) | |
|---|---|---|---|---|
| | Clinical response | Clinical remission | Clinical response | Clinical remission |
| UPA 45mg induction, 15mg maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| ADA 160/80/40mg [b] induction, 40mg maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| VDZ 300mg induction, 300mg Q8W maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| VDZ 300mg induction, 108mg Q2W maintenance | [redacted] | [redacted] | [redacted] | [redacted] |
| UST 90mg Q12W | [redacted] | [redacted] | [redacted] | [redacted] |
| UST 6mg/kg | [redacted] | [redacted] | [redacted] | [redacted] |

Green shading indicates that the point estimate of the risk ratio favours etrasimod; red shading indicates that the point estimate of the risk ratio favours the comparator; no shading indicates that the point estimate is 1. Statistically significant results are in bold (95% CrIs do not cross 1)
[a]Fixed-effects model was associated with reasonable model fits in terms of DIC and residual deviance; the random-effects model did not converge
[b]160mg at Week 0, 80mg at Week 2, 40mg at Weeks 4 and 6
[c]200mg at Week 0, 100mg at Week 2
[d]Model fit statistics suggested that the random-effects model was associated with an improved fit, given the residual deviance was lower and the DIC was substantially lower (>5 points) than the fixed-effects model
ADA=adalimumab; Bio=biologics; CrI=credible interval; DIC=deviance information criterion; ETR=etrasimod; FIL=filgotinib; GOL=golimumab; IFX=infliximab; IV=intravenous; OZN=ozanimod; PBO=placebo; SC=subcutaneous; TOF=tofacitinib; UPA=upadacitinib; UST=ustekinumab; VDZ=vedolizumab
Source: CS Appendix F, Table 48 and Table 49

The EAG has grouped comparators as follows: S1P (ozanimod); JAKi (filgotinib, tofacitinib, upadacitinib); TNFi (adalimumab, golimumab, infliximab); other biologic agents (vedolizumab, ustekinumab).

**Etrasimod versus S1P**

[redacted]

**Etrasimod versus JAKi**

[redacted]

**Etrasimod versus TNFi**

[redacted]

**Etrasimod versus other biologic agents**

[redacted]

[redacted]

**Results of inconsistency assessments**

The company provided results of inconsistency assessments for the NMAs of clinical response and clinical remission in their response to clarification question A4.

Comparing the fixed-effects unrelated mean effects (UME) model and the fixed-effects NMA model for each analysis, differences in the residual deviance values were all less than 5. However, there were some differences in the deviance information criterion (DIC) values. For the analysis of efficacy outcomes in the biologic-naïve population during the induction phase, the difference in DIC values between the fixed-effects UME model and the NMA model was 6.17. For the analysis of efficacy outcomes in the biologic-experienced population during the maintenance phase, the difference in DIC values between the fixed-effects UME model and the NMA model was 9.4. For each of these analyses, the company examined deviance contribution points, noting no points fell significantly below the line of equality.

To supplement the assessment of inconsistency, the company also measured heterogeneity for each pairwise comparison to which more than one study contributed. In the biologic-naïve population during the induction phase, the company noted moderate heterogeneity for several pairwise treatment comparisons. In the biologic-experienced population during the maintenance phase, no heterogeneity was detected.

Considering the assessment of the inconsistency in conjunction with the assessment of heterogeneity for pairwise comparisons, the company concluded that they did not “expect there to be any significant inconsistency among the analyses”.

### 3.7.6 Results of the network meta-analyses: serious infections

The networks of evidence for the analyses of serious infections are provided in Figure 8 of the CS. A summary of the results from the company’s NMA for serious infections is provided in Table 11. The EAG has not presented results for each comparator versus placebo, or the probabilities of experiencing a serious infection for each treatment, or SUCRA values. These results are available in the CS (CS, Appendix F, Table 50). The EAG has only presented results for comparator doses that are used in NHS clinical practice (and in the company economic model).

Table 11 Summary of the company’s NMA results: serious infections (fixed-effects model)

| Comparator | Etrasimod vs comparator |
|---|---|
| | Risk ratio, median (95%CrI) |
| PBO | [redacted] |
| ETR 2mg | [redacted] |
| OZN 1mg | [redacted] |
| TOF 10mg | [redacted] |
| FIL 200mg | [redacted] |
| UPA 45mg | [redacted] |
| ADA 160/80/40mg [a] | [redacted] |
| GOL 200/100mg [b] | [redacted] |
| IFX 5mg/kg | [redacted] |
| VDZ 300mg | [redacted] |
| UST 6mg/kg | [redacted] |

Green shading indicates that the point estimate of the risk ratio favours etrasimod; red shading indicates that the point estimate of the risk ratio favours the comparator; no shading indicates that the point estimate is 1
Model fit statistics suggested that the random-effects model was associated with an improved fit. However, due to the rarity of the event the uncertainty in the treatment effects generated by the random-effects model lacked face validity. For this reason, primary results for serious infections during the induction periods were derived from the fixed-effects model
[a]160mg at Week 0, 80mg at Week 2, 40mg at Weeks 4 and 6
[b]200mg at Week 0, 100mg at Week 2
ADA=adalimumab; Bio-=biologics; CrI=credible interval; ETR=etrasimod; FIL=filgotinib; GOL=golimumab; IFX=infliximab; OZN=ozanimod; PBO=placebo; TOF=tofacitinib; UPA=upadacitinib; UST=ustekinumab; VDZ=vedolizumab; SUCRA=surface under cumulative ranking curve

Results from the company's NMA for serious infections were mixed; several relative efficacy estimates favoured comparator treatments over etrasimod, and several favoured etrasimod over comparator treatments. All credible intervals were very wide, with no statistically significant differences observed.

**Results of inconsistency assessments**

The company provided results of inconsistency assessments for the NMAs of serious infections in their response to clarification question A4. Comparing the fixed-effects unrelated UME model and the NMA model, differences in the residual deviance and DIC values were less than 5. The company concluded that they do not "expect there to be any significant inconsistency among the analyses".

### 3.7.7 EAG comment on NMA methods

Generally, the EAG considers that the NMAs were well-conducted. However, the EAG considers that the company's assessment of inconsistency was limited in the following ways:

- It is not clear how the assessment of heterogeneity for pairwise comparisons (clarification question A4) was conducted as only the name of one treatment was provided for each comparison. Most pairwise treatment comparisons in the networks of evidence were comparisons with placebo but, in the biologic-naïve induction phase network, there was one comparison of two different doses of adalimumab (two studies contributed data).
- The company compared the fixed-effects UME model and the fixed-effects NMA model for each analysis. The EAG considers that, for the biologic-experienced population during the induction phase, it would have been more appropriate to compare the random-effects UME model with the random-effects NMA model as the results presented in the CS for this network of evidence were from the random-effects model.
- The company did not compare estimated treatment effects from the UME model with estimated treatment effects from the NMA model.

The EAG agrees with the company that the results of the inconsistency assessments suggested no strong evidence of inconsistency. However, it is not clear how the results of the inconsistency assessments would be impacted if the previously discussed limitations were addressed. Furthermore, the EAG highlights guidance from NICE DSU TSD4,[50] which states that "while tests for inconsistency must be carried out, they are inherently underpowered, and will often fail to detect it. Investigators must therefore also ask whether, if inconsistency is not detected, conclusions from combining direct and indirect evidence can be relied upon".

The EAG notes the following sources of heterogeneity which should be considered when interpreting the company's NMA results:

- patients in the placebo arms had received and responded to different induction treatments (including various active treatments and placebo) with potentially different

persistent effects after treatment has ended (relevant to analyses of maintenance phases)

- different definitions of biologic-exposure status (see Section 3.7.4)
- variation between trials in terms of patient characteristics, including disease duration, the proportion of patients with extensive colitis or pan-colitis, and levels of concurrent corticosteroid use (see Section 3.7.2)

### 3.7.8 EAG comment on NMA results

In the main body of the CS (CS, p51), the company considers that, for both biologic-naïve and biologic-experienced patients, NMA results underpin the claim that etrasimod is likely to provide similar or greater health benefits compared to treatment with adalimumab, infliximab and vedolizumab, the three most widely used NHS comparator treatments. The company presents NMA results for etrasimod versus the remaining comparators in CS, Appendix F; however, these results are not discussed in the text and claims of treatment similarity have not been explicitly made by the company. For completeness (and because the full set of treatments is included in the cost comparison analysis), the EAG has commented on the full set of NMA results.

**Biologic-naïve patients: efficacy**

[redacted]

**Biologic-experienced patients: efficacy**

[redacted]

**All patients: safety (serious infections)**

[redacted]

# 4 EAG CRITIQUE OF COMPANY COST COMPARISON EVIDENCE

## *4.1 Introduction*

In the final scope issued by NICE, it is stated that etrasimod has been selected to be appraised as a cost comparison analysis. The company considered that NMA efficacy and safety results demonstrated that treatment with etrasimod was likely to provide similar or greater health benefits than the three comparator treatments most commonly used in the NHS (i.e., adalimumab, infliximab and vedolizumab) and carried out a cost comparison analysis.

## *4.2 Company cost comparison model*

The company model was developed in MS Excel. As part of the company clarification response, the company provided a model that included additional scenario analysis. These results that were generated using the assumption that 10% of patients (all treatments) would have a complete response at 12 months and discontinue treatment (Section 4.5.2).

### 4.2.1 Population

The company performed separate efficacy NMAs (etrasimod versus comparators) to consider clinical response and clinical remission for four populations:

- biologic-naïve: induction
- biologic-naïve: maintenance
- biologic-experienced: induction
- biologic-experienced: maintenance

The company also performed a safety NMA (etrasimod versus comparators) to consider serious infections using overall population data (i.e., biologic-naïve and biologic-experienced patients).

For the purposes of the cost comparison analysis, the company has assumed that treatment costs for biologic-experienced patients are the same as those for biologic naïve patients.

### 4.2.2 Intervention and comparators

Cost comparison results have been provided for the comparison of etrasimod versus adalimumab, filgotinib, golimumab, infliximab, ozanimod, tofacitinib, upadacitinib, ustekinumab and vedolizumab. The company included the comparison of etrasimod versus mirikizumab in the company model (but not in the CS); mirikizumab is currently under NICE evaluation and is therefore not relevant to this appraisal.

### 4.2.3 Perspective, time horizon and discounting

The company's base case analysis comprised drug acquisition and administration costs, pre-initiation ECG (for etrasimod and ozanimod) and concomitant treatment costs over a 5-year period. The company did not discount costs.

## *4.3 Treatment costs*

The analysis considered the cost of (i) induction (part of first year costs only) and (ii) maintenance treatment over a 5 year period.

### 4.3.1 Drug costs

The dosing schedules used in the company model are presented in Table 12. Drug acquisition costs and administration costs are presented in Table 13 and Table 14, respectively. Where different drug prices are available, the company has used the lowest price to estimate drug costs.

Table 12 Drug costs: dosing schedules used in the company model

| Drug | Route of Administration | Dosing | |
|---|---|---|---|
| | | Initiation | Maintenance |
| **Company model** | | | |
| Etrasimod | Oral | 2mg once daily | |
| Adalimumab | SC | 160mg at Week 0<br>80mg at Week 2 | 40mg every other week |
| Infliximab then Remsima* | Initiation: IV<br>Maintenance: SC | 5mg/kg at Weeks 0 and 2 | 120mg every 2 weeks from Week 6 |
| Infliximab* | IV | 5mg/kg at Weeks 0, 2 and 6 | 5mg/kg every 8 weeks |
| Vedolizumab | IV | 300mg at Weeks 0, 2 and 6 | 300mg every 8 weeks |
| Vedolizumab | Initiation: IV<br>Maintenance: SC | 300mg at Weeks 0 and 2 | 108mg at Week 6 and every other week thereafter |
| Golimumab | SC | 200mg at Week 0, 100mg at Week 2 | 50mg every 4 weeks thereafter |
| Tofacitinib | Oral | 10mg twice daily for 8 weeks | 5mg twice daily |
| Filgotinib | Oral | 200mg once daily | |
| Upadacitinib | Oral | 45mg once daily for 8 weeks | 15mg once daily |
| **Company clarification model** | | | |
| Ozanimod | Oral | Dose escalation from day 1 to day 7 (0.23mg once daily for days 1 to 4 then 0.46mg once daily for days 5 to 7) | 0.92mg once daily |
| Ustekinumab | IV and SC | Assume patient weight 56-85kgs; 390mg (IV) then 90mg after 8 weeks (SC) | 90mg every 12 weeks (SC) |

*Average weight of 78.5kgs was used to calculate required dose
IV=intravenous; SC=subcutaneous injection
Source: company model

Table 13 Drug acquisition costs

| Drug | Total cost | |
|---|---|---|
| | Year 1 | Subsequent years |
| **Company model** | | |
| Etrasimod (oral) | £11,000 [redacted] | £11,000 [redacted] |
| Adalimumab (SC) | £9,820.80 | £8,236.80 |
| Infliximab then Remsima (IV then SC) | £11,643.75 | £9,819.16 |
| Infliximab (IV only) | £11,830.26 | £8,872.70 |
| Vedolizumab (IV then SC) | £16,400.00 | £12,300.00 |
| Vedolizumab (IV only) | £16,400.00 | £13,325.00 |
| Golimumab (SC) | £11,826.04 | £9,918.61 |
| Tofacitinib (oral) | £10,350.45 | £8,970.39 |
| Filgotinib (oral) | £10,472.28 | £10,472.28 |
| Upadacitinib (oral) | £13,035.36 | £10,472.28 |
| **Company clarification model** | | |
| Ozanimod (oral) | [redacted] | [redacted] |
| Ustekinumab (IV and SC) | [redacted] | [redacted] |

Source: CS, Table 20 and company clarification model

Table 14 Drug administration costs

| Administration method | Cost | Reference |
|---|---|---|
| IV | £133.40 | Average of consultant led and non-consultant led, non-admitted face-to-face attendance, follow-up, WF01A[51] |
| SC | £0.00 | Assume patients self-administer and therefore there is no administration cost. Additionally, it has been assumed that the one off nurse training cost to teach patients how to self-administer the injection is covered by the manufacturer in line with previous TAs (TA856[9] and TA547[7]) |
| Oral | £0.00 | Assumed no administration cost |

IV=intravenous; SC=subcutaneous
Source: CS, Table 22

The concomitant medications included in the model are shown in

Table 15. Concomitant medication usage is assumed to stay constant over the model time horizon.

Table 15 Concomitant medications

| Drug | Total annual cost | Utilisation | |
|---|---|---|---|
| | | S1Ps[10] | All other treatments[5,6] |
| Balsalazide | £341.64 | 0% | 0% |
| Mesalazine | £201.66 | 13% | 13% |
| Olsalazine | £1,958.83 | 0% | 0% |
| Sulfasalazine | £87.86 | 0% | 0% |
| Prednisolone | £1.47 | 36% | 36% |
| Hydrocortisone | £40.03 | 0% | 0% |
| Azathioprine* | £9.52 | 0% | 0% or 39% |
| 6-mercaptopurine* | £502.09 | 0% | 15% |
| Methotrexate* | £15.77 | 0% | 9.0% |
| Budesonide | £126.10 | 1% | 1% |

S1P=sphingosine-1-phosphate
*Patients receiving etrasimod and 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
Source: CS, Table 23

### 4.3.2 Monitoring and pre-initiation costs

A single ECG is required prior to treatment with an S1P (etrasimod and ozanimod). The company has assumed that the cost of an ECG is £74.91 (EY51Z,[51] Directly Accessed Diagnostic Services, Electrocardiogram, Monitoring or Stress Testing).

Monitoring costs were assumed similar for etrasimod and existing treatments and were not included in the model.

## *4.4 Adverse events*

Company safety (serious infection) NMA results (etrasimod versus existing treatments) demonstrated that there were no statistically significant differences between treatments (CS, p70). Therefore, the company did not include AE-related costs in the analysis.

## *4.5 Company cost comparison results*

### 4.5.1 Base case results

The company base case results are presented in Table 16.

Table 16 Company cost comparison base case results

| Treatment | Total 5-year cost per patient | Current market share |
|---|---|---|
| Etrasimod | £55,215 ([redacted] | - |
| Adalimumab | £43,308 | [redacted] |
| Infliximab (IV then SC) | £52,527 | [redacted] |
| Infliximab (IV only) | £52,129 | |
| Vedolizumab (IV then SC) | £70,506 | [redacted] |
| Vedolizumab (IV only) | £70,408 | |
| Golimumab | £52,040 | [redacted] |
| Tofacitinib | £46,753 | [redacted] |
| Filgotinib | £52,901 | [redacted] |
| Upadacitinib | £55,464 | [redacted] |
| Ozanimod | £89,460 | [redacted] |
| Ustekinumab | £54,348 | [redacted] |

Source: company model and company clarification response, Table 2; CS, Table 24

### 4.5.2 Company cost comparison scenario results

The company carried out three scenario analyses; results are provided in Table 17.

Table 17 Company cost comparison scenario analysis results

| Technology | 2-year time horizon | 5-year time horizon Infliximab (IV only) and vedolizumab (IV only) for initiation and maintenance | Positive stopping rule of 10% at 12 months applied to all treatments |
|---|---|---|---|
| Etrasimod | £22,131 (list) [redacted] | £55,215 (list) [redacted] | £50,804 (list) [redacted] |
| Adalimumab | £18,273 | £43,308 | £39,970 |
| Infliximab (IV then SC) | £22,746 | - | £48,556 |
| Infliximab (IV only) | £22,786 | £52,129 | £48,217 |
| Vedolizumab (IV only) | £30,783 | £70,408 | £65,125 |
| Vedolizumab (IV then SC) | £30,208 | - | £65,133 |
| Golimumab | £21,960 | £52,040 | £48,029 |
| Tofacitinib | £19,529 | £46,753 | £43,123 |
| Filgotinib | £21,160 | £52,901 | £48,669 |
| Upadacitinib | £23,723 | £55,464 | £51,232 |
| Ozanimod | £35,829 | £89,460 | £82,309 |
| Ustekinumab | £26,113 | £54,348 | £50,583 |

IV=intravenous; PAS=Patient Access Scheme
Source: company model

## *4.6 EAG critique of company cost comparison analysis*

### 4.6.1 Company approach to cost comparison analysis

In the final scope issued by NICE, it is stated that etrasimod has been selected to be appraised as a cost comparison. To establish that the clinical efficacy and safety of etrasimod is similar to comparator treatments, the company carried out several NMAs. The company has focussed its discussion and presentation of the NMA results on the comparison of etrasimod versus adalimumab, infliximab and vedolizumab. For completeness (and because all treatments were included in the company cost comparison analysis), the EAG has commented on the full set of company NMA results.

Company **biologic-naïve** NMA results showed that:

- in the induction and maintenance phases, etrasimod is only statistically significantly superior to adalimumab (clinical response and clinical remission)
- in the induction phase, etrasimod is statistically significantly inferior to upadacitinib (clinical response and clinical remission)
- for all other comparisons, the difference between treatments is not statistically significant (clinical response and clinical remission).

Company **biologic-experienced** NMA results showed that:

- induction and maintenance phases, etrasimod is not statistically significantly superior to any of the drugs in the network (clinical response and clinical remission)
- for the comparison of etrasimod versus infliximab or golimumab, there was no relevant evidence available to include in the NMA

In the absence of non-inferiority or equivalence testing, the EAG considers that only statistically significant NMA results favouring etrasimod can provide conclusive evidence that etrasimod is likely to provide similar or greater health benefits versus comparator treatments.

The EAG considers that the results from previous NMAs[4-10] conducted as part of similar NICE appraisals are mixed and it is often difficult to determine whether the new intervention is likely to provide greater health benefits than comparator treatments. Previous NICE appraisals[4-10] of comparator drugs have all included cost utility analyses, except for mirikizumab which used a cost comparison approach.

### 4.6.2 Minor errors

The EAG identified and corrected the following minor errors in the company cost comparison model):

- the number of infliximab, adalimumab, golimumab, ustekinumab and vedolizumab maintenance doses in the first year were slightly overestimated

- the unit cost for simple delivery of chemotherapy (NHS Reference Cost SB12Z: £286.17) should have been applied for the IV administration cost

The EAG has identified the following three issues that may require further consideration: drug acquisition costs, duration of treatment and time horizon.

### 4.6.3 Drug acquisition costs

The EAG notes that some dosing regimens described in the CS were incorrect; however, the correct regimens were applied in the company cost comparison model.

The company analyses have been conducted using list prices for all other drugs; confidential discounted prices (PAS and CMU) are available for comparator drugs. Cost comparison results generated using all discounted prices are available in a confidential appendix.

### 4.6.4 Duration of treatment

**Subsequent treatments**

Clinical advice to the EAG is that for patients who do not relapse and have no tolerability issues, it may be reasonable to assume treatment on the same drug continues for 5 years; however, for patients who relapse, this assumption may not be appropriate as these patients will receive one or more subsequent treatments.

Subsequent treatment costs are not included in the company cost comparison analysis; implicitly, therefore, the company has assumed that first-line treatment does not influence choice of subsequent treatments. Clinical advice to the EAG is that choice of subsequent treatment will be influenced by prior treatment. Further, company NMA results suggest that the efficacy, and therefore (implicitly) treatment duration, of UC treatments may differ according to setting (biologic-naïve/biologic-experienced). For example, if a biologic-experienced patient had previously failed on a TNFi, a non-TNFi is likely to be considered; results from the company NMA and a published NMA[52] suggest JAKis could be one of the most effective treatment options in this setting.

Clinical advice to the EAG is that a significant proportion of patients fail first-line treatment and subsequent lines of treatment; therefore, it may be important to consider subsequent treatment costs.[52] The EAG acknowledges that the high number of available subsequent treatment options and lack of sequential efficacy data are likely to present challenges for modelling and therefore subsequent treatment costs remain an area of uncertainty.

**Treatment discontinuation due to benefit**

In line with NICE recommendations,[4,5] patients who have a complete response at 12 months may pause or withdraw from treatment. In response to clarification question B4, the company has attempted to address the uncertainty associated with the impact on cost effectiveness results of some patients discontinuing treatment due to benefit by presenting results from a scenario analysis in which 10% of patients stopped treatment at 12 months; information about the source of this proportion were not provided.

There is a lack of long-term data informing the proportions of patients who relapse or pause treatment (and when this happens); therefore, the extent to which patients pause and receive subsequent treatments is unknown. Clinical advice to the EAG is that NHS patients in complete remission with no tolerability issues rarely discontinue treatment at 12 months and are likely to continue longer-term treatment, especially if disease history is well established. The EAG therefore considers that if all treatments have equal efficacy and safety, assuming equivalent time on treatment is reasonable.

**Time horizon**

The annual costs of each treatment are the same from Year 2 onwards, the EAG considers that it may be more appropriate to use results from a 2-year, rather than a 5-year, time horizon to inform decision making (company scenario analysis [CS, Table 25]).

## 4.7 EAG cost comparison results

After implementing the minor corrections described in Section 4.6.2, the EAG's updated cost comparison results are presented in Table 18. The EAG corrections had a minimal impact on the company cost comparison results. Details of the EAG's minor corrections to the company model are presented in Appendix 6.4.

Table 18 EAG cost comparison base case results (etrasimod PAS price, list price all other drugs)

| Treatment | Total 5-year cost per patient | Total 2-year cost per patient | 5-year difference (etrasimod vs comparator) | 2-year difference (etrasimod vs comparator) |
|---|---|---|---|---|
| Etrasimod | [redacted] | [redacted] | | |
| Adalimumab | £42,991 | £17,957 | [redacted] | [redacted] |
| Filgotinib | £52,901 | £21,160 | [redacted] | [redacted] |
| Golimumab | £51,659 | £21,579 | [redacted] | [redacted] |
| Infliximab (IV then SC) | £53,754 | £23,972 | [redacted] | [redacted] |
| Infliximab (IV only) | £57,035 | £24,933 | [redacted] | [redacted] |
| Ozanimod | £89,460 | £35,829 | [redacted] | [redacted] |
| Tofacitinib | £46,753 | £19,529 | [redacted] | [redacted] |
| Upadacitinib | £55,464 | £23,723 | [redacted] | [redacted] |
| Ustekinumab | £53,070 | £24,835 | [redacted] | [redacted] |
| Vedolizumab (IV then SC) | £70,301 | £30,002 | [redacted] | [redacted] |
| Vedolizumab (IV only) | £75,314 | £32,930 | [redacted] | [redacted] |

IV=intravenous; SC=subcutaneous

## 4.8 Conclusions

Clinical advice to the EAG is that etrasimod, as an oral drug, is a valuable addition to the currently available basket of treatments for patients with moderately to severely active UC. In addition, clinical advice to the EAG is that current NICE recommended treatments for moderately to severely active UC are generally considered to have similar efficacy and safety and that choice of treatment depends on several factors, including patient preferences and cost.

In the final scope issued by NICE, it is stated that etrasimod has been selected to be appraised as a cost comparison. The company (via the NMAs) has shown that etrasimod is statistically significantly superior to adalimumab (biologic-naïve patients, induction and maintenance phases, clinical remission/clinical response) and statistically significantly inferior to upadacitinib (biologic-naïve patients, induction, clinical remission/clinical response). For all other comparisons, company NMA results did not show that etrasimod was statistically significantly superior/inferior to any of the other drugs listed in the final scope.

If the NICE AC considers that etrasimod and comparator drugs are similar and any differences in patient outcomes can be ignored, then the EAG considers that the company cost comparison analysis may produce a robust estimate of the likely cost savings for patients treated with etrasimod, provided the following assumptions are considered reasonable:

- subsequent treatment costs are likely to be similar irrespective of the first-line treatment received
- treatment costs for biologic-experienced patients are assumed to be the same as those for biologic-naïve patients

Due to an absence of treatment sequencing data, the EAG considers that a cost utility analysis may not reduce the uncertainty around comparative effectiveness, treatment duration and subsequent treatments.

# 5 REFERENCES


1. National Institute for Health and Care Excellence (NICE). Etrasimod for treating moderately to severely active ulcerative colitis ID5091. Final scope. 2023 [October 2023]; Available from: https://www.nice.org.uk/guidance/gid-ta10991/documents/final-scope.
2. National Institute for Health and Care Excellence. Ulcerative colitis: management [NG130]. 2019. Available at: https://www.nice.org.uk/guidance/ng130/resources/ulcerative-colitis-management-pdf-66141712632517 [cited 2023 30 May]; Available from: https://www.nice.org.uk/guidance/ng130/resources/ulcerative-colitis-management-pdf-66141712632517.
3. 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; 68:s1-s106.
4. National Institute for Health and Care Excellence. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy. TA329. 2015 [October 2023]; Available from: https://www.nice.org.uk/guidance/ta329.
5. National Institute for Health and Care Excellence (NICE). Vedolizumab for treating moderately to severely active ulcerative colitis. TA342. 2015 [October 2023]; Available from: https://www.nice.org.uk/guidance/ta342.
6. National Institute for Health and Care Excellence (NICE). Ustekinumab for treating moderately to severely active ulcerative colitis. TA633. 2020 [October 2023]; Available from: https://www.nice.org.uk/guidance/ta633.
7. National Institute for Health and Care Excellence (NICE). Tofacitinib for moderately to severely active ulcerative colitis. TA547. 2018 [October 2023]; Available from: https://www.nice.org.uk/guidance/ta547.
8. National Institute for Health and Care Excellence (NICE). Filgotinib for treating moderately to severely active ulcerative colitis. TA792. 2022 [October 2023]; Available from: https://www.nice.org.uk/guidance/ta792.
9. National Institute for Health and Care Excellence (NICE). Upadacitinib for treating moderately to severely active ulcerative colitis. TA856. 2023 [October 2023]; Available from: https://www.nice.org.uk/guidance/ta856.
10. National Institute for Health and Care Excellence (NICE). Ozanimod for treating moderately to severely active ulcerative colitis. TA828. 2022 [October 2023]; Available from: https://www.nice.org.uk/guidance/ta828.
11. Ordás I, Eckmann L, Talamini M, Baumgart DC, Sandborn WJ. Ulcerative colitis. Lancet. 2012; 380:1606-19.
12. Office for National Statistics. Estimates of the population for the UK, England, Wales, Scotland and Northern Ireland. 2021 [October 2023]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland.
13. Sandborn WJ, Vermeire S, Peyrin-Biroulet L, Dubinsky MC, Panes J, Yarur A, *et al.* Etrasimod as induction and maintenance therapy for ulcerative colitis (ELEVATE): two randomised, double-blind, placebo-controlled, phase 3 studies. The Lancet. 2023; 401:1159-71.
14. ClinicalTrials.gov. NCT03950232: An Extension study for treatment of moderately to severely active ulcerative colitis (ELEVATE UC OLE). 2023 [October 2023]; Available from: https://clinicaltrials.gov/study/NCT03950232.
15. 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; 158:550-61.

16. ClinicalTrials.gov. NCT02536404: Extension study of APD334-003 in patients with moderately to severely active ulcerative colitis. 2021 [October 2023]; Available from: https://clinicaltrials.gov/study/NCT02536404.
17. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. Process and methods [PMG9]. 2013 [cited October 2023]; Available from: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781.
18. Pfizer. Clinical Study Report APD334-302 – Report Body: A phase 3, randomized, double-blind, placebo-controlled, 12-week study to assess the efficacy and safety of etrasimod in subjects with moderately to severely active ulcerative colitis. Data on File. 2022
19. Pfizer. Clinical Study Report APD334-301 – Report Body: A phase 3, randomized, double-blind, placebo-controlled, 52-week study to assess the efficacy and safety of etrasimod in subjects with moderately to severely active ulcerative colitis. Data on File. 2022
20. Pfizer. IQVIA ALHP. Data on file. 2023.
21. ClinicalTrials.gov. ELEVATE UC 52 Clinical Study Protocol. 2020 [cited October 2023]; Available from: https://storage.googleapis.com/ctgov2-large-docs/88/NCT03945188/Prot_000.pdf.
22. ClinicalTrials.gov. ELEVATE UC 12 Clinical Study Protocol. 2021 [cited October 2023]; Available from: https://storage.googleapis.com/ctgov2-large-docs/69/NCT03996369/Prot_000.pdf.
23. Pfizer. Etrasimod 2 mg film-coated tablets SmPC. Data on file. 2022. 16 Aug 2023.
24. Feagan BG, Sandborn WJ, D'Haens G, Hanauer S, Wolf DC, Vermeire S*, et al.* Clinical remission demonstrated with oral ozanimod in the overall population and across multiple subgroups of patients with moderately to severely active ulcerative colitis in the TOUCHSTONE trial. J Crohns Colitis. 2019; 13:S464-S.
25. Sandborn WJ, Ghosh S, Panes J, Vranic I, Su C, Rousell S*, et al.* Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. NEJM. 2012; 367:616-24.
26. Sands BE, Hanauer SB, Colombel JF, Sandborn W, Schreiber S, Danese S*, et al.* Subcutaneous infliximab (CT-P13 SC) as maintenance therapy for ulcerative colitis: A Phase 3, randomized, placebo-controlled study: Results of the LIBERTY-UC study. J Crohns Colitis. 2023; 17(Supplement 1):i623-i4.
27. 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; 60:780-7.
28. Rubin DTD, I. DuVall, A. Bouhnik, Y. Radford-Smith, G. Higgins, P. D. R. Mishkin, D. S. Arrisi, P. Scalori, A. Oh, Y. S. Tole, S. Chai, A. Chamberlain-James, K. Lacey, S. McBride, J. Panes, J. Hibiscus Study Group. Etrolizumab versus adalimumab or placebo as induction therapy for moderately to severely active ulcerative colitis (HIBISCUS): two phase 3 randomised, controlled trials. Lancet Gastroenterol Hepatol. 2022; 7:17-27.
29. 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; 142:257-65 e1-3.
30. 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. J Gastroenterol. 2014; 49:283-94.
31. Sands BE, Peyrin-Biroulet L, Loftus EV, Danese S, Colombel JF, Törüner M*, et al.* Vedolizumab versus adalimumab for moderate-to-severe ulcerative colitis. N Engl J Med. 2019; 381(13): Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01990858/full.
32. Feagan BG, Danese S, Loftus EV, Jr., Vermeire S, Schreiber S, Ritter T*, et al.* Filgotinib as induction and maintenance therapy for ulcerative colitis (SELECTION): a

phase 2b/3 double-blind, randomised, placebo-controlled trial. The Lancet. 2021; 397:2372-84.
33. 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). J Gastroenterol. 2017; 52:1101-11.
34. 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; 146:85-95; quiz e14-5.
35. 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; 146:96-109 e1.
36. Jiang XL, Cui HF, Gao J, Fan H. Low-dose infliximab for induction and maintenance treatment in Chinese patients with moderate to severe active ulcerative colitis. J Clin Gastroenterol. 2015; 49:582-8.
37. 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. J Gastroenterol. 2016; 51:241-51.
38. NCT01551290. A phase 3, multicenter, randomized, double-blind, placebo-controlled study evaluating the efficacy and safety of infliximab in Chinese subjects with active ulcerative colitis. . 2014 [October 2023]; Available from: https://classic.clinicaltrials.gov/ct2/show/NCT01551290.
39. 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; 353:2462-76.
40. Sandborn WJ, Feagan BG, D'Haens G, Wolf DC, Jovanovic I, Hanauer SB*, et al.* Ozanimod as induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2021; 385:1280-91.
41. Sandborn WJ, Su C, Sands BE, D'Haens GR, Vermeire S, Schreiber S*, et al.* Tofacitinib as induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2017; 376:1723-36.
42. Danese S, Vermeire S, Zhou W, Pangan AL, Siffledeen J, Greenbloom S*, et al.* Upadacitinib as induction and maintenance therapy for moderately to severely active ulcerative colitis: results from three phase 3, multicentre, double-blind, randomised trials. The Lancet. 2022; 399:2113-28.
43. Sandborn WJ, Ghosh S, Panes J, Schreiber S, D'Haens G, Tanida S*, et al.* Efficacy of upadacitinib in a randomized trial of patients with active ulcerative colitis. Gastroenterology. 2020; 158:2139-49.e14.
44. 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; 381:1201-14.
45. Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel J-F, Sandborn WJ*, et al.* Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2013; 369:699-710.
46. 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:e0212989.
47. 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; 158:562-72.e12.
48. Feagan BG, Rubin DT, Danese S, Vermeire S, Abhyankar B, Sankoh S*, et al.* Efficacy of vedolizumab induction and maintenance therapy in patients with

ulcerative colitis, regardless of prior exposure to tumor necrosis factor antagonists. Clin Gastroenterol Hepatol. 2017; 15:229-39. e5.
49. Danese S, Vermeire S, Zhou W, Pangan A, Siffledeen J, Hebuterne X*, et al.* OP24 efficacy and safety of upadacitinib induction therapy in patients with moderately to severely active ulcerative colitis: results from the phase 3 U-ACHIEVE study. J Crohns Colitis. 2021; 15:S022-S4.
50. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades A. NICE DSU Technical Support Document 4: Inconsistency in networks of evidence based on randomised controlled trials. 2011.
51. 2021/22 National schedule of NHS costs [database on the Internet]2022. Accessed 14 Aug 2023. [Internet] Available at: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/. Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/.
52. Panaccione R, Collins EB, Melmed GY, Vermeire S, Danese S, Higgins PDR*, et al.* Efficacy and safety of advanced therapies for moderately to severely active ulcerative colitis at induction and maintenance: an indirect treatment comparison using Bayesian network meta-analysis. Crohn's & Colitis 360. 2023; 5.

# 6 APPENDICES

## 6.1 Appendix 1: EAG summary and critique of the company's methodological approach in the ELEVATE trials

Table 19 EAG assessment of statistical approaches used in the ELEVATE UC 12 and ELEVATE UC 52 trials

| Item | EAG assessment | Statistical approach with EAG comments |
|---|---|---|
| Were all analysis populations clearly defined and pre-specified? | Yes | All key efficacy outcomes are reported for the primary efficacy analysis set (PEAS) in the ELEVATE UC trials. The PEAS population includes only patients with a baseline modified Mayo score (MMS) of 5 to 9 who received at least one dose of study drug or placebo. The safety analysis set was defined as all randomised patients who received ≥1 dose of the study drug or placebo (CS, p35)<br><br>The EAG notes that the inclusion criteria of the ELEVATE UC trials allowed for the recruitment of patients with a MMS of 4. The EAG is satisfied that the PEAS populations was clearly defined and pre-specified in the TSAP for each of the ELEVATE UC trials (TSAP Table 1). |
| Was an appropriate sample size calculation pre-specified? | Yes | A trial sample size calculation was pre-specified in the TSAP for ELEVATE UC 12 (p19). For the primary endpoint analysis of clinical remission at Week 12, the company estimated that a sample size of 330 patients (220 etrasimod, 110 placebo) was required to achieve at least 90% power to detect a difference of 12.5% between the etrasimod treatment group (18.5%) and the placebo treatment group (6.0%).<br><br>A trial sample size calculation was pre-specified in the TSAP for ELEVATE UC 52 (p23). For the primary endpoint analysis of clinical remission, the company estimated that a sample size of 420 patients (280 etrasimod, 140 placebo) was required to achieve 93.4% power to detect a difference of 13.5% at Week 52 between the etrasimod treatment group (23.5%) and the placebo treatment group (10.0%). With this sample size, there was 96% power to detect a difference of 12.5% in the other primary endpoint of clinical remission at Week 12, assuming a placebo rate at 6.0%. Since the two primary endpoints were expected to be at least moderately positively correlated, the actual overall power to reject both of their null hypotheses was likely >90%.<br><br>The EAG is satisfied that the sample size calculations were appropriate. |
| Were all changes in the conduct of the trial or planned analysis made prior to analysis? | Yes | Changes in the conduct of the trial are listed in the CSR for each of the ELEVATE UC trials (Table 2). |

| Item | EAG assessment | Statistical approach with EAG comments |
|---|---|---|
| Were all primary and secondary efficacy outcomes pre-defined and analysed appropriately? | Yes | The primary and secondary efficacy endpoints are listed in the CSR for the primary data analysis of the ELEVATE UC trials (CSR, Table 1). Definitions and analysis approaches for these endpoints were pre-specified in the TSAPs for each of the ELEVATE UC trials (Section 16).<br><br>The ELEVATE UC trials used a gatekeeping procedure to account for multiple testing of hypotheses. The procedure was pre-specified in the TSAPs (Figure 2).<br><br>See text in Section 3.2.4 of this EAG report for further discussion of the analysis approach for the primary and secondary efficacy endpoints |
| Was the analysis approach for PROs appropriate and pre-specified? | Yes | All PROs in the ELEVATE UC trials were listed as supportive efficacy outcomes (TSAP, Section 17).<br>The EAG considers that the analysis approach for the PROs was prespecified and appropriate. |
| Was the analysis approach for AEs appropriate and pre-specified? | Yes | Safety data presented in the CS for the ELEVATE UC trials included a summary of TEAEs, SAEs, AEs leading to treatment discontinuation, or interruption and AEs of special interest (CS, Section B.3.10).<br>Safety analyses were descriptive only and were pre-specified in the TSAP for each of the ELEVATE UC trials (Section 17). |
| Was a suitable approach employed for handling missing data? | Yes | The company's approach to handling missing data is outlined in the TSAP for each of the ELEVATE UC trials (Section 16.1.2).<br>The EAG is satisfied that the approach described was appropriate. |
| Were all subgroup and sensitivity analyses pre-specified? | Yes | Primary and key secondary efficacy outcomes, including clinical remission, symptomatic remission, endoscopic improvement-histologic remission, clinical response, at week 12 and week 52 were analysed according to the key pre-specified subgroups in each of the ELEVATE UC studies:<br>•Naïve to biologic or JAK inhibitor therapy at study entry (yes or no)<br>•Baseline corticosteroid use (yes or no)<br>•Baseline disease activity (MMS: 4 to 6 or 7 to 9)<br>For ELEVATE UC 52, subgroup analyses on sustained clinical remission and steroid-free clinical remission were also conducted.<br>The EAG is satisfied that the subgroup analyses presented in the CS were prespecified in the TSAP for each of the ELEVATE UC trials (Section 8.5) |

AE=adverse event; CSR=clinical study report; JAK=Janus kinase; MMS=modified Mayo score; PEAS=primary efficacy analysis set; PROs=patient-reported outcomes; SAE=serious adverse event; TEAE=treatment-emergent adverse event; TSAP=trial statistical analysis plan
Source: CS, CSR, TSAP

## 6.2 Appendix 2: Company and EAG quality assessment of trials included in the company NMAs

Table 20 Company and EAG quality assessment of trials included in the company NMAs

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| ELEVATE UC 12[13] | Yes, central randomisation using IWRS | Yes, central randomisation using IWRS | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, treatment discontinuation was approximately similar in both groups (ETR: 10.5%, PBO 11.2%) | No, outcomes were reported as per the protocol | No, full analysis set (FAS) |
| EAG comment | | | Yes (except for prior treatment with 5-ASA) | | | | The company analysed the results from the primary efficacy analysis set. This was appropriate |
| ELEVATE UC 52[13] | Yes, central randomisation using IWRS | Yes, central randomisation using IWRS | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, treatment discontinuation was not similar in both groups (ETR: 44.3%, PBO: 68.05%) | No, outcomes were reported as per the protocol | No, full analysis set (FAS) |
| EAG comment | | | Yes (except for duration of UC, which is longer in the ETR than PBO arm {7.5 years vs 5.9 years) | | | | The company analysed the results from the primary efficacy analysis set. This was appropriate |
| OASIS[15] | Yes, randomisation was performed | Yes, study drug were supplied as capsules with the | Yes, baseline characteristics were balanced between | Yes, double blind | No, treatment discontinuation was approximately | No, outcomes were reported as per the | Yes, ITT population |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| | centrally with a block size of 6 | same appearance | treatment arms | | similar in all groups (ETR 1mg: 9.6%, ETR 2mg: 8%, PBO: 11.11%) | protocol | |
| EAG comment | | Yes (randomisation codes were generated by a statistician not directly involved with the study) | Yes (except duration of UC which is longer in the PBO arm than the ETR arm [8.6 years vs 6.2 years]) | | | | |
| TRUE NORTH[40] | Yes, IVRS/IWRS | Yes, patients were assigned to treatment/randomised using the IVRS/IWRS | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, drop-out with the PBO arm having twice as many drop-outs as OZA in the induction (11% vs 6%) and maintenance (45% vs 20%) period | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | | | Yes (pre-specified "other efficacy endpoints" including change in Mayo score from baseline to Week 10 were not reported) | |
| U-ACHIEVE[43,49] | Yes, IWRS; block randomisation schedules (block size of 3) | Yes, IWRS | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, PBO had a twice higher dropout rate (12%) than UPA (4%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG | | | | | | Yes (pre-specified | |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| comment | | | | | | additional outcomes including PROs were not reported) | |
| U-ACCOMPLISH[49] | Yes, IWRS; block randomisation schedules (block size of 3) | Yes, IWRS | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, PBO had a twice higher dropout rate (65%) than UPA 15mg (33%) and UPA 30mg (21%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | | | Yes (pre-specified additional outcomes including PROs were not reported) | |
| SELECTION[32] | Yes, central randomisation using IWRS | Yes, IWRS | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, higher discontinuation rates in the PBO (6.5%) compared to FIL 100mg (6%) and FIL 200mg (3%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | | Yes (discontinuation rates appear low in all treatment arms) | | Partly (the trial definition for FAS was consistent with an ITT population for the induction phase but not the maintenance phase) |
| OCTAVE Induction 1[41] | Yes, central randomisation using TRS | Yes, central randomisation | Yes, baseline characteristics were balanced between | Yes, double blind | No, slightly lower proportion of patients | No, outcomes were reported as per the protocol | Yes, ITT population |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| | | | treatment arms | | discontinued PBO (3%) than TOF 10mg (7%) | | |
| EAG comment | | | | | | | |
| OCTAVE Induction 2[41] | Yes, central randomisation using TRS | Yes, central randomisation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, slightly higher proportion of patients discontinued PBO (13%) than TOF 10mg (8%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | Yes (except for a higher percentage of males in the TOF arm compared with PBO [60.4% vs 49.1%] ) | | | | |
| OCTAVE Sustain[41] | Yes, central randomisation using TRS | Yes, central randomisation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, higher discontinuation rates in PBO (73%) compared to TOF 5mg (44%) and TOF 10mg (36%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | Yes (except for never smoker status which was greater in the TOF 5mg arm than in the PBO arm [71.7% vs 57.1%]) | Unclear (no mention of who was blinded to treatment) | | | |
| UNIFI[44] | Yes, | Yes, permuted | Yes, baseline | Yes, double blind | No, higher drop-out | No, outcomes were | Yes, ITT population |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| | randomisation was performed with the use of permuted blocks | blocks | characteristics were balanced between treatment arms | | was observed in PBO than the intervention (UST 6mg/kg 4%, UST 130mg 4%, PBO 5%) | reported as per the protocol | |
| EAG comment | | | | Unclear (no mention of who was blinded to treatment) | | | |
| GEMINI 1[45] | Yes, randomisation was performed centrally with the use of computer-generated randomisation schedules | Yes, NR | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, higher proportion of PBO discontinued treatment compared to VED in ind. phase (9% vs 2%) and maintenance phase (PBO 62%, VED Q8W 37%, VED Q4W 33%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | | | | |
| VISIBLE 1[47] | Yes, IWRS | Unclear, No information | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, treatment discontinuation was not similar across the group PBO 64.2%, VED SC 29.2%, VED IV 27.7% | No, outcomes were reported as per the protocol | No, full analysis set (FAS) |
| EAG comment | | Yes (IWRS) | | | | | Yes (the trial definition for FAS |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| | | | | | | | was consistent with an ITT population) |
| VARSITY[31] | Yes, IVRS/IWRS | Yes, investigational pharmacist or designee will mask the IV bags after preparation in order to maintain the study blind | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, treatment discontinuation was not similar across the group ADA: 43.7% VED: 29.8% | No, outcomes were reported as per the protocol | No, full analysis set (FAS) |
| EAG comment | | Yes (IWRS) | | | | | Yes (the trial definition for FAS was consistent with an ITT population) |
| Motoya 2019[46] | Yes, randomisation schedules were generated by sponsor-designated personnel (dynamic randomisation was performed with the previous TNFα antagonist use) | Yes, NR | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, higher proportion of drop-outs in the PBO arm compared to VED in the induction (5% vs 5%) and maintenance (57% vs 27%) period | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | | Partly (a higher proportion of patients in the placebo arm discontinued treatment during the | | |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| | | | | | maintenance period than in the vedolizumab arm) | | |
| ULTRA 1[27] | Yes, randomisation done by central randomisation scheme generated by the study sponsor | Yes, central randomisation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, drop-out between treatment group are almost similar (PBO 7%, ADA 160/80/40mg 7%, ADA 80/40mg 9%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | | | | |
| ULTRA 2[29] | Yes, randomisation was performed centrally | Yes, central randomisation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, unexpected imbalance in the dropout between two treatment group (PBO 48%, ADA 37%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | | No, discontinuations appear similar between groups | | |
| Suzuki 2014[30] | Yes, randomised based on centrally designed randomisation table | Yes, central randomisation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, unexpected imbalance in the dropout between two treatment group (PBO 23%, ADA 33%) | No, outcomes were reported as per the protocol | No, full analysis set (FAS) |
| EAG comment | | Unclear (the EAG has no information on how | Yes (except sex) | Unclear (no mention of who was blinded to | | Unclear (unable to access protocol) | Yes (the trial definition for FAS was consistent with |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| | | randomisation table was accessed) | | treatment) | | | an ITT population) |
| HIBISCUS I[28] | Yes, permuted block randomisation using IVRS/IWRS | Yes, permuted blocks | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, unexpected imbalance in the dropout at induction and maintenance phases | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | | Partly (treatment discontinuation rate was high [>75%] in all treatment arms but was highest in the placebo arm) | | Yes (modified ITT population) defined as all randomly assigned patients who received at least one dose of study drug |
| HIBISCUS I[28] | Yes, permuted block randomisation using IVRS/IWRS | Yes, permuted blocks | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, unexpected imbalance in the dropout at induction and maintenance phases | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | | Partly (treatment discontinuation rate was high [>75%] in all treatment arms but was highest in the placebo arm) | | Yes (modified ITT population) |
| PURSUIT-SC[34] | Yes, central randomisation using IVRS | Yes, central randomisation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, 2.3% of patients withdrew from each study arm | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | Unclear (no mention of who was | | | |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| | | | | blinded to treatment) | | | |
| PURSUIT-M[35] | Yes, ARP | Yes, ARP | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, slightly higher in GOL 100mg (11%) and PBO (15%) than GOL 50mg (10%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | Unclear (no mention of who was blinded to treatment) | No (discontinuation rates were similar between treatment arms) | | |
| PURSUIT-J[33] | Yes, a computer-generated randomisation (PBR) | Yes, computer-generated randomisation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | No, PBO had a twice higher dropout rate (39%) than GOL (16%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | Unclear (no mention of who was blinded to treatment) | | | |
| NCT01551290[38] | Yes, NR | Unclear, no information | Unclear, no information | Yes, double blind | Unclear, no information | Unclear, no information | Yes, ITT population |
| EAG comment | Unclear (randomisation method not given) | Unclear (randomisation method not given) | | Unclear (no mention of who was blinded to treatment) | | | Partly (ITT but no mention of methods for missing data handling) |
| Jiang 2015[36] | Yes, central randomisation | Yes, central randomisation with a dynamic treatment allocation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, more than twice as many patients in the PBO group as in the other 2 groups | No, outcomes were reported as per the protocol | Yes, ITT population |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| | | | | | prematurely discontinued the infusions | | |
| EAG comment | | | | Unclear (no mention of who was blinded to treatment) | | | Partly (ITT but no mention of methods for missing data handling) |
| Kobayashi 2016[37] | Yes, randomisation was performed centrally with the use of CGRS | Yes, central randomisation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Unclear, no information | No, outcomes were reported as per the protocol | No, full analysis set (FAS) |
| EAG comment | | | | Unclear (no mention of who was blinded to treatment) | | | Partly (FAS but no definition provided and no mention of methods for missing data handling) |
| ACT-1[39] | Yes, central randomisation | Yes, central randomisation with a dynamic treatment allocation | Yes, baseline characteristics were balanced between treatment arms | Yes, double blind | Yes, higher proportion of PBO (47%) discontinued treatment compared to INF (INF 5mg 32% and INF 10mg 32%) | No, outcomes were reported as per the protocol | Yes, ITT population |
| EAG comment | | | | Unclear (no mention of who was blinded to treatment) | | | Partly (ITT but no mention of methods for missing data handling) |
| ACT-2[39] | Yes, central randomisation | Yes, central randomisation with a dynamic | Yes, baseline characteristics were balanced between | Yes, double blind | Yes, higher proportion of PBO (40%) discontinued | No, outcomes were reported as per the protocol | Yes, ITT population |

| Study[a] | Was the randomisation method adequate? | Was allocation adequately concealed? | Were baseline characteristics similar between treatment arms? | Were participants and investigators blind to exposure & comparison? | Were discontinuations dissimilar between groups? | Were any outcomes measured but not reported? | Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data? |
|---|---|---|---|---|---|---|---|
| | | treatment allocation | treatment arms | | treatment compared to INF (INF 5mg 20% and INF 10mg 20%) | | |
| EAG comment | | | | Unclear (no mention of who was blinded to treatment) | | | Partly (ITT but no mention of methods for handling missing data) |

[a] An EAG comment is provided where either the EAG assessment differs from the company assessment or where extra information was required
ADA=adalimumab; AE=adverse event; ARP=adaptive randomisation procedure; CGRS=computer generated randomisation schedule; discontinuation=discontinuation; ETR=etrasimod; FIL=filgotinib; GOL=golimumab; INF=infliximab; ITT=intention to treat; IVRS=interactive voice response system; IWRS=interactive web response system; OZA=ozanimod; PBO=placebo; PBR=permuted block randomisation; PRO=patient reported outcome; TOF=tofacitinib; TRS=tele randomisation system; UPA=upadacitinib; UST=ustekinumab; VED=vedolizumab
Source: CS, Appendix F, Table 35 with EAG comment

## 6.3 Appendix 3: EAG summary and critique of the company's methodological approach for NMAs

Table 21 EAG summary and critique of the company's methodological approach for NMAs

| Item | EAG assessment | EAG comment |
|---|---|---|
| Were appropriate outcomes synthesised in NMAs? | Yes | The company conducted NMAs for the following outcomes:<br>• clinical response<br>• clinical remission<br>• serious infections<br>For the NMAs of clinical response and clinical remission, the company preferentially extracted data for centrally read endoscopic outcomes; if these were not available, the company extracted data for locally read endoscopic outcomes. Clinical response and clinical remission could be defined using either the full or modified Mayo score. Clinical advice to the EAG is that including trials in the NMAs reporting either full or adapted Mayo Score is not of concern. |
| Was an appropriate model used to conduct the NMAs? | Yes | The NMAs were conducted under a Bayesian framework using MCMC sampling. All analyses were implemented in WinBUGS version 1.4.3 statistical software with non-informative priors. An initial burn-in of at least 20,000 simulations was used, and convergence was confirmed through visual inspection of the Brook-Gelman-Rubin diagnostic and history plots. This was followed by 50,000 simulations on 3 chains, thinned by a factor of 10, to estimate the sampled parameters. Convergence was assessed by visual inspection of the history, kernel density and autocorrelation plots as well as the Brooks Gelman-Rubin diagnostic plot.<br><br>For clinical response and clinical remission, the company synthesised data using a multinomial model with probit link. It was assumed that the numbers of patients who were reported in the trial publications as being in clinical response also included those patients who were in clinical remission. Trials could be included in the analysis if they provided data for only one of these outcomes (i.e., clinical response, or clinical remission).<br><br>The EAG notes that the company provided treatment effects on the probit scale, which are difficult to interpret. Using a logit link would have overcome this problem. However, the EAG notes that the use of the probit link was pre-specified in the NMA SAP (p17). Furthermore, in addition to the treatment effects expressed on the probit scale, the company also provides risk ratios, and SUCRA values, which are comparatively easy to interpret.<br><br>For the proportion of patients experiencing serious infections, the company synthesised data using a binomial model with logit link. |
| Were the methods of selection between fixed-effects and random-effects models appropriate? | Yes | The company selected whether to use fixed-effects or random-effects based on a combination of statistical and clinical considerations. The company considered whether each network of evidence consisted primarily of single-trial connections, as in this scenario, fixed-effects models may be more suitable than random-effects models, due to a lack of information available to estimate between trial heterogeneity. The company also examined DIC and residual deviance values. The EAG considers that the company's methods to select between fixed-effects and random-effects models were appropriate. |
| Were any additional analyses pre-specified and conducted appropriately? | Yes | Fixed-effects and random-effects models were fitted and run using both an unadjusted relative effects analysis, as well as incorporating a meta-regression adjustment to account for variation in baseline risk. Both analyses were pre-specified in the NMA SAP (p18). However, the analyses including an adjustment to account for cross-trial variation in |

| | | |
|---|---|---|
| | | baseline risk failed to converge, and results from these analyses are not presented in the CS. |
| Were appropriate methods used to assess inconsistency? | Partial | A UME model was run to assess and identify any sources of inconsistency among the analyses. Three chains were run for the UME model. For each analysis, the posterior median of total residual deviance and the DIC were recorded and a deviance contribution plot comparing the NMA model with the UME model was produced. The company considered differences of more than 5 (in either the DIC or residual deviance values) between models to be potentially meaningful differences that should be investigated further by examining the deviance contribution plot. The EAG considers that a comparison of the estimated treatment effects from the UME model with the estimated treatment effects from the NMA model would have been a useful addition to the assessment of inconsistency. |

DIC=deviance information criterion; MCMC=Markov Chain Monte Carlo; NMA=network meta-analysis; SAP=statistical analysis plan; UME=unrelated mean effects; SUCRA=surface under cumulative ranking curve

## 6.4 Appendix 4: EAG revisions to the company model

This appendix contains details of the changes that the EAG made to the company model.

Table 22 EAG revisions to the cost comparison model

| EAG revisions | Implementation instructions |
|---|---|
| Corrections to first year maintenance doses | **Insert sheet "EAG Revisions"**<br><br>Set value in cell C3 = "C1"<br>Set value in cell D3 = 1<br><br>In Sheet 'Cost_Drug'<br><br>Set value in cell O14 =(WeeksInYear-IF('EAG Revisions'!D3=1,4,2))/2<br><br>Set value in cell O21 =WeeksInYear-IF('EAG Revisions'!D3=1,6,2)<br>Set value in cell P21 =N21*IF('EAG Revisions'!D3=1,ROUNDUP(O21/4,0),O21/4)<br>Set value in cell R21 =IF('EAG Revisions'!D3=1,N21,O21)*WeeksInYear/4<br><br>Set value in cell O40 =(WeeksInYear-IF('EAG Revisions'!D$3=1,20,8))/12<br>Set value in cell S39 =J39*3+J40+J40*IF('EAG Revisions'!D3=1,ROUNDUP(O40,0),O40)<br><br>Set value in cell S43 =2*K41*L41+SUM(IF('EAG Revisions'!D3=1,0,M43),P43)*L43 |
| Correction to IV administration cost | In Sheet 'EAG Revisions'<br>Set value in cell C4 = "C2"<br>Set value in cell D4 = 1<br><br>In Sheet 'Cost_Drug'<br><br>Set value in cell E51 =IF('EAG Revisions'!D4=1,286.71,"")<br><br>Set value in cell H55 =F55*IF('EAG Revisions'!$D$3,$D$51, $F$51)<br><br>Copy formula in cell H55<br>Paste to range H55:I60 |