EUnetHTA Joint Action 3 WP4

Relative effectiveness assessment of pharmaceutical technologies

USTEKINUMAB FOR THE TREATMENT OF ADULT PATIENTS WITH MODERATELY TO SEVERELY ACTIVE ULCERATIVE COLITIS (UC) WHO HAVE HAD AN INADEQUATE RESPONSE WITH, LOST RESPONSE TO OR WERE INTOLERANT TO EITHER CONVENTIONAL THERAPY OR A BIOLOGIC OR HAVE MEDICAL CONTRAINDICATIONS TO SUCH THERAPIES

Project ID: PTJA07

Version 1.0, 22/10/2019

DOCUMENT HISTORY AND CONTRIBUTORS

Version	Date	Description
V0.1	**04/09/19**	First draft
V0.2	**04/10/19**	Input from dedicated reviewers has been processed.
V0.3	**17/10/19**	Input from manufacturer and medical editor has been processed.
V1.0	**22/10/19**	Final version

Disclaimer

The assessment represents a consolidated view of the EUnetHTA assessment team members and is in no case the official opinion of the participating institutions or individuals.

EUnetHTA Joint Action 3 is supported by a grant from the European Commission. The sole responsibility for the content of this document lies with the authors and neither the European Commission nor EUnetHTA are responsible for any use that may be made of the information contained therein.

Assessment team

Author(s)	Ministry of Health of the Republic of Croatia (MIZ Croatia), Croatia
Co-Author(s)	Agency for Health Technology Assessment and Tariff System (AOTMiT), Poland Dental and Pharmaceutical Benefits Agency (TLV), Sweden
Dedicated Reviewer(s)	SNHTA, Switzerland UCSC Gemelli NIPN, Hungary SESCS and FIISC, Spain
Observer	NCPHA, Bulgaria

Further contributors

External experts	
Dr. V. Pittet	Answer specific question during the assessment
Manufacturer(s) [v0.2]	
Janssen-Cilag International NV	Preparation of the submission dossier Factual accuracy check
Medical editor [v0.2]	
CompuScript	Medical editing of the 2nd draft assessment
Patient(s) / patient organisation(s) / citizens	
Pacienti IBD z.s European Federation of Crohn's and Ulcerative Colitis Associations (EFCCA)	Provided input in response to the Open Call for Patient Input published on February 13th, 2019

Conflict of interest

All authors and dedicated reviewers involved in the production of this assessment have declared they have no conflicts of interest in relation to the technology assessed according to the EUnetHTA Declaration of Interest and Confidentiality Undertaking form.

How to cite this assessment

Please cite this assessment as follows:

Ministry of Health of the Republic of Croatia (MIZ Croatia), Agency for Health Technology Assessment and Tariff System (AOTMiT) and Dental and Pharmaceutical Benefits Agency (TLV). Relative effectiveness assessment of pharmaceutical technologies. Ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (UC) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a biologic, or have medical contraindications to such therapies. EUnetHTA Project ID: PTJA07. 2019.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

LIST OF ABBREVIATIONS

AE	Adverse event
AEoSI	Adverse event of special interest
ASA	Acetylsalicylic acid
ATC	Anatomical Therapeutic Chemical [classification system]
ATMP	Advanced therapy medicinal product
BF	Biologic failure
BID	Twice daily
CD	Crohn's disease
CI	Confidence interval
CrI	Credible interval
CRP	C-reactive protein
CSR	Clinical study report
CTR	Clinical trial register
DIC	Deviance information criterion
ECCO	European Crohn's and Colitis Organisation
EL	Evidence level
EMA	European Medicines Agency
EOW	Every other week
EQ-5D	EuroQol Questionnaire 5 Dimensions
EQ-VAS	EuroQol Visual Analogue Scale
GOL	Golimumab
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
HRQoL	Health-related quality of life
HTAi	Health Technology Assessment International
IBDQ	Inflammatory Bowel Disease Questionnaire
IFX	Infliximab
IL	Interleukin
IPD	Individual patient data
ITT	Intention to treat
IV	Intravenous
MAH	Marketing authorisation holder
MCS	Mental component summary
MD	Mean difference
MeSH	Medical Subject Headings
MTX	Methotrexate
NA	Not applicable
NBF	Nonbiologic failure
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
PBO	Placebo
PCS	Physical component summary
PsA	Psoriatic arthritis
PsO	Psoriasis

OR	Odds ratio
qXw	Every X weeks
QoL	Quality of life
RCT	Randomised controlled trial
RR	Relative risk
RWE	Real-world evidence
SAE	Serious adverse event
SC	Subcutaneous
SD	Standard deviation
SF-36	36-item Short Form survey
SLR	Systematic literature review
SmPC	Summary of product characteristics
SR	Systematic review
SUCRA	Surface under the cumulative ranking
TNF	Tumour necrosis factor
TNFi	Tumour necrosis factor inhibitor
TOC	Tofacitinib
UC	Ulcerative colitis
UST	Ustekinumab
VDZ	Vedolizumab

EXECUTIVE SUMMARY OF THE ASSESSMENT OF USTEKINUMAB

Introduction

Ulcerative colitis

Ulcerative colitis (UC) is a lifelong, chronic, idiopathic inflammatory bowel disease with a remitting and relapsing course that can progress from asymptomatic mild inflammation to extensive inflammation of the colon. Mucosal inflammation can extend from the rectum to the more proximal colon, with variable extents. A patient may have proctitis, left-sided colitis (the proximal limit being below the splenic flexure), extensive colitis (involving the transverse colon) or pancolitis. The majority (80%) of adult patients newly diagnosed with UC present with disease limited to the distal or left side of the colon. Approximately 20% of UC patients present with extensive colitis or pancolitis.

The symptoms of UC are diverse, depending on the extent of the disease, and can be severe, with a profound impact on patients' lives. Signs and symptoms may include frequent bloody stools, colonic motility dysfunction, potentially permanent fibrosis, tissue damage and systemic symptoms such as tiredness, weight loss and fever. In general, greater severity and extent of UC are associated with worsening bloody diarrhoea and the development of systemic signs. In patients with proctitis, urgency and tenesmus might be more prominent. Patients with pancolitis may experience more frequently bloody diarrhoea and abdominal pain. Up to 10% of patients with proctitis or left-sided colitis can suffer from paradoxical constipation. In extensive or severe disease, more general symptoms such as weight loss and fatigue, which can be accompanied by nausea, vomiting and fever, are often present. Nearly 70% of patients experience UC symptom flares every few months. Three-quarters of patients (75.6%) have reported that symptoms limit their ability to enjoy leisure activities, while 68.9% have admitted that UC symptoms affect their ability to perform at work. Progression of disease extent is associated with poor prognosis.

UC affects approximately the same number of women and men. The peak age for UC onset is between 30 and 40 years, but UC can occur at any age. The highest annual UC incidence has been reported as 24.3 per 100,000 person-years in Europe, 6.3 per 100,000 person-years in Asia and the Middle East and 19.2 per 100,000 person-years in North America.

Standardised mortality rates among patients with UC do not appear to be elevated overall compared with the general population. However, elderly patients and those with more comorbidities have a higher risk of mortality. Perioperative and postoperative mortality are also associated with a higher total mortality risk among patients with UC. Furthermore, it has been demonstrated that mortality among UC patients correlates with infection-related hospitalisations and increasing extent of colitis, and patients with UC are more likely to have a disability preventing them from working.

As in many chronic diseases, an appropriate management plan must be tailored to the individual patient when treating UC. The choice of treatment depends on disease activity status, distribution (proctitis, left-sided or extensive colitis), previous medications and the course of the disease. The ultimate goal of UC treatment is to induce and then maintain remission and to reduce the risk of long-term complications and surgery. Medications used to treat UC include: aminosalicylates (5-ASAs), corticosteroids, immune modifiers (immunomodulators), biologics, and over-the-counter medications such as antidiarrhoeal agents and analgesics. Despite better treatment options, long-term colectomy rates have not declined over a 10-year period, highlighting the urgent need for new therapeutic agents.

At the European level, the European Crohn's and Colitis Organisation (ECCO) developed a third version of their consensus guideline in 2017 that addresses UC.

Description of the technology and comparators

Ustekinumab

Ustekinumab is a human monoclonal antibody that binds with specificity to the cytokines interleukin (IL)-12 and IL-23. Ustekinumab may exert its clinical effects in UC through interruption of the T-helper 1 and T-helper 17 cytokine pathways, which are central to the pathology of this disease. Extension of the indication for ustekinumab was approved to include treatment of adult patients

with moderately to severely active UC who have had an inadequate response with, lost response to or were intolerant to either conventional therapy or a biologic, or have medical contraindications to such therapies.

Ustekinumab concentrate for solution for infusion is intended for use under the guidance and supervision of physicians experienced in the diagnosis and treatment of UC. Ustekinumab concentrate for solution for infusion should only be used for the intravenous induction dose. Ustekinumab treatment is to be initiated with a single intravenous dose based on body weight (approximately 6 mg/kg). The first subcutaneous dose (90 mg) should be given at week 8 following the intravenous dose. Details can be found in the summary of product characteristics (SmPC) [29].

Comparators

Biological therapies are treatment options for patients with moderately to severely active UC for whom conventional therapy has failed and for patients for whom prior biological treatment failed. Several possible relevant comparators in the European setting have been identified on the basis of recommendations in guidelines.

Adalimumab (Humira), infliximab (Remicade) and golimumab (Simponi) are three tumour necrosis factor (TNF) inhibitors (TNFi) currently approved for the treatment of adult patients with moderately to severely active UC who have had an inadequate response with, lost response to or were intolerant to either conventional therapy or a biologic or have medical contraindications to such therapies. There are also several agents biosimilar to infliximab and adalimumab available for the treatment of UC.

Vedolizumab (Entyvio; anti-integrin therapy) and tofacitinib (Xeljanz; oral Janus kinase inhibitor) are the only treatments approved for both the biologic-naïve population and for patients who have experienced biologic failure (BF). The label for vedolizumab states that it is indicated for patients for whom a previous TNFi has failed, while for tofacitinib any prior biologic therapy is allowed.

Objective and scope

Objective

The aim of this EUnetHTA Joint Relative Effectiveness Assessment is to compare the clinical effectiveness and safety of ustekinumab in the target patient populations with relevant comparators. The target patient populations and relevant comparators (according to the requirements of the EUnetHTA partners) are defined in the project scope in Table 2.1.

Methods

The assessment is based on the data and analyses included in the submission dossier prepared by the primary marketing authorisation holder (MAH). During the assessment, the completeness of the data and analyses in the submission dossier was verified. Furthermore, the methods for data analysis and synthesis applied by the primary MAH were checked against the requirements for the submission dossier and applicable EUnetHTA guidelines (https://www.eunethta.eu/methodology-guidelines/) and assessed with regard to scientific validity.

Literature search and assessment approach

A systematic literature review (SLR) was performed by the MAH to identify randomised controlled trials (RCTs) among patients with moderate to severe UC. The last search of bibliographical databases was performed on 28 March 2019, and of trial registries on 19 July 2019. The evidence base with regard to the drug under assessment provided by the MAH was reviewed by the authoring team. Search strategies were checked for appropriateness and the results for information retrieval included in the MAH submission dossier were checked for completeness against a search of study registries and against the studies included in the regulatory assessment report. Since no flaws were identified in the MAH search strategy, no supplementary searches were performed by the Information Specialist.

Information used to assess the clinical effectiveness and safety was extracted from the submission dossier and verified against the clinical study reports (CSRs) and other original documentation provided in the submission dossier. The ustekinumab dose of ~6 mg/kg (in the induction trial) is more relevant for this assessment then fixed doses of, for example, 130 mg.

The quality rating tool developed by the Cochrane Collaboration (version 5.1.0; March 2011) was used to assess the risk of bias in randomised trials. The results for the risk of bias assessment are presented at the study level only.

For rating the quality of the evidence of network meta-analyses, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) method was not applied, neither by the MAH nor by the authors of this Joint Assessment.

Indirect treatment comparison

The efficacy and safety of ustekinumab in patients with moderately to severely active UC have been evaluated in the non–head-to-head UNIFI trial, a phase 3, randomised, double-blind, placebo-controlled, parallel-group, multicentre study. Since no head-to-head studies of ustekinumab versus other active therapies in UC have been conducted, an indirect treatment comparison comprising a Bayesian network meta-analyses (NMAs) was performed and submitted by the MAH. The list of comparators includes all the biologics currently approved for moderate to severe UC, and two unlicensed doses of infliximab. The analyses were conducted using clinical response, clinical remission and mucosal healing as endpoints and were performed separately for induction trials (range 6–8 weeks) and maintenance trials (range 44–54 weeks; defined as the time from the end of induction to the end of maintenance) and two different patient populations: those for whom conventional care failed (nonbiologic failure [NBF] population) and patients with BF (defined by the MAH as a population of patients for whom one or more prior biologics failed).

The NMAs were conducted within a Bayesian framework with a Bayesian hierarchical model, using WinBUGS statistical software. All the endpoints were binary, so odds ratios (ORs; defined as relative treatment effects for ustekinumab versus a comparator) were estimated as the treatment effects. For each NMA, the intention of the MAH was to build both fixed effects and random effects models and to select the one associated with the lowest Deviance Information Criterion (DIC) and with a difference of at least three points in DIC. The approach was justified by the MAH using the explanation that the DIC allows choice of the model with the best compromise between adequacy and complexity. However, the DIC was not used for maintenance networks. Regarding the NMA statistical approaches to assess heterogeneity, assessments of heterogeneity and inconsistency in the networks were not feasible because in the main a single trial contributed to each comparison and comparisons between active treatments were indirect. Where manageable, the MAH performed a few pairwise meta-analyses of trials using the same comparison to assess their heterogeneity.

As statistical assessment of heterogeneity in the network or between individual trials was not possible, clinical and methodological heterogeneity was qualitatively assessed instead. The MAH considered heterogeneity among trials due to various issues that could impact the interpretation of effectiveness and safety outcomes, such as differences in patient populations, comparability at baseline, time of assessment, study design and analytical methodology.

Two major sources of heterogeneity were identified by the MAH: prior TNFi exposure and trial design. To minimise the effect of prior TNFi exposure on the treatment effect, the MAH divided patients within each trial into two broad categories—those without a biologic therapy failure (biologic-naïve or NBF) and those for whom one or more biologic therapies had failed (BF)—and built NMAs separately for each subgroup. Trial design was identified as an important source of heterogeneity in maintenance trials but not induction trials. To make the maintenance endpoint data comparable between the study designs, the MAH proposed an intention-to-treat (ITT) approach mimicking a treat-through trial design (also called a base case approach) in which conditional probabilities were used to recalculate maintenance endpoints in a re-randomisation trial (clinical response, remission or mucosal healing) to mimic those from a treat-through design.

Placebo maintenance endpoint data (generated under the treatment sequence of placebo at induction followed by placebo at maintenance) were imputed where missing for both induction responders and nonresponders to placebo using individual patient data (IPD) from available trials.

Data for the active treatment maintenance endpoint were also imputed for induction nonresponders in the GEMINI 1 trial and induction responders in the OCTAVE trial.

For NBF patients in the maintenance treatment arms, the MAH did not identify dose response relationship (higher dose/shorter interval between doses did not lead to higher clinical responses) and concluded that pooling of doses for the same treatment was appropriate. No formal dose-

response relationship testing was done. In the base-case NMAs, doses were pooled or not according to prespecified criteria. In the cases for which it was considered appropriate to pool the doses, additional NMAs were conducted in which the doses were not pooled to assess the impact of pooling.

To evaluate the robustness of the results, sensitivity analyses with regard to methodological factors presented in the submission dossier and the corresponding methods applied were evaluated. The MAH preplanned the following sensitivity analyses: include/exclude open-label trials in induction NMAs; analyse the results of a fixed-effects model if the random-effects model was selected in the base case; and assess the effect of the alternative prior in random-effects models. The MAH also performed further analyses: trials on Asian populations were additionally included in induction and maintenance networks to check if the results were replicated in a more broadly defined population; an alternative approach to analysis of maintenance endpoints was applied by mimicking the response-based trial design to apparently check the robustness of results; and a multiple imputation approach was simulated to assess the impact of using fixed values for imputation of missing placebo data.

Patient involvement

EUnetHTA conducted an open call asking general questions to elicit patient views on living with UC, the important outcomes considered in this assessment and expectations for the drug being assessed. Questions were based on the Health Technology Assessment International (HTAi) questionnaire template (https://www.htai.org/interestgroups/patient-and-citizen-involvement/resources/for-patients-and-patient-groups/).

European and national patient organisations had to provide an organisational perspective on the questions in English. For all parts of the open call, the term “patient” refers to anyone living with, or who has lived with, the condition for which the new medicine is indicated.

Two patient organisations completed the survey: Pacienti IBD z.s. (Czech Republic) and the European Federation of Crohn's and Ulcerative Colitis Associations. The information gathered from the open call was used to inform the scope of this assessment and in particular the outcomes to be considered. In addition, the information was used to help in the assessment of the health technology, specifically to provide the summary text related to the impact of the condition; experience with currently available pharmaceutical treatments, and experiences with and expectations of the pharmaceutical treatment under assessment.

Results

A total of 49 publications were identified in the SLR (including 32 full articles, 15 abstracts and 2 posters) for six drugs: ustekinumab, adalimumab, infliximab, golimumab, tofacitinib and vedolizumab. For ustekinumab, six abstracts were identified via an electronic search, but the clinical results for ustekinumab were primarily extracted from the CSRs provided by Janssen. A total of 21 trials were identified. Of these 21 trials, data were extracted from 19 trials and contributed to the analysis of the results and from 18 trials data were extracted for the NMA. In addition to studies related to ustekinumab (UNIFI-I and UNIFI-M), further studies were included in the NMA: four studies related to adalimumab (ULTRA 1 and ULTRA 2; NCT00853099, Suzuki et al. [1]; VARSITY; five studies related to infliximab (Jiang et al. [2]; Probert et al. [3]; Japic CTI-060298; ACT 1; ACT 2; two studies related to golimumab (PURSUIT-M; PURSUIT-SC); four studies related to tofacitinib (NCT00787202, Sandborn et al. [4]; OCTAVE-I1; OCTAVE-I2; OCTAVE-Sustain); and three studies related to vedolizumab (GEMINI 1; NCT02039505; VARSITY). For the VARSITY RCT, only the abstract was published; the registered protocol was available in the ClinicalTrials.gov register, without results posted. Of these studies, further studies were included in the induction NMAs (ACT 1, ACT 2, GEMINI 1, Probert et al., PURSUIT-SC phase 2, PUSUIT-SC phase 3, OCTAVE 1, OCTAVE 2, ULTRA 1, ULTRA 2, UNIFI, NCT00787202) and seven were included in the 1-year NMAs (ACT 1, GEMINI 1, PURSUIT [PURSUIT-SC and PURSUIT-M], OCTAVE [OCTAVE 1, 2 and Sustain], ULTRA 2, UNIFI and VARSITY). All trials compared active treatment to placebo with the exception of VARSITY, which compared vedolizumab to adalimumab.

All studies were conducted among subjects with moderately to severely active UC for whom nonbiologic therapy failed and/or subjects for whom prior biologic treatment(s) failed, with results reported separately for the two subpopulations.

Most studies that continued to evaluate the maintenance effect of the drug after the induction phase re-randomised the patients who responded to active arms at the beginning of the maintenance phase (e.g., UNIFI, PURSUIT, OCTAVE and GEMINI 1); other trials did not re-randomise the patients for the maintenance phase and patients continued to receive the same treatment as in the induction phase; in other words, patients were treated-through (e.g., ACT 1, ULTRA 2 and VARSITY). The main reason for the change in clinical trial design was that it was not considered ethical to expose patients to ineffective placebo treatments. The definition of the efficacy endpoints assessed and their time of assessment slightly change across trials. The primary efficacy endpoint most commonly reported was remission (defined as a Mayo score of ≤2 with no subscore >1) but the time of assessment varied across studies from week 6 to week 8 in induction trials and from week 30 to week 54 after the end of induction in the maintenance trials. The sample size varied across trials from 20 patients per arm to more than 400 per arm in induction trials, and from 30 patients per arm to more than 300 per arm in maintenance trials. At the trial outset there were no major discrepancies in baseline demographic characteristics (including age, weight and the proportion of males at baseline) and the baseline disease characteristics (including duration of disease, C-reactive protein [CRP] level and Mayo score) across studies. All studies included patients with moderate to severe UC (defined as a Mayo score between 6 and 12) for the induction phase. The mean Mayo score at baseline reported for patient subgroups ranged from 8 to 9 across all induction phase studies and the differences were limited across trials. In relation to disease duration, a general trend for shorter disease duration was observed for the NBF group than for the overall patient cohort. At the start of the induction phase, 15 studies reported a mean disease duration varying from 4 to 8 years for NBF patients, and from 6 to 11 years for the overall patient cohort. The mean CRP level at baseline reported in 15 studies for the induction phase ranged from 4 to 17 mg/l for NBF patients and from 5 to 19 mg/l for the overall patient cohort.

Direct evidence

Ustekinumab versus active treatment

No head-to-head studies of ustekinumab versus other active therapies in UC have been conducted.

Ustekinumab versus placebo

In the UNIFI induction trial of ustekinumab versus placebo in UC (Table 0.1), results were statistically significant for both induction doses of ustekinumab (~6 mg/kg and 130 mg) for all primary and major secondary efficacy endpoints, including clinical remission, clinical response, endoscopic healing and mucosal healing (combination of endoscopic and histologic evidence of healing) at week 8 (all $p < 0.001$). The ustekinumab dose of ~6 mg/kg is more relevant for this assessment.

In the UNIFI maintenance trial of ustekinumab versus placebo in UC (Table 0.2), results were statistically significant for both dose regimens (90 mg q8w and 90 mg q12w) for all primary and major secondary efficacy endpoints, including clinical remission, maintenance of clinical response, endoscopic healing, corticosteroid-free remission and maintenance of clinical remission through week 44 (all $p < 0.05$ except for ustekinumab q8w for maintenance of clinical remission, for which $p = 0.69$). Improvements in clinical outcomes were accompanied by reductions in inflammatory biomarkers and improvements in HRQoL measures (IBDQ, EuroQol Questionnaire 5 Dimensions [EQ-5D] and 36-item Short Form survey [SF-36]).

Table 0.1. UNIFI-I trial of ustekinumab versus placebo: summary of results for outcomes (clinical response, clinical remission, endoscopic healing) at week 8 in the ITT analysis

Study reference/ID	Ustekinumab ~6 mg/kg		Ustekinumab 130 mg		Placebo		Adjusted treatment difference vs. placebo (95% CI)	
	N	*n* (%)	*N*	*n* (%)	*N*	*n* (%)	Ustekinumab ~6 mg/kg	Ustekinumab 130 mg
UNIFI-I								
Clinical response	322	199 (61.8)	320	164 (51.3)	319	100 (31.3)	30.5 (23.2–37.8); $p < 0.001$	19.9 (12.5–27.3); $p < 0.001$
BF patients	166	95 (57.2)	164	74 (45.1)	161	44 (27.3)	$p < 0.001$	$p < 0.001$
NBF patients	156	104 (66.7)	156	90 (57.7)	158	56 (35.4)	$p < 0.001$	$p < 0.001$
Clinical remission	322	50 (15.5)	320	50 (15.6)	319	17 (5.3)	10.2 (5.6–14.8); $p < 0.001$	10.3 (5.7–14.9); $p < 0.001$
BF patients	166	21 (12.7)	164	19 (11.6)	161	2 (1.2)	11.4 (6.1–16.7); $p < 0.001$	10.4 (5.2–15.5); $p < 0.001$
NBF patients	156	29 (18.6)	156	31 (19.9)	158	15 (9.5)	9.0 (1.4–16.5); $p = 0.022$	10.4 (2.7–18.1); $p = 0.009$
Endoscopic healing	322	87 (27.0)	320	84 (26.3)	319	44 (13.8)	13.7 (7.3–19.3); $p < 0.001$	12.4 (6.5–18.4); $p < 0.001$
BF patients	156	52 (33.3)	164	30 (18.3)	161	11 (6.8)	$p < 0.001$	$p = 0.002$
NBF patients	156	52 (33.3)	156	54 (34.6)	158	33 (20.9)	$p = 0.014$	$p = 0.006$

Abbreviations: BF = biologic failure; CI = confidence interval; ITT = intention to treat; *n* = number of patients with at least one event; *N* = number of patients analysed; NBF = nonbiologic failure.

Table 0.2. UNIFI-M trial of ustekinumab versus placebo: summary of results for outcomes (clinical response, clinical remission, endoscopic healing) at week 44 in the ITT analysis

Study reference/ID	Ustekinumab SC 90 mg q12w		Ustekinumab SC 90 mg q8w		Placebo SC		Adjusted treatment difference vs. placebo (95% CI)	
	N	*n* (%)	*N*	*n* (%)	*N*	*n* (%)	Ustekinumab 90 mg q12w	Ustekinumab SC 90 mg q8w
UNIFI-M								
Subjects maintaining clinical response	172	117 (68.0)	176	125 (71.0)	175	78 (44.6)	23.5 (13.7–33.3); $p < 0.001$	26.4 (16.6–36.1); $p < 0.001$
BF patients	70	39 (55.7)	91	59 (64.8)	88	34 (38.6)	$p < 0.01$	$p < 0.001$
NBF patients	102	78 (76.5)	85	66 (77.6)	87	44 (50.6)	$p < 0.001$	$p < 0.001$
Subjects maintaining clinical remission	172	66 (38.4)	176	77 (43.8)	175	42 (24.0)	14.5 (5.5–23.6); $p < 0.01$	19.7 (10.3–29.0); $p < 0.001$
BF patients	70	16 (22.9)	91	36 (39.6)	88	15 (17.0)	$p < 0.01$	$p < 0.001$
NBF patients	102	50 (49.0)	85	41 (48.2)	87	27 (31.0)	$p < 0.01$	$p < 0.01$
Subjects in clinical remission and not receiving concomitant corticosteroids	172	65 (37.8)	176	74 (42.0)	175	41 (23.4)	14.5 (5.5–23.6); $p < 0.01$	18.5 (9.3–27.8); $p < 0.001$
BF patients	70	16 (22.9)	91	34 (37.4)	88	14 (15.9)	$p < 0.01$	$p < 0.001$
NBF patients	102	49 (48.0)	85	40 (47.1)	87	27 (31.0)	$p < 0.01$	$p < 0.01$
Subjects with endoscopic healing	172	75 (43.6)	176	90 (51.1)	175	50 (28.6)	15.2 (5.8–24.6); $p < 0.01$	22.5 (12.8–32.2); $p < 0.001$
BF patients	70	18 (25.7)	91	41 (45.1)	88	20 (22.7)		$p < 0.001$
NBF patients	102	57 (55.9)	85	49 (57.6)	87	30 (34.5)	$p < 0.01$	$p < 0.01$

Abbreviations: BF = biologic failure; CI = confidence interval; ITT = intention to treat; *n* = number of patients with at least one event; *N* = number of patients analysed; NBF = nonbiologic failure; qXw = every X weeks; SC = subcutaneous.

QoL

Ustekinumab improved irritable bowel disease (IBD)-specific and general HRQoL outcomes as evaluated using the IBDQ, SF-36 and EQ-5D questionnaires at week 8, with statistically significant results for both induction doses.

Through to week 44, ustekinumab 90 mg at q8w and q12w was generally able to maintain improvements in patient-reported HRQoL as assessed using the IBDQ, SF-36 and EQ-5D instruments when compared to placebo.

Hospitalisations and UC-related surgeries

Through to week 8 of UNIFI-I, the proportion of subjects with UC-related hospitalisations was significantly lower in the ustekinumab recommended dose group (1.6%, 5/322) than in the placebo group (4.4%, 14/319), and no subjects receiving ustekinumab at the recommended induction dose underwent UC-related surgeries, compared to 0.6% (2/319 subjects) in the placebo group.

Through to week 44 of UNIFI-M, a significantly lower number of UC-related hospitalisations was observed for subjects in the combined ustekinumab group (2.0%, 7/348) than in the placebo group (5.7%, 10/175). In addition, a lower number of subjects in the ustekinumab group (0.6%, 2/348) underwent UC-related surgeries than in the placebo group (1.7%, 3/175).

Indirect evidence: NMAs

Comparisons of active treatments are more relevant at long-term time points. The NMAs were conducted to assess the continued year-long treatment regimens to provide a more complete picture of the long-term relative efficacy of treatments received in clinical practice.

The evidence on which all the networks were based was scarce, indirect data in a setting for which the heterogeneity or inconsistency of a network could not be statistically assessed and assessment of statistical heterogeneity between trials informing each comparison was limited and probably underpowered. Even for the networks that included a head-to-head comparison, a closed-loop inconsistency assessment between different sources of evidence was not performed because of data imputation in one of the trials. As a network is most justifiable under an assumption of consistency between different sources of evidence, the lack of this data largely increases uncertainty regarding NMA results.

The MAH stated that clinical and methodological heterogeneity of prior TNFi exposure and trial design were mitigated by the procedures previously described, and identified no other sources of heterogeneity.

The results of the indirect comparison are presented in Table 0.3–Table 0.10. The credible interval (CrI) for most outcomes was quite wide, indicating high uncertainties in the results. While this is partly expected as the assessment is based on indirect comparisons, which contribute additional uncertainty, in some models the CrI width hampers interpretation. For example, in the case of the NMA results for clinical remission in BF patients at induction, the upper limit of the CrI for the OR reaches a value of 95, which seriously undermines confidence in the results from such a model.

Regarding the ranking of treatments, the MAH presented both surface under the cumulative ranking (SUCRA) values and the probability of ustekinumab being better than a comparator. All the networks were associated with largely uncertain treatment effects, so the probability of a treatment being better (which is known to underperform in such cases) was not used for interpretation. The SUCRA values were likewise derived from low-certainty evidence under the circumstances of large uncertainness in effect sizes, allowing for the possibility that a treatment with the higher rank is better than another treatment only by a small amount that is not clinically meaningful. Consequently, while the ustekinumab-based interventions in the maintenance NMAs seemed to consistently be ranked top, the results could easily be misleading.

Induction NMA

The statistical heterogeneity of trials informing a direct comparison to placebo was assessed in three induction NMAs in which such an analysis was possible (networks including at least 1 comparison with at least 2 trials). All pairwise meta-analyses were based on two trials (except 1 with 3 trials) and were thus probably underpowered for exploration of heterogeneity. Nevertheless, sig-

nificant statistical heterogeneity was identified in networks for NBF patients: in trials reporting on clinical remission for infliximab 5 mg/kg versus placebo (n = 3; I^2 = 63%; Q-test p = 0.07; significant at a level of 0.1) and in trials reporting on clinical response for infliximab 10 mg/kg versus placebo (n = 2; I^2 = 70%; Q-test p = 0.07).

For all of the induction NMAs, comparison of DIC values revealed that the fixed-effects model was preferred over the random-effects model. However, the DIC values observed for fixed-effects and random-effects models were very similar and did not provide an evidence basis for preference of the fixed-effects model. Given that there is a good intuitive clinical rationale for why heterogeneity may be expected but insufficient studies are available to detect it, the random-effects model should have been selected.

Pooling of data from the two OCTAVE trials for NMAs for clinical response and mucosal healing in the NBF population and for clinical remission in the BF population increased the risk of bias. This was especially true under the fixed-effects model as pooling artificially enlarged the sample size of trials and created a trial with an active arm at least three times the size of active arms in other trials.

Comparative effectiveness NMA for the NBF population

Patients who received ustekinumab (~6 mg/kg) had a significantly higher odds of achieving a clinical response, remission and mucosal healing in comparison to placebo after an induction period of 6–8 weeks. In addition, ustekinumab at ~6 mg/kg had a significantly higher odds of achieving clinical response than adalimumab 160/80 mg (Table 0.3). The point estimates for these comparisons were high, with ORs ranging from 1.9 to 3.7. However, estimates were very uncertain and it was therefore not possible to state if the effect size was small, moderate or high. For clinical remission and mucosal healing endpoints we could not exclude the possibility that the true effect of ustekinumab ~6 mg/kg in comparison to placebo was not clinically important.

The clinical response SUCRA value of 77% for this ustekinumab dose seemed also higher than values for other active treatments except for infliximab interventions (Table 0.4). According to the SUCRA values, interventions with ustekinumab are ranked better than placebo and adalimumab 160/80 mg for all endpoints and better than golimumab 200/100 mg for clinical response and mucosal healing. However, it was not possible to conclude if these differences are clinically important.

.

Table 0.3. NMA results for clinical response, clinical remission and mucosal healing outcomes: NBF patients on induction treatment

Clinical response		Clinical remission		Mucosal healing	
Treatment*	**Median OR (CrI)**	**Treatment***	**Median OR (CrI)**	**Treatment***	**Median OR (CrI)**
	Ustekinumab ~6 mg/kg vs.		**Ustekinumab ~6 mg/kg vs.**		**Ustekinumab ~6 mg/kg vs.**
Infliximab 5 mg/kg	0.89 (0.49–1.63); Pr = 36%	Infliximab 5 mg/kg	0.49 (0.22–1.14); Pr = 5%	Infliximab 5 mg/kg	0.57 (0.30–1.10); Pr = 5%
Infliximab 10 mg/kg	0.96 (0.53–1.76); Pr = 45%	Vedolizumab 300mg	0.48 (0.13–1.58); Pr = 12%	Infliximab 10 mg/kg	0.59 (0.32–1.13); Pr = 6%
Ustekinumab 6 mg/kg	–	Infliximab 10 mg/kg	0.64 (0.28–1.48); Pr = 15%	Vedolizumab 300 mg	0.64 (0.29–1.45); Pr = 14%
Vedolizumab 300 mg	1.14 (0.52–2.47); Pr = 63%	Golimumab 200/100 mg	0.74 (0.31–1.78); Pr = 25%	Tofacitinib 10 mg	0.85 (0.41–1.72); Pr = 32%
Tofacitinib 10 mg	1.36 (0.74–2.53); Pr = 84%	Tofacitinib 10 mg	0.90 (0.35–2.24); Pr = 41%	Ustekinumab 130 mg	0.94 (0.59–1.52); Pr = 41%
Ustekinumab 130 mg	1.47 (0.93–2.34); Pr = 95%	Ustekinumab 130 mg	1.47 (0.44–4.93); Pr = 39%	Ustekinumab 6 mg/kg	–
Golimumab 200/100 mg	1.60 (0.90–2.84); Pr = 95%	Ustekinumab 6 mg/kg	–	Golimumab 200/100 mg	1.06 (0.57–1.98); Pr = 57%
Adalimumab 160/80 mg	1.94 (1.10–3.45); Pr = 99%	Adalimumab 160/80 mg	0.99 (0.43–2.30); Pr = 49%	Adalimumab 160/80 mg	1.26 (0.68–2.35); Pr = 77%
Placebo	3.66 (2.31–5.88); Pr = 100%	Placebo	2.19 (1.14–4.39); Pr = 99%	Placebo	1.90 (1.15–3.20); Pr = 99%

Abbreviations: CrI = credible interval; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR > 1 indicates a higher likelihood of a response with ustekinumab than with the comparator, suggesting that ustekinumab performs better). *Treatments are ordered by SUCRA value from most to least favourable.

Table 0.4. Ranking of treatments from most to least favourable with corresponding SUCRA values

Clinical response		Clinical remission		Mucosal healing	
Infliximab 5 mg/kg	86%	Infliximab 5 mg/kg	88%	Infliximab 5 mg/kg	87%
Infliximab 10 mg/kg	79%	Vedolizumab 300 mg	81%	Infliximab 10 mg/kg	84%
Ustekinumab ~6 mg/kg	77%	Infliximab 10 mg/kg	67%	Vedolizumab 300 mg	76%
Vedolizumab 300 mg	64%	Golimumab 200/100 mg	58%	Tofacitinib 10 mg	56%
Tofacitinib 10 mg	49%	Tofacitinib 10 mg	43%	Ustekinumab 130 mg	47%
Ustekinumab 130 mg	41%	Ustekinumab 130 mg	42%	Ustekinumab ~6 mg/kg	41%
Golimumab 200/100 mg	34%	Ustekinumab ~6 mg/kg	36%	Golimumab 200/100 mg	36%
Adalimumab 160/80 mg	20%	Adalimumab 160/80 mg	35%	Adalimumab 160/80 mg	22%
Placebo	0%	Placebo	0%	Placebo	0%

Abbreviations: SUCRA = surface under the cumulative ranking.

Comparative effectiveness NMA for the BF population

In terms of the number of trials, sample size and number of events, the data for NMAs for BF patients were even more scarce than for NMAs for NBF patients. In addition, a potential effect of data pooling in the OCTAVE trials on the fixed-effects and random-effects estimates was observed in the pairwise meta-analysis for tofacitinib 10 mg versus placebo.

The NMA model for clinical remission with induction resulted in unrealistically high ORs, at 13.41 (CrI 3.62–94.58) for ustekinumab ~6 mg/kg versus placebo and 9.97 (CrI 1.77–88.37) for ustekinumab ~6 mg/kg versus adalimumab, indicating that the data are probably too scarce for valid estimates. Therefore, the quality of the evidence for the estimates of NMA model for clinical remission is too weak to allow any inferences.

For clinical response and mucosal healing the NMAs also show very wide CrIs but the values are somewhat less extreme. Patients who received ustekinumab at ~6 mg/kg for induction had significantly higher odds of achieving mucosal healing or a clinical response (remission was not discussed as explained above) in comparison to placebo after an induction period of 6–8 weeks, with an increase of at least 86% and 127% in odds, respectively (Table 0.5). For comparison of ustekinumab at ~6 mg/kg to adalimumab, significantly higher odds of achieving mucosal healing and a clinical response were reported in patients treated with ustekinumab showing by at least 33% and 17% increase in odds, respectively. While point estimates and the upper CrI limits for these comparisons indicate that the true effect may be higher, we cannot exclude the possibility that differences to mucosal healing or clinical response for comparison to adalimumab are not clinically important. So we conclude that ustekinumab may be more effective than placebo in acchiving mucosal healing and clinical reponse, while we are uncertain if ustekinumab improves these endpoints in comparison to adalimumab. For comparison of ustekinumab ~6 mg/kg to other active treatments, CrI estimates were not informative (effect size ranged from highly/moderately harmful to highly beneficial).

The SUCRA values show that ustekinumab ~6 mg/kg was rated as the best or second best treatment for a clinical response and mucosal healing, but again the ranking could not be supported by the effect size estimates, which were not informative except for placebo and adalimumab (Table 0.6).

In summary, scarce data with even lower sample sizes and event counts than in NBF induction NMAs, particularly for clinical remission, together with previously stated limitations including the general comments on indirectness, and expected heterogeneity accompanied with the use of fixed-effects models, suggest that there is more uncertainty and a higher risk of bias in the results for induction networks obtained for the BF group than for the NBF group.

Table 0.5. NMA results for clinical response, clinical remission and mucosal healing outcomes: BF patients on induction treatment

Clinical response		Clinical remission		Mucosal healing	
Treatment*	**Median OR (CrI)**	**Treatment***	**Median OR (CrI)**	**Treatment***	**Median OR (CrI)**
	Ustekinumab ~6 mg/kg vs.		**Ustekinumab ~6 mg/kg vs.**		**Ustekinumab ~6 mg/kg vs.**
Tofacitinib 10 mg	1.05 (0.55–1.98); Pr = 56%	Tofacitinib 10 mg	0.59 (0.02–7.92); Pr = 35%	Tofacitinib 10 mg	0.87 (0.29–2.46); Pr = 40%
Ustekinumab 130 mg	1.63 (1.06–2.52); Pr = 99%	Ustekinumab 130 mg	1.11 (0.57–2.17); Pr = 62%	Ustekinumab 130 mg	1.20 (0.69–2.08); Pr = 74%
Vedolizumab 300 mg	1.43 (0.58–3.43); Pr = 78%	Vedolizumab 300 mg	3.60 (0.32–40.71); Pr = 86%	Vedolizumab 300 mg	2.19 (0.76–6.41); Pr = 93%
Adalimumab 160/80 mg	2.48 (1.17–5.31); Pr = 99%	Adalimumab 160/80 mg	9.97 (1.77–88.37); Pr = 100%	Adalimumab 160/80 mg	3.42 (1.33–9.12); Pr = 99%
Placebo	3.58 (2.27–5.74); Pr = 100%	Placebo	13.41 (3.62–94.58); Pr = 100%	Placebo	3.73 (1.86–8.04); Pr = 100%

Abbreviations: BF = biologic failure; CrI = credible interval; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR >1 indicates a higher likelihood of a response with ustekinumab than with the comparator, suggesting that ustekinumab performs better); SUCRA = surface under the cumulative ranking. *Treatments are ordered by SUCRA value from most to least favourable.

Table 0.6. Ranking of treatments from most to least favourable with corresponding SUCRA values

Clinical response		Clinical remission		Mucosal healing	
Ustekinumab ~6 mg/kg	86%	Tofacitinib 10 mg	85%	Tofacitinib 10 mg	86%
Tofacitinib 10 mg	82%	Ustekinumab ~6 mg/kg	77%	Ustekinumab ~6 mg/kg	81%
Vedolizumab 300 mg	59%	Ustekinumab 130 mg	71%	Ustekinumab 130 mg	68%
Ustekinumab 130 mg	47%	Vedolizumab 300 mg	44%	Vedolizumab 300 mg	40%
Adalimumab 160/80 mg	24%	Adalimumab 160/80 mg	18%	Adalimumab 160/80 mg	17%
Placebo	2%	Placebo	6%	Placebo	9%

Abbreviations: SUCRA = surface under the cumulative ranking.

Clinical effectiveness in 1-year NMA

Assessments of heterogeneity and inconsistency in the maintenance networks or between the trials were not feasible because only a single trial informed each comparison and even in networks on clinical remission that included one head-to-head comparison there was no closed evidence loop informed by exclusively direct data to assess agreement between direct and indirect data. Consequently, all the networks were based on scarce indirect data. In addition, only the fixed-effects models were assessed, as a random-effects model could not be built because of the convergence problem. This was done for data for which it is reasonable to assume heterogeneity.

Trial design was identified as an important source of heterogeneity and the MAH performed recalculation and imputation for four out of seven trials to make the endpoints of re-randomised trials comparable to those of treat-through designs. This resulted in heavy imputation of data.

NBF population

To increase the statistical power in this model, the MAH pooled doses for treatment arms as no dose–response relationship was observed. However, this decision should not be based on the lack of statistical significance in a data-scarce model, but rather on knowledge from the literature. It was recently suggested that use of more stringent endpoints (e.g., combining clinical remission with steroid-free remission) made it possible to see differentiation between dose groups emerging in maintenance trials. In addition, pooling of doses results in overprecise results (as data from two active treatment arms of a trial are pooled together and the sample size of the active arm in that trial is increased twice), which could affect the fixed-effect NMA results.

In the NBF population, patients who received ustekinumab ~6 mg/kg during induction followed by ustekinumab 90 mg had significantly higher odds of achieving a clinical response, remission and mucosal healing after a 1-year-long treatment regimen in comparison to placebo, golimumab and adalimumab (Table 0.7). The point estimates for comparisons to placebo were high, with ORs of three endpoints ranging from 5.1–8.7, and althgough the CrIs were wide, the lower CrI bound was ≥2.8, indicating that ustekinumab may increases the odds of clinical response, remission, and mucosal healing compared to placebo. As for comparisons to golimumab or adalimumab, while OR point estimates were ≥2.4, for primary endpoints of clinical remission and mucosal healing the lower CrI bound for estimates of the effect size were such that there was a possibility that the true effect was not clinically important (The results show that ustekinumab increases odds of clinical remission in comparison to adalimumab 10% and more and in comparison to golimumab 13% and more, while odds of mucosal healing in comparison to the same drugs increased at least 30 and 39% and more).

In addition, a clinical response was significantly more likely for patients treated with ustekinumab than for those treated with infliximab or tofacitinib, but again the credibility interval cannot exclude the possibility that ustekinumab increases odds of clinical response in comparison to tofacitinib by just 6% (OR 2.27, CrI 1.06–4.86) or to infliximab by just 22% (2.62, 1.22–5.60). The SUCRA values seemed to consistently show that ustekinumab pooled doses ranked first for all the endpoints. In addition, placebo, golimumab and adalimumab treatments, for which responses were significantly worse than for ustekinumab, seemed to be consistently ranked bottom for all the endpoints as well. However, given that clear interpretation of effect sizes between active treatments is hindered, we are uncertain about the validity of the ranking presented (Table 0.8).

Overall, given that various methodological issues related to maintenance models for the NBF population were raised and discussed, such as the paucity of the data, indirectness and the fact that inconsistency between different sources of evidence was not estimated (this includes NMA for clinical remission where inconsistency could be estimated due to a closed loop but the data of GEMINI I trial within the loop were imputed and the loop was apparently used to cross-check the imputation approach by comparing the re-calculated efficacy from GEMINI-1 with the results from VARSITY and ULTRA-II), that data were recalculated and imputed in majority of trials (either for induction responders and/or non-responders in placebo or active arm), that fixed effect model was used in a setting where heterogeneity is expected but could not be tested, and that there is a potential impact of bias due to pooling of doses, we are uncertain if ustekinumab improves clinical response, clinical remission, and mucosal healing after a 1-year treatment regimen when compared with golimumab and adalimumab, as well as clinical response when compared with infliximab and tofacitinib.

Table 0.7. NMA results for clinical response, clinical remission and mucosal healing for pooled doses in the NBF population over 1 year: base case mimicking a treat-through approach

Treatment sequence for induction to maintenance	Median OR (CrI) for pooled ustekinumab 6 mg/kg and 90 mg		
	Clinical response	Clinical remission	Mucosal healing
Placebo – Placebo	8.70 (5.03–15.40); Pr = 100%	5.11 (2.83–9.52); Pr = 100%	5.57 (3.19–9.92); Pr = 100%
Infliximab pooled – Infliximab pooled	2.62 (1.22–5.60); Pr = 99.31%	1.89 (0.83–4.29); 93.59%	1.43 (0.66–3.09); Pr = 81.59%
Adalimumab 160/80/40mg – Adalimumab 40mg EOW	4.76 (2.25–10.16); Pr = 100%	2.43 (1.10–5.42); Pr = 98.59%	2.91 (1.33–6.39); Pr = 99.62%
Golimumab 200/100mg – Golimumab pooled	3.76 (1.90–7.57); Pr = 99.99%	2.40 (1.13–5.22); Pr = 98.84%	2.79 (1.39–5.69); Pr = 99.81%
Tofacitinib 10mg - Tofacitinib pooled	2.27 (1.06–4.86); Pr = 98.21%	1.51 (0.64–3.51); Pr = 82.97%	1.94 (0.88–4.25); Pr = 95.11%
Vedolizumab 300mg – Vedolizumab 300mg q8w	1.93 (0.75–4.82); Pr = 91.45%	1.47 (0.65–3.33); Pr = 82.38%	1.60 (0.69–3.77); Pr = 86.24%

Abbreviations: CrI = credible interval; EOW = every other week; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR > 1 indicates higher odds of a response with ustekinumab than with the comparator, suggesting that ustekinumab performs better); q8w = every 8 weeks.

Table 0.8. Ranking of treatments from most to least favourable with corresponding SUCRA values: pooled doses

Clinical response		Clinical remission		Mucosal healing	
UST 6-UST pooled	98%	UST 6-UST pooled	89%	UST 6-UST pooled	91%
VDZ 300-VDZ 300 pooled	73%	VDZ 300-VDZ 300 pooled	66%	IFX pooled-IFX pooled	69%
TOC 10-TOC pooled	66%	TOC 10-TOC pooled	62%	VDZ 300-VDZ 300 q8w	63%
IFX pooled-IFX pooled	57%	IFX pooled-IFX pooled	46%	TOC 10-TOC pooled	49%
GOL 200/100-GOL pooled	34%	GOL 200/100-GOL pooled	29%	GOL 200/100-GOL pooled	26%
ADA 160/80/40-ADA 40 EOW	22%	ADA 160/80/40-ADA 40 EOW	27%	ADA 160/80/40-ADA 40 EOW	23%

Abbreviations: ADA = adalimumab; EOW = every other week; GOL = golimumab; IFX = infliximab; q8w = every 8 weeks; SUCRA = surface under the cumulative ranking curve; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

BF population

According to results in the individual trials, a potential dose–response relationship was observed for clinical response to tofacitinib so it was deemed inappropriate to pool the doses for the same treatment. A dose–response relationship was observed in the data for clinical remission with tofacitinib and ustekinumab so it was deemed inappropriate to pool the doses for the same treatment. Only the results for unpooled doses are therefore presented below (Table 0.9).

Results are not available for mucosal healing in the BF population as no data from GEMINI 1 were available to inform the imputation of induction nonresponder data (required for UNIFI, PURSUIT and OCTAVE).

The 1-year NMA for efficacy for the BF group showed significantly higher odds of achieving a clinical response and clinical remission with both ustekinumab regimens versus placebo (~6 mg/kg followed by 90mg q8w for both induction responders and nonresponders; ~6 mg/kg followed by 90 mg q12q for induction responders or 90 mg q8w for induction nonresponders).

While the point estimates for these comparisons were high, with ORs of 5.3–6.9 for clinical remission and 4.8 for clinical response, the reported CrIs ranged from ~2 to ≥9. Nevertheless, the lower CrI bound was ~2, indicating that ustekinumab may increase the odds of clinical response and

remission twofold when compared to placebo. All other comparisons for efficacy in terms clinical response or clinical remission between ustekinumab and any active treatments were nonsignificant, with the wide CrIs demonstrating that little is known about the effect and that further information is needed. For this reason, as discussed for the maintenance NMAs for NBF patients, the fact that all the point estimates for all the comparisons to active treatments, although not significant, point towards increase in odds of achieving clinical response or clinical remission in patients treated with ustekinumab can not be used to support the absolute ustekinumab efficacy.

The CrIs for the estimated treatment effects are also wide because of the smaller patient counts in this subgroup, lower placebo efficacy rates and no pooling of the doses.

As for the SUCRA ranking, both the ustekinumab doses were ranked first and second for clinical response although their SUCRA values were similar to that for the treatment sequence tofacitinib 10 mg → tofacitinib 10 mg early + tofacitinib 10 mg delayed (Table 0.10). Regarding clinical remission, only the dosing regime ustekinumab ~6 mg/kg → ustekinumab 90 mg q8w early + ustekinumab 90 mg SC q8w delayed was ranked high (first place). While these rankings seem to show consistency for UST ranking 1st, they can also be misleading and the ranking results cannot be taken at face value. As previously stated, SUCRA does not consider the magnitude of differences in effects between treatments, which are hindered in these networks by very wide CrIs including an OR of 1.

Overall, given that NMA of mucosal healing in the BF group was not feasible and taking into account various methodological issues raised and discussed for maintenance models for BF patients: even more scarce data than in NMAs in NBF patients, indirectness and the fact that inconsistency between different sources of evidence was not estimated (this includes the NMA for clinical remission where inconsistency could be estimated due to a closed loop but the data of GEMINI I trial within the loop were imputed and the loop was apparently used to cross-check the imputation approach by comparing the re-calculated efficacy from GEMINI-1 with the results from VARSITY and ULTRA-II), the fact that data were recalculated and imputed in majority of trials, that fixed effect model was used in a setting where heterogeneity is expected but could not be tested, we are uncertain in these results.

Table 0.9. NMA results for clinical response and clinical remission outcomes, unpooled doses – BF patients – 1-year – base case mimicking a treat-through approach

	Median OR (CrI); Pr			
	Clinical response		Clinical remission	
Treatment sequence	**UST 6mg/kg → UST 90 mg q8w for induction responders and nonresponders vs.**	**UST 6 mg/kg → UST 90 mg q12w for induction responders + UST 90 mg q8w for induction nonresponders vs.**	**UST 6mg/kg → UST 90 mg q8w for induction responders and nonresponders vs.**	**UST 6mg/kg → UST 90 mg q12w for induction responders + UST 90 mg q8w for induction nonresponders vs.**
Placebo → placebo	4.83 (2.56–9.25); 100%	4.82 (2.28–10.30); 100%	6.89 (2.98–16.90); 100%	5.34 (1.97–14.62); 99.94%
VDZ 300 mg → VDZ 300 mg q8w for induction responders + VDZ 300 mg q4w for induction nonresponders	1.76 (0.51–6.00); 81.45%	1.75 (0.48–6.35); 80.04%	1.26 (0.31–4.91); 62.87%	0.97 (0.22–4.11); 48.53%
VDZ 300 mg → VDZ 300 mg q4w for induction responders and nonresponders	1.89 (0.53–6.69); 83.94%	1.88 (0.50–7.06); 82.54%	1.32 (0.26–6.63); 63.48%	1.02 (0.19–5.48); 51.07%
ADA 160/80/40 mg → ADA 40 mg EOW	2.03 (0.70–5.72); 90.52%	2.02 (0.65–6.14); 88.85%	1.71 (0.42–6.55); 77.63%	1.32 (0.29–5.48); 64.31%
TOC 10 mg → TOC 5 mg for induction responders + TOC 10 mg for induction nonresponders	1.66 (0.69–3.94); 87.24%	1.65 (0.63–4.28); 84.72%	1.57 (0.44–5.36); 76.05%	1.21 (0.31–4.52); 60.94%
TOC 10 mg → TOC 10 mg for induction responders and nonresponders	1.21 (0.51–2.83); 66.49%	1.20 (0.46–3.08); 64.70%	1.08 (0.31–3.61); 54.80%	0.83 (0.21–3.05); 39.18%
UST 6 mg/kg → UST 90 mg q12w for induction responders + UST 90 mg q8w for induction nonresponders	1.00 (0.45–2.25); 50.38%	-	1.29 (0.53–3.32); 70.88%	-
UST 6mg/kg → UST 90 mg q8w for induction responders and nonresponders	-	1.00 (0.44–2.23); 49.62%	-	0.77 (0.30–1.90); 29.12%

Abbreviations: CrI = Credible interval; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR >1 indicates higher odds of a response with ustekinumab than with the comparator suggesting that ustekinumab performs better); ADA = adalimumab; EOW = every other week; GOL = golimumab; IFX = infliximab; qXw = every X weeks; SUCRA = surface under the cumulative ranking curve; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Table 0.10. Ranking of treatments from most to least favourable with corresponding SUCRA values: unpooled analysis

Clinical response		Clinical remission	
UST 6 → UST 90 SC q8w early + UST 90 SC q8w delayed	80%	UST 6 → UST 90 SC q8w early + UST 90 SC q8w delayed	72%
UST 6 → UST 90 SC q12w early + UST 90 SC q8w delayed	79%	TOC 10 → TOC 10 early + TOC 10 delayed	70%
TOC 10 → TOC 10 early + TOC 10 delayed	71%	VDZ 300 → VDZ 300 q8w early + VDZ 300 q4w delayed	60%
VDZ 300 → VDZ 300 q8w early + VDZ 300 q4w delayed	46%	VDZ 300 → VDZ 300 q4w early + VDZ 300 q4w delayed	56%
TOC 10 → TOC 5 early + TOC 10 delayed	45%	UST 6 → UST 90 SC q12w early + UST 90 SC q8w delayed	56%
VDZ 300 → VDZ 300 q4w early + VDZ 300 q4w delayed	42%	TOC 10 → TOC 5 early + TOC 10 delayed	44%
ADA 160/80/40 → ADA 40 EOW	37%	ADA 160/80/40 → ADA 40 EOW	41%
Placebo → placebo	1%	Placebo → placebo	0%

Abbreviations: ADA = adalimumab; EOW = every other week; GOL = golimumab; IFX = infliximab; qXw = every X weeks; SC = subcutaneous; SUCRA = surface under the cumulative ranking curve; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Safety results for ustekinumab according to UNIFI

AE rates in the UNIFI studies were similar in the ustekinumab and placebo arms. Serious AE (SAE) rates were not significantly different between treatment groups in the induction and maintenance studies; event rates were higher for placebo than for ustekinumab groups in the induction phase (6.6% placebo vs 3.7% ustekinumab 130 mg and 3.1% ustekinumab ~6 mg/kg).

AEs of special interest were similar across treatment groups. Infections of any severity were more common in the placebo group than in the ustekinumab groups and the rate of serious infections was low and similar across all treatment groups. In the final safety analysis of the induction study, only two subjects reported malignancies (prostate cancer and rectal adenocarcinoma); in the maintenance study, six subjects reported malignancies (5 ustekinumab-treated subjects and 1 placebo-only subject). Through to week 8 of the induction study there was one SAE of ischaemic stroke in the placebo group; among all treated subjects in the maintenance study, a major adverse cardiac event (nonfatal myocardial infarction, nonfatal stroke or cardiovascular death) was reported for two subjects.

Table 0.11. Summary of adverse events in UNIFI: safety analysis set

Safety endpoint	UNIFI full patient population					
	Induction (week 8)			Maintenance (week 44)		
	PBO (*N* = 319)	UST 130 mg (N = 321)	UST 6 mg/kg (*N* = 320)	PBO sc (*N* = 175)	UST 90 mg q12w (*N* = 172)	UST 90 mg q8w (*N* = 175)
Adverse events, *n* (%)	153 (48.0)	133 (41.4)	160 (50.0)	138 (78.9)	119 (69.2)	136 (77.3)
Serious adverse events, *n* (%)	22 (6.6)	12 (3.7)	10 (3.1)	17 (9.7)	13 (7.6)	15 (8.5)
Most frequent adverse events, *n* (%)						
Worsening ulcerative colitis	18 (5.6)	9 (2.8)	7 (2.2)	50 (28.6)	19 (11.0)	18 (10.2)
Nasopharyngitis	NR	NR	NR	28 (16.0)	31 (18)	26 (14.8)
Headache	14 (4.4)	22 (6.9)	13 (4.1)	7 (4.0)	11 (6.4)	18 (10.2)
Arthralgia	2 (0.6)	3 (0.9)	6 (1.9)	15 (8.6)	15 (8.7)	8 (4.5)
Any infection	48 (15.0)	51 (15.9)	49 (15.3)	81 (46.3)	58 (33.7)	86 (48.9)
Serious infection	4 (1.3)	2 (0.6)	1 (0.3)	4 (2.3)	6 (3.5)	3 (1.7)
Adverse events of special interest, *n* (%)						
Malignancies (excluding nonmelanoma skin cancer)	0	0	0	0	1 (0.6)	1 (0.6)
Possible anaphylactic reaction and possible delayed hypersensitivity	1 (0.3)	0	0	0	0	0
Cardiovascular event	1 (0.3)	0	0	0	0	0
Death	0	0	1 (0.3)	0	0	0
Adverse events leading to discontinuation	NR	NR	NR	20 (11.4)	9 (5.2)	5 (2.8)
Abnormal laboratory results	NA	NA	NA	1	0	0

Abbreviations: NA = not applicable; NR = not reported; PBO = placebo; qXw = every X weeks; SC = subcutaneous; UST = ustekinumab

Safety comparison for ustekinumab: NMA

NMAs were conducted for safety outcomes in the induction phases of clinical trials. These demonstrated that overall, ustekinumab ~6 mg/kg was associated with similar likelihoods of overall AEs, SAEs and overall infections as for other active comparators. Owing to the low event counts, the NMA results for serious infections were associated with high uncertainty. The relevance of the results from these NMAs is limited because of the short time period for assessment and inherent uncertainty given low event counts for some endpoints and small sample sizes.

NMA on safety after 1 year of treatment was not conducted because of various methodological challenges. Real-world evidence (RWE) from large registries and regulatory safety data from SmPCs published by the European Medicines Agency (EMA) offer detailed assessments of safety based on data obtained across indications and may provide more meaningful insights into the comparative safety profile.

The integrated safety summary for ustekinumab demonstrated a consistent safety profile across indications (in Crohn's disease [CD], psoriasis [PsO], and psoriatic arthritis [PsA]), with low incidences of major adverse cardiovascular events, malignancies and deaths through to 1 year (≤0.5/100 patient-years) and with a comparable safety profile to placebo across indications.

A large PsO registry (PSOLAR study) has demonstrated that ustekinumab was not associated with a higher risk of serious infections and that the cumulative risk of serious infections for patients who received ustekinumab was lower than for patients who received etanercept, adalimumab or infliximab. Findings were similar for the subpopulation with concomitant IBD.

While comparing safety profiles across products is inherently difficult and subject to qualitative interpretation and weighing of different risk profiles, comparison of the SmPCs obtained after assessment of detailed data from regulators suggests that ustekinumab has an acceptable safety profile.

Patient input

Moderately to severely active UC affects a patient's quality of life in many ways. The greatest challenges involved in living with moderately to severely active UC are to remain in a good mood and to believe in better days. Current therapies are inadequate because many patients still require surgery. New therapies are needed to ensure that there are several options to choose from in cases such as intolerance and loss of response so that ever UC patient can still work and have a social and family life despite their illness. This new medicine will be important for patients because it gives new hope for better QoL.

Discussion

The included endpoints in this report are the ones that are recommended by guidelines. The evidence base for ustekinumab for the treatment of UC is supported mainly by the UNIFI clinical trial programme.

The UNIFI trials included patients reflective of a population with moderately to severely active UC who had previously experienced failure or were intolerant of conventional and/or biologic therapies, including TNF antagonists and/or vedolizumab. Specifically, UNIFI was the first trial to include vedolizumab failures in the study population. This reflects the current relevant patient population (NBF, TNFi failure and vedolizumab failure) in the real world. Compared to placebo, ustekinumab demonstrates favourable efficacy in terms of remission, response and mucosal healing for the whole patient population, independent of conventional therapy failure or BF. Mucosal healing could not be assessed at 1 year for BF patients. As with many other UC clinical trials, the UNIFI studies lack a direct comparison to active comparators (i.e., other nonconventional therapies). Moreover, the trial contained patients refractory (or intolerant) to both anti-TNF agents and vedolizumab, for which the inclusion of an active comparator is questionable. NMAs were done to allow an assessment of the relative efficacy of ustekinumab compared to active comparators.

The evidence drawn from all the NMAs was based on indirect data: Bayesian NMAs in a setting in which network heterogeneity or inconsistency could not be statistically assessed because of data paucity, and assessment of statistical heterogeneity between trials informing a particular comparison was limited and probably underpowered. Even in the networks that included a head-to-head comparison, inconsistency assessment between different sources of evidence was not performed

because of data imputation in one of the trials informing a closed loop. As a network is most justifiable under an assumption of consistency between different sources of evidence, lack of these data largely increases uncertainty regarding the NMA results. In addition, given the reasons stressed below, it is not certain if the key assumption for NMA (that the different sets of studies included in the analysis should be on average similar in all important factors that may affect the relative effects) is met.

The NMA included all relevant comparators used for the treatment of UC. Additionally, the MAH included infliximab 10mg/kg comparator which is not the recommended regimen for UC, but the agent is recommended for Crohn's disease and a dose escalation with infliximab to 10mg/kg is apparently common practice in Europe.

The list of endpoints available for the NMA could be drawn from the table in the submission file reporting on characteristics of the studies included, but judging from the data presented for the UNIFI trial, not all the endpoints are listed. Specifically, the major secondary endpoint of corticosteroid-free clinical remission was not considered for NMAs, which increases the risk of bias in selection of the result reported for the indirect comparison analysis (as discordant pieces of evidence possibly drawn from a relevant secondary endpoint may not be able to inform decision-making if available but excluded from analysis).

A potential source of heterogeneity due to differences in outcome measurements was observed for the mucosal healing endpoint, as the endoscopic score for the efficacy analyses used in the NMA was assessed by a local endoscopist in six out of eight studies (the VARSITY trial did not report who performed the readings, while readings were performed centrally in OCTAVE), and no information on whether standardisation of local reading was performed in these studies was presented by the MAH.

The results for indirect comparisons, performed via Bayesian NMA for different endpoints and subgroups of patients and run separately for induction and maintenance trials, showed that the CrIs for most endpoints were quite wide, indicating high uncertainties in the results. While this is partly expected as the assessment is based on indirect comparisons, which contribute additional uncertainty, in some models the CrI width hampers interpretation. For example, in the case of NMA results for clinical remission among BF patients during induction, the upper CrI bound for the OR reaches a value of 95, which seriously undermines confidence in results for such a model.

Regarding the ranking of treatments, the MAH presented both SUCRA values and the probability of ustekinumab being better than a comparator. However, in this submission all the networks had very wide CrIs. It is known that the probability of a treatment being better underperforms when the treatment effect of a intervention is largely uncertain, as an intervention with the most uncertainty can have the greatest probability. Therefore, the probability of ustekinumab being better was not used in interpretation. The Cochrane handbook recommends the use of SUCRA rankings to avoid biased data analysis. Nonetheless, even with SUCRA values there were concerns regarding the interpretation of rankings (due to scarce data, heavy imputation for 1-year NMAs, indirect data, statistically confirmed or probable heterogeneity analysed via fixed-effects NMA) and are by definition based on any differences among the treatment effects, no matter how small. Therefore, because of the large uncertainty for effect sizes and the few significant effects, it is possible that a treatment with a higher rank is better than another treatment by only a small, not clinically meaningful amount. Consequently, while the ustekinumab-based interventions in the maintenance NMAs seemed to be consistently ranked at the top, we are uncertain about the described ranking as they could be misleading.

Overall, trials included in the body of evidence for NMAs were judged as comparable at baseline in terms of patient characteristics. However, owing to the paucity of the data, statistical assessment of heterogeneity in the distribution of potential effect modifiers across comparisons was not possible. Instead, heterogeneity between individual trials in terms of several baseline demographic and clinical characteristics was assessed by comparing descriptive statistics, while data for mixed populations (including BF and NBF patients) and different outcome measures (mean and median) were also reported. In addition, data were not reported for all trials and for both subgroups, which increases uncertainty regarding the conclusion. The data for BF patients were particularly under-represented in these analyses, so it was not possible to draw valid conclusion on the comparability of baseline characteristics across trials for this subgroup. The same was true for the maintenance trials, as only two studies (OCTAVE-S and UNIFI) reported baseline characteristics at the onset of the maintenance phase. Therefore, evidence of homogeneity among the trials at baseline is limited and is applicable only to the NBF population in induction trials.

Time points differed across trials for induction and maintenance phases. As partial Mayo scores were consistent within these time ranges in several trials, it was concluded that these time points are equivalent in terms of clinical efficacy. This was shown for the maintenance phase in the UNIFI, PURSUIT and GEMINI trials, and for CD patients in the UNITI (ustekinumab) and CHARM (adalimumab) trials, as well as for the induction phase in the UNIFI trial. However, deviation from this pattern was in an induction-phase trial among CD patients treated with adalimumab, as the largest effect for clinical remission was observed at 6 weeks for 40 mg every other week (EOW) and at 8 weeks for 40 mg weekly. As adalimumab was also one of few drugs that demonstrated significantly worse efficacy than ustekinumab in induction-phase NMAs for both subgroups, there is a possibility that dose and/or length of the induction period (6–8 weeks) may be suboptimal for adalimumab for assessment of its true clinical remission in the induction phase.

The MAH identified prior TNFi exposure as the source of heterogeneity and performed analyses separately for NBF and BF patients to minimise this. However, increasingly selective criteria for RCTs of new biologics has been changing the definition of target patient populations, which may affect the adequacy of a common comparator in UC trials. Patients have been enrolled in these trials as biologic-naïve or after failure of a particular TNFi agent, any TNFi agent or any TNFi and/or vedolizumab. In other words, more recent trials enrolled more specific patient cohorts (those with failure/intolerance to a particular TNFi drug, any TNFi drug, or any TNFi agent and/or vedolizumab) with increasing probability that patients have experienced failure of more than one biologic treatment, which means that BF patient populations might not be comparable. Moreover, considering the observation that TNFi agents have lower clinical efficacy and mucosal healing rates with increasing disease duration [5,6,7] several theories have been postulated that corroborate this view. This impels heterogeneity among BF populations.

Moreover, the issue of population heterogeneity among UC trials and sample representativness has recently become even more critical, as many trials investigating new biologics in UC fail to study representative patient populations. Specifically, in a market in which several drugs are available for prescription, assignment to the placebo arm of a long-term study is not very appealing to patients. Consequently, the proportion of patients available in centres that are recruiting to newer clinical trials is so low that the results can no longer be regarded as representative. Ghosh et al. [8] illustrated this difference in populations of patients with CD, a closely related disease, by comparing the time needed to enrol and/or complete a given IBD trial, which has changed dramatically: “In 1999, the 54-week ACCENT 1 trial enrolled 573 anti-TNF therapy naïve patients with CD from 55 centres in 12 countries and was completed in approximately 2 years. In contrast, by 2008 the 52-week GEMINI 2 trial required 1115 patients with CD (62% prior TNFi exposed) from 285 centres in 39 countries and took almost 3.5 years to complete”. A similar difference in patient populations probably applies to UC. With that said, the data from earlier clinical trials on biologics are not likely to be comparable with more recently conducted RCTs, as they include different populations of patients.

The designs for induction-phase studies were consistent across the trials, allowing a standard NMA approach to be taken.

In the maintenance phase, however, study design was identified as a substantial source of heterogeneity. To make the maintenance endpoint data between the designs comparable, treatment effects in re-randomised trials were re-calculated using a treat-through approach. This approach allows randomisation to be maintained as the treatment arms are based on induction therapy received by the randomised patients. However, recalculation of the total number of patients in the re-randomised responder trials reduced the sample sizes.

The data from earlier clinical trials on biologics are not likely to be comparable with more recently conducted RCTs as they include different populations of patients. Since a valid NMA relies on the assumption that the different sets of studies included in the analysis are on average similar in all important factors that may affect the relative effects, for these NMAs it is very uncertain if the results are valid or are due to bias/artefacts of analysis.

The NMA for 1-year regimens provides ambiguous results from evidence of very low certainty of the treatment effects via treat-through comparisons. Adding to concerns about meeting the key assumption of a valid meta-analysis, the NMA models for both subgroups were built on scarce, indirect and heavily imputed data in a setting in which heterogeneity is expected but could not be tested owing to paucity of data, and by using fixed-effects models as the only available model option. Differences observed among registrational clinical trials of biological therapies in adults with IBDs and in meta-analysis of placebo remission rates in UC patients with active disease

showed significant heterogeneity of these rates, ranging from 0% to 40% [8,9], it is reasonable to assume some level of heterogeneity between the trials included in NMAs. It is acknowledged that a random approach was not feasible because of the data paucity, but the MAH should nevertheless clearly indicate which method was planned. Finally, there is inconsistency in the submission: while the Methods and Results sections state that DIC was used to select models, the Data limitations section states that random-effects models are preferred over fixed-effects models.

The 1-year NMA estimates for both subgroups of patients had very wide CrIs that hindered interpretation of the results. In addition, owing to high uncertainty and the inability to precisely estimate differences in effect sizes between the treatments, the probability of ustekinumab being better and the SUCRA rankings could not be used as valid indicators of treatment efficacy.

For the NMAs for NBF population in the maintenance phase, additional limitation was identified. As apparently no dose–response relationship was observed, the data for interventions with different doses of the same drug were pooled for treatment arms. The claim that no dose–response relationship exists in these particular networks (whereas the opposite was observed in all the others, including BF patients at 1 year, suggesting an underlying biological phenomenon), which is supported only by observation that higher dose/shorter interval between doses led to higher clinical rates (no statistical testing was performed), does not represent convincing evidence. It is important to note that while pooling of doses might increase the analysis power, as stressed by the MAH, it will also result in overprecise results (as data from two active treatment arms of a trial are pooled together and the sample size of then active arm is increased twice) which could affect comparison with a study having just one active arm (ULTRA 2) and the fixed-effect NMA results [10].

For the NMAs for the NBF population in the maintenance phase, all the point estimates for all the comparisons, irrespective of their significance, point towards higher odds of meeting an endpoint with ustekinumab compared to other active treatments or placebo. While the results seemed to support at least moderate effect of ustekinumab to placebo for all three endpoints, for comparisons to other active treatments, only for comparison of clinical response to golimumab or adalimumab the effect was found significant and the trivial effect could be excluded based on CrIs. Nevertheless, taking into consideration limitations of 1-year NMAs stressed above and uncertainity of the results stemming from non-significant and/or wide CrIs including trivial the evidence base is too limited and we are uncertain if ustekinumab improves clinical response when compared to golimumab or adalimumab.

For BF patients, in addition to the general limitations of maintenance NMAs already stated, the 1-year NMA models were based on even more scarce data with smaller sample sizes and smaller numbers of events in study arms compared to the NBF NMAs. At the same time, in line with what has been discussed on sources of heterogeneity, it was more likely that this population included heterogeneous trials. The eligibility criteria for BF included failure to TNFi agents or vedolizumab for UNIFI, but for the other trials the criteria only included failure to TNFi agents. However, the trials for this body of evidence were not assessed for initial imbalances in BF patients since they mainly presented baseline data for mixed populations. In addition, placebo imputations for the 1-year NMAs were based on less robust data compared to the data used in the NBF group (only the UNIFI trial was used to impute induction responders to placebo in this population), while results for mucosal healing in the BF population were not available as no data from GEMINI 1 were available to inform imputation of data for induction nonresponders (required for UNIFI, PURSUIT and OCTAVE). For the other two endpoints, 1-year NMA of efficacy showed significantly higher odds of achieving a clinical response and clinical remission when compared to placebo for both ustekinumab regimes (~6 mg/kg → 90 mg q8w for induction responders and nonresponders; and ~6 mg/kg → 90 mg q12w for induction responders + 90 mg q8w for induction nonresponders). The point estimates were high, with ORs of 5.3–6.9 for clinical remission and 4.8 for clinical response, while the CrIs excluded trivial effects (OR ranged from ~2 to ≥9, low certainty evidence. None of the active treatments demonstrated a significant difference to ustekinumab. Although, similar to NBF patients, all point estimates for BF patients pointed towards higher odds of achieving the endpoint with ustekinumab, the main conclusion is that the evidence base is insufficient to make inferences.

Additional limitations were also observed. The UNIFI trial grouped patients according to BF and NBF, whereas other trials grouped patients according to TNFi-naïve/experienced or biologic-naïve/experienced. However, only a small proportion of patients in the UNIFI trial were NBF with previous exposure to a biologic therapy (27 out of 961 patients [2.8%] across the induction study

arms in the primary efficacy analysis). A number of limitations are associated with the analysis of maintenance data for the 1-year NMA specifically. Some of the limitations faced in the analysis were considered to bias results against ustekinumab; OCTAVE trial re-randomises both placebo and tofacitinib induction responders and estimated clinical remission rate from their combined induction arms; delayed responders had to respond at week 14 in PRSUIT and 16 in UNIFI which potentially underestimates the number of delayed responders included in the golimumab and ustekinumab trial arms; OCTAVE included an open-label treatment for induction non-responders. However the fact that only UST data (UNIFI trial) with the highest placebo response rate among non-biologic failure patients at 1-year were used to impute the data for biologic failure population might have also driven the results the other way.

Conclusion

Direct evidence

Conclusions from trials of ustekinumab versus active treatment

No head-to-head studies of ustekinumab versus other active therapies in UC have been conducted.

Conclusions from trials of ustekinumab versus placebo

Clinical efficacy

The UNIFI randomised-control trials (induction and maintenance) provided statistically significant results that ustekinumab at both doses was effective at inducing and maintaining a clinical response, clinical remission, endoscopic healing and mucosal healing (a combination of endoscopic and histologic healing), reducing the inflammatory burden and improving health-related quality of life in a population of subjects with moderately to severely active UC who had previously experienced failure or were intolerant of conventional and/or biologic therapy. This included a large proportion of patients who responded between week 8 and week 16. Overall, 56% of patients without a clinical response at week 8 had a clinical response at week 16 after receiving ustekinumab 90 mg SC 8 weeks after induction.

We are uncertain in the long-term efficacy since no efficacy data is available beyond 1 year.

Treatment of subjects with prior failure to biological therapy represents a key goal and an unmet need. Maintenance of efficacy has been demonstrated for subjects who have experienced failure of either anti-TNF or vedolizumab but there were very few patients enrolled in the pivotal trial with failure of both. For patients with failure to anti-TNF and vedolizumab the clinical remission was achieved in 22.7%, 33.3% and 14.8% of those on ustekinumab 90 mg q12w, ustekinumab 90 mg q8w and placebo, respectively but the results did not reach statistical significance. The study was not powered to detect a difference in this subgroup and the numbere were fewer than ten patients in each group.

Clinical safety

Adverse event rates in the UNIFI studies were similar in the ustekinumab and placebo arms over the induction and maintenance study phases in the population of adult subjects with moderately to severely active ulcerative colitis. The MAH identified a new non-serious adverse drug reaction of sinusitis (at a frequency of common [≥1/100 and <1/10]). The rate of serious adverse event did not relevantly differ between treatment groups in the induction and maintenance studies. AEs of special interest were similar across treatment groups, including infections, malignancy and cardiovascular events.

It should be mentioned that the sample was too small to conclude that these frequencies were similar, and follow-up was too short to draw conclusion on the risk on malignancy.

The integrated safety summary for ustekinumab revealed a similar safety profile across indications (in patients with psoriasis, psoriatic arthritis, Crohn's disease and ulcerative colitis). Incidences of major cardiovascular AE, malignancies and deaths through to 1 year were ≤0.5/100 patient-years and a comparable safety profile to placebo across indications. A large registry in

psoriasis (PSOLAR study) has demonstrated that ustekinumab was not associated with a higher risk of serious infections and that the cumulative risk of serious infections for patients who received ustekinumab was lower than for patients who received etanercept, adalimumab or infliximab. Findings were similar for the subpopulation with concomitant IBD.

Indirect evidence: NMAs

For rating the quality or certainty of the evidence (as high, moderate, low and very low certainty) from a network meta-analysis, the Grading of Recommendations Assessment, Development and Evaluation (GRADE)-method was not applied, by either the MAH or the authors of this joint assessment.

Conclusions on clinical effectiveness from 1-year NMA

Non-biologic failure (NBF) population

Ustekinumab versus placebo, golimumab and adalimumab

The NMAs in the NBF population suggested that patients who received ustekinumab ~6 mg/kg for induction followed by ustekinumab 90 mg had statistically significant higher odds of achieving a clinical response, clinical remission, and mucosal healing after a 1-year long treatment regimen when compared to placebo, golimumab and adalimumab.

While the point estimates for these comparisons were high, ranging from OR of 2.4 to 8.7, the findings need to be viewed with caution. Because the evidence base was very limited and we could not exclude that effects on clinical remission and mucosal healing for comparisons to golimumab and adalimumab are trivial, we are uncertain if ustekinumab is more effective than these two active treatments mentioned above.

Ustekinumab versus infliximumab and tofacitinib

In addition, NMAs suggested that a statistically significant difference was observed in a clinical response, but not in endpoints of clinical remission and mucosal healing, in patients treated with ustekinumab than for those treated with infliximab or tofacitinib. However, it should be stressed that trivial effects could not be excluded and that evidence based was very limited, which make us uncertain if ustekinumab is more effective in improving clinical response than these two drugs.

Ustekinumab versus vedolizumab

NMAs suggested that no statistically significant difference was observed in a clinical response, clinical remission, and mucosal healing in patients treated with ustekinumab when compared with vedolizumab.

Overall, given that various methodological issues related to maintenance models for non-biological failure patients were raised and discussed, such as the paucity of the data, indirectness and the fact that inconsistency between different sources of evidence was not estimated, that data were recalculated and imputed in majority of trials (either for induction responders and/or non-responders in placebo or active arm), that fixed effect model was used in a setting where heterogeneity is expected but could not be tested, and that there is a potential impact of bias due to pooling of doses, we are uncertain if ustekinumab improves clinical response, clinical remission, and mucosal healing after a 1-year treatment regimen when compared with golimumab and adalimumab, as well as clinical response when compared with infliximab and tofacitinib.

Biological failure (BF) population

Ustekinumab versus placebo

The 1-year NMA of effectiveness for the BF group suggests significantly higher odds of achieving a clinical response and clinical remission in comparison to placebo with both ustekinumab regimes. While the point estimates for these comparisons were high, with ORs of 5.3-6.9 for clinical remission and 4.8 for clinical response endpoint, and wide CrIs (which ranged from OR ~2 to ≥9) suggested that at ustekinumab may have moderate or higher effect, the findings need to be viewed in light of very limited evidence base.

Ustekinumab versus golimumab, adalimumab, infliximab, vedolizumab and tofacitinib

All other comparisons for effectiveness in terms of clinical response or clinical remission between ustekinumab and any active treatments were nonsignificant with wide CrIs, demonstrating that the evidence base is limited. Although SUCRA values suggested that ustekinumab ~6 mg/kg-ustekinumab 90 mg SC q8W early + ustekinumab 90 mg SC q8W delayed is more effective than other active treatments, we are uncertain in this result due to various methodological issues already raised and discussed.

Overall, given that NMA of mucosal healing in the BF group was not feasible as imputation data needed for placebo were not available in this population, and taking into account various methodological issues for the maintenance models for BF patients (even more scarce data than in NMAs in non-biologic failure patients, indirectness and the fact that inconsistency between different sources of evidence was not estimated, the fact that data were recalculated and imputed in majority of trials, that fixed effect model was used in a setting where heterogeneity is expected but could not be tested), we are uncertain in these results.

Conclusions on clinical safety from NMA

Induction NMA

NMAs were conducted for safety outcomes in the induction phases of clinical trials. These suggested that ustekinumab 6mg/kg was associated with similar likelihoods of overall adverse events, serious adverse events and overall infections as for other active comparators. Owing to the low event counts, we are uncertain in the NMA results for serious infections.

The relevance of the results from these NMAs is limited by the short time period for assessment and the low event counts for some endpoints and small sample sizes.

1-year NMA

No data on long-term relative safety after 1 year are available because various methodological challenges limit the ability to conduct an NMA on safety after 1 year of treatment.

For the long-term safety (> 1 year) of ustekinumab in UC, further evidence is needed.

1 BACKGROUND

1.1 Overview of the disease or health condition

Pathoaetiology

UC (International Classification of Diseases-10 code K51.90) is a lifelong, chronic, idiopathic inflammatory bowel disease with a remitting and relapsing course that can progress from asymptomatic mild inflammation to extensive inflammation of the colon [11]

The exact aetiology of UC remains unknown, but it has been suggested that a combination of genetic factors, abnormal immune responses and environmental triggers can play an important role in its development. Approximately 8–14% of patients with UC have a family history of IBD. First-degree relatives have four times the risk of developing the disease [12]. To date, more than 200 genetic loci have been identified in genome-wide association studies as contributing to susceptibility to IBDs, including UC. However, genetics cannot completely explain the high incidence and prevalence of UC. The increasing incidence of UC worldwide suggests the importance of environmental factors in its development [13].

The pathogenesis of the inflammatory lesions appears to be due to dysregulation of the gut immune system. Inflammation of the colonic mucosa is responsible for the signs, symptoms, course and complications of UC [14].

Mucosal inflammation can extend from the rectum to the more proximal colon, with variable extents. A patient may have proctitis, left-sided colitis (with a proximal limit below the splenic flexure), extensive colitis (involving the transverse colon) or pancolitis. The majority (80%) of adult patients with newly diagnosed UC present with disease limited to the distal or left side of the colon [15]. Approximately 20% of UC patients present with extensive colitis or pancolitis [16].

Symptoms

The symptoms of UC are diverse, depending on the extent of the disease, and can be severe, with a profound impact on patients' lives. Signs and symptoms may include frequent bloody stools, colonic motility dysfunction, potentially permanent fibrosis, tissue damage and systemic symptoms such as tiredness, weight loss and fever [11]. In general, greater UC severity and extent are associated with worsening bloody diarrhoea and the development of systemic signs. In patients with proctitis, urgency and tenesmus might be more prominent. Patients with pancolitis may experience more frequently bloody diarrhoea and abdominal pain. Up to 10% of patients with proctitis or left-sided colitis can suffer from paradoxical constipation. In extensive or severe disease, more general symptoms such as weight loss and fatigue are often present and can be accompanied by nausea, vomiting and fever [11]. Nearly 70% of patients experience UC symptom flares every few months. Three-quarters of patients (75.6%) reported that symptoms limit their ability to enjoy leisure activities, while 68.9% admitted that UC symptoms affect their ability to perform at work [17]. Progression of disease extension is associated with poor prognosis [18].

Diagnosis

No gold standard method for diagnosing UC is available. Diagnosis is based on the combination of medical history, endoscopic findings on colonoscopy and histological findings. Endoscopic evidence of continuous colonic inflammation and histological findings of distortion of crypt architecture with shortening and disarray of the crypts or crypt atrophy are the key features and are essential for proper diagnosis of UC. Furthermore, endoscopy plays a substantial role in assessing disease severity because endoscopic healing is related to better remission rates and a lower risk of colectomy [11]. A variety of laboratory tests are also used, including blood tests, stool samples, comprehensive metabolic profile and CRP measurement. Laboratory abnormalities are more frequent with increasing severity and extent of disease; however, normal levels do not exclude disease activity. Many studies have assessed the utility of serologic markers in the diagnosis of UC, but the data indicate that no serologic markers are sensitive or specific enough to allow a diagnosis of UC [19,20].

Prevalence

UC affects about the same number of women and men. The peak age for UC onset is between 30 and 40 years, but UC can occur at any age [13]. The highest annual incidence of UC has been reported as 24.3 per 100,000 person-years in Europe, 6.3 per 100,000 person-years in Asia and the Middle East, and 19.2 per 100,000 person-years in North America [21,22]. However, it has been demonstrated that the incidence and prevalence of UC are highest in western countries, specifically in Northern Europe and Canada [21,22]. A time-trend analysis showed that 60% of studies on UC revealed a rising incidence that was statistically significant [23].

Mortality

Standardised mortality rates among patients with UC do not appear to be elevated overall when compared with the general population. A meta-analysis of 22 studies could not detect a higher risk of death among UC patents compared with community controls. The authors demonstrated that 14 out of 22 studies showed similar survival for patients with UC and the general population [24]. However, mortality is higher for elderly patients and those who have more comorbidities. Perioperative and postoperative mortality are also associated with a higher total mortality risk for patients with UC [18,25]. Furthermore, it has been demonstrated that mortality among UC patients correlates with infection-related hospitalisations and increasing extent of colitis, and that patients with UC are more likely to have a disability that prevents them from working [18,26,27].

Treatment

As in many chronic diseases, an appropriate plan of management must be tailored to the individual patient when treating UC. The choice of treatment depends on disease activity status, distribution (proctitis, left-sided or extensive colitis), previous medications and the course of the disease [15]. The ultimate goal of UC treatment is to induce remission and then maintain remission, as well as to reduce the risk of long-term complications and surgery [17]. Medications to treat UC include aminosalicylates (5-ASAs), corticosteroids, immune modifiers (immunomodulators), biologics, and over-the-counter medications such as antidiarrhoeals and pain relievers. Despite better treatment options, long-term colectomy rates have not declined over a 10-year period [14], highlighting the urgent demand for new therapeutic agents.

1.2 Current clinical practice

At the European level, European Crohn's and Colitis Organisation (ECCO) developed a third version of their consensus guideline in 2017 that addresses UC [28]. This latest version of the European UC recommendations was drafted by 28 ECCO members from 14 European countries. Evidence levels (ELs) are based on the Oxford Centre for Evidence-Based Medicine.

According to the guidelines, the treatment strategy for UC is mainly based on the severity, distribution (proctitis, left-sided, extensive) and pattern of disease. The latter includes relapse frequency, disease course, response to previous medications, side effects of medication and extra-intestinal manifestations. Age at onset and disease duration are also important factors. The guidelines highlight that it is important to distinguish patients with severe UC necessitating hospital admission from those with mildly or moderately active disease who can be managed as outpatients. The technology under assessment is dedicated to patents with moderately to severely active UC.

In the case of proctitis, a once-daily suppository of mesalamine 1 g is the preferred initial treatment for mildly or moderately active disease [EL1]. Refractory proctitis may require treatment with systemic steroids, immunosuppressants and/or biologics [EL4].

Mildly to moderately active left-sided UC should initially be treated with an ASA enema at ≥1 g/day [EL1] combined with oral mesalamine at ≥2.4 g/day [EL1], which is more effective than oral or topical ASAs or topical steroids alone [EL1]. Topical mesalamine is more effective than topical steroids [EL1].

In patients with mild to moderate active extensive UC, treatment should be initiated with an ASA enema at 1 g/day [EL1] combined with oral mesalamine at ≥2.4 g/ day [EL1]. Systemic corticosteroids are appropriate for patients with moderate to severe activity and those with mild activity who do not respond to mesalamine.

The initial treatment recommended for severe active UC is intravenous steroids [EL1]. Monotherapy with intravenous cyclosporin [EL2] is an alternative, especially in cases of SAEs due to steroids. Patients with a severe condition (bloody diarrhoea ≥6 times/day and any signs of systemic toxicity, i.e. pulse >90 beats/min, temperature >37.8°C, haemoglobin <105 g/l, erythrocyte sedimentation rate >30 mm/h or CRP >30 mg/l) should be admitted to hospital for intensive treatment [EL4]. Patients with comorbidities or aged >60 years have a higher risk of mortality [EL3].

In patients who do not respond to intravenous steroids, treatment options including cyclosporin [EL1], infliximab [EL1], tacrolimus [EL2] and surgery should be considered. Colectomy is recommended if there is no improvement following 4–7 days of salvage therapy [EL4]. Patients with steroid-dependent disease should be treated with a thiopurine [EL2], TNFi agent [EL1] (preferably combined with thiopurines, at least for infliximab [EL2]), vedolizumab [EL2] or methotrexate (MTX) [EL2]. In the case of treatment failure, second-line medical therapy with an alternative TNFi drug [EL4] or vedolizumab [EL2] or colectomy [EL5] should be considered.

Moderate disease refractory to oral steroids should be treated with either intravenous steroids [EL4] or a TNFi [EL1], preferably combined with thiopurines, at least for infliximab [EL2], vedolizumab [EL2], or tacrolimus [EL2]. Second-line medical therapy with a different TNFi agent [EL4] or vedolizumab [EL2] may be an option; colectomy should also be considered.

Patients with moderate colitis refractory to thiopurines should be treated with a TNFi [EL1], preferably combined with thiopurines, at least for infliximab [EL2], or vedolizumab [EL2]. In cases of treatment failure, a different TNFi drug [EL4] or vedolizumab [EL2] should be considered, and colectomy is recommended if further medical therapy does not yield a clear clinical benefit [EL5].

Maintenance of remission

The goal of maintenance therapy in UC is to maintain steroid-free remission, defined clinically [EL1] and endoscopically [EL2]. The choice of maintenance treatment is determined by disease extent [EL1], disease course (frequency and intensity of flares) [EL5], failure and AEs of previous maintenance treatment [EL5], severity of the most recent flare [EL5], treatment used for inducing remission during the most recent flare [EL5], safety of maintenance treatment [EL1] and cancer prevention [EL2].

Options for a stepwise escalation of maintenance therapy include dose escalation of oral/rectal ASAs [EL1], addition of thiopurines [EL2] and TNFi therapy or vedolizumab [EL1]. For patients responding to mesalamine or steroids (oral or rectal) [EL1] mesalamine compounds are the first-line maintenance treatment.

For patients responding to a TNFi agent, maintaining remission by continuing TNFi therapy with or without thiopurines [EL1] is appropriate. Thiopurine maintenance is an alternative option [EL3].

A TNFi or vedolizumab may be used as first-line biologic therapy. Vedolizumab is effective in patients with TNFi failure [EL2]. For patients responding to vedolizumab, maintenance therapy with vedolizumab is appropriate [EL2].

For thiopurine-naïve patients with severe colitis responding to steroids, cyclosporin or tacrolimus, thiopurines are appropriate for maintaining remission [EL2]. Patients responding to infliximab should continue infliximab with or without thiopurines [EL2]; thiopurine maintenance is an alternative option [EL4].

The ECCO guidelines do not mention tofacitinib or ustekinumab; however, market authorisation was only granted for tofacitinib in March 2017 and for ustekinumab in September 2019.

According to the marketing authorisation, ustekinumab is to be used at the same treatment stages as TNFi therapy and vedolizumab. The anticipated approach for using ustekinumab (and tofacitinib) if it is incorporated into the ECCO clinical care pathway is presented in Figure 1.1 and Figure 1.2.

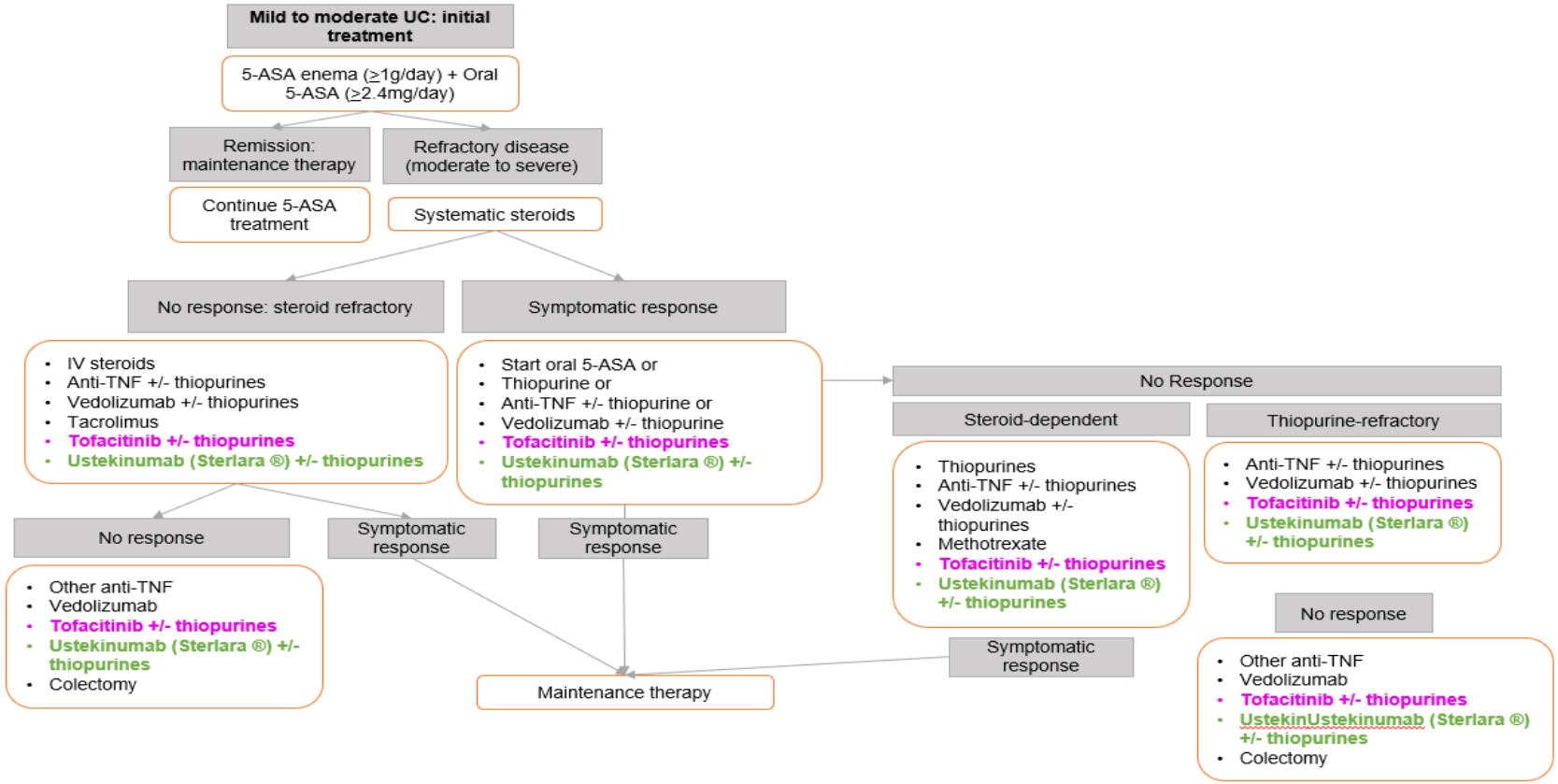


Figure 1.1. Anticipated approach for using ustekinumab (and tofacitinib; differentiated by colour) if added to the ECCO clinical care pathway for mild to moderate UC.

Abbreviations: ASA = aminosalicylic acid; ECCO = European Crohn's and Colitis Organisation; IV = intravenous, TNF = tumour necrosis factor; UC = ulcerative colitis.
Colour code: purple text = comparator products that have not been included in the latest ECCO guidelines; green text = technology under assessment that has not been included in the latest EC-CO guidelines.

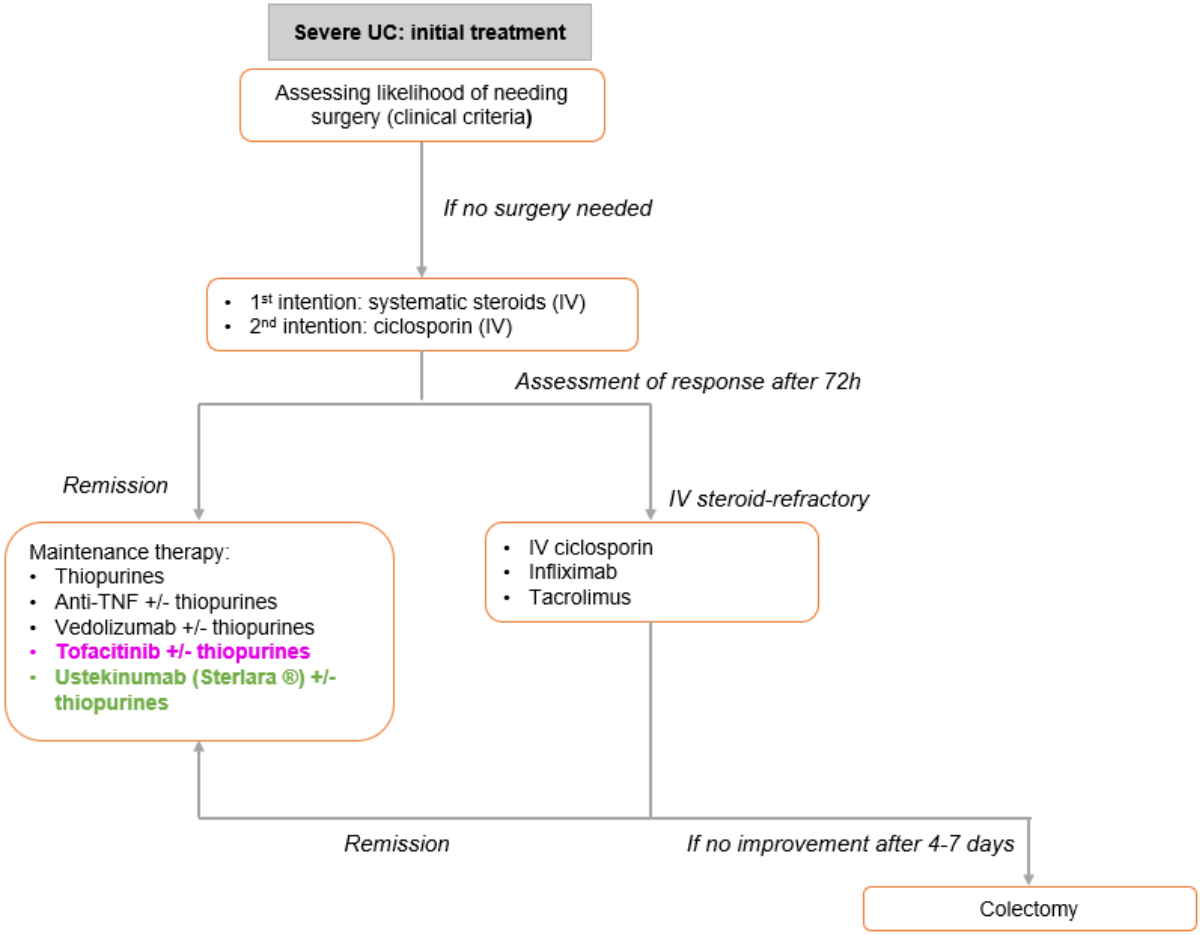


Figure 1.2. Anticipated approach for using ustekinumab (and tofacitinib; differentiated by colour) if added to the ECCO clinical care pathway for severely active UC.

Abbreviations: ASA = aminosalicylic acid; ECCO = European Crohn's and Colitis Organisation; IV = intravenous; TNF = tumour necrosis factor; UC = ulcerative colitis.
Colour code: purple text = comparator products that have not been included in the latest ECCO guidelines; green text = technology under assessment that has not been included in the latest ECCO guidelines
Note: the position of treatments listed in the boxes are not in any particular order.

Global guidelines updated by the World Gastroenterology Organisation in 2015 are widely consistent with the European approach. This guideline lists ASAs and corticosteroids as options to induce UC remission, while thiopurines such as azathioprine and mercaptopurine are mentioned as effective for maintenance of remission induced by corticosteroids. The use of methotrexate in patients with UC is a matter of debate. Use of cyclosporin or tacrolimus should be limited to acute (corticosteroid-refractory) severe colitis and tacrolimus should be used only if other therapies have failed. According to the guidelines, TNFi agents such as infliximab and adalimumab can be considered for acute severe colitis or moderately severe corticosteroid-dependent or corticosteroid-resistant colitis. Infliximab or vedolizumab intravenously and adalimumab or golimumab SC are also options for ambulatory patients with moderate to severe disease.

Ustekinumab is to be available as a treatment in parallel to TNFi agents and vedolizumab. Tofacitinib was not available at the time of publication of the World Gastroenterology Organisation guidelines.

An overview of country-specific guidelines presented by the company and available in the English language (from the UK, Spain, Switzerland and Italy) proves that the general rules for UC management are similar in these countries and differences seem to arise from the availability of different treatments at the time of guideline publication (Appendix, Table A1).

1.3 Features of the intervention

The features of the intervention and its comparators are summarised in Table 1.1. Administration and dosing of the technology and its comparators are summarised in Table 1.2.

Ustekinumab

Ustekinumab is a human monoclonal antibody that binds with specificity to the cytokines IL-12 and IL-23. Ustekinumab may exert its clinical effects in UC via interruption of the T-helper 1 and T-helper 17 cytokine pathways, which are central to the pathology of this disease. Extension of the indication for ustekinumab was approved to include treatment of adult patients with moderately to severely active UC who have had an inadequate response with, lost response to or were intolerant to either conventional therapy or a biologic or have medical contraindications to such therapies [29].

Comparators

Biologic therapies are treatment options for patients with moderately to severely active UC for whom conventional therapy or prior biologic treatment has failed. Several possible relevant comparators in the European setting have been identified according to recommendations in guidelines.

Adalimumab (Humira), infliximab (Remicade) and golimumab (Simponi) are three TNF inhibitors currently approved for the treatment of adult patients with moderately to severely active UC who have had an inadequate response with, lost response to or were intolerant to either conventional therapy or a biologic or have medical contraindications to such therapies [30,31,32]. There are also several agents biosimilar to infliximab and adalimumab available for the treatment of UC.

Vedolizumab (Entyvio; anti-integrin therapy) and tofacitinib (Xeljanz; oral Janus kinase inhibitor) are the only treatments approved for both bio-naïve and BF populations [33,34]. The label for vedolizumab label states that it is indicated for patients with previous TNFi failure, while for tofacitinib any prior biologic therapy is allowed.

Table 1.1. Features of the intervention and its comparators

Nonproprietary name	Ustekinumab	Adalimumab	Infliximab	Golimumab	Vedolizumab	Tofacitinib
Proprietary name	Stelara	Humira (reference product) Biosimilars: Amgevita Imraldi Hyrimoz Kromeya Idacio Halimatoz Hefiya Hulio	Remicade (reference product) Biosimilars: Remsina Zessly Flixabi Inflectra	Simponi	Entyvio	Xeljanz
Registered EMA indication	UC: for the treatment of moderately to severely active UC in adult patients who have had an inadequate response with, lost response to or were intolerant to either conventional therapy or a biologic or have medical contraindications to such therapies Ustekinumab is also indicated for: Plaque psoriasis in adult and adolescent patients Psoriatic arthritis Crohn's disease	UC: for the treatment of moderately to severely active UC in adult patients who have had an inadequate response to conventional therapy including corticosteroids and 6-MP or AZA, or who are intolerant to or have medical contraindications for such therapies. Adalimumab is also indicated for: Rheumatoid arthritis Juvenile idiopathic arthritis: polyarticular juvenile idiopathic arthritis and active enthesitis Psoriatic arthritis Plaque psoriasis in adult and paediatric patients	UC: for the treatment of moderately to severely active UC in adult patients who have had an inadequate response to conventional therapy including corticosteroids and 6-MP or AZA, or who are intolerant to or have medical contraindications for such therapies. Infliximab is also indicated for: Rheumatoid arthritis Adult Crohn's disease Paediatric Crohn's disease Paediatric UC Ankylosing spondylitis Psoriatic arthritis Psoriasis	UC: for the treatment of moderately to severely active UC in adult patients who have had an inadequate response to conventional therapy including corticosteroids and 6-MP or AZA, or who are intolerant to or have medical contraindications for such therapies. Golimumab is also indicated for: Polyarticular juvenile idiopathic arthritis Axial spondyloarthritis Psoriatic arthritis Rheumatoid arthritis	UC: for the treatment of moderately to severely active UC in adult patients who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a TNFα antagonist. Vedolizumab is also indicated for Crohn's disease	UC: for the treatment of moderately to severely active UC in adult patients who have had an inadequate response, lost response, or were intolerant to either conventional therapy or a biologic agent. Tofacitinib is also indicated for: Rheumatoid arthritis Psoriatic arthritis

Nonproprietary name	Ustekinumab	Adalimumab	Infliximab	Golimumab	Vedolizumab	Tofacitinib
		Axial spondyloarthritis Crohn's disease in adult and paediatric patients Hidradenitis suppurativa (acne inversa) Uveitis				
Prospective marketing authorisation holder	Janssen-Cilag International NV	AbbVie Deutschland GmbH & Co. KG Biosimilars: Amgen Europe B.V., Samsung Bioepis NL B.V., Sandoz GmbH, Fresenius Kabi Deutschland GmbH, Boehringer Ingelheim International GmbH	Janssen Biologics B.V. Biosimilars: Celltrion Healthcare Hungary Kft, Sandoz GmbH, Samsung Bioepis NL B.V., Pfizer Europe MA EEIG	Janssen Biologics B.V.	Takeda Pharma A/S	Pfizer Europe MA EEIG
Contraindications	Hypersensitivity to the active substance or to any of the excipients Clinically important, active infection (e.g., active tuberculosis)	Hypersensitivity to the active substance or to any of the excipients Active tuberculosis or other severe infections such as sepsis or opportunistic infections Moderate or severe heart failure (New York Heart Association class III/IV)			Hypersensitivity to the active substance or to any of the excipients Active severe infections such as tuberculosis, sepsis, cytomegalovirus, listeriosis or opportunistic infections such as progressive multifocal leukoencephalopathy	Hypersensitivity to the active substance or to any of the excipients Active tuberculosis, serious infections such as sepsis or opportunistic infections Severe hepatic impairment Pregnancy and lactation
Drug class	Interleukin inhibitor	TNF-α inhibitor	TNF-α inhibitor	TNF-α inhibitor	Gut-selective immunosuppressive biologic	Janus kinase inhibitor
Active substance(s)	Ustekinumab	Adalimumab	Infliximab	Golimumab	Vedolizumab	Tofacitinib citrate
Pharmaceutical formulation(s)	Concentrate for solution for infusion	Solution for injection	Powder for concentrate for	Solution for injection	Concentrate for solution for infusion	Film-coated tablets

Nonproprietary name	Ustekinumab	Adalimumab	Infliximab	Golimumab	Vedolizumab	Tofacitinib
	Solution for injection Solution for injection in prefilled syringe		solution for infusion			
ATC code	L04AC05	L04AB04	L04AB02	L04AB06	L04AA33	L04AA29
In vitro diagnostics required	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
Monitoring required	Periodic safety update reports are required (1-year frequency)	Periodic safety update reports are required (3-year frequency). All biosimilar products are subject to additional monitoring (black triangle products)	Periodic safety update reports are required(3-year frequency). Biosimilars Zessly and Flixabi are subject to additional monitoring (black triangle products)	Periodic safety update reports are required (3-year frequency).	Periodic safety update reports are required (1-year frequency).	Periodic safety update reports are required (1-year frequency). Subject to additional monitoring (black triangle product)
Orphan designation	No	No	No	No	No	No
ATMP	No	No	No	No	No	No

Abbreviations: UC = ulcerative colitis; ATC = Anatomical Therapeutic Chemical; TNF = tumour necrosis factor; AZ = azathioprine; 6-MP = 6-mercaptopurine.
Sources: Summary of product chracteristics for Stelara [29], Humira [30], Amgevita [35], Imraldi [36], Kromeya [37], Halimatoz [38], Idacio [39], Hefiya [40], Hulio [41], Remicade [31], Remsina [42], Zessly [43], Flixabi [44], Inflectra [45], Simponi [32], Entyvio [33] and Xeljanz [34].

Table 1.2. Administration and dosing of the technology and its comparators

	Ustekinumab	Adalimumab	Infliximab	Golimumab	Vedolizumab	Tofacitinib
Method of administration	Infusion followed by SC injection	SC injection	IV infusion	SC injection	IV infusion	Oral use
Doses	130 mg: concentrate for solution for infusion 90 mg: solution for injection 45 mg: solution for injection	80 mg: solution for injection 40 mg: solution for injection	100 mg: powder for concentrate for solution for infusion	50 mg: solution for injection 100 mg: solution for injection	300 mg: powder for concentrate for solution for infusion After reconstitution, each ml contains 60 mg of vedolizumab.	5 mg: film-coated tablets 10 mg: film-coated tablets
Dosing frequency	Treatment is to be initiated with a single IV dose based on BW. The infusion solution is to be composed of the number of vials of 130 mg specified below: BW of patient at time of dosing / Recommended dose* / No. of 130-mg vials ≤55 kg / 260 mg / 2 >55 kg to ≤85 kg / 390 mg / 3 >85 kg / 520 mg / 4 * Approximately 6 mg/kg The first SC administration of 90 mg should take place at week 8 after the IV. After this, dosing every 12 weeks is recommended.	The recommended induction dose regimen for adult patients with moderate to severe UC is 160 mg at week 0 (given as four 40-mg injections in 1 day or as two 40-mg injections per day for 2 consecutive days) and 80 mg at week 2 (given as two 40-mg injections in 1 day). After induction treatment, the recommended dose is 40 mg every other week via SC injection.	5 mg/kg given as an IV followed by additional 5 mg/kg infusion doses at 2 and 6 weeks after the first infusion, then every 8 weeks thereafter.	For adults (BW <80 kg): initial dose of 200 mg, followed by 100 mg at week 2. Patients who have an adequate response should receive 50 mg at week 6 and every 4 weeks thereafter. For adults (BW ≥80 kg): initial dose of 200 mg, followed by 100 mg at week 2, then 100 mg every 4 weeks thereafter.	The recommended dose regimen is 300 mg administered by IV infusion at 0, 2 and 6 weeks and then every 8 weeks thereafter.	The recommended dose is 10 mg given orally BID for induction for 8 weeks and 5 mg given BID for maintenance.
Standard length of a course of treatment	Chronic/continuous due to the chronic character of UC As long as required by the prescribing healthcare professional.	Chronic/continuous due to the chronic character of UC As long as required by the prescribing	Chronic/continuous due to the chronic character of UC As long as required by the prescribing	Chronic/continuous due to the chronic character of UC As long as required by the prescribing	Chronic/continuous due to the chronic character of UC As long as required by the	Chronic/continuous due to the chronic character of UC As long as required by the prescribing healthcare professional.

	Ustekinumab	Adalimumab	Infliximab	Golimumab	Vedolizumab	Tofacitinib
		healthcare professional.	healthcare professional.	healthcare professional.	prescribing healthcare professional.	
Standard interval between courses of treatments	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
Standard number of repeat courses of treatments	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
Dose adjustments	Patients who have not shown adequate response at 8 weeks after the first SC dose may receive a second SC dose at this time. Patients who lose response on dosing every 12 weeks may benefit from an increase in dosing frequency to every 8 weeks. Patients may subsequently be dosed every 8 weeks or every 12 weeks according to clinical judgment.	Some patients who experience a decrease in their response to 40 mg every other week may benefit from an increase in dosage to 40 mg every week or 80 mg every other week.	No dose adjustment is indicated for UC patients.	Adult (BW <80 kg) patients who have an inadequate response may benefit from continuing with 100 mg at week 6 and every 4 weeks thereafter.	Some patients who have experienced a decrease in their response may benefit from an increase in dosing frequency to 300 mg every 4 weeks.	For patients who do not achieve adequate therapeutic benefit by week 8, the induction dose of 10 mg BID can be extended for an additional 8 weeks (16 weeks total), followed by 5 mg BID for maintenance. For some patients, such as those with TNFi failure, continuation of the 10-mg BID dose should be considered to maintain a therapeutic effect. Patients who experience a decrease in response on tofacitinib 5 mg BID maintenance therapy may benefit from an increase to tofacitinib 10 mg administered BID

Abbreviations: BID = twice daily; BW = body weight; IV = intravenous; UC = ulcerative colitis; SC = subcutaneous; TNF = tumour necrosis factor.
Sources: Summary of product characteristics for Stelara [29], Humira [30], Remicade [31], Simponi [32], Entyvio [33] and Xeljanz.

2 OBJECTIVE AND SCOPE

The aim of this EUnetHTA Joint Relative Effectiveness Assessment is to compare the clinical effectiveness and safety of ustekinumab in the target patient populations with relevant comparators. The target patient populations and relevant comparators (based on the requirements of EUnetHTA Partners) are defined in the project scope below.

The assessment was based on a dossier submitted by the primary MAH Janssen [46].

Table 2.1 provides the scope according the project plan published for the assessment of ustekinumab [47]. Comparators and outcomes were selected according to relevant clinical guidelines, the EUnetHTA guidelines (https://www.eunethta.eu/methodology-guidelines/), health technology assessments and clinical studies.

Table 2.1. Scope of the assessment

Description	Assessment scope
PICO 1	
Population	Adult patients with moderately to severely active ulcerative colitis who have had an inadequate response with, lost response to, or were intolerant to conventional therapy and to at least one biologic therapy or have medical contraindications to such therapies.
Intervention	Ustekinumab
Comparison	Adalimumab, infliximab, golimumab, vedolizumab, tofacitinib
Outcomes	**Efficacy**: • Clinical response after induction • Clinical response after 1 year • Clinical remission after induction • Clinical remission after 1 year • Inflammatory Bowel Disease Questionnaire (IBDQ) response • Generic health-related quality of life HRQoL • Steroid-free remission • Mucosal healing (endoscopic healing) after induction • Mucosal healing (endoscopic healing) after 1 year • Surgery required • Hospitalisation **Safety**: • Overall adverse events (AEs) • Serious AEs • Discontinuations due to AEs • Severe AEs • Fatal AEs • Infections • Severe infections
PICO 2	
Population	Adult patients with moderately to severely active ulcerative colitis who have had an inadequate response with, lost response to, or were intolerant to conventional therapy or have medical contraindications to such therapy.
Intervention	Ustekinumab
Comparison	Adalimumab, infliximab, golimumab, vedolizumab, tofacitinib
Outcomes	**Efficacy**: • Clinical response after induction • Clinical response after 1 year • Clinical remission after induction • Clinical remission after 1 year • IBDQ response • Generic HRQoL

	• Steroid-free remission • Mucosal healing (endoscopic healing) after induction • Mucosal healing (endoscopic healing) after 1 year • Surgery required • Hospitalisation **Safety**: • Overall AEs • Serious AEs • Discontinuations due to AEs • Severe AEs • Fatal AEs • Infections • Severe infection

3 METHODS

The assessment is based on the data and analyses included in the submission dossier prepared by the primary MAH [46]. During the assessment, the completeness of data and analyses in the submission dossier was verified. Furthermore, the methods for data analysis and synthesis applied by the primary MAH were checked against the requirements of the submission dossier and applicable EUnetHTA guidelines and assessed with regard to scientific validity.

3.1 Information retrieval

The evidence base with regard to the drug under assessment provided by the MAH was reviewed by the authoring team. Search strategies were checked for appropriateness and the results of information retrieval included in the MAH submission dossier [46] were checked for completeness against a search in study registries and against the studies included in the regulatory assessment report. Since no flaws were identified in the MAH search strategy, no supplemental searches were performed by the Information Specialist.

Table 3.1. Summary of information retrieval and study selection

Elements	Details
List of studies submitted by the marketing authorisation holder	• Jiang et al. [2] • Probert et al. [3] • Japic CTI060298 [48] • ACT 1 [49] • ACT 2 [49] • Silva et al. [50] • OCTAVE Induction 1 (OCTAVE-I 1) [51] • OCTAVE Induction 2 (OCTAVE-I 2) [51] • OCTAVE Sustain (OCTAVE-S) [51] • NCT00787202 [4] • ULTRA 1 [52,82] • ULTRA 2 [53] • NCT00853099 [1] • VARSITY [54] • PURSUIT-SC [55] • PURSUIT-J [56] • PURSUIT-M [55] • GEMINI 1 [57] • Kobayashi et al. [58] • NCT02039505 [59] • UNIFI [60,61]
Databases and trial registries searched	Databases: Medline, Medline-In-Process, Embase, The Cochrane Library Trial registries: ClinicalTrials.gov, EU Clinical Trials Register
Search date	First run: 14 August 2018 Updates: 22 January 2019, 28 March 2019
Keywords	Covering “Population”, “Intervention and Comparators” and “Study Type”
Inclusion criteria	Population: Patients with moderate to severe active UC with failure of conventional therapy, and patients with failure of prior biologic(s) Intervention(s): • Ustekinumab • Infliximab • Adalimumab • Golimumab • Vedolizumab • Tofacitinib Comparator(s): Same as for Intervention(s)

Elements	Details
	Outcomes: *Efficacy:* • Clinical response • Durable clinical response • Clinical remission • Durable clinical remission • IBDQ response • Steroid-free remission • Mucosal healing • Durable mucosal healing *Safety:* • Surgery required • Hospitalisations • Overall AEs • Serious AEs • Discontinuations due to AEs • Severe AEs and fatal AEs • Infections and severe infections Settings (if applicable): • Peer reviewed published in journals or retrieved via hand searches • Abstracts and posters Study design: Randomised control trials Language restrictions: Full text version available in English Other search limits or restrictions applied: NA
Exclusion criteria	Population: Naïve patients, mild active UC only Intervention(s): All other treatments Comparator(s): All other treatments Outcomes: Not efficacy, safety or QoL related to IBDQ response Settings (if applicable): Letters and editorials Study design: • Single-arm trials • Observational studies • Case-control studies • Cohort studies • Cross-sectional studies • Case series/reports • Systematic literature reviews • Economic evaluations • Background information/expert opinion Language restrictions: Not available in English Other search limits or restrictions applied: Not applicable
Date restrictions	See above “Search date” element
Other search limits or restrictions	Language restriction: English

The study pool for the assessment was compiled on the basis of the following information:

Sources provided by the company in the submission dossier:

- Study list provided by the MAH on ustekinumab (status: 28 March 2019)
- Bibliographical databases (last search on 28 March 2019)
- Trial registries (last search on 28 March 2019 for ClinicalTrials.gov and EU Clinical Trials Register; date unknown for the WHO International Clinical Trials Registry Platform added by the company after the completeness check)

Check of the completeness of the study pool:

- Trial registries (last search on 19 July 2019)

The check identified no additional relevant study.

3.2 Data extraction

Information used for the assessment of clinical effectiveness and safety was extracted from the submission dossier and verified against the CSRs or other original documentation provided in the submission dossier [46,60,61].

3.3 Risk of bias assessment

The quality rating tool developed by the Cochrane Collaboration (version 5.1.0; March 2011 [62]) was used to assess the risk of bias in randomised trials. The risk of bias at study level was assessed for six different domains:

- Method used to generate the sequence for randomisation (random sequence generation);
- Method used to mask the sequence for allocation to treatment (allocation concealment);
- Measures used to ensure the blindness of the study with respect to treatment assignment (blinding of participants, medical personnel and outcome assessors);
- Completeness of the data for each outcome considered (incomplete outcome data);
- Selective description of the results (selective outcome reporting); and
- Other sources of bias (e.g., bias due to early interruption of the study because of the benefits without an appropriate stopping rule, use of a nonvalidated measurement instrument, incorrect statistical analysis).

For each domain, two independent assessors judged the risk of bias (low risk, high risk or unclear) on the basis of the information retrieved from the full-text publications, the protocols and the submission dossier. The results for the risk of bias assessment are presented at the study level only.

3.4 Results from and analyses of the studies included

The information in the submission dossier [46] on the study design, study methods, populations, endpoints (patient relevance, validity and operationalisation) and study results was evaluated. The results of this evaluation are presented and were used for identification of relevant analyses and considered for the conclusions of the assessment report.

Indirect comparisons

The methods for indirect comparisons applied and, if applicable, the justifications in the event of deviations from the required approaches were evaluated [63].

Bayesian NMAs were conducted by the MAH to provide comparative effectiveness estimates for ustekinumab compared with other biologic treatments in treating patients with moderately to severely active UC. The list of comparators included all the biologics currently approved in the EU for moderate to severe UC, including biosimilar agents. The analyses were done for the endpoints of clinical response, clinical remission and mucosal healing, and were performed separately for induction phase alone (at 6–8 weeks) and induction followed by a maintenance phase corresponding to approximately 1 year of treatment (at 44–54 weeks). To allow for heterogeneity in the prior BF status of patients, the networks were stratified and performed independently for failure of conventional care (NBF population) and the BF population (defined by the MAH as the population of patients with failure of one or more prior biologics).

The NMAs were conducted based on a selection of a specific set of studies identified as part of the systematic literature review (the selection are presented in the appendix, A2.2.1. The selection criteria

for inclusion of clinical studies in NMA). Only RCTs assessing the efficacy of at least one intervention with biologics for UC were included in the NMA.

As described above, separate NMAs were conducted to assess the relative effectiveness of ustekinumab versus active comparators for the following regimens and time points:

- Induction regimen: ustekinumab 6 mg/kg and 130 mg (8 weeks).
- Maintenance or 1-year regimen: ustekinumab 6 mg/kg induction (8 weeks) followed by 90 mg q8w and q12w maintenance (44 weeks).

For the analysis of 1-year outcomes, the 6 mg/kg induction regimen was assessed in the base-case analyses as this reflects the expected licence for ustekinumab. As a sensitivity analysis on request from EUnetHTA, the 130 mg induction regimen was assessed for 1-year outcomes.

Comparators

Licensed doses for each comparator were included in the NMA based on the EMA licensing. The MAH also included the following unlicensed doses: infliximab 10 mg/kg intravenous (IV) at weeks 0, 2 and 6 [46] and infliximab 10 mg/kg IV every 8 weeks in maintenance [46]. Although these were not in line with EMA licensing, these treatments were included to enrich the data and allow analysis of induction-to-maintenance treatment strategies. Infliximab 10mg/kg is not the recommended regimen for UC, however it is recommended for Crohn's disease and included in the Canadian product monograph for UC. Furthermore, dose escalation with infliximab to 10mg/kg is common practice in Europe [64]. Therefore, this dose was included in the NMA of 1-year regimens.

All trials compared an active treatment to a placebo arm, with the exception of the VARSITY trial which was a head-to-head study of vedolizumab versus adalimumab [54].

NMA models

The NMAs were conducted within a Bayesian framework using Bayesian hierarchical model, in the WinBUGS statistical software [65]. The MAH preferred a Bayesian approach over the frequentists' strategy as a more suitable option for decision-making in cases for which no formal head-to-head comparison exists and there is a need for ranking of interventions. As all the endpoints were binary, ORs (defined as relative treatment effects for ustekinumab versus a comparator) were estimated as the treatment effects using a binomial likelihood and logit link function. Since the outputs generated by WinBUGS were posterior distributions of a parameter of interest, the MAH used the following summary statistics for these distributions to present the effects: median OR and corresponding 95% CrI and the probability that the ustekinumab treatment would perform better than the comparator (Appendix 2, A2.2.2. Outputs generated by WinBUGS). gives a more detailed interpretation of each of these statistics). In addition to ORs, alternative scales for risk ratios and risk differences were also assessed for maintenance networks only.

For each NMA, the intention of the MAH was to build both fixed-effects and random-effects models and to select the one associated with the lowest DIC and with a difference of at least three points in DIC. The approach was justified by MAH with the explanation that the DIC allows for choice of the model with the best compromise between adequacy and complexity.

Vague prior distributions were used in all analyses for the treatment effects of interest (normal distributions with mean of 0 and variance of 10,000) and in random-effects models for setting the between-trial standard deviation (uniform distribution with a large enough range of [0,2] for binary outcomes), as recommended by the National Institute for Health and Care Excellence (NICE) [66]. In addition, in both fixed-effects and random-effects models, trial-specific baselines were also given a noninformative prior distribution of ~normal(0, 100^2), whereas the treatment effect for the reference treatment was set to 0.

Subgroup analysis and other effect modifiers

During the assessment, the subgroup analyses examining potential effect modifiers presented in the submission dossier and the corresponding methods applied were evaluated.

As statistical assessment of heterogeneity in the network or between individual trials was not possible because in most cases a single trial contributed to each comparison, clinical and methodological het-

erogeneity was assessed instead. The MAH considered heterogeneity among trials due to various issues that could impact interpretation of the effectiveness and safety outcomes, such as differences in patient populations, comparability at baseline, time of assessment, study design and analytical methodology.

Prior TNFi exposure

In line with literature data, the MAH identified prior TNFi exposure as a potential source of heterogeneity since the exclusion criteria for prior exposure or failure to biologic therapy varied across the trials. To minimise this heterogeneity the MAH identified two broad categories of patient populations across the trials: the subgroup of patients who had not failed biologic therapy (biologic-naïve or NBF patients) and the subgroup of patients for whom one or more biologic therapies had failed (BF patients). NMAs were built separately for each subgroup. The subgroup definitions from the UNIFI trial were used to recognise the closest corresponding subgroup in the comparator trials (Table 3.2). The UNIFI trial is a phase 3, double-blind, placebo-controlled, parallel-group, multicentre trial sponsored by the MAH that assessed the efficacy and safety of ustekinumab for the UC indication and consisted of an 8-week induction period (UNIFI Induction), after which responders to ustekinumab were re-randomised to a 44-week maintenance period financed by the MAH (UNIFI Maintenance). In UNIFI, the BF cohort was defined as patients who were treated with one or more TNFi treatments or vedolizumab, whereas the NBF cohort was defined as patients who may be biologic-naïve or may have been exposed to biologic therapy but did not demonstrate an inadequate response or intolerance to treatment with a biologic agent. For safety endpoints, however, stratification of data by prior TNFi exposure was not possible as only the endpoint data for the full population were available for most studies. Therefore, the full population was analysed for the safety NMAs.

Table 3.2. Subgroup definitions by prior therapy across trials and the population corresponding to UNIFI

Trial	Regimens	Design	Population	Definition	Population corresponding to UNIFI
UNIFI [60,61]	Placebo Ustekinumab 130 mg IV at week 0 Ustekinumab 6 mg/kg IV at week 0	Parallel, double-blind	BF NBF	Patients who have received treatment with ≥1 TNFi and/or vedolizumab Subjects who may be biologic-naïve or may have been exposed to biologic therapy but not demonstrated an inadequate response or intolerance to treatment with a biologic agent	–
ACT 1 & 2 [49]	Placebo Infliximab 5mg/kg at week 0, 2 and 6 Infliximab 10 mg/kg at week 0, 2 and 6	Parallel, double-blind	Bionaïve	Patients previously exposed to infliximab or any other TNFi were excluded	Corresponds to NBF patients from UNIFI
GEMINI 1 [57]	Cohort 1: Placebo Cohort 1: VDZ 300 mg at day 1 and 15 Cohort 2: VDZ 300 mg at day 1 and 15	Parallel, double-blind for cohort 1, open-label for cohort 2	TNFi-naïve Prior TNFi failure No prior TNFi failure	No prior TNFi therapy Prior failure of TNFi therapy No prior failure of TNFi therapy	Prior TNFi failure corresponds to BF patients in UNIFI No prior TNFi failure corresponds to NBF patients in UNIFI
Probert et al. [3]	Infliximab 5mg/kg at week 0 and 2	Parallel, double-blind	Bionaïve	No prior therapeutic agent used to directly reduce TNF	Corresponds to NBF patients in UNIFI
PURSUIT [55]	Placebo Golimumab 100/50	Parallel, double-	Bionaïve	No prior biologic TNFi agent(s)	Corresponds to NBF patients in

	mg at week 0 and 2 Golimumab 200/100 mg at week 0 and 2 Golimumab 400/200mg at week 0 and 2	blind			UNIFI
OCTAVE 1 & 2 [51]	Placebo Tofacitinib 10 mg daily for 8 weeks	Parallel, double-blind	No prior TNFi treatment Prior TNFi failure No prior TNFi failure	No previous TNFi treatment TNFi treatment failure Previous TNFi treatment and no failure	Prior TNFi failure corresponds to BF patients from UNIFI No prior TNFi failure corresponds to NBF patients from UNIFI
ULTRA 1 & 2 [52,53]	ULTRA 1: Placebo ADA 80 mg and placebo at week 0, ADA 40 mg and placebo at week 2, ADA 40 mg at week 4 and 6 ADA 160 mg at week 0, ADA 80 mg at week 2, ADA 40 mg at week 4 and 6 ULTRA 2: Placebo ADA 160 mg at week 0, 80 mg at week 2 and then 40 mg EOW beginning at week 4	Parallel, double-blind	TNFi naïve patients ULTRA 2: TNFi experienced patients	No previous receipt of any TNFi or any biologic agent ULTRA 2: previous receipt of any TNFi	TNFi naïve corresponds to NBF patients in UNIFI TNFi experienced corresponds to BF patients in UNIFI
VARSITY [54]	VDZ IV 300 mg at weeks 0, 2 and 6 and every 8 weeks after until week 46 ADA 160 mg at week 0, 80 mg at week 2 and 40 mg q2w until week 50	Parallel, double-blind	TNFi naïve TNFi exposure/failure	NR	TNFi naïve corresponds to NBF patients in UNIFI TNFi experience/failure corresponds to BF patients in UNIFI

Abbreviations: ADA = adalimumab; EOW = every other week; IV = intravenous; q2w = every 2 weeks; VDZ = vedolizumab.

Time of assessment

As for other sources of heterogeneity, the MAH discussed differences in time points for assessment in induction (range 6–8 weeks) and maintenance trials (range 44–54 weeks; defined as the time from the end of induction to the end of maintenance). The consistency of partial Mayo scores within the said time frames was used to assess the comparability of clinical endpoints estimated in induction and maintenance trials.

Various time points for assessment were found across induction trials, so the most similar times for assessment were selected for each intervention. This means that for golimumab and vedolizumab the MAH used endpoints reported at 6 weeks, while for ustekinumab, tofacitinib, adalimumab and infliximab the endpoints reported at 8 weeks were used as inputs (Table 3.3 and Table 4.3). The MAH presented details on either the change in partial Mayo score from baseline at week 4 and week 6/week 8, or the partial Mayo scores at these time points.

These outcome measures were recorded in the majority of the induction trials and apparently showed consistency (Table 3.5).

For maintenance trials, the time point corresponding to the end of maintenance and prior therapy exposure also differed across the trials. The time for assessment for the maintenance phase (defined as the time from the end of induction to the end of maintenance) varied from week 44 (UNIFI and ULTRA 2 trials) to week 54 (PURSUIT-M). Partial Mayo scores within the trials were shown to be fairly consistent around the 1-year time frame, indicating that clinical outcomes are similar. To support this claim, the MAH presented changes over time in the median partial Mayo score in the maintenance phase of PURSUIT (up to 54 weeks), GEMINI (52 weeks) and two similar trials with biologics performed in CD patients. The MAH also presented these data for the UNIFI trial (up to just 44 weeks) although the placebo data seemed to increase by one point between 40–44 weeks.

Table 3.3. Induction studies and time points for assessment

Time point	Trial	Clinical response		Clinical remission		Mucosal healing		Included in BCA
		NBF	BF	NBF	BF	NBF	BF	
8 weeks	OCTAVE Induction 1 [51]	O	O	O (no prior TNFi treatment)	O (prior TNFi failure)	O (no prior TNFi)	O (prior TNFi failure)	Yes
8 weeks	OCTAVE Induction 2 [51]	O	O	O (no prior TNFi treatment)	O (prior TNFi failure)	O (no prior TNFi)	O (prior TNFi failure)	Yes
6 weeks	PURSUIT-SC (phase 2) [55]	O (bionaïve)	X	O (bio-naïve)	X	O (bio-naïve)	X	Yes
6 weeks	PURSUIT-SC (phase 3) [55]	O (bionaïve)	X	O (bio-naïve)	X	O (bio-naïve)	X	Yes
8 weeks	ULTRA 1 [52]	O (bionaïve)	X	O (bio-naïve)	X	O (bio-naïve)	X	Yes
8 weeks	ULTRA 2 [53]	O (no prior TNFi)	O (prior TNFi)	O (no prior TNFi)	O (prior TNFi)	O (no prior TNFi)	O (prior TNFi)	Yes
6 weeks	GEMINI 1 [57] *	O (TNF-naïve)	O (TNFi failure)	O (TNF-naïve)	O (TNFi failure)	O (TNF-naïve)	O (TNFi failure)	Yes
8 weeks	NCT00787202 [4]	O (no prior TNFi)	O (prior TNFi)	X	X	X	X	Yes
8 weeks	ACT 1 [49]	O (bionaïve)	X	O (bio-naïve)	X	O (bio-naïve)	X	Yes
8 weeks	ACT 2 [49]	O (bionaïve)	X	O (bionaïve)	X	O (bio-naïve)	X	Yes
6 weeks	Probert et al. [3]	X	X	O (bionaïve) **	X	X	X	Yes
8 weeks	UNIFI [60,61]	O	O	O	O	O	O	Yes
8 weeks	Suzuki et al. [1] ***	X	X	X	X	X	X	No
6 weeks	PURSUIT-J [56] ***[a]	O (bionaïve)	X	O (bionaïve)	X	O (bionaïve)	X	No
8 weeks	Japic CTI060297 [48] ***	O (bionaïve)	X	O (bionaïve)	X	O (bionaïve)	X	No
8 weeks	Jiang et al. [2] ***	O (TNFi-naïve)	X	O (TNFi-naïve)	X	O (TNFi-naïve)	X	No
NR	Silva et al. [50] ***[b]	X	X	X	X	X	X	No
NR	VARSITY [54]	X	X	X	X	X	X	No

Abbreviations: ADA = adalimumab; BCA = base-case analysis; EOW = every other week; IV = intravenous; NR = not reported; O = data available; q2w = every 2 weeks; VDZ = vedolizumab; X = data not available.
Note: Time points correspond to time spent on induction therapy in the studies. * Cohort 1 in GEMINI 1 consist of a double-blind group with randomised patients receiving placebo or vedolizumab at week 0 and 2. ** Definition of clinical remission different for Probert et al. (Ulcerative Colitis Symptom Score ≤2). *** Trials excluded from the BCA: [a] PURSUIT-J was excluded from the analysis: all patients received golimumab in the induction period. [b] Silvia et al. was excluded from the sensitivity analysis: no time of assessment was reported.

Table 3.4. Maintenance placebo and active treatment data available for approach including induction responders and nonresponders (treat-through approach)

Time point (weeks) data available								
Treatment	**Period**	**ACT 1 (infliximab)**	**ULTRA 2 (adalimumab)**	**VARSITY (vedolizumab, adalimumab)**	**UNIFI (ustekinumab)**	**PURSUIT-SC (golimumab)**	**GEMINI 1 (vedolizumab)**	**OCTAVE (tofacitinib)**
Placebo arm	Induction	8 weeks	8 weeks	–	8 weeks	6 weeks	6 weeks	8 weeks
	Maintenance	54 weeks	52 weeks		52 weeks (induction responders only)	52 weeks (induction responders only)	52 weeks (induction nonresponders only for full population)	–
Active arm	Induction	8 weeks	8 weeks	52 weeks * (remission and mucosal healing only)	8 weeks	6 weeks	6 weeks	8 weeks
	Maintenance	54 weeks	52 weeks		52 weeks	52 weeks	52 weeks (induction responders; induction nonresponders only for full population)	60 weeks ** (induction nonresponders)
Imputations required								
Imputation required	Induction	None	None	None	None	None	None	None
	Maintenance	None	None		Placebo nonresponders at end of induction	Placebo nonresponders at end of induction	1) Placebo responders at end of induction 2) Population-specific induction nonresponder data for active and placebo arms	1) Placebo responders at end of induction 2) Placebo nonresponders at end of induction

Table 3.5. Partial Mayo score at weeks 4 6 and 8 across trials

Trial	Treatment	Mean baseline partial Mayo score (SD)	Mean change in partial Mayo score from baseline (SD)			Partial Mayo score (SD)		
			At 4 weeks	At 6 weeks	At 8 weeks	At 4 weeks	At 6 weeks	At 8 weeks
UNIFI	**Placebo**	6.2 (1.46)	−1.4 (1.86)	NR	−1.5 (2.07)	NR	NR	NR
	Ustekinumab 6 mg/kg	6.2 (1.33)	−2.5 (1.93)	NR	−2.9 (2.20)	NR	NR	NR
	Ustekinumab 130 mg	6.2 (1.42)	−2.1 (1.86)	NR	−2.6 (2.31)	NR	NR	NR
PURSUIT-SC (all randomised patients)	**Placebo**	5.9 (1.37)	−1.2 (1.83)	–1.2 (2.04)	NR	NR	NR	NR
	Golimumab 100/50 mg	5.8 (1.15)	−2.0 (1.96)	−1.8 (2.21)	NR	NR	NR	NR
	Golimumab 200/100 mg	6.1 (1.35)	−2.3 (2.07)	−2.3 (2.21)	NR	NR	NR	NR
	Golimumab 400/200 mg	6.1 (1.29)	−2.3 (2.13)	−2.3 (2.37)	NR	NR	NR	NR
OCTAVE 1	**Placebo**	6.5 (1.2)	−1.6 (0.2)	NR	−1.6 (0.2)	NR	NR	NR
	Tofacitinib 10 mg	6.3 (1.2)	−2.8 (0.1)	NR	−3.1 (0.1)	NR	NR	NR
OCTAVE 2	**Placebo**	6.4 (1.2)	−1.5 (0.2)	NR	−1.7 (0.2)	NR	NR	NR
	Tofacitinib 10 mg	6.4 (1.3)	−2.7 (0.1)	NR	−3.0 (0.1)	NR	NR	NR
ACT 1 *	**Placebo**	6.0 (5.0-7.0)	NR	NR	NR	NR	5.0 (3.0–6.0)	5.0 (3.0–6.0)
	Infliximab 5 mg	6.0 (5.0-7.0)	NR	NR	NR	NR	3.0 (2.0–5.0)	2.0 (1.0–4.0)
	Infliximab 10 mg	6.0 (5.0-7.0)	NR	NR	NR	NR	3.0 (2.0–5.0)	3.0 (1.0–5.0)
GEMINI **	**Placebo**	6.12 (0.45)	NR	NR	NR	5.20 (0.67)	5.19 (0.76)	NR
	Vedolizumab 300 mg	6.01 (0.44)	NR	NR	NR	4.39 (0.61)	4.09 (0.60)	NR

Abbreviations: NR = not reported; SD = standard deviation.
* For ACT 1, values were reported as median (interquartile range). ** For GEMINI, values were reported as mean (SD) and were digitised from a graph.

Heterogeneity due to trial design and imputation methods

Trial design was identified as a significant source of heterogeneity in maintenance but not induction trials.

For induction trials, the data from GEMINI 1 included open-label active treatment, whereas in other trials all the patients were blinded to the induction treatment. As GEMINI 1 also generated double-blind data for the same dosage regime (dose vedolizumab 300 mg IV, Cohort 1), the MAH did not include open-label data in efficacy analyses of induction and all the data used in induction NMAs were generated from randomised, placebo-controlled, double-blind trials.

For the maintenance-phase trials, two study designs were identified. These designs have evolved over time from standard treat-through designs for TNFi therapies to designs that re-randomised patients on the basis of response to treatment for the newer therapies. In the standard treat-through design, patients who are randomised to active or placebo arms at induction continue with the same treatment to the maintenance phase (Appendix 2, A2.2.3. Treat-trough trial design; includes infliximab in ULTRA 2, adalimumab in ACT 1 and vedolizumab vs. adalimumab in VARSITY). In re-randomised designs, for an active arm at induction, responders to active treatment during the induction phase are re-randomised to the treatment or placebo arm for the maintenance phase, while nonresponders are treated-through for the maintenance phase up to 1 year; for a placebo arm at induction, patients either remain on placebo from induction to maintenance regardless of response, or continue to receive placebo maintenance treatment on the basis of response at the end of induction (Appendix 2, A2.2.4. Response-based re-randomised trial design; includes vedolizumab in GEMINI 1, tofacitinib in OCTAVE, golimumab in PURSUIT and ustekinumab in UNIFI) [46]. The response-based trial design thus rather focuses on the question of benefit from continued treatment for induction responders. While these trial designs are considered to be more ethical as they limit patient exposure to ineffective treatment, they also introduce substantial heterogeneity in analyses of comparative efficacy. The main obstacles to directly conducting NMAs on maintenance data generated by the two designs were: a) inconsistent definitions of active maintenance arms across trials; b) a high risk of selection bias in the available evidence from trials that only included in the maintenance phase the patients who responded to induction therapy (and therefore excluded late responders); and c) a lack of appropriate common comparators, as placebo arms in the re-randomised design were not true placebo arms but instead represented re-randomised responders from active induction therapy who were subject to carryover effects (Appendix 2, A2.2.5. Carry-over effect) of induction therapy. A summary of the trial designs for studies included in the maintenance NMAs is provided in Table 3.6. Of the seven trials included, three had treat-through designs and four had re-randomised response-based designs.

Table 3.6. Study design for the maintenance trials included

Treatments	Maintenance study	Study design
Ustekinumab vs. placebo	UNIFI	Re-randomised response-based
Golimumab vs. placebo	PURSUIT-M	Re-randomised response-based
Adalimumab vs. placebo	ULTRA 2	Treat-through
Infliximab vs. placebo	ACT 1	Treat-through
Tofacitinib vs. placebo	OCTAVE Sustain	Re-randomised response-based
Vedolizumab vs. placebo	GEMINI 1	Re-randomised response-based
Vedolizumab vs. adalimumab	VARSITY	Treat-through

Abbreviations: NMA = network meta-analysis; SC = subcutaneous.
Notes: ULTRA 1 was included in the induction NMA but excluded from the maintenance NMA as the maintenance data were based on a single open-label arm of adalimumab 40mg SC; ACT 2 was included in the induction NMA but excluded from the maintenance NMA as the end-of-study time point was 30 weeks.

Dealing with study design heterogeneity in maintenance trials

To make the maintenance endpoint data between the study designs comparable, the MAH proposed an ITT approach mimicking a treat-through trial design (also called a base-case approach) in which conditional probabilities were used to recalculate maintenance endpoints in a re-randomisation trial (clinical response, remission or mucosal healing) so those could be mimicked from a treat-through design. In more detail, the percentage response at 1 year for the ITT population was calculated using formula (A × C) + (B × D) where A is the percentage response at the end of induction for the ITT population, B is the percentage no-response at the end of induction for the ITT population, C is the percentage response at the end of maintenance for induction responders, and D is the percentage response at the end of maintenance for induction nonresponders (Appendix 2, A2.2.5.1. Approaches for comparisons at 1-year and choice of base case). This approach takes into account the response from both induction responders and nonresponders for the estimation. However, this required imputation of data for four out of seven trials included in the maintenance NMAs. Namely, while 1-year endpoint data were directly available from the trials with a standard treat-through design (ACT 1, ULTRA 2 and VARSITY) and induction data were available from the re-randomised response-based trials, maintenance data from the re-randomised response-based trials needed to be recalculated using the above-mentioned formula and the data inputs were missing in some trials. The availability of placebo and active treatment data for each trial is detailed in Table 3.7 for the base-case approach (mimicking a treat-through design).

Table 3.7. Maintenance placebo and active treatment data available for approach including induction responders and induction non-responders (treat-through based approach)

Time point (weeks) data available								
Treatment	**Phase**	**ACT 1 (infliximab)**	**ULTRA 2 (adalimumab)**	**VARSITY (vedolizumab, adalimumab)**	**UNIFI (ustekinumab)**	**PURSUIT-SC (golimumab)**	**GEMINI 1 (vedolizumab)**	**OCTAVE (tofacitinib)**
Placebo arm	**I**	8 weeks	8 weeks	–	8 weeks	6 weeks	6 weeks	8 weeks
	M	54 weeks	52 weeks		52 weeks (IRs only)	52 weeks (IRs only)	52 weeks (INRs only for full population)	–
Active arm	**I**	8 weeks	8 weeks	52 weeks* (remission and mucosal healing only)	8 weeks	6 weeks	6 weeks	8wks
	M	54 weeks	52 weeks		52 weeks	52 weeks	52 weeks (IRs; INRs only for full population)	60 weeks ** (INRs)
Imputations required								
Imputation required	**I**	None	None	None	None	None	None	None
	M	None	None		PNRs at end of I	PNRs at end of I	1) PRs at end of I 2) Population-specific INR data for active and placebo arms	1) PRs at end of I 2) PNRs at end of I

Abbreviations: I = induction; INR = induction nonresponder; IR = induction responder; M = maintenance; PNR = placebo nonresponder; PR = placebo responder..
Notes: for VARSITY results are only reported for the end of 52 weeks which can be used directly in the NMA. Wks, weeks; *in VARSITY trial both arms are active arms; **open-label for induction non-responders

Imputation of placebo maintenance endpoint data in ITT approach mimicking a treat-through trial design

Placebo maintenance endpoint data (generated under the treatment sequence of placebo at induction → placebo at maintenance) were imputed where missing for both induction responders and nonresponders to placebo using individual patient-level data (IPD) from available trials [46].

For induction responders to placebo, imputations were required for the OCTAVE and GEMINI trials. For NBF induction responders to placebo, a weighted average of the UNIFI, ACT 1 and PURSUIT IPD data was estimated and used as the imputed placebo value, whereas for BF induction responders only the UNIFI trial informed the imputed placebo value. A summary of the data used for imputation of maintenance endpoint data in placebo induction responders, stratified by prior TNFi exposure, is provided in Table 3.8. It is evident from the table that in trials used to inform imputation, the placebo maintenance rates for clinical response and clinical remission endpoints among induction responders to placebo varied by ≥10%.

Table 3.8. Efficacy outcomes with placebo at the end of maintenance in UNIFI, ACT 1 and PURSUIT for induction responders to placebo

	N		Outcome at the end of maintenance placebo (%) *					
			Clinical response		Clinical remission		Mucosal healing	
Trial	**NBF**	**BF**	**NBF**	**BF**	**NBF**	**BF**	**NBF**	**BF**
UNIFI (IPD)	57	46	47.4	43.5	26.3	13.0	31.6	17.4
ACT (IPD)	45	NA	40.0	NA	35.6	NA	35.6	NA
PURSUIT (IPD)	103	NA	35.9	NA	25.2	NA	34.0	NA

Abbreviations: BF = biologic failure; IPD = individual patient data; NBF = nonbiologic failure; NA=not applicable.
Note: * Within each population, endpoints are not mutually exclusive, so rates will not add up to 100%.

For induction nonresponders to placebo, imputations were required for the PURSUIT, OCTAVE and UNIFI trials, as previously described. For NBF induction nonresponders to placebo, ACT 1 IPD data were used to inform the imputation, whereas for BF patients the MAH used GEMINI 1 published data from the Federal Joint Committee (G-BA) document [67]. These sources were used to impute missing placebo data for clinical response and clinical remission. However, mucosal healing data were not available from GEMINI 1 and therefore for the BF population only it was not possible to obtain these estimates. An additional issue with the GEMINI 1 trial as the data source was that endpoint data (clinical response and clinical remission) for induction nonresponders in both arms were only provided for the mixed population (including BF and NBF patients). To estimate percentage by population, the relative risk (RR) for patients responding after induction for NBF versus BF was approximated by arm. For placebo, the RR was extrapolated from UNIFI data for placebo induction responders, whereas for the active vedolizumab arm the RR was based on GEMINI 1 maintenance data for vedolizumab q4w induction responders. Imputation maintenance data derived from ACT 1 and from GEMINI 1 for placebo induction nonresponders and used in the maintenance NMAs are shown in Table 3.9.

Table 3.9. Imputation maintenance data from ACT 1 and GEMINI 1 for placebo induction nonresponders used in the NMA

Treatment sequence	Population	Clinical response (%)	Clinical remission (%)
Placebo induction → placebo maintenance	NBF patients (ACT 1)	7.89	5.26
	BF patients (GEMINI 1)*	8.15	3.17

* Data are approximated from RR of patients responding after induction for NBF vs. BF observed in UNIFI

Imputation of active-treatment maintenance endpoint data in the ITT approach mimicking a treat-through trial design

For the active treatment arms of the response-based studies (GEMINI, OCTAVE, PURSUIT and UNIFI), induction and maintenance data for clinical response, clinical remission and mucosal healing endpoints were available for both induction responders and nonresponders. The exceptions were the data from GEMINI 1 and OCTAVE trials. As already described above, for the GEMINI trial it was necessary to estimate maintenance endpoint values for induction nonresponders in the active treatment arm separately for the BF and NBF groups. This was done by using the RR for patients responding after induction for NBF versus BF, which was extrapolated from GEMINI 1 maintenance data for vedolizumab q4w induction responders. However, in the OCTAVE trial, instead of induction responders to active treatment as in other trials, OCTAVE re-randomised induction responders from both the tofacitinib and placebo arms. Nevertheless, maintenance results in this trial were not reported by the induction treatment received (placebo or tofacitinib 10 mg twice daily). Regardless, the MAH justified inclusion of these data on the assumption that placebo and tofacitinib induction responders were comparable, as: a) patients who responded to placebo at induction in OCTAVE were then given active therapy (tofacitinib) in maintenance; and b) subgroup analyses assessing the impact of induction treatment on clinical remission showed minimal differences of up to 2% between the rates for the tofacitinib + placebo and tofacitinib induction arms. Still, 12% of induction responders were responders to placebo.

Alternative approach for comparing maintenance endpoints

In addition to the ITT approach mimicking a treat-through trial design, the MAH also considered an alternative approach to making maintenance endpoints comparable: an ITT approach conditional on induction response in which endpoints from treat-through trials were recalculated to correspond to a response-based design. While the MAH acknowledged that the latter approach required less imputation, the fact that it does not take into account late responders, which is an important phenomenon in interventions with biologics (these patients are considered as nonresponders after induction) was one of the reasons why the MAH chose an ITT approach mimicking a treat-through trial design to recalculate endpoints for primary analysis. The ITT approach conditional on induction response was then conducted as a sensitivity analysis, but as discussed in the section on sensitivity analyses it cannot be considered as a proper control for an ITT approach mimicking a treat-through trial design. The ITT approach mimicking a treat-through trial design was considered to more closely reflect clinical practice and allows a clearer interpretation of the treatment effects and the inclusion of head-to-head trial data.

Data sources used in these calculations were as follows. IPD from ACT, PURSUIT and UNIFI were available and used in the base-case analysis (for ACT 1 these data corresponded exactly to the published results since the trial is a treat-through trial). For the remaining studies, published data from the SLR or identified from publicly accessible documents were included. The data required for maintenance are divided into induction responders and nonresponders, as both are required for the base-case approach.

Pooling of doses

For the maintenance treatment arms for NBF patients, the MAH did not identified a dose–response relationship for the treatments with different dosage regimes and concluded that pooling of doses for the same treatment was appropriate. No formal dose-response relationship testing was done. The evidence for a dose–response relationship was defined as follows:

- Dose–response relationship: a higher dose/shorter interval between doses led to higher clinical response rates when compared to a lower dose/longer interval between doses (as expected).
    - In this case doses were not pooled.
- No dose–response relationship: a lower dose/longer interval between doses led to higher clinical response rates when compared to a higher dose/shorter interval between doses (opposite to what is expected).
    - In this case doses were pooled.

In the base-case NMAs, doses were pooled or not on the basis of these criteria. In the cases for which it was considered appropriate to pool the doses, additional NMAs were conducted in which the doses were unpooled to assess the impact of pooling (full details on the results are provided in the NMA report from the MAH) [46].

Baseline

Regarding the distribution of potential effect modifiers across comparisons, according to the literature the key disease characteristics for patients who should be evaluated at baseline in IBD trials to mitigate the risk of an effect modifier affecting the results include: endoscopic evidence of inflammation; disease attributes (duration, severity, extent and location); complications (such as fistulas); biomarkers of inflammation (CRP or faecal calprotectin); and prior and/or concomitant medications [68]. In the submission file the following characteristics were assessed: duration of disease; age and weight at baseline; proportion of males/females; CRP level; and Mayo score at baseline [46]. Since the data were too scarce to use meta-regression techniques (NMAs included from 4 to a maximum of 11 trials and compared from 4 to 8 different interventions), the comparability of studies was assessed at the level of descriptive statistics in the presence of missing data and different outcome measures (median vs mean). Full details on the baseline characteristics across studies included in the NMAs are provided in the NMA report of the MAH core submission file [46]. The MAH concluded that baseline characteristics for patients across the trials were similar enough for the findings to be pooled.

Statistical assessment of heterogeneity

NMA statistical approaches

Statistical methods to assess heterogeneity

The MAH tested statistical heterogeneity between studies for a particular pairwise comparison (exclusively comparisons to placebo). Owing to the paucity of the data, the majority of comparisons in NMAs were informed by just one trial and pairwise meta-analyses were feasible only in four out of eleven NMAs that were built. Analyses were predominantly informed by two studies (in only one case the analysis included three trials) and thus were probably underpowered, meaning that the absence of statistical heterogeneity could not be considered as a proof of the absence of heterogeneity.

As described in the paragraph above, the NMAs submitted were too small to obtain reliable estimates of heterogeneity in the NMAs or to use meta-regression techniques.

Statistical methods to assess inconsistency

The treatment effects estimated using NMAs in the submission file were predominantly based on indirect evidence derived from a limited number of trials ($n \leq 11$ trials for assessment of the clinical remission endpoint). The induction-phase NMAs do not include any head-to-head comparison of different interventions, whereas in the maintenance-phase NMAs, which do include one such comparison (VARITY trial), only a single study per comparison was available as a source of direct evidence, which increases the likelihood of a closed loop to show inconsistency [69]. Moreover, one of the studies in the closed loop informing on comparison to placebo (GEMINI 1) was partly based on imputed data due to recalculation mimicking a treat-through approach and thus did not contain direct data per se. It therefore seems that the inconsistency could not be tested even in a network containing a head-to-head comparison and a closed loop. The MAH used the data from this closed loop with partly imputed data to “cross-check the approach by comparing the recalculated efficacy from GEMINI-1 with the results from VARSITY and ULTRA 2” but this was merely done to check if the imputed value is in agreement with direct data and cannot be considered as an inconsistency analysis. Of note, it is not clear how this comparison was carried out, as no results were reported.

Sensitivity analysis

To evaluate the robustness of the results, sensitivity analyses with regard to methodological factors presented in the submission dossier and the corresponding methods applied were evaluated. These methodological factors arise from decisions made within the framework for retrieval and assessment of information, such as the specification of cutoffs for the time points for data collection and the choice of effect measures.

While the sensitivity analyses were not explicitly preplanned in the project plan [47], the MAH envisaged a series of sensitivity analyses in their Methods section, of which several were not feasible. In

particular, sensitivity analysis of including/excluding open-label trials in induction NMAs was not conducted, as the only trial investigating open-label data (GEMINI 1) also presented double-blind data for the same intervention. Furthermore, a sensitivity analysis reporting the results from the fixed-effects model if the random-effects model was selected in the base-case analysis was not performed as the fixed-effect models were always selected. For the same reason, a test of how the gamma prior distribution (0.001,0.001) for the between-trial precision in the random-effects model impacted the results was not preformed. However, while the random-effects models for the maintenance NMAs were not built because of the paucity of the data, for induction NMAs both fixed-effects and random-effects models were built and no model should have been preferred, as the DIC difference between them was negligible. It is therefore not clear why the last two planned sensitivity analyses were not performed for the induction NMAs.

Regarding the sensitivity analyses that were performed, the MAH primarily focused on replicating "directional similarity" between the primary and the sensitivity analyses by commenting on the direction of point estimates, whereas the effect sizes and uncertainty were not discussed. For example, the Janssen study report [70] states that ustekinumab 6 mg/kg showed higher odds of a clinical response than vedolizumab 300 mg, although the CrI for the median OR was largely uncertain (1.37, CrI 0.71–2.68; Pr = 82%).

The following sensitivity analysis were reported:

1. By additionally including trials on Asian populations in induction and maintenance NMAs the MAH intended to test whether including a broader population provided similar results to the base-case analysis. For induction NMAs, sensitivity analyses included the following trials focusing on Japanese and Chinese populations: Jiang et al [2], Japic CTI060297 [48], NCT02039505 [59]; and Suzuki et al. [1]. For maintenance NMAs at 1 year the Suzuki trial [1] was also included in the analysis. Details on the results for these sensitivity analyses are reported in the MAH NMA report [46].
2. The MAH also planned to asses the impact of the method used to reduce heterogeneity for maintenance endpoint data caused by different study designs; by assessing the results of NMAs based on ITT approach conditional on induction response mimicking a response based design and comparing them to results of NMAs based on ITT approach mimicking a treat-through trial design (Section 3.4.3). In particular, the MAH sought to replicate the evidence of superiority of ustekinumab over placebo and other active treatments. However, as the responder-based analyses do not address the same underlying question as the base-case analysis, they could not be considered as a sensitivity analysis. Namely, even if the results replicate a superiority pattern for ustekinumab there is an issue with the directness of the data and such evidence would not be able to "upgrade" the strength of the evidence.
3. An additional sensitivity analysis was conducted to assess the impact of using fixed values for the imputation of missing placebo data. A multiple imputation approach, based on the assumption that values used for imputation of placebo outcomes are sampled from the same underlying distribution of the data, was taken and beta and normal distributions with the central tendency corresponding to the weighted average of data used in imputation were implemented. After rerunning NMAs with sampled values for placebo, the posterior simulations for the treatment effects were combined across the NMAs to attain a pooled posterior distribution for each treatment effect, which represented the treatment effects accounting for variation in the imputed placebo data, and then the results were compared to the base case. Details on the methodology and results of the multiple imputation analysis are provided in the MAH submission file [46].
4. Given that ustekinumab 6mg/kg is expected to be the licensed dose, only the ustekinumab 6 mg/kg induction dose followed by each maintenance dose (ustekinumab 6 mg/kg or ustekinumab 90 mg) was used for the ustekinumab treatment sequences of interest in the main body of evidence. For that reason the maintenance NMAs were built containing the alternative ustekinumab treatment sequence (ustekinumab 130 mg/kg at induction → ustekinumab 90 mg) that was also reported in UNIFI.

Statistical assessment of publication bias

The full list of available endpoints for the NMA analyses was not provided in the submission file. Endpoints could be drawn from the table of characteristics of the studies included that is presented in the submission file, but judging from the data presented for the UNIFI trial, not all the endpoints are listed. Specifically, the major secondary endpoint of corticosteroid-free clinical remission from the UNIFI trial was not reported and was not considered for NMA analyses.

3.5 Patient involvement

EUnetHTA conducted an open call asking general questions to elicit patients' views on living with ulcerative colitis, the important outcomes considered in this assessment, and expectations of the assessed drug. Questions were based on the HTAi questionnaire template (https://www.htai.org/interestgroups/patient-and-citizen-involvement/resources/for-patients-and-patient-groups/).

European and national patient organisations had to provide an organisational perspective on the questions in English. In all parts of the open call, the term "patient" refers to anyone living with, or who has lived with, the condition for which the new medicine is indicated.

Two patient organisations completed the survey: Pacienti IBD z.s. (Czech Republic) and European Federation of Crohn's and Ulcerative Colitis Associations. The information gathered from the open call was used to inform the scope of this assessment and in particular the outcomes to be considered. In addition, the information was used with the aim of helping in the assessment of health technology, specifically to provide the summary text related to impact of condition; experience with currently available pharmaceuticals; and experiences with and expectations for the drug under assessment.

4 RESULTS

4.1 Information retrieval

The MAH provided a systematic literature search and review of the evidence, which was critically assessed by the authors of this assessment. The following electronic databases were included in the MAH literature search: Medline, Medline-In-Process, Embase and the Cochrane Library. Medline and Medline-In-Process were searched through the PubMed platform, while Embase was searched through the Embase.com platform. These databases include both published studies and conference abstracts. In addition to the database searches, manual searches were undertaken to identify relevant publications and posters not captured in the electronic searches. Clinical trial websites reporting information on interventions for non-naïve patients with moderately to severely active UC were also searched. Congresses and conferences of interest such as the ECCO, United European Gastroenterology Week and Digestive Disease Week were not searched manually since they are indexed in Embase. The final update search was performed on 28 March 2019.

The search protocol and a Preferred Reporting Items for Systematic Reviews and Meta-analyses flow chart were included as part of the submission file [46,71].

Overall, the authors considered that the search followed EUnetHTA guidelines [72] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement. Excluded studies with reasons could be found in Appendix 2, A2.1 Excluded studies with reasons.

4.2 Studies included in the assessment

A total of 49 publications were identified in the SLR (including 32 full articles, 15 abstracts and 2 posters) for six drugs: ustekinumab, infliximab, tofacitinib, adalimumab, golimumab, and vedolizumab. For ustekinumab, six abstracts were identified via electronic searches, although the clinical results for ustekinumab were primarily extracted from the CSRs provided by Janssen [60,61]. A total of 21 trials were identified. Table 4.1 gives an overview of the trials organised by treatment.

Table 4.1. Overview of the number of relevant studies by treatment

Treatment	Abbreviation	*N* *	Trial names
Ustekinumab	UST	1	UNIFI [60,61]
Adalimumab	ADA	5 *	ULTRA 1 [52,82], ULTRA 2 [53], NCT00853099 [1], Silva et al [50]**, VARSITY [54]
Infliximab	IFX	6 *	Jiang et al. [2], Probert et al. [3], Japic CTI060298 [48], ACT 1 [49], ACT 2 [49], Silva et al. [50]*
Golimumab	GOL	3	PURSUIT-SC [55], PURSUIT-J [56], PURSUIT-M [55]
Tofacitinib	TOC	4	OCTAVE Induction 1 (OCTAVE 1) [51], OCTAVE Induction 2 (OCTAVE 2) [51], OCTAVE Sustain (OCTAVE-S) [51], NCT00787202 [4]
Vedolizumab	VDZ	4 *	GEMINI 1 [57], Kobayashi et al. [58], NCT02039505 [59], VARSITY [54] ***

Notes: *There are a total of 21 trials, but these add up to 23 in the table because two head-to-head trials were assigned to IFX and ADA, and VDZ and ADA, respectively. ** Head-to-head comparison between IFX and ADA. *** VARSITY conducted head-to-head comparison between VDZ and ADA.

Among the 21 trials identified, 19 were extracted and contributed to the analysis of the results and 18 trials to the NMA. The two studies excluded from the analysis of outcomes were those by Silva et al. [50] and Kobayahsi et al. [58]. Silva et al. [50] conducted a head-to-head trial comparing IFX and ADA with only an abstract available. The study was excluded because of missing information on the time point for assessment for the clinical endpoints reported, and a small sample size (ADA, n = 21; IFX, n = 10; dose information is missing in the abstract). The study by Kobayashi et al. [58] was excluded because this was an open-label phase 1 trial with only nine patients involved (VDZ 150 mg, n = 3; VDZ 300 mg, n = 6), and the primary objective was to assess the pharmacokinetics of the treatment. In addition, the observation for clinical efficacy endpoints was assessed on days 17, 155 and 239

(roughly corresponding to weeks 10, 31 and 34), which are not comparable with the other trials. Therefore, the efficacy and safety endpoints reported in this study were incomparable with the other phase 2–4 studies included for analysis.

PURSUIT-J [56] was the only study included in the narrative synthesis but excluded from the NMAs, as all patients in the trial received golimumab 200 mg at induction and were not randomised to induction treatment.

The studies listed in Table A4 (Appendix 2, A2.3. Study pool – list of relevant studies used for the assessment) were included in the assessment.

4.3 Studies excluded

Table 4.2 lists the studies that were included in the submission dossier provided by the MAH but excluded from further consideration in this assessment.

Table 4.2. Studies excluded

Study reference/ID	Reason for exclusion
Silva et al. [50]	A head-to-head trial comparing IFX and ADA with only an abstract available; missing information on the assessment time point for the clinical endpoints reported and a small sample size (ADA, *n* = 21; IFX, *n* = 10; dose information missing in the abstract)
Kobayashi et al. [58]	An open-label phase 1 trial with only nine patients involved (VDZ 150 mg, *n* = 3; VDZ 300 mg, *n* = 6) for which the primary objective to assess the pharmacokinetics of the treatment. The observations for clinical efficacy endpoints on days 17, 155 and 239 (roughly corresponding to weeks 10, 31 and 34) were not comparable with the other trials; the efficacy and safety endpoints reported were not comparable with the other phase 2–4 studies included for analysis.

Abbreviations: ADA = adalimumab, IFX = infliximab, VDZ = vedolizumab.

As already stated, PURSUIT-J [56] was excluded from the NMAs as all patients in the trial received golimumab 200 mg at induction and were not randomised to induction treatment.

4.4 Characteristics of the studies included

Direct comparison

Ustekinumab versus active treatment

No such trials were identified.

Ustekinumab versus placebo

The submission for ustekinumab for the treatment of moderately to severely active UC in patients who have demonstrated an inadequate response or failure to tolerate nonbiologic therapy (i.e. corticosteroids, immunomodulators) or biologic therapy is based on the results of the UNIFI phase 3, multicentre, randomised, double-blind, placebo-controlled, parallel-group trial that compared the efficacy and safety of ustekinumab to placebo.

The UNIFI clinical trial programme consisted of an 8-week induction period (UNIFI Induction) with responders re-randomised to a 44-week maintenance period (UNIFI Maintenance), and included a long-term extension phase during which eligible subjects were followed for an additional 3 years in a long-term efficacy trial under the same prior protocol. A visual overview of the UNIFI trial is shown in Figure 4.1.

The induction study included subjects with moderately to severely active UC who had demonstrated an inadequate response or failure to tolerate non-biologic or biologic therapy.

The maintenance study was a randomised withdrawal study targeting subjects with moderately to severely active UC who had demonstrated a clinical response to induction treatment with IV ustekinumab.

Subjects who were randomised to ustekinumab and did not show a clinical response at week 8 received a single SC injection of 90 mg ustekinumab, while those randomised to placebo received a single IV infusion of ustekinumab (~6 mg/kg) to identify if a delayed response could be achieved in some subjects. These “delayed responders” were assessed for response at week 16, with responders eligible to enter the study and receive 90 mg ustekinumab SC q8w.

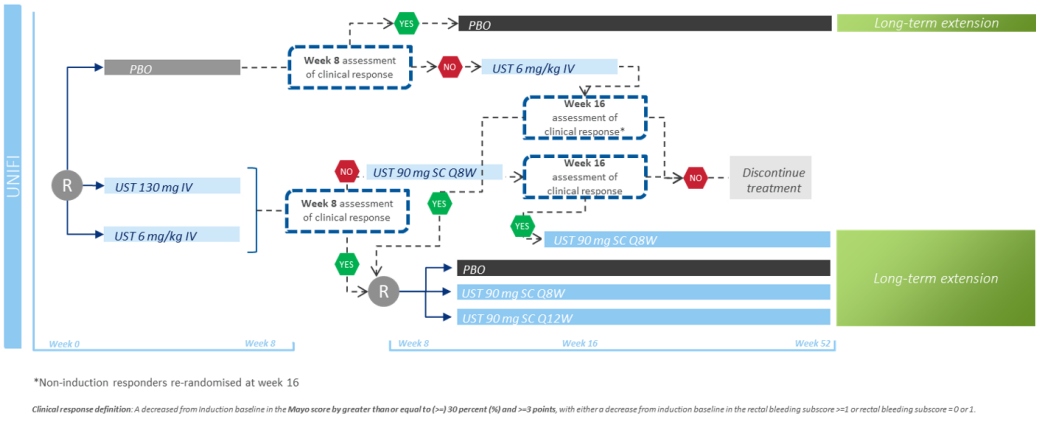


Figure 4.1. Overview of the UNIFI phase 3 trial.

In summary, the programme evaluated ustekinumab treatment in subjects with moderately to severely active UC through at least 1 year of induction and maintenance therapy; after completion of the maintenance study through to week 44 (i.e., 52 weeks after the initial baseline for the induction period), a long-term efficacy trial will follow eligible subjects for an additional 3 years.

UNIFI-I induction study

The UNIFI-I induction study [60] was a phase 3, multicentre, placebo-controlled, parallel-group, double-blind, randomised trial that included 961 patients. The primary endpoint for the induction study was the proportion of subjects in clinical remission at week 8. Patients were randomised to receive a single IV administration of either the recommended tiered dose of approximately 6 mg/kg, a fixed dose of 130 mg ustekinumab or placebo at week 0. Concomitant doses of oral corticosteroids, immunomodulators and ASAs were permitted and 90% of patients continued to receive at least one of these medications. Enrolled patients had to have failed conventional therapy (corticosteroids or immunomodulators) or at least one biologic (TNFi and/or vedolizumab). Some 49% of the patients had experienced failure of conventional therapy but not a biologic (of whom 94% where biologic-naïve), while 51% had experienced failure of or were intolerant to a biologic. For approximately 50% of the patients at least one prior TNFi therapy had failed (of whom 48% were primary nonresponders) and for 17% at least one TNFi therapy and vedolizumab had failed. A schematic overview of UNIFI-I is presented in Figure 4.2.

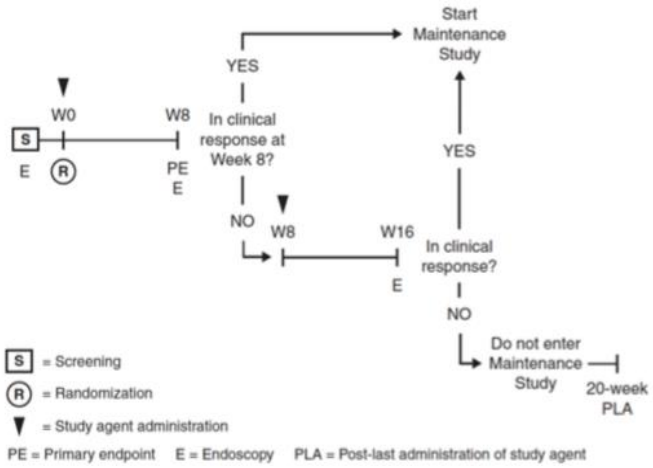


Figure 4.2. Overview of UNIFI-I.

UNIFI-M maintenance study

The UNIFI-M maintenance study [61] was a phase 3, multicentre, placebo-controlled, parallel-group, double-blind, randomised-withdrawal study designed to evaluate the efficacy and safety of SC regimens of ustekinumab maintenance therapy in the primary population of subjects with moderately to severely active UC for whom IV ustekinumab induced a clinical response.

Although all subjects who were responders to the study agent in the induction study were eligible for enrolment in the maintenance study, only subjects with a clinical response to IV ustekinumab during induction comprised the primary population in the maintenance study. The following subjects from the induction study were included in the primary population for the maintenance study: subjects who were randomised to receive IV ustekinumab at week 0 of the induction study and showed a clinical response at week 8 of the induction study; and subjects who were randomised to receive IV placebo at week 0 of the induction study and showed no clinical response at week 8 but had a clinical response at week 16 of the induction study after receiving an induction dose of IV ustekinumab (~6 mg/kg) at week 8.

A schematic overview of UNIFI-M is presented in Figure 4.3.

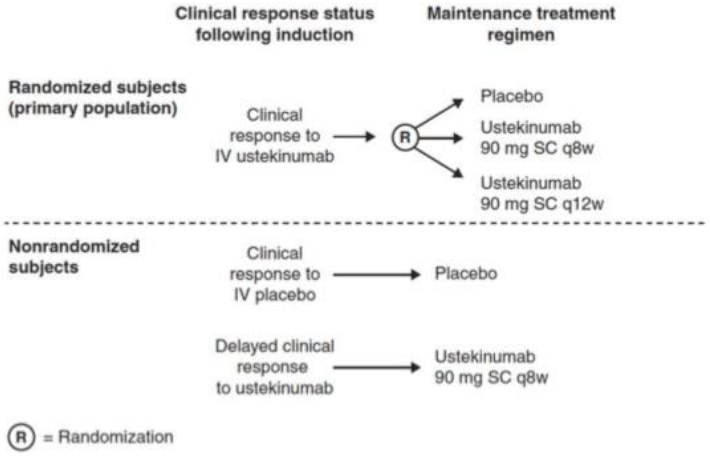


Figure 4.3. Overview of UNIFI-M.

A total of 783 subjects who completed the induction study and had a clinical response to the induction study agent were enrolled in the maintenance study. Of these, 523 subjects were in the targeted primary population for the maintenance study and were randomised (i.e., had a clinical response to IV ustekinumab induction; randomised subjects) and 260 subjects were not part of the primary population for the maintenance study and were not randomised (i.e., placebo induction responders and ustekinumab induction delayed responders; nonrandomised subjects). Subjects were randomised per clinical remission and steroid use at maintenance baseline. The proportions of patients in each arm remained similar between the induction and maintenance phases for important variables such as UC disease characteristics and BF status. Some 24.6%, 14% and 10.2% of subjects in the placebo, q12w and q8w arms, respectively, discontinued the study treatment. These rate do not appear to be higher than in phase 3 studies for other therapies in UC.

Table A5 in Appendix 2, A2.4. Characteristics of the studies included, Induction UNIFI-I and Maintance UNIFI-M studies lists the details for the induction and maintained studies under the scope of this assessment.

Studies included in the NMA

Indirect comparison

Adalimumab, infliximab, golimumab, tofacitinib and vedolizumab

Summary of studies in the NMA

In addition to the UNIFI-I and UNIFI-M studies related to ustekinumab [60,61] , further studies were included in NMA: four studies related to adalimumab [1,52,53,54,82]; five studies related to infliximab [2,3,48,49]; two studies related to golimumab (PURSUIT-M; PURSUIT-SC [55]); four studies related to tofacitinib (NCT00787202 Sandborn et al, 2012; OCTAVE-I1; OCTAVE-I2; OCTAVE-Sustain) [4,51]; and three studies related to vedolizumab (GEMINI 1; NCT02039505; VARSITY) [54,57,59]. For the VARSITY RCT [54] only the abstract was published; a registered protocol was available in the ClinicalTrials.gov register, without results posted. Of these studies, further studies were included in the induction NMAs (ACT 1, ACT 2, GEMINI 1, Probert et al., PURSUIT-SC phase 2, PUSUIT-SC phase 3, OCTAVE 1, OCTAVE 2, ULTRA 1, ULTRA 2 and UNIFI, NCT00787202) [1,3,4,49,51,52,53,55,57,60] and seven in the 1-year NMAs (ACT 1, GEMINI 1, PURSUIT [PURSUIT-SC and PURSUIT-M], OCTAVE [OCTAVE 1, 2 and Sustain], ULTRA 2, UNIFI and VARSITY) [1,4,49,51,53,54,55,57,61]. All trials compared active treatment to placebo with the exception of VARSITY, which compared vedolizumab and adalimumab [54].

All studies were conducted among subjects with moderately to severely active UC with failure of non-biologic therapy and/or failure of prior biological treatment(s), with results reported separately for both subpopulations. Most studies that continued to evaluate the maintenance of the drug effect after the induction phase re-randomised the patients who responded to active arms at the beginning of the maintenance phase (e.g., UNIFI, PURSUIT, OCTAVE and GEMINI 1) [4,55,57,60,61]; other trials did not re-randomise the patients for the maintenance phase and patients continued to receive the same treatment as in the induction phase, that is, patients were treated-through (e.g., ACT 1, ULTRA 2 and VARSITY) [49,53,54]. The main reason for the change in clinical trial design was that it was not considered ethical to expose patients to ineffective placebo treatments. The definition of efficacy endpoints assessed and their corresponding time of assessment slightly change across trials. The most commonly reported primary efficacy endpoint was remission (defined as a Mayo score of ≤2 with no subscore >1) but the time of assessment varied across studies from week 6 to week 8 in induction trials and from week 30 to week 54 after the end of induction in the maintenance trials. Sample size varied across trials from 20 patients per arm to more than 400 per arm in induction trials, and from 30 patients per arm to more than 300 per arm in maintenance trials. At the trial outset, there were no major discrepancies in the baseline demographic characteristics (including age, weight and the proportion of males at baseline) and the baseline disease characteristics (including duration of disease, CRP level and Mayo score) across studies. All studies included patients with moderate to severe UC, defined as a Mayo score between 6 and 12 for the induction phase. The mean Mayo score at baseline reported for subgroups of patients ranged from 8 to 9 across all induction phase studies and the differences are limited across trials. In relation to disease duration, a general trend for shorter disease duration was observed for NBF patients than for the full patient group. At the start of the induction phase, 15 studies reported a mean disease duration varying from 4 to 8 years for NBF patients, and from 6 to 11 years for the full patient group. The mean CRP level at baseline reported for 15 studies for the induction phase ranged from 4 to 17 mg/l for NBF patients and from 5 to 19 mg/l for the full patient group. The characteristic of the studies included in the NMA are listed in Table A6 in Appendix 2. A summary of studies included in the NMAs by time point is provided in Table 4.3.

Table 4.3. Summary of studies included in the NMAs by time point

Trial	Comparators	Included in NMAs	
		Induction NMA	1-year NMA
OCTAVE Induction 1 [51] OCTAVE Induction 2 [51] OCTAVE 1 and 2 combined [51]	Induction: PBO TOC 10 mg BID	✓	✓
OCTAVE Sustain [51]	Maintenance: PBO TOC 5 mg BID TOC 10 mg BID		✓
PURSUIT-SC (phase 2) [55] PURSUIT-SC (phase 3) [55]	Induction: PBO GOL 200/100 mg	✓	✓
PURSUIT-M [55]	Maintenance: PBO-PBO GOL 100 mg q4w GOL 50 mg q4w		✓
ULTRA 1 [52]	Induction: PBO ADA 160/80 mg Maintenance: ADA 160/80 mg	✓	
ULTRA 2 [53]	Induction: PBO ADA 160/80/40 mg Maintenance: PBO ADA 40 mg EOW	✓	✓
GEMINI 1 [57]	Induction: PBO VDZ 300 mg Maintenance: PBO VDZ 300 mg q8w VDZ 300 mg q4w	✓	✓
NTC00787202 [4	Induction: PBO TOC 10 mg BID	✓	
ACT 1 [49]	Induction: PBO IFX 5 mg IFX 10 mg Maintenance: PBO IFX 5mg q8w IFX 10mg q8w	✓	✓
ACT 2 [49]	Induction: PBO IFX 5 mg IFX 10 mg Maintenance: PBO IFX 5 mg q8w IFX 10mg q8w	✓	
Probert et al. [3]	Induction: PBO IFX 5 mg	✓	

Trial	Comparators	Included in NMAs	
		Induction NMA	1-year NMA
UNIFI [60,61]	Induction: PBO UST 130 mg UST 6 mg/kg Maintenance: PBO UST 90 mg SC q8w UST 90 mg SC q12w	✓	✓
Suzuki et al. [1]	Induction: PBO ADA 160/80 mg ADA 80/40 mg Maintenance: PBO ADA 40 mg EOW	✓ (SA)	✓ (SA)
Japic CTI060297 [48]	Induction: PBO IFX 5 mg	✓ (SA)	
Jiang et al [2]	Induction: PBO IFX 5 mg	✓ (SA)	
VARSITY [54]	Induction: ADA 160/80/40 mg VDZ 300 mg Maintenance: ADA 40 mg EOW VDZ 300 mg q8w		✓
NCT02039505 [59]	Induction: PBO VDZ 300 mg	✓ (SA)	

Abbreviations: ADA = adalimumab; BID = twice daily; EOW = every other week, GOL = golimumab; IFX = infliximab; PBO = placebo; qXw = every X weeks; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab; SA = included in the sensitivity analysis with Asian populations only; SC = subcutaneous.

UNIFI-I

For the induction study [60], 60.6% of the trial subjects were male and 76.0% were white, the median age was 41.0 years and the median weight was 71.2 kg. The median duration of disease was 5.97 years, 45.7% of subjects had extensive disease, the median Mayo score was 9.0, 84.4% of subjects had moderate UC (Mayo score ≥6 and ≤10) and 15.3% had severe disease (Mayo score >10), 51.1% of subjects had a history of BF and all subjects had a history of UC medication use before study entry. At baseline, 90.2% of subjects were receiving a concomitant UC medication, with similar proportions of subjects across all treatment groups. More than half of the subjects(51.8%) were using corticosteroids at baseline and 28.2% were using immunomodulatory drugs (AZA, 6-MP or MTX). The proportion of subjects using ASAs was 68.7%. Among the 961 randomised subjects, 51.1% had a documented history of BF (50.5% had failure with at least one TNFi agent and 16.6% had failure for both a TNFi and vedolizumab) and 48.9% did not have a history of BF (46.1% were biologic-naïve and 2.8% were biologic-experienced but not documented as BF). Infliximab, golimumab and adalimumab were the TNFi therapies that subjects were previously exposed to and exposure was similar across the three groups. Vedolizumab was the other biologic that patients may have been exposed to. BF to vedolizumab (regardless of TNFi) was reported for 17.3% (note: 6 subjects had BF to only vedolizumab). More specifically, the results at baseline for disease characteristics show that they were generally similar across all treatment groups.

UNIFI-M

A total of 783 subjects who completed the induction study and had a clinical response to induction ustekinumab were enrolled in the maintenance study [61]. Of these, 523 subjects were in the targeted primary population for the maintenance study and were randomised (i.e., had a clinical response to IV ustekinumab induction; randomised subjects) and 260 subjects were not part of the primary population for the maintenance study and were not randomised (i.e., placebo induction responders and ustekinumab induction delayed responders; nonrandomised subjects). All subjects in this maintenance study entered from the induction study. The primary population is the population of randomised subjects.

Baseline patient characteristics in studies used in the NMA are listed in Table A7 (Appendix 2).

4.5 Outcomes included

In the UNIFI clinical trial programme for ustekinumab in UC, outcomes were measured for disease activity, HRQoL and health utility using different instruments and scoring systems (Table 4.4).

A key component of the efficacy outcomes of clinical remission, clinical response and endoscopic and mucosal healing is the Mayo score. The Mayo score is calculated as the sum of four subscores (stool frequency, rectal bleeding, Physician Global Assessment score and endoscopy findings) and ranges from 0 to 12 points. Scores of 3–5 points indicate mildly active disease, 6–10 points indicate moderately active disease and 11–12 points indicate severe disease. The partial Mayo score is the Mayo score without the endoscopy subscore and ranges from 0 to 9 points.

AEs were also recorded as safety endpoints.

Outcome definitions in the UNIFI maintenance phase were identical to those in the induction study. There were additional outcomes in the UNIFI maintenance study regarding corticosteroid use; however, all technical outcome definitions are identical to those in the induction study (Table 4.4).

Table 4.4. Matrix of outcomes in the UNIFI trials [60,61]

Outcome	Definition
Efficacy	
Clinical remission	Mayo score ≤2 points, with no individual subscore >1.
Clinical response	A decrease from induction baseline in Mayo score by ≥30% and ≥3 points, with either a decrease in the rectal bleeding subscore ≥1 or a rectal bleeding subscore of 0 or 1.
Endoscopic healing	Mayo endoscopy subscore of 0 or 1.
Safety	
Adverse events	
Patient-reported outcomes	
IBDQ	The Inflammatory Bowel Disease Questionnaire (IBDQ) is a 32-item questionnaire for subjects with IBD that will be used to evaluate the disease-specific health-related quality of life across 4 dimensional scores: bowel (loose stools, abdominal pain), systemic (fatigue, altered sleep pattern), social (work attendance, need to cancel social events), and emotional (anger, depression, irritability) Scores range from 32 to 224, with higher scores indicating better HRQoL.
SF-36	The Short Form 36 questionnaire (SF-36) consists of eight multi-item scales: limitations in physical functioning due to health problems; limitations in usual role activities due to physical health problems; bodily pain; general mental health (psychological distress and well-being); limitations in usual role activities due to personal or emotional problems; limitations in social functioning due to physical or mental health problems; vitality (energy and fatigue); and general health perception. Scales are scored from 0 to 100, with higher scores indicating better health.

EQ-5D	The EuroQol Questionnaire 5 Dimensions (EQ-5D) is a self-administered, non–disease-specific measure of health status that provides a simple descriptive profile and a single index value that can be used in the clinical and economic evaluation of health care and in population health surveys. Specifically, the EQ-5D assesses health outcomes from a wide variety of interventions on a common scale for purposes of economic evaluation, resource allocation and monitoring.

Outcomes included in the NMA

According to the submission file, the outcomes of interest analysed in the NMAs were based on a feasibility assessment and a review of outcomes assessed in recent publications and HTA submissions.

Both efficacy and safety outcomes were considered. The efficacy endpoints were analysed in both induction and maintenance NMAs. However, for safety endpoints only the induction data were assessed with NMAs, whereas for the maintenance safety endpoints the MAH proposed that comparison of conclusions by regulators may provide a more adequate comparison of the safety profile than an NMA.

Efficacy

Among the endpoint data extracted from the SLR, the following efficacy outcomes were considered feasible for both BF and NBF subgroups: clinical response, clinical remission and mucosal healing. Each of these outcomes was assessed at the end of induction and the end of maintenance (corresponding to 1 year of treatment). However, the full list of available efficacy outcomes for NMA analyses was not provided in the submission file, nor did the MAH report details on any feasibility assessment to determine which of these efficacy endpoints were selected for NMAs. To father support the choice of the efficacy outcomes, the MAH stressed that inclusion of these outcomes was aligned with recently conducted NMAs identified in the literature (e.g., Singh et al. [76]) and HTA submissions to NICE for tofacitinib [77] and vedolizumab [78] in UC. Nevertheless, the choices of outcomes that were derived from different bodies of evidence (HTA submissions to NICE for tofacitinib and vedolizumab, recent publications) could not be considered as supporting evidence in its own right.

Safety

The induction phases of RCTs in UC include standard treat-through designs from which it is possible to compare treatments in terms of safety endpoints. Given the request from EUnetHTA to compare safety outcomes across studies, an NMA of the following safety outcomes in induction were conducted: overall AEs; SAEs; overall infections; and serious infections.

However, the findings of the safety NMA in induction are considered to be limited owing to the short follow-up in the induction phases of trials (6–8 weeks) and low event counts for AEs such as serious infections, leading to uncertainty in the results and potential heterogeneity between the trial populations because of the inclusion of both NBF and BF patients.

For the maintenance-phase safety endpoints no NMA was performed as the MAH indicated a number of factors that could influence safety results. These included different definitions of the maintenance placebo safety population in re-randomised trials: a) only placebo induction responders continued on placebo (UNIFI and PURSUIT); b) induction placebo responders were re-randomised and placebo nonresponders were treated separately (OCTAVE); and c) placebo responders and nonresponders continued on placebo (GEMINI). There were also different efficacy and exposure levels in nonhomogeneous placebo arms and different inclusion criteria that may affect results for infections and could result in spurious conclusions about safety for both SAEs and infections. Typically, safety comparisons between the active arms and the withdrawal arm (placebo after active induction) during the maintenance phase were reported. However, there are limitations to this approach and withdrawal may result in additional (gastrointestinal) AEs. Given the limitations of conducting NMA on long-term safety outcomes, no NMA of safety was considered to be feasible for the 1-year time point. Further justification for why an NMA of 1-year outcomes is not appropriate is provided in the submission file [46].

Instead, the long-term safety profile for ustekinumab has been evaluated indirectly through integrated publication of safety across indications, assessment of regulatory submission documents, a comparison of the SmPCs for treatments and consideration of the PSOLAR study [79].

Outcome definitions

For the endpoints of clinical remission and clinical response, the MAH reported no significant discrepancies in definitions of the endpoints analysed across the trials forming the evidence base for NMAs. However, a potential source of heterogeneity due to differences in outcome measurements was observed for the mucosal healing endpoint, as the endoscopic score for the efficacy analyses used in NMAs was assessed by a local endoscopist in all but two trials (VARSITY did not report who performed the readings, while readings in OCTAVE were performed centrally), and no information on whether standardisation of local reading was performed in these studies was presented by the MAH. The EMA guideline [80] recommends that adjudication of endoscopic evidence of activity should be preferably done by central reading of the examinations, and if local reading of examination is performed, standardisation of reading should be demonstrated.

The definitions for each efficacy endpoint reported across studies are as follows:

- **Clinical response** was defined as a reduction in Mayo score of ≥3 points and ≥30% from the baseline score, with a decrease of ≥1 point on the rectal bleeding subscale or an absolute rectal bleeding score of 0 or 1.
- **Clinical remission** was most commonly defined as a Mayo score of ≤2 points, with no individual subscore >1 at the assessment time point. There were some discrepancies in the definition used in the following trials:

  OCTAVE-I 1, OCTAVE-I 2 and OCTAVE-S [51] defined remission as a full Mayo score of ≤2 with no subscore >1 and a rectal bleeding subscore of 0.

  Probert et al. [3] used a different scoring system, namely the Ulcerative Colitis Symptom Score (USCC), with a score of ≤2 defining the clinical remission.
- **Mucosal healing** was defined as achieving a Mayo endoscopy score of 0 or 1 across the studies.

There was a discrepancy in the definition of mucosal healing between UNIFI [60,61] and the other studies. UNIFI defined mucosal healing as “having both endoscopic healing and histologic healing”, with endoscopic healing defined as a Mayo endoscopy subscore of 0 or 1 and histologic healing defined as 0% to <5% neutrophils in epithelium, no crypt destruction and no erosions or ulcerations or granulations. Since the “endoscopic healing” definition in UNIFI was the same as the “mucosal healing” definition in the other studies, to maintain consistency for comparisons, the endoscopic healing results from UNIFI are reported in this study for mucosal healing. The term “mucosal healing” is used throughout the rest of the report. As shown in Table 4.5, the endoscopic score was assessed by a local endoscopist in most of the studies, except for OCTAVE, in which the endoscopic reading for both eligibility screening and efficacy analyses was based on centrally assessed Mayo subscores. In UNIFI, both a local reading and a central reading (performed by a readier who reviewed the video of endoscopy) were conducted; the efficacy analyses were based on the local reading data in UNIFI.

- **Central vs. local endoscopic readings:** A summary of the methods of central versus local endoscopic readings where reported is summarised in Table 4.5.

Table 4.5. Central versus local endoscopic reading

Trial name	Treatment	Endoscopic measurement	Data used for efficacy analysis
OCTAVE (I1, I2 and Sustain) [51]	Tofacitinib	Local and central	Central
ACT (1 and 2) [49] *	Infliximab	Local	Local
PURSUIT-SC [55]	Golimumab	Local	Local
PURSUIT-M* [55]	Golimumab	Local	Local
ULTRA 1 and 2 [52,53] *	Adalimumab	Local	Local
GEMINI 1 [**57**] *	Vedolizumab	Local	Local
UNIFI [60,61]	Ustekinumab	Local and central	Local

Notes: * Reference used: NICE tofacitinib submission [77] since information regarding local versus central reading was not discussed in the papers published for these trials. ** Both a local reading and a central reading performed by a reader who reviewed the endoscopy video were conducted; but the efficacy analyses were based on the local reading data.

Additional comparisons of the description and measurement of the main study outcomes are presented in Table 4.6, Table 4.7 and Table 4.8.

Table 4.6. Methods for data collection and analysis for clinical remission

Study reference/ID	Endpoint definition	Method of analysis
All trials except those below	Clinical remission	A Mayo score ≤2 points, with no individual subscore >1 at the time of assessment
OCTAVE trials	Clinical remission	A Mayo score of ≤2 with no subscore >1 and a rectal bleeding subscore of 0
UNIFI trial	Clinical remission	A Mayo score ≤2 points, with no individual subscore >1
Probert et al. [3]	Clinical remission	Ulcerative Colitis Symptom Score (USCC) ≤2

Table 4.7. Methods for data collection and analysis for clinical response

Study reference/ID	Endpoint definition	Method of analysis
All trials	Clinical response	A reduction in Mayo score of ≥3 points and ≥30% from the baseline score, with a decrease of ≥1 point on the rectal bleeding subscale or an absolute rectal bleeding score of 0 or 1

Table 4.8. Methods for data collection and analysis for mucosal healing

Study reference/ID	Endpoint definition	Method of analysis
All trials except UNIFI	Mucosal healing	A Mayo endoscopy score of 0 or 1
UNIFI trial	Mucosal healing	Both endoscopic healing (Mayo endoscopy score of 0 or 1) and histologic healing (0% to <5% neutrophils in epithelium, no crypt destruction and no erosions or ulcerations or granulations)

4.6 Risk of bias

In terms of study design, sequence generation was rated as having a low risk of bias for the majority of the studies except for ULTRA 1 Maintenance [52] and VARSITY [54]. The majority of the studies reported adequate methods for concealment of treatment allocation, with a low risk of bias, except for the ULTRA 1 Maintenance [52], Probert et al. [3] and VARSITY [54] studies.

Unclear or high risk of bias for blinding of participants and study personnel was assigned to the ULTRA 1 Maintenance [52], Suzuki et al. [1], PURSUIT-M [55] and PURSUIT-SC [55] studies.

Further studies did not describe how outcome assessment was blinded and rated as having an unclear risk of bias (ULTRA 1 Induction [82], ULTRA 2 [53], NCT00853099 [1], Probert et al [3], ACT 1 and 2 [49], PURSUIT-M [55], PURSUIT-SC [55] and GEMINI 1 [57]).

The presence or absence of method to adjust for missing data was unclear and rated as an unclear risk of bias or a high risk of bias in several studies [2,48,49,51,54].

The majority of the studies are rated as having a low risk of bias because of selective reporting and other sources of bias. Table 4.9 describe the risk of bias at the study level.

Table 4.9. Risk of bias in randomised studies at study level

Study reference/ID	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)*	Blinding of outcome assessment (detection bias)**	Incomplete outcome data addressed (attrition bias)	Selective reporting (reporting bias	Other potential sources of bias
UNIFI [60,61]	L	L	L	L	L	L	L	L
ULTRA 1 Induction [82]	L	L	L	U[2]	U[2]	L	U[3]	L
ULTRA 1 Maintenance [52]	H[1]	H[1]	H[1]	H[1]	H[1]	L	L	Patients switched to open-label adalimumab after completing the placebo-controlled induction period at week 8
ULTRA 2 [53]	L	L	L	U[2]	U[2]	L	L	L
NCT00853099 [1]	L	L	U[3]	U[2]	U[2]	L	U [3]	L
Jiang et al [2]	L	L	L	L	L	U [2]	L	L
Probert et al [3]	L	U[3]	L	U [2]	U [2]	L	L	Open-label IFX given to patients at 6 weeks for those continuing to receive active therapy in maintenance
Japic CTI060297 [48]	L	L	L	L	L	U [3]	L	L
ACT 1 and 2 [49]	L	L	L	U [2]	U [2]	H [4]	L	L
PURSUIT-M [55]	L	L	U [3]	U [2]	U [2]	L	L	L
PURSUIT-SC [55]	L	L	U [3]	U [2]	U [2]	L	L	L
NCT00787202 [4]	L	L	L	L	L	L	L	L
OCTAVE (I1, I2 and S) [51]	L	L	L	L	L	L (H in S trial) [5]	L	L
GEMINI 1 [57]	L	L	L	U[2]	U[2]	L	L	Open-label vedolizumab arm included in addition to the re-randomised

								arm to provide sufficient patient numbers for entering maintenance
NCT02039505 [59]	L	L	L	L	L	L	L	Includes open-label cohort 2 that received vedolizumab in addition to the randomised cohort
VARSITY [54]	U [2]	U [2]	L	L	L	U [2]	U [2]	Only conference presentation available

Abbreviations: L = Low risk of bias (RoB); U = unclear RoB; H = high RoB; IFX = infliximab.
Notes: * For self-reported outcomes including pain, function and global assessment. ** For assessor-reported outcomes. [1] Open-label follow-up of a phase 3, randomised, placebo-controlled 8-week induction trial evaluating the efficacy and safety of adalimumab for maintenance of clinical remission for patients with moderately to severely active ulcerative colitis. Results at 52 weeks are reported. Patients who completed the induction phase were eligible to continue in the follow-up phase of the study. [2] Not described in the methods or results sections. [3] Methods not described or further details not provided. [4] Intention to treat reported but >50% in placebo and >30% in IFX 5 and 10 mg arms did not complete. [5] Discontinuation rates differed between placebo (73.2%), tofacitinib 5 mg (43.9%) and tofacitinib 10 mg (35.7%): the main reason for discontinuation was lack of clinical response (66.7%; 35.4%; 27.0%); few discontinued due to adverse events (<5% in each arm).

External validity

The external validity of the trials included was assessed using the EUnetHTA guideline on the applicability of evidence in a relative effectiveness assessment of pharmaceuticals considering the following elements: population, intervention, comparator, outcomes and setting (PICO[S]) [83].

The evidence base for ustekinumab for the treatment of UC patients is supported mainly by the UNIFI clinical trial programme [60,61] and the NMAs conducted by the MAH [46].

All PICO elements generally have validity in clinical practice; most importantly, the population is identifiable in practice and the outcomes are of real importance to patients. Some specific comments on external validity for some of the PICO elements are as follows.

Population: The population was clinically relevant in the UNIFI trials and included patients reflective of a population of patients with moderately to severely active UC who have experienced failure of or intolerance to conventional and/or biologic therapies, including TNFi agents and/or vedolizumab. Specifically, UNIFI is the first trial to include vedolizumab failures in the study population. This reflects the current relevant patient population (NBF, TNFi failure and vedolizumab failure) in the real world.

NMAs: The issue of the adequacy of a common comparator in UC trials has recently become even more critical, as many new trials investigating new biologics in UC fail to study representative patient populations. Specifically, in a market in which several drugs are available for prescription, assignment to a placebo arm of a long-term study is not very appealing to patients. Consequently, the proportion of patients available in centres who are recruited to newer clinical trials is so low that the results can no longer be regarded as representative. Ghosh et al. [8] illustrated this difference in populations of CD patients by comparing time needed to enrol and/or complete a given IBD trial, which has changed dramatically: “In 1999, the 54-week ACCENT 1 trial enrolled 573 anti–tumor necrosis factor (TNF)-therapy naïve patients with CD from 55 centres in 12 countries and was completed in approximately 2 years. In contrast, by 2008 the 52-week GEMINI 2 trial required 1115 patients with CD (62% prior TNFi exposed) from 285 centres in 39 countries and took almost 3.5 years to complete”. A similar difference in patient populations probably applies to UC. Furthermore, as described earlier, patient populations selected for new trials are also changing. Patients could be selected because of failure to a particular TNFi, any TNFi, any TNF and/or vedolizumab, any biologic except X, among others, implying a difference in patient populations. Nevertheless, the data from earlier clinical trials on biologics are not likely to be comparable with more recent RCTs as they include different populations of patients. Since a valid NMA relies on the assumption that the different sets of studies included in the analysis are on average similar in all important factors that may affect the relative effects, this NMA does not adhere to best practices and is not supported as a source of evidence informing the ranking of treatments.

Intervention: The intervention probably reflected the doses that will be used in clinical practice.

Comparators: There is no head-to-head study of active comparators for direct comparisons in relative effectiveness assessments. The appropriateness of comparators in NMA is supported by recent clinical practice guidelines and these were aligned with previous comparator trials in this disease area.

Outcomes: Outcomes were aligned with previous comparator trials in this disease area; AEs were also recorded as safety endpoints. Endpoints for 1 year of treatment were considered most relevant.

Setting: Different regions of the world were included, such as Western and Eastern European countries, Asia, Australia, New Zealand and the USA.

4.7 Results for clinical effectiveness and safety

Ustekinumab versus placebo: UNIFI-I

Of the 961 randomised subjects, 912 (94.9%) completed their study participation: 783 (81.5%) subjects entered the maintenance phase and 129 (13.4%) who did not enter the maintenance phase completed the final safety visit [60]. Forty-nine subjects (5.1%) discontinued participation during the study. A total of 20 subjects (2.1%) discontinued participation before week 8. The most common reason for discontinuation before week 8 was withdrawal of consent, reported for 14 subjects (1.5%; no subjects in the 6 mg/kg group, 5 [1.6%] in the 130 mg group, and 9 [2.8%] in the placebo group). Of the remaining 29 subjects who discontinued participation, four discontinued at week 8 and 25 after week 8. Three subjects were unblinded during the study (2 in the 6 mg/kg group and 1 in the placebo group). None of these subjects completed the week 8 visit but all three completed the final safety visit.

The results demonstrate that at week 8, significantly greater proportions of subjects in the ~6 mg/kg and 130 mg groups achieved a clinical response (61.8% and 51.3%) and clinical remission (15.5% and 15.6%, respectively) compared with subjects in the placebo group (31.3%; $p < 0.001$ for both comparisons and 5.3%; $p < 0.001$ for both comparisons).

At week 8, significantly greater proportions of subjects in the 6 mg/kg (27.0%) and 130 mg (26.3%) groups achieved endoscopic healing compared with subjects in the placebo group (13.8%; $p < 0.001$ for both comparisons).

At baseline, median IBDQ scores were similar across all treatment groups. At week 8, the median improvement in IBDQ score from baseline was significantly greater in the 6 mg/kg (31.0) and 130 mg (31.5) groups compared with the placebo group (10.0; $p < 0.001$ for both comparisons).

HRQoL

The SF-36 evaluates eight individual subscales (physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health) and the physical component summary (PCS; calculated from the physical functioning, role-physical, bodily pain and general health subscales) and the mental component summary (MCS; calculated from the vitality, social functioning, role-emotional and mental health subscales) scores. Each of the SF-36 subscales and two summary scores were derived and were scaled to scores with a mean of 50 and SD of 10. Higher scores indicate better general health status.

At week 8, significantly greater median increases in PCS, MCS and all eight SF-36 scale scores were observed for both ustekinumab groups compared with the placebo group. In addition, greater proportions of subjects in both ustekinumab groups achieved a clinically meaningful improvement (≥5 points) from baseline in both PCS and MCS scores, indicating greater improvement in general HRQoL.

At week 8, median changes from baseline in the EQ-5D index and the health state EuroQol Visual Analogue Scale (EQ-VAS) were significantly greater in both ustekinumab groups than in the placebo group.

Endpoints according to BF status

Clinical endpoints by BF status

For subjects with and without a history of BF, the proportions of subjects who achieved clinical response, clinical remission and endoscopic healing were significantly greater in the ~6 mg/kg and 130 mg groups compared with subjects in the placebo group ($p < 0.025$ for both comparisons).

Although the treatment differences were generally similar between the BF and NBF groups, the proportions of subjects who achieved each endpoint were consistently lower across all treatment groups for those with a history of BF than for those with no history of BF.

Of the subjects with a history of BF, significantly greater proportions in the 6 mg/kg (13.3%) and 130 mg (11.6%) groups achieved clinical remission at week 8 (US definition) than in the placebo group (2.5%; $p < 0.001$ and $p = 0.001$, respectively).

Of the subjects without a history of BF, significantly greater proportions in the 6 mg/kg (25.0%) and 130 mg (21.8%) groups achieved clinical remission at week 8 than in the placebo group (10.1%; $p < 0.001$ and $p = 0.005$, respectively).

Endoscopic healing at week 8 by BF status

Of the subjects with a history of BF, significantly greater proportions in the 6 mg/kg (21.1%) and 130 mg (18.3%) groups achieved endoscopic healing at week 8 than in the placebo group (6.8%; $p < 0.001$ and $p = 0.002$, respectively).

Of the subjects who did not have a history of BF, significantly greater proportions in the 6 mg/kg (33.3%) and 130 mg (34.6%) groups achieved endoscopic healing at week 8 than in the placebo group (20.9%; $p = 0.014$ and $p = 0.006$, respectively).

Clinical response at week 8 by BF status

Of the subjects with a history of BF, significantly greater proportions in the 6 mg/kg (57.2%) and 130 mg (45.1%) groups achieved a clinical response at week 8 than in the placebo group (27.3%; $p < 0.001$ for both comparisons).

Of the subjects with no history of BF, significantly greater proportions in the 6 mg/kg and 130 mg groups (66.7% and 57.7%, respectively) achieved a clinical response at week 8 than in the placebo group (35.4%; $p < 0.001$ for both comparisons).

Consistent with the clinical response rate in the primary analysis population, the proportion of subjects achieving a clinical response at week 8 was higher in the 6 mg/kg group than in the 130 mg group regardless of BF history.

Clinical endpoints stratified by BF status

Post hoc analyses were conducted to evaluate the efficacy of the clinical endpoints of clinical remission (global and US definitions), endoscopic healing, clinical response and mucosal healing stratified by BF status as follows: subjects who were biologic-naïve, subjects with BF to at least one TNFi (regardless of vedolizumab) and subjects with BF to both TNFi and vedolizumab.

Subjects who were biologic-naïve comprised 94.3% of the group with no history of BF (443 of 470 subjects). In this subpopulation, the proportions of subjects who achieved each of the clinical endpoints of clinical remission (global and US definitions), endoscopic healing clinical response, and mucosal healing were significantly greater in the 6 mg/kg and 130 mg groups than in the placebo group ($p < 0.05$ for both comparisons).

Subjects with BF to at least one TNFi (regardless of vedolizumab) comprised 98.8% of those with a history of BF (485 of 491 subjects). In this subpopulation, the proportions who achieved each of the clinical endpoints of clinical remission (global and US definitions), endoscopic healing, clinical response and mucosal healing were significantly greater in the 6 mg/kg and 130 mg groups than in the placebo group ($p \leq 0.002$ for all comparisons).

Subjects with BF to both TNFi and vedolizumab comprised 16.6% of those randomised in this study (160 of 961 subjects) and 32.6% of those with a history of BF (160 of 491 subjects). In this subpopula-

tion, the proportions who achieved clinical remission (global definition) were significantly greater in the ~6 mg/kg and 130 mg groups than in the placebo group (p = 0.033 and p = 0.019, respectively).

For the other clinical endpoints (clinical remission [US definition], endoscopic healing, clinical response and mucosal healing) statistical significance was not reached. However the proportions of subjects who achieved each endpoint were greater in the 6 mg/kg and 130 mg groups than in the placebo group. The exception is the result for the clinical response endpoint for the 6 mg/kg group, for which statistical significance was reached (p = 0.036).

Hospitalisations and UC-related surgeries

Through week 8 of UNIFI-I, the proportion of subjects with UC-related hospitalisation was significantly lower in the ustekinumab recommended dose group (1.6%, 5/322) than in the placebo group (4.4%, 14/319). No subjects in group receiving the recommended ustekinumab induction dose underwent UC-related surgeries compared to 0.6% of subjects(2/319) in the placebo group.

Table 4.10 and Table 4.11 summarise results for comparison of ustekinumab with placebo in the UNIFI-I [60] and UNIFI-M [61] trials.

Table 4.10. UNIFI-I induction trial of ustekinumab versus placebo: summary of results for clinical response, clinical remission and endoscopic healing outcomes at week 8 in the ITT analysis for the full population and BF and NBF groups

Study reference/ID	Ustekinumab 6 mg/kg		Ustekinumab 130 mg		Placebo		Adjusted treatment difference vs. placebo (95% CI)	
	N	*n* (%)	*N*	*n* (%)	*N*	*n* (%)	Ustekinumab 6 mg/kg	Ustekinumab 130 mg
UNIFI-I [60]								
Clinical response	322	199 (61.8)	320	164 (51.3)	319	100 (31.3)	30.5 (23.2–37.8); p < 0.001	19.9 (12.5–27.3); p < 0.001
BF patients	166	95 (57.2)	164	74 (45.1)	161	44 (27.3)	p < 0.001	p < 0.001
NBF patients	156	104 (66.7)	156	90 (57.7)	158	56 (35.4)	p < 0.001	p < 0.001
Clinical remission	322	50 (15.5)	320	50 (15.6)	319	17 (5.3)	10.2 (5.6–14.8); p < 0.001	10.3 (5.7–14.9); p < 0.001
BF patients	166	21 (12.7)	164	19 (11.6)	161	2 (1.2)	11.4 (6.1–16.7); p < 0.001	10.4 (5.2–15.5); p < 0.001
NBF patients	156	29 (18.6)	156	31 (19.9)	158	15 (9.5)	9.0 (1.4–16.5); p = 0.022	10.4 (2.7–18.1); p = 0.009
Endoscopic healing	322	87 (27.0)	320	84 (26.3)	319	44 (13.8)	13.7 (7.3–19.3); p < 0.001	12.4 (6.5–18.4); p < 0.001
BF patients	156	52 (33.3)	164	30 (18.3)	161	11 (6.8)	p < 0.001	p = 0.002
NBF patients	156	52 (33.3)	156	54 (34.6)	158	33 (20.9)	p = 0.014	p = 0.006

Abbreviations: BF = biologic failure; CI = confidence interval; *n* = number of patients with at least one event; *N* = number of patients analysed; NBF = nonbiologic failure.

Ustekinumab versus placebo: UNIFI-M

A total of 783 subjects who completed the induction study and achieved a clinical response to the induction study agent were enrolled in the maintenance phase [61]. Of these, 523 subjects were in the targeted primary population for the maintenance study and were randomised (i.e., clinical response to IV ustekinumab induction; randomised subjects) and 260 were not part of the primary population for the maintenance study and were not randomised (i.e., placebo induction responders and ustekinumab induction delayed responders; nonrandomised subjects).

All subjects in this maintenance study entered from the induction study. The primary population is the population of randomised subjects. Therefore, presentation of data focuses on the randomised subjects, while data for the nonrandomised subjects are also provided with a focus on the ustekinumab induction delayed-responder group.

In general, the treatment groups were well balanced in terms of demographics, disease characteristics and concomitant and previous therapies. No clinically relevant imbalances were observed. The inflammatory burden, judged in terms of faecal calprotectin, was slightly higher in the ustekinumab than in the placebo group. The difference is of doubtful relevance and in any case is not in favour of ustekinumab and thus would not induce any bias in favour of the test drug.

Subjects were randomised according to clinical remission and steroid use at maintenance baseline. The proportions of patients in each arm remained similar between induction and maintenance for important variables such as UC disease characteristics and prior BF status. In 24.6%, 14% and 10.2% of subjects in the placebo, q12w and q8w arms, respectively, study treatment was discontinued. These rates do not appear to be higher than in phase 3 studies for other therapies in UC.

The primary efficacy analysis set consisted of 523 subjects with a clinical response to IV ustekinumab induction who were randomised to ustekinumab 90 mg SC q8w or 90 mg SC q12w or placebo. All efficacy analyses were based on the ITT principle; therefore, subjects were analysed according to the group to which they were assigned regardless of the treatment received.

The primary endpoint of clinical remission was met, with a treatment difference of 14.5% for the q12w group and 19.7% for the q8w group. The secondary endpoint of steroid-free clinical remission was considered the other key endpoint in view of the current Committee for Medicinal Products for Human Use guidelines. This endpoint was also met and the proportions of patients meeting each guideline are >95% concordant. The other controlled secondary endpoints were met. There was a very modest treatment benefit for ustekinumab 90 mg SC q8w versus q12w except for sustained clinical remission; although this had the weakest power for detection of changes among the major secondary endpoints.

Overall, ustekinumab SC maintenance therapy achieved significantly greater clinical remission and response through to week 44 compared to placebo and achieved greater endoscopic healing and corticosteroid-free remission at week 44 in this cohort of subjects with moderately to severely active UC who had responded to ustekinumab IV induction therapy.

HRQoL

HRQoL was assessed using the IBDQ, SF-36 and EQ-5D questionnaires. Improvements achieved at week 8 (significantly greater and clinically meaningful improvements in IBDQ total score, EQ-5D and EQ-5D VAS, and SF-36 MCS and PCS scores compared to placebo) were maintained in ustekinumab-treated patients in UNIFI-M through to week 44.

Primary and major secondary results by BF status

The primary endpoint and the key secondary endpoint of steroid-free clinical remission were met in subjects with and without a BF history. The other controlled secondary endpoints were met. There was a treatment benefit of >10% for ustekinumab 90 mg SC q8w versus q12w among BF subjects. The q8w and q12w results were similar among NBF patients.

It is noted that subjects who were biologically naïve accounted for 94.2% of the BNF group. Subjects who received at least one TNFi accounted for 99.2% of the BF group regardless of vedolizumab, as only 2/249 had received vedolizumab alone in this group.

Week 16 responders to ustekinumab induction

Ustekinumab-treated patients without a response at week 8 of UNIFI-I received 90 mg SC ustekinumab at week 8 (36% of patients). Of those subjects, 9% of patients who were initially randomised to the recommended induction dose achieved clinical remission and 58% achieved a clinical response at week 16. Patients without a clinical response to ustekinumab induction at week 8 of the UNFI-I study but had a response at week 16 (157 patients) entered the nonrandomised UNIFI-M subgroup and continued to receive maintenance dosing every 8 weeks; among these patients, a response was maintained for the majority (62%) and 30% achieved remission at week 44.

Hospitalisations and UC-related surgeries

Through to week 44 of UNIFI-M, a significantly lower number of UC-related hospitalisations was observed in the combined ustekinumab group (2.0%, 7/348) than in the placebo group (5.7%, 10/175). A lower number of subjects in the ustekinumab group (0.6%, 2/348) underwent UC-related surgeries than in the placebo group (1.7%, 3/175) through to week 44.

Table 4.11. UNIFI-M maintenance trial of ustekinumab versus placebo: summary of results for clinical response, clinical remission and endoscopic healing outcomes at week 44 in the ITT analysis for the full population and BF and NBF groups

Study reference/ID	Ustekinumab SC 90 mg q12 w		Ustekinumab SC 90 mg q8w		Placebo sc		Adjusted treatment difference (95% CI) vs. placebo	
	N	*n* (%)	*N*	*n* (%)	*N*	*n* (%)	Ustekinumab SC 90 mg q12w	Ustekinumab SC 90 mg q8w
UNIFI-M [61]								
Maintained clinical response	172	117 (68.0)	176	125 (71.0)	175	78 (44.6)	23.5 (13.7–33.3); $p < 0.001$	26.4 (16.6–36.1); $p < 0.001$
BF patients	70	39 (55.7)	91	59 (64.8)	88	34 (38.6)	$p < 0.01$	$p < 0.001$
NBF patients	102	78 (76.5)	85	66 (77.6)	87	44 (50.6)	$p < 0.001$	$p < 0.001$
Maintained clinical remission	172	66 (38.4)	176	77 (43.8)	175	42 (24.0)	14.5 (5.5–23.6); $p < 0.01$	19.7 (10.3–29.0); $p < 0.001$
BF patients	70	16 (22.9)	91	36 (39.6)	88	15 (17.0)	$p < 0.01$	$p < 0.001$
NBF patients	102	50 (49.0)	85	41 (48.2)	87	27 (31.0)	$p < 0.01$	$p < 0.01$
Subjects in clinical remission not receiving concomitant corticosteroids	172	65 (37.8)	176	74 (42.0)	175	41 (23.4)	14.5 (5.5–23.6); $p < 0.01$	18.5 (9.3–27.8); $p < 0.001$
BF patients	70	16 (22.9)	91	34 (37.4)	88	14 (15.9)	$p < 0.01$	$p < 0.001$
NBF patients	102	49 (48.0)	85	40 (47.1)	87	27 (31.0)	$p < 0.01$	$p < 0.01$
Subject with endoscopic healing	172	75 (43.6)	176	90 (51.1)	175	50 (28.6)	15.2 (5.8–24.6); $p < 0.01$	22.5 (12.8–32.2); $p < 0.001$
BF patients	70	18 (25.7)	91	41 (45.1)	88	20 (22.7)		$p < 0.001$
NBF patients	102	57 (55.9)	85	49 (57.6)	87	30 (34.5)	$p < 0.01$	$p < 0.01$

Abbreviations: BF = biologic failure; CI = confidence interval; *n* = number of patients with at least one event; *N* = number of patients analysed; NBF = nonbiologic failure.

Results from the clinical effectiveness NMAs

For rating the quality of the evidence from a network meta-analysis, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach was not applied, by either the MAH or the authors of this joint assessment [73].

Results from both the induction and the maintenance NMAs are discussed here, although comparisons of active treatments are considered more relevant at long-term time points than at induction. The maintenance NMAs were conducted to assess the continued year-long treatment regimens to provide a more complete picture of the long-term relative efficacy of treatments received in clinical practice.

The evidence drawn from all the networks was based on indirect data in a setting in which heterogeneity or inconsistency of a network could not be statistically assessed, and assessment of statistical heterogeneity between trials informing each comparison was limited and likely underpowered. Even in the networks which included head-to-head comparison, a closed-loop inconsistency assessment was not feasible because of data imputation in one of the trials. As a network is most justifiable under an assumption of consistency between different sources of evidence, the lack of this data largely increases uncertainty regarding NMA results.

The results of indirect comparisons, performed via Bayesian NMA for different endpoints and patient subgroups and run separately for induction and maintenance trials, are presented in detail in Table 4.12–Table 4.24 and Figure 4.4–Figure 4.25. In general, the CrIs for most endpoints were quite wide, with few comparisons showing statistical significance of ustekinumab versus another active treatment (significance at all three endpoints: adalimumab in induction networks for both NBF and BF populations; adalimumab and golimumab in 1-year maintenance networks for NBF patients. For details see the results of individual NMAs. In all the networks ustekinumab was better than placebo). While this is partly expected as the assessment is based on indirect comparisons, which add additional uncertainty, in some models the CrI width hampers interpretation. For example, in the case of NMA results for clinical remission for BF patients at induction, the upper CrI bound for the OR reaches a value of 95, which seriously undermines confidence in results from such a model.

Regarding the ranking of treatments, the MAH presented both SUCRA values [70,46] and the probability of ustekinumab being better than a comparator. However, in this submission all the networks had very wide CrIs. It is known that the probability of a treatment being better underperforms when the treatment effect of a intervention is largely uncertain, as an intervention with the most uncertainty can have the greatest probability. Therefore, the probability of ustekinumab being better was not used in interpretation. To avoid biased data analysis, the Cochrane handbook recommends use of SUCRA rankings [85,86]. Nonetheless, even with SUCRA values there were concerns with interpretation of the rankings, as these were established from low-certainty evidence (i.e., scarce data, heavy imputation, indirect, statistically confirmed and probable heterogeneity analysed via fixed-effects NMA) and are by definition based on any differences among the treatment effects, no matter how small. Therefore, under the conditions of large uncertainness in effect sizes and few significant effects, it is possible that a treatment with the higher rank is better than another treatment only by a small amount that is not clinically meaningful [87]. Consequently, while the ustekinumab interventions in the maintenance NMAs seemed to be consistently be ranked top, caution is required in interpreting the results as they could be misleading.

Induction-phase NMA

Overall, treatment effects estimated via the induction NMAs were entirely based on indirect evidence as the networks did not include any head-to-head comparison of different biologics. In addition, the evidence was derived from a limited number of trials ($n \leq 11$) informing from five to eight comparisons to placebo. The statistical heterogeneity of trials informing a direct comparison to placebo was assessed in all induction NMAs in which such an analysis was possible (networks including at least 1 comparison with at least 2 trials): three NMAs were performed on the endpoints of clinical remission, clinical response and mucosal healing for NBF patients and one NMA on clinical response for BF patients with only one comparison informed by multiple trials. All pairwise meta-analyses were based on two trials (except for infliximab 5mg/kg vs. placebo in the NMA for clinical remission for NBF patients, with three trials) and were thus probably underpowered for exploration of heterogeneity. Nevertheless, significant statistical heterogeneity was identified using Cochran's Q test at a significance level of 0.1 in the two networks for NBF patients: in trials reporting on clinical remission comparing infliximab 5mg/kg versus placebo ($n = 3$; $I^2 = 63\%$; Q test $p = 0.07$) and trials reporting on clinical response comparing infliximab 10 mg/kg versus placebo ($n = 2$; $I^2 = 70\%$; Q test $p = 0.07$). As Cochran's Q test has low power in a meta-analysis with only a few studies, a p value of 0.10, rather than the conventional level of 0.05, is used to determine statistical significance [62]. This also means that while a statistically significant result may indicate a problem with heterogeneity, a nonsignificant result must not be taken as evidence of no heterogeneity.

For all of the induction NMAs, the fixed-effects models were preferred over the random-effects models after comparison of the DIC values and assessment of convergence for the random-effects models. However, the DIC values for the fixed-effects and random-effects models were very similar. The difference in DIC, as shown in Appendix 6 of the NMA report [46] was less than 3 and ranged from 0.07 for the clinical remission endpoint in the induction phase (base-case analysis, BF patients) to 2.05 for mucosal healing in the induction phase (sensitivity analysis, NBF patients). Given that there is a good intuitive clinical rationale why heterogeneity may be expected but insufficient studies are available to detect it, the random-effects model should have been selected. This view is further supported by detection of statistical heterogeneity in pairwise meta-analyses based on direct comparisons, as described in the paragraph above.

An additional point of concern in interpreting the induction NMAs was the fact that in all the induction networks except for the one for clinical remission among NBF patients, data on tofacitinib from the OCTAVE 1 and OCTAVE 2 trials were merged. This was because the endpoints of interest were not reported separately for each trial (clinical response and mucosal healing) or no events in the placebo group in OCTAVE 1 were recorded (clinical remission in BF trial). As a consequence, a mega trial with less direct data was created for which the sample size in the tofacitinib arm (465 BF patients) was almost three times the size of the active treatments arms in the second largest trial (UNIFI BF patients, ustekinumab arms of 164 and 166 patients), which affected the results of the fixed-effects models. In the tofacitinib 10 mg versus placebo pairwise comparison in the clinical response network for BF patients, the fixed-effects results (OR 3.38, 95% CrI 2.18–5.23) and random-effects results (OR 1.07, 95% CrI 0.62–1.87) demonstrate how the mega trial significantly affects the fixed-effects model estimates. Moreover, such a pooled analysis yields overprecise results (artificially reduces the width of CrIs) and may lead to bias. A treat-as-one-trial approach for BF patients in the OCTAVE trials resulted in 124 patients in the placebo arm versus 465 in the active treatment arm.

NMA comparative effectiveness of ustekinumab for the NBF population at induction

Details for the network of evidence, data inputs and sample size of the studies included are presented below for each endpoint individually. Patients without a previous BF who received UST induction ustekinumab (6 mg/kg or 130 mg) had significantly higher odds of achieving a clinical response, remission and mucosal healing after an induction period of 6–8 weeks in comparison to placebo (Table 4.12, Table 4.13 and Table 4.14). While the point estimates for these comparisons were high, ranging from OR of 1.9 to 3.6, the findings need to be viewed with caution as the certainty of the estimated effect sizes was low, as evident from wide CrIs (ranging between 1.14 and 4.39 for clinical remission for ustekinumab 6 mg/kg vs. placebo). Moreover, for some of the comparisons the lower CrI bound for the estimated effect size was such that there was a possibility that the true effect was not clinically important. Results for ustekinumab 6 mg/kg versus placebo comparison allow for only a 14% increase in clinical remission (OR 2.19, CrI 1.14–4.39) and 15% in mucosal healing (OR 1.90, CrI 1.15–3.20). Regarding other active treatments, ustekinumab 6 mg/kg had significantly higher odds of achieving a clinical response than adalimumab 160/80 mg, with the point estimate indicating a twofold increase (OR 1.94, CrI 1.10–3.45). However, the lower bound for this CrI allows for a scenario with an increase of just 10%, which might not be clinically important.

The SUCRA value for the clinical response endpoint assigned to ustekinumab 6 mg/kg was 77% and seemed to be higher than the values for other active treatments except for infliximab interventions, with SUCRA of 79–86% (Table 4.15). However, for the clinical remission and mucosal healing endpoints the SUCRA values for the ustekinumab 6 mg/kg and 130 mg doses were grouped in the lower half of the ranks. Overall, according to SUCRA values, both interventions with ustekinumab seemed to be positioned better than placebo and adalimumab 160/80 for all endpoints, and better than golimumab 200/100 mg for clinical response and mucosal healing. These results should be interpreted with caution as the SUCRA rankings, as explained above, might be misleading.

Table 4.12. NMA results for clinical response for NBF patients at induction

Treatment	Median OR (CrI)	
	Ustekinumab 130 mg vs.	Ustekinumab 6 mg/kg vs.
Infliximab 5 mg/kg	0.61 (0.33–1.10); Pr = 5%	0.89 (0.49–1.63); Pr = 36%
Infliximab 10 mg/kg	0.65 (0.36–1.19); Pr = 8%	0.96 (0.53–1.76); Pr = 45%
Ustekinumab 6 mg/kg	0.68 (0.43–1.08); Pr = 5%	–
Vedolizumab 300 mg	0.78 (0.36–1.68); Pr = 26%	1.14 (0.52–2.47); Pr = 63%
Tofacitinib 10 mg	0.92 (0.50–1.71); Pr = 40%	1.36 (0.74–2.53); Pr = 84%
Ustekinumab 130 mg	–	1.47 (0.93–2.34); Pr = 95%
Golimumab 200/100 mg	1.09 (0.61–1.93); Pr = 61%	1.60 (0.90–2.84); Pr = 95%
Adalimumab 160/80 mg	1.32 (0.75–2.33); Pr = 83%	1.94 (1.10–3.45); Pr = 99%
Placebo	2.49 (1.58–3.96); Pr = 100%	3.66 (2.31–5.88); Pr = 100%

Abbreviations: CrI = credible interval; OR = odds ratio; NBF = nonbiologic failure; Pr = Bayesian probability of ustekinumab being better than the comparator.

Table 4.13. NMA results for clinical remission for NBF patients at induction

Treatment	Median OR (CrI)	
	Ustekinumab 130 mg vs.	Ustekinumab 6 mg/kg vs.
Infliximab 5 mg/kg	0.54 (0.24–1.22); Pr = 7%	0.489 (0.22–1.14); Pr = 5%
Vedolizumab 300 mg	0.52 (0.14; 1.70); Pr = 14%	0.48 (0.13–1.58); Pr = 12%
Infliximab 10 mg/kg	0.70 (0.31–1.62); Pr = 20%	0.64 (0.28–1.48); Pr = 15%
Golimumab 200/100 mg	0.80 (0.34–1.93); Pr = 31%	0.74 (0.31–1.78); Pr = 25%
Tofacitinib 10 mg	0.98 (0.38–2.42); Pr = 48%	0.90 (0.35–2.24); Pr = 41%
Ustekinumab 130 mg	–	1.47 (0.44–4.93); Pr = 39%
Ustekinumab 6 mg/kg	1.09 (0.62–1.92); Pr = 61%	–
Adalimumab 160/80 mg	1.08 (0.47–2.49); Pr = 57%	0.99 (0.43–2.30); Pr = 49%
Placebo	2.38 (1.24–4.78); Pr = 100%	2.19 (1.14–4.39] , Pr = 99%

Abbreviations: CrI = credible interval; OR = odds ratio; NBF = nonbiologic failure; Pr = Bayesian probability of ustekinumab being better than the comparator.

Table 4.14. NMA results for mucosal healing for NBF patients at induction

Treatment	Median OR [CrI]	
	Ustekinumab 130 mg vs.	Ustekinumab 6 mg/kg vs.
Infliximab 5 mg/kg	0.61 (0.32–1.15); Pr = 6%	0.57 (0.30–1.10); Pr = 5%
Infliximab 10 mg/kg	0.63 (0.33–1.20); Pr = 8%	0.59 (0.32–1.13); Pr = 6%
Vedolizumab 300 mg	0.68 (0.30–1.52); Pr = 17%	0.64 (0.29–1.45); Pr = 14%
Tofacitinib 10 mg	0.90 (0.44–1.81); Pr = 38%	0.85 (0.41–1.72); Pr = 32%
Ustekinumab 130 mg	–	0.94 (0.59–1.52); Pr = 41%
Ustekinumab 6 mg/kg	1.06 (0.66–1.70); Pr = 59%	–
Golimumab 200/100 mg	1.12 (0.60–2.09); Pr = 64%	1.06 (0.57–1.98); Pr = 57%
Adalimumab 160/80 mg	1.33 (0.72–2.49); Pr = 82%	1.26 (0.68–2.35); Pr = 77%
Placebo	2.01 (1.22–3.40); Pr = 100%	1.90 (1.15–3.20); Pr = 99%

Abbreviations: CrI = credible interval; OR = odds ratio; NBF = nonbiologic failure; Pr = Bayesian probability of ustekinumab being better than the comparator.

Table 4.15. Ranking of treatments by SUCRA value from most to least favourable for clinical response, clinical remission and mucosal healing for NBF patients at induction

Clinical response		Clinical remission		Mucosal healing	
Infliximab 5 mg/kg	86%	Infliximab 5 mg/kg	88%	Infliximab 5 mg/kg	87%
Infliximab 10 mg/kg	79%	Vedolizumab 300 mg	81%	Infliximab 10 mg/kg	84%
Ustekinumab 6 mg/kg	77%	Infliximab 10 mg/kg	67%	Vedolizumab 300 mg	76%
Vedolizumab 300 mg	64%	Golimumab 200/100 mg	58%	Tofacitinib 10 mg	56%
Tofacitinib 10 mg	49%	Tofacitinib 10 mg	43%	Ustekinumab 130 mg	47%
Ustekinumab 130 mg	41%	Ustekinumab 130 mg	42%	Ustekinumab 6 mg/kg	41%
Golimumab 200/100 mg	34%	Ustekinumab 6 mg/kg	36%	Golimumab 200/100 mg	36%
Adalimumab 160/80 mg	20%	Adalimumab 160/80 mg	35%	Adalimumab 160/80 mg	22%
Placebo	0%	Placebo	0%	Placebo	0%

Abbreviations: NBF = nonbiologic failure; SUCRA = surface under the cumulative ranking.

NBF: networks of evidence and data inputs

Clinical response

The network of evidence for clinical response among NBF patients at induction is shown in Figure 4.4. Ten studies were included in the analysis.

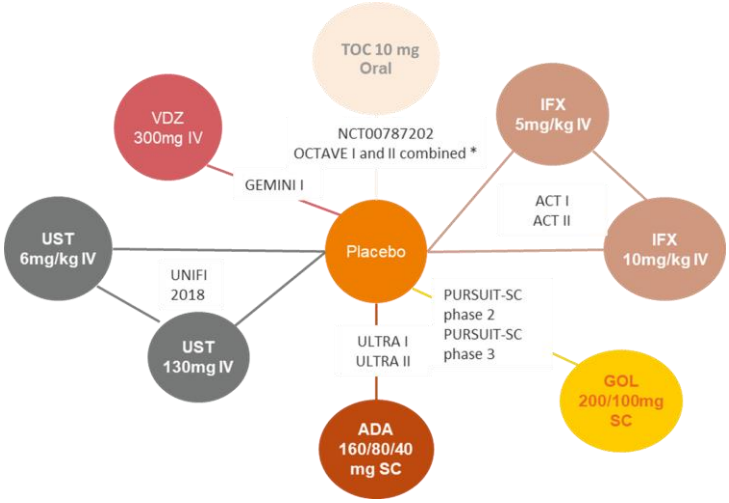


Figure 4.4. Network of evidence for clinical response for NBF patients at induction

* No results for clinical response reported for OCTAVE 1 and OCTAVE 2 separately.
Abbreviations: IFX = infliximab; ADA = adalimumab; VDZ = vedolizumab; UST = ustekinumab; GOL = golimumab; TOC = tofacitinib; IV = intravenous; SC = subcutaneous.

The data inputs from these studies are shown in Figure 4.5.

The sample size for the studies ranged from 56 (NTC00787202) to 550 (OCTAVE 1 and 2 combined). The response rates ranged from 43.9% in PURSUIT-SC phase 2 (GOL 200/100 mg; *n* = 41) to 69.4% in ACT 1 (IFX 5 mg; *n* = 121). The clinical response rate in the placebo arms across all trials ranged from 26.3% in GEMINI 1 (*n* = 76) to 45.5% in NTC00787202 (*n* = 33).

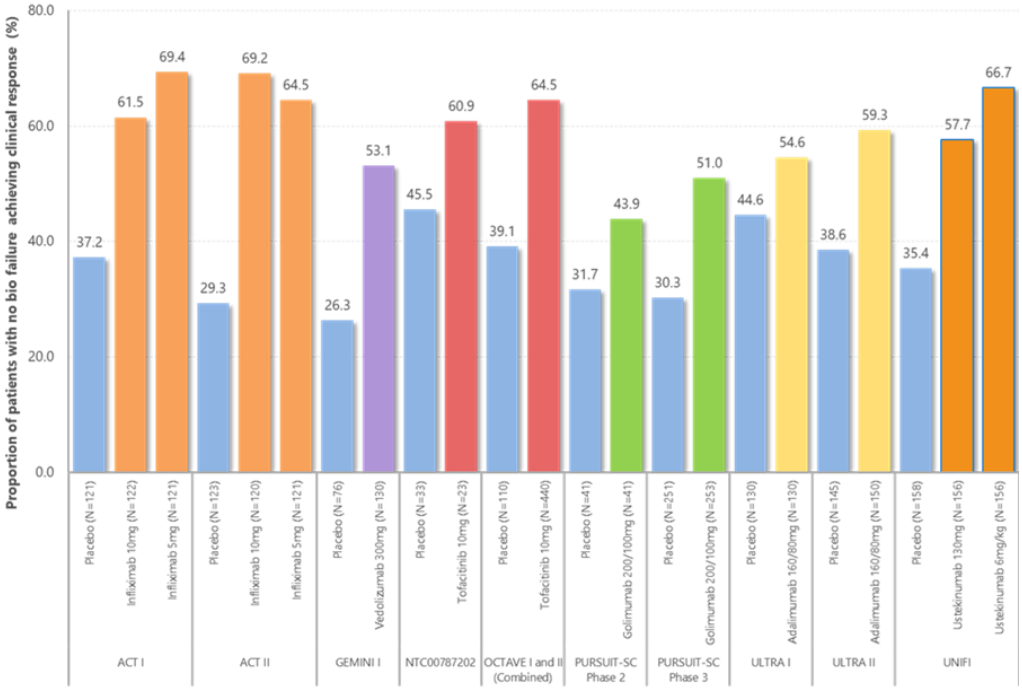


Figure 4.5. Data inputs for clinical response for NBF patients at induction.

Clinical remission

The network of evidence for clinical remission among NBF patients is shown in Figure 4.6.

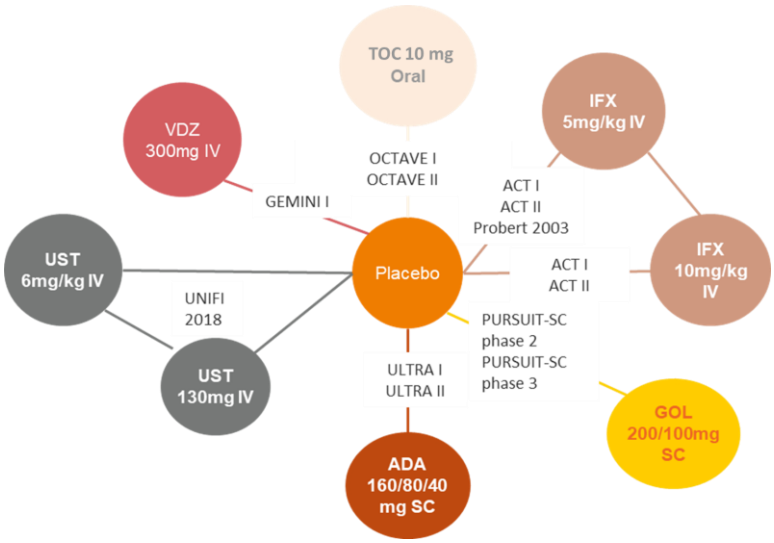


Figure 4.6. Network of evidence for clinical remission for NBF patients at induction

Abbreviations: ADA = adalimumab; GOL = golimumab; IFX = infliximab, IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab'

The data inputs from these studies are shown in Figure 4.7.

The sample size for the studies ranged from 43 (Probert et al.) to 254 (PURSUIT-SC phase 3). The remission rate ranged from 17.1% in PURSUIT-SC phase 2 (GOL 200/100 mg; *n* = 41) to 39% in Probert et al. (IFX 5 mg/kg; *n* = 23). UNIFI reported a remission rate of 9.5% in the placebo arm [46].

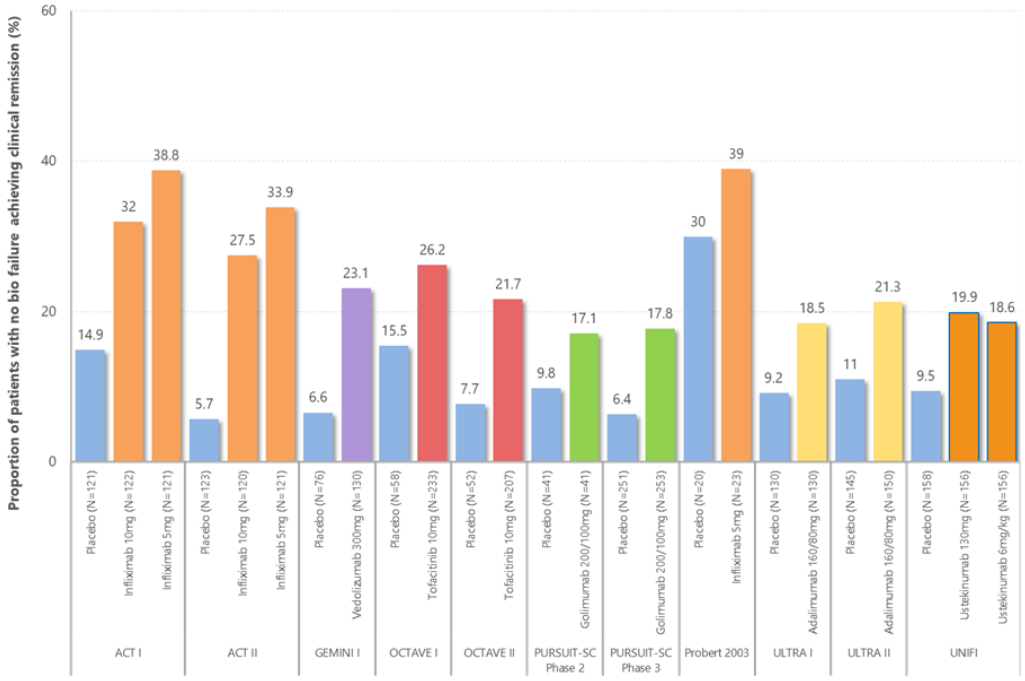


Figure 4.7. Data inputs for clinical remission for NBF patients at induction

Mucosal healing

The network of evidence for mucosal healing among NBF patients is shown in Figure 4.8. Nine studies were included in the analysis.

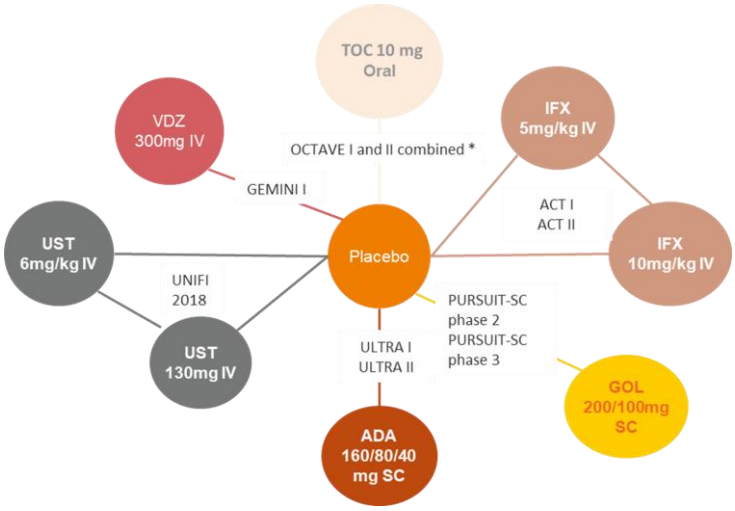


Figure 4.8. Network of evidence for mucosal healing for NBF patients at induction

* No results for mucosal healing reported for OCTAVE 1 and OCTAVE 2 separately.
Abbreviations: ADA = adalimumab; GOL = golimumab; IFX = infliximab; IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

The data inputs from these studies are shown in Figure 4.9. The sample size for the studies ranged from 82 (PURSUIT-SC phase 2) to 550 (OCTAVE 1 and 2 combined). The mucosal healing rate ranged from 33.3% in UNIFI (UST 6 mg/kg; *n* = 156) to 62% in ACT 1 (IFX 5 mg/kg; *n* = 121). The mucosal healing rate in the placebo arms across all trials ranged from 20.9% in UNIFI (*n* = 158) to 41.5% in ULTRA 1 (*n* = 130).

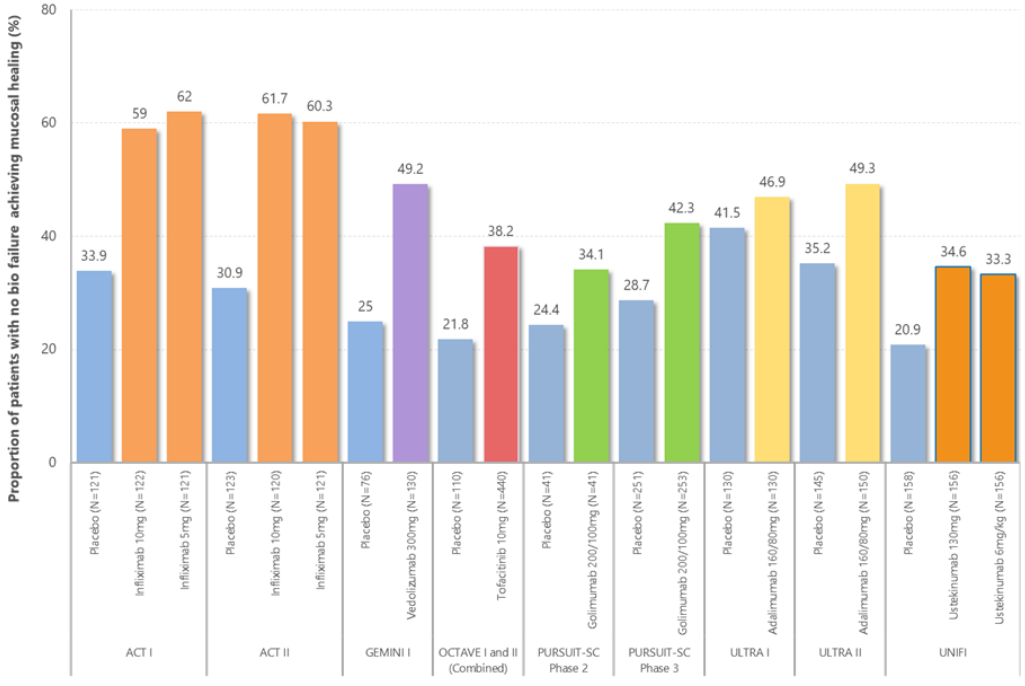


Figure 4.9. Data inputs for mucosal healing for NBF patients at induction

NMA comparative effectiveness of ustekinumab for the BF population at induction

Details for the network of evidence, data inputs and sample size for the studies included are presented below for each endpoint individually.

Overall, the data for induction NMAs for BF patients were even more scarce than for NMAs for NBF patients, with four to five trials informing up to five comparisons, which further affected the certainty of the evidence. For example, the induction NMA model for clinical remission in this population resulted in unrealistically high ORs, with the treatment effect estimated at 13.41 (CrI 3.62–94.58) for ustekinumab 6 mg/kg versus placebo and 9.97 (CrI 1.77–88.37) for ustekinumab 6 mg/kg versus adalimumab. This indicates that the data are too scarce for reliable estimates and points to scarce data bias. For this reason, the clinical remission model is reported but not further discussed. For the clinical response and mucosal healing NMAs, the estimates also had wide CrIs, but the estimates were more informative and thus were interpreted with caution.

The effect of merging of the OCTAVE 1 and OCTAVE 2 trials on estimates from fixed-effects versus random-effects model was evident. The fixed-effects results (OR 3.38, 95% CI 2.18–5.23) for clinical response in the tofacitinib 10 mg versus placebo pairwise comparisons point to quite different effects than the random-effects results (OR 1.07, 95% CI 0.62–1.87). Therefore, the NMA results for the fixed-effects model should be viewed in light of this finding.

Statistical heterogeneity in BF induction networks could only be assessed in the network for clinical response for tofacitinib 10 mg versus placebo, which included two trials: NTC00787202 with 25 patients (10 placebo + 15 active arm) and pooled OCTAVE data with 589 patients (124 placebo + 465 active arms). No heterogeneity was identified (n = 2; I^2 = 0; Q test p = 0.885) but as pointed out earlier, owing to the low power, this cannot be taken as evidence of a lack of heterogeneity. All the other comparisons to placebo were informed by just one trial.

The results show that patients who received ustekinumab 130 mg or 6 mg/kg at induction had significantly higher odds of achieving mucosal healing (OR 3.1, CrI 1.53–6.78; OR 3.7, CrI 1.86–8.04) or a clinical response (OR 2.2, CrI 1.39–3.53; OR 3.6, CrI 2.27–5.74, respectively) after a 6–8-week induction period in comparison to placebo. In addition, for those receiving ustekinumab 6 mg/kg at induction the odds of a clinical response (OR 2.48, CrI 1.17–5.31) and mucosal healing (OR 3.42, CrI 1.33–9.12) were significantly higher compared to adalimumab. For those receiving the lower ustekinumab dose of 130 mg, only the odds of achieving mucosal healing were significantly higher when compared to adalimumab (OR 2.85, CrI 1.10–7.68), but the lower CrI bound allows for the possibility that this increase is just 10%. Finally, ustekinumab 6 mg/kg had significantly higher odds for achieving a clinical response compared to the lower dose of 130 mg (OR 1.63, CrI 1.06–2.52), although again the lower CrI limit represents a difference of just 6%. All the CrIs were wide. For all the other comparisons of ustekinumab doses to other active treatments, the effects were nonsignificant and had very wide CrIs.

The SUCRA values showed that ustekinumab 6 mg/kg was rated as the best treatment for clinical response (SUCRA of 86%) and the second best treatment after tofacitinib for mucosal treatment (81%; Table 4.19). However, the effect size estimates in terms of the OR between ustekinumab 6 mg/kg and tofacitinib 10 mg for these endpoints were noninformative because of very wide, nonsignificant CrIs; results were similar for the CrIs for ustekinumab comparisons to other active treatments, except for adalimumab. For this reason, as explained before, the SUCRA rankings cannot be viewed as firm evidence for the ranking of clinical efficacy. The ustekinumab 130 mg dose seemed to be positioned worse than ustekinumab 6 mg/kg for clinical response and mucosal healing, which is in line with the significant difference in OR for clinical response between ustekinumab 6 mg/kg and 130 mg, but again the true effect size of this comparison ranged from clinically unimportant (6% increase) to a substantial effect (152% increase).

In summary, the lower sample sizes and event counts, particularly for clinical remission, together with previously stated limitations including the general comments on indirectness, expected heterogeneity and the use of fixed-effects models, as well as possible bias due to merging of the OCTAVE trials, suggest that there is more uncertainty and a higher risk of bias in the induction NMA results obtained for the BF compared to the NBF group.

Table 4.16. NMA results for clinical response for BF patients at induction

Treatment*	Median OR (CrI)	
	Ustekinumab 130 mg vs.	Ustekinumab 6 mg/kg vs.
Ustekinumab 6 mg/kg	0.61 (0.40–0.95); Pr = 1%	–
Tofacitinib 10 mg	0.64 (0.34–1.21); Pr = 8%	1.05 (0.55–1.98); Pr = 56%
Vedolizumab 300 mg	0.87 (0.35–2.11); Pr = 38%	1.43 (0.58–3.43); Pr = 78%
Ustekinumab 130 mg	–	1.63 (1.06–2.52); Pr = 99%
Adalimumab 160/80 mg	1.52 [0.71–3.25); Pr = 86%	2.48 (1.17–5.31); Pr = 99%
Placebo	2.20 (1.39–3.53); Pr = 100%	3.58 (2.27–5.74); Pr = 100%

Abbreviations: CrI = credible interval; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR >1 indicates greater odds of a response with ustekinumab than with the comparator, suggesting that ustekinumab performs better). Treatments are ordered by SUCRA values from most to least favourable.

Table 4.17. NMA results for clinical remission for BF patients at induction

Treatment*	Median OR (CrI)	
	Ustekinumab 130 mg vs.	Ustekinumab 6 mg/kg vs.
Tofacitinib 10 mg	0.54 (0.02–7.18); Pr = 32%	0.59 [0.02–7.92); Pr = 35%
Ustekinumab 6 mg/kg	0.90 (0.46–1.76); Pr = 38%	–
Ustekinumab 130 mg	–	1.11 (0.57–2.17); Pr = 62%
Vedolizumab 300 mg	3.27 (0.29–36.81); Pr = 84%	3.60 (0.32–40.71); Pr = 86%
Adalimumab 160/80 mg	9.01 (1.58–80.08); Pr = 99%	9.97 (1.77–88.37); Pr = 100%
Placebo	12.12 (3.24–86.24); Pr = 100%	13.41 (3.62–94.58); Pr = 100%

Abbreviations: CrI = credible interval; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR >1 indicates greater odds of a response with ustekinumab than with the comparator, suggesting that ustekinumab performs better). Treatments are ordered by SUCRA values from most to least favourable.

Table 4.18. NMA results for mucosal healing - biologic failure patients (induction)

Treatment*	Median OR [CrI]	
	Ustekinumab 130 mg vs.	Ustekinumab 6 mg/kg vs.
Tofacitinib 10 mg	0.73 (0.24–2.07); Pr = 28%	0.87 (0.29–2.46); Pr = 40%
Ustekinumab 6 mg/kg	0.84 (0.48–1.44); Pr = 26%	
Ustekinumab 130 mg		1.20 (0.69–2.08); Pr = 74%
Vedolizumab 300 mg	1.83 (0.63–5.40); Pr = 86%	2.19 (0.76–6.41); Pr = 93%
Adalimumab 160/80 mg	2.85 (1.10–7.68); Pr = 98%	3.42 (1.33–9.12); Pr = 99%
Placebo	3.12 (1.53–6.78); Pr = 100%	3.73 (1.86–8.04); Pr = 100%

Abbreviations: CrI = credible interval; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR >1 indicates greater odds of a response with ustekinumab than with the comparator, suggesting that ustekinumab performs better). Treatments are ordered by SUCRA values from most to least favourable.

Table 4.19. Ranking of treatments from most to least favourable with corresponding SUCRA values

Clinical response		Clinical remission		Mucosal healing	
Ustekinumab 6 mg/kg	86%	Tofacitinib 10 mg	85%	Tofacitinib 10 mg	86%
Tofacitinib 10 mg	82%	Ustekinumab 6 mg/kg	77%	Ustekinumab 6 mg/kg	81%
Vedolizumab 300 mg	59%	Ustekinumab 130 mg	71%	Ustekinumab 130 mg	68%
Ustekinumab 130 mg	47%	Vedolizumab 300 mg	44%	Vedolizumab 300 mg	40%
Adalimumab 160/80 mg	24%	Adalimumab 160/80 mg	18%	Adalimumab 160/80 mg	17%
Placebo	2%	Placebo	6%	Placebo	9%

BF: networks of evidence and data inputs

Clinical response

The network of evidence for clinical response among BF patients at induction is shown in Figure 4.10. A total of five studies were included in the analysis.

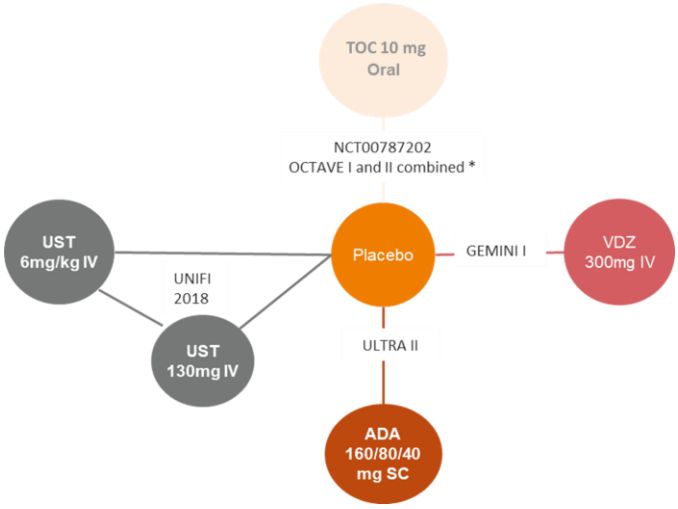


Figure 4.10. Network of evidence for clinical response for BF patients at induction.

* No results for clinical response reported for OCTAVE 1 and OCTAVE 2 separately.
Abbreviations: ADA = adalimumab; IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

The data inputs from these studies included in the network of evidence for clinical response among BF patients are shown in Figure 4.11.

The sample size for the studies ranged from 25 (NTC00787202) to 589 (OCTAVE 1 and 2 combined). The response rate ranged from 36.7% in ULTRA 2 (ADA 160/80 mg; *n* = 98) to 60% in NTC00787202 (TOC 10 mg; *n* = 10). The clinical response rate in the placebo arms across all trials ranged from 20.6% in GEMINI 1 (*n* = 63) to 33.3% in NTC00787202 (*n* = 15).

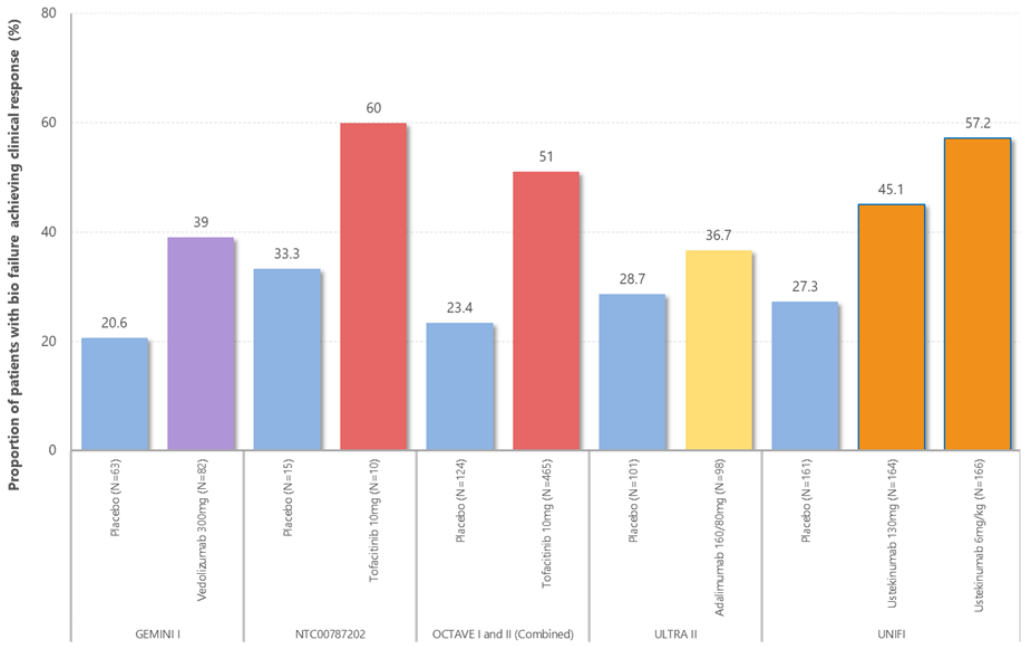


Figure 4.11. Data inputs for clinical response for BF patients at induction

Clinical remission

The network of evidence for clinical remission among BF patients at induction is shown in Figure 4.12. A total of four studies were included in the analysis.

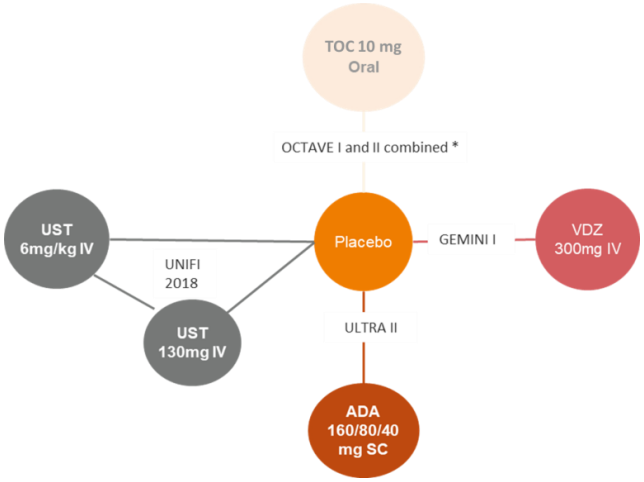


Figure 4.12. Network of evidence for clinical remission for BF patients at induction.

* OCTAVE 1 showed no events in the placebo group so the pooled study data was used as input.
Abbreviations: ADA = adalimumab; IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

The data inputs from these studies are shown in Figure 4.13. The sample size for the studies ranged from 145 (GEMINI 1) to 589 (OCTAVE 1 and 2 combined). The remission rate ranged from 9.2% in ULTRA 2 (ADA 160/80 mg; *n* = 98) to 12.7% in UNIFI (UST 6 mg/kg; *n* = 166). The clinical remission rate in the placebo arms across all trials ranged from 0.8% in OCTAVE 1 and 2 combined (*n* = 124) to 6.9% in ULTRA 2 (*n* = 101).

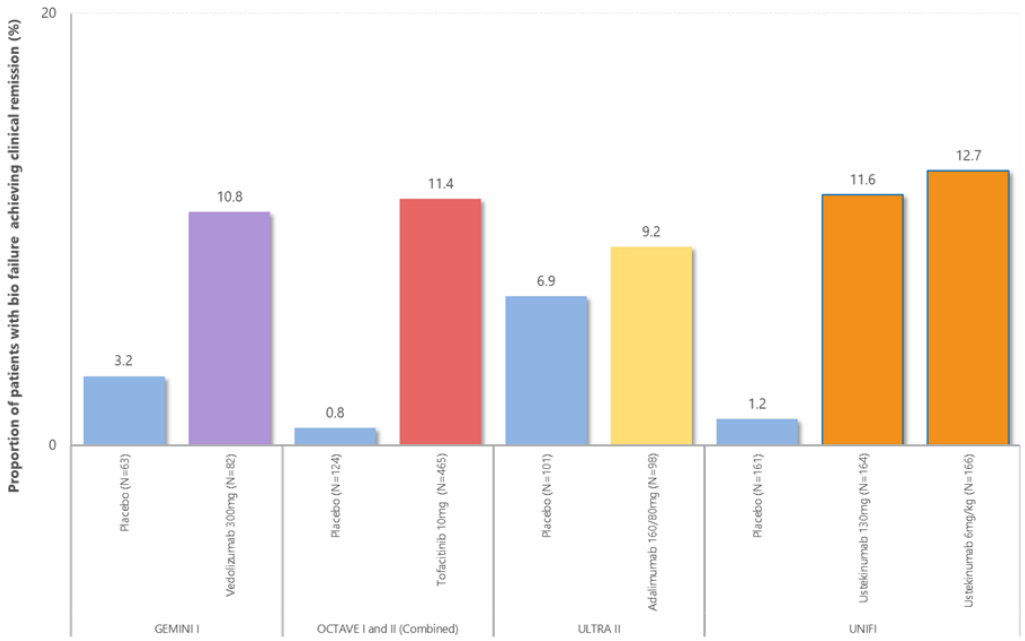


Figure 4.13. Data inputs for clinical remission for BF patients at induction

Mucosal healing

The network of evidence for mucosal healing among biologic failure patients is shown in Figure 4.14. A total of four studies were included in the analysis.

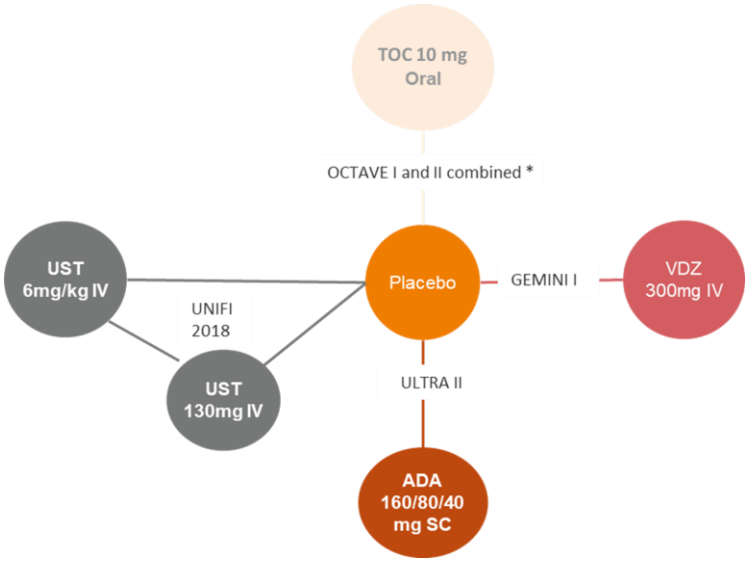


Figure 4.14. Network of evidence for mucosal healing for BF patients at induction

* No results for mucosal healing reported for OCTAVE 1 and OCTAVE 2 separately.
Abbreviations: ADA = adalimumab; IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

The data inputs from the four studies included in the network of evidence for mucosal healing among BF patients are shown in Figure 4.15.

The sample size for the studies ranged from 145 (GEMINI 1) to 589 (OCTAVE 1 and 2 combined). The mucosal healing rate ranged from 18.3% in UNIFI (UST 130 mg; *n* = 164) to 30.5% in GEMINI 1

(VDZ 300 mg; *n* = 82). The mucosal healing rate in the placebo arms across all trials ranged from 6.5% in OCTAVE 1 and 2 combined (*n* = 124) to 26.7% in ULTRA 2 (*n* = 101).

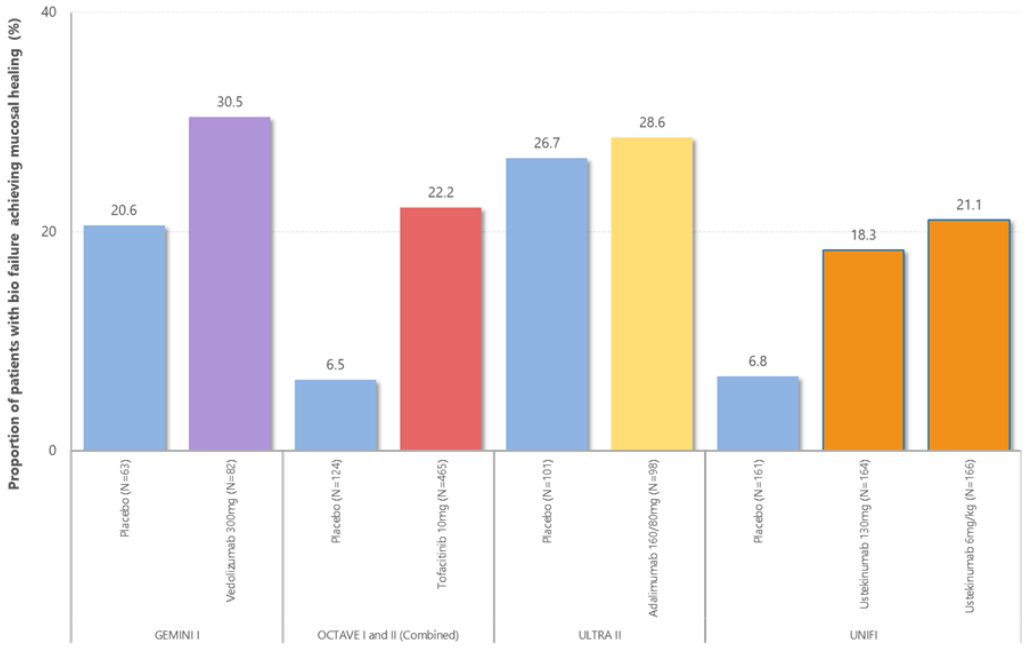


Figure 4.15. Data inputs for mucosal healing for BF patients at induction

One-year NMA

The results of the 1-year NMAs are presented for the base-case approach mimicking a treat-through design. Data for clinical efficacy at the end of 1 year are presented below. Full details on the calculations for these data are provided in the NMA report [46].

Treatment effects estimated via the maintenance NMAs were also mainly based on indirect evidence as the networks either did not include any head-to-head comparison of different biologics (clinical response networks for both NBF and BF patients) or did include one head-to-head comparison but no data were available on a comparison of one of the treatments to placebo (mucosal healing network for NBF patients). Only for the clinical remission networks for both NBF and BF patients was there a closed loop informed by VARSITY, GEMINI 1 and ULTRA 2 trials that would allow comparison of direct and indirect evidence, but the statistical agreement of direct and indirect treatment effects was not tested as the efficacy data in GEMINI 1 were partly imputed (see Appendix 3 on GEMINI data imputation in the NMA report [46]). There were no data to build the mucosal healing network in biologic failure patients.

Overall, the evidence was derived from a limited number of trials (from 4 to 7) informing from four to seven comparisons, with only one study informing each pair of treatments in all the maintenance networks. As there were not enough data to inform the between-trial standard deviation, the random-effects models did not converge since posterior distributions were dominated by the vague prior distribution. Therefore, only the fixed-effects models were discussed for the maintenance networks. To demonstrate lack of convergence in a random-effects maintenance network, the MAH provided results for the model of clinical response for 1-year NMA in induction responders with a prior biologic [46]. For the same reason (one trial per comparison) statistical heterogeneity among trials informing a direct comparison could not be assessed. However, while the maintenance random-effects models described may well be unreliable, the MAH decided not to use alternative methods such as external data or clinician estimates to inform the parameter as is usually done in the case of a limited network. Although the networks were too small to obtain reliable estimates of heterogeneity, based on the differences observed among registrational clinical trials of biological therapies in adults with IBDs, and in meta-analysis of placebo remission rates among UC patients with active disease that showed signifi-

cant heterogeneity, ranging from 0% to 40% [8,9], it is reasonable to assume there are differences between the trials included in the NMAs. While a random-effects model would explicitly model these differences and capture the uncertainty in the true treatment effect, fixed-effect models are likely to underestimate the uncertainty and possibly come to the wrong conclusion.

To allow for combining of data from different study designs in 1-year NMAs, the MAH recalculated maintenance endpoints for both placebo and active treatment arms of re-randomised trials using conditional probabilities and, when needed, performed imputation mainly from IPD data from other trials. As a result, endpoint data in four out of seven maintenance trials, mainly for the placebo arms, were imputed. This was all done in a situation in which, as explained above, the heterogeneity could not be assessed statistically and it was reasonable to assume it. In addition, along with the heavy imputation under the conditions of scarce data and expected heterogeneity, some of the data used for imputation were not direct evidence but were themselves extrapolated. For GEMINI 1, for which data are used for placebo–placebo imputation for BF induction nonresponders, data for clinical response and clinical remission for induction nonresponders were only provided for the mixed population (including BF and NBF patients). To estimate the percentage by population, the RR for patients responding after induction for NBF versus BF was calculated by arm and used to estimate the end response. However, a true sensitivity analysis to control for the effect of imputation was not applied. Although imputation for NBF induction responders was based on IPD data from UNIFI, ACT 1 and PURSUIT, the MAH provided simulations that investigated the change in effects when trial values are sampled randomly. The assumption was that trial data are drawn from the same underlying distribution, with the central tendency value corresponding to a weighted average for the data actually used in imputation, and in the majority of the runs it was the imputed value or a value close to the imputed value that was generated. The MAH also stated that for the clinical remission network for NBF patients it compared the corresponding arms of the VARSITY head-to-head trial of vedolizumab versus adalimumab with the results from ULTRA 2 for adalimumab and imputed the GEMINI 1 results for vedolizumab, which apparently corroborated the imputation, but no details on this analysis are provided. As for the sensitivity analysis mimicking a response-based approach, the analysis does not tackle the same underlying question as the base-case approach mimicking a treat-through design (it targets only induction responders) so it cannot be considered a sensitivity analysis.

NBF patients

Unlike for the induction NMAs and the maintenance NMAs for BF patients, for the maintenance NMAs for NBF patients, data on interventions with different doses were pooled for treatment arms, as apparently no dose–response relationship was observed according to results in the individual trials. Details on this analysis were not provided, but the MAH did provide results without the pooling for reference (Tables 32, 34 and 36 in the NMA report [46] give effect size estimates). No SUCRA values were provided for the unpooled analysis in the Missing points addendum. There were some discrepancies between the pooled and unpooled analyses, but given the very wide CrIs for treatment effects observed in all the maintenance NMAs for NBF patients, most of which include OR of 1, indicating that we have little knowledge about the effect and that further information is needed, interpretation of these discrepancies and of the results from pooled analyses is hampered. As a general comment, the claim that no dose–response exists in these particular networks (whereas the opposite was observed in all the others, suggesting an underlying biological phenomenon), which is supported only by statistical testing results, does not represent convincing evidence. Moreover, while pooling of doses might increase the analysis power as stressed by the MAH, it will also result in overprecise results (as data from two active treatment arms of a trial are pooled together and the sample size of the active arm in that trial is increased twice) which could affect the fixed-effects NMA results [10].

Overall, NBF patients who received ustekinumab 6 mg/kg at induction followed by ustekinumab 90 mg either q8w or q12w had significantly higher odds of achieving a clinical response, remission and mucosal healing after 1-year-long treatment regimen than those receiving placebo or the treatment sequences adalimumab 160/80/40mg → adalimumab 40 mg EOW or golimumab 200/100 mg → golimumab pooled (Table 4.20). While the point estimates for these comparisons were high, ranging from OR of 2.4 to 8.7, the findings need to be viewed with caution as the certainty for the estimated effect sizes was low, as evident from wide CrIs. For some of the said comparisons, the lower CrI bound for estimates of the effect sizes was such that there was a possibility that the true effect was not clinically important. For example, the NMA for clinical remission suggested only a 10–13% increase in this endpoint when the ustekinumab treatment sequence was compared to adalimumab (OR 2.43, CrI 1.10–5.42) or golimumab (OR 2.40, CrI 1.13–5.22). In addition to treatment sequences showing consistent results for all three endpoints, a clinical response was also significantly more likely

for patients treated with ustekinumab than those treated with infliximab or tofacitinib. Again, the interpretation of findings was not unambiguous, as the CrI for the OR for ustekinumab versus tofacitinib comparison allows for the possibility that the true effect size is below the minimal clinically important difference (OR 2.27, CrI 1.06–4.86). While all the point estimates for all the comparisons, respectable of their significance, point towards higher odds of achieving an endpoint for patients treated with ustekinumab, this fact ca not be used to support the absolute ustekinumab efficacy as all the CrIs are very wide, and for the nonsignificant ones we can only conclude that not much is known about the effect and that more data are needed (i.e., comparison of clinical response between UST and TOC: OR 1.51, CrI 0.64–3.51).

The SUCRA rankings seemed to consistently show ustekinumab pooled doses at the top of a rank for all the endpoints, with high SUCRA values between 89% and 98% (Table 4.21). In addition, the treatments placebo, golimumab and adalimumab, which were significantly worse than ustekinumab for all the endpoints, seemed to be consistently positioned at the bottom of the rank for all the endpoints. However, given that clear interpretation of effect sizes between active treatments is hindered, we are uncertain about the validity of the ranking presented.

Overall, given that various methodological issues related to maintenance models for the NBF population were raised and discussed, such as the paucity of the data, indirectness and the fact that inconsistency between different sources of evidence was not estimated (this includes NMA for clinical remission where inconsistency could be estimated due to a closed loop but the data of GEMINI I trial within the loop were imputed and the loop was apparently used to cross-check the imputation approach by comparing the re-calculated efficacy from GEMINI-1 with the results from VARSITY and ULTRA-II), that data were recalculated and imputed in majority of trials (either for induction responders and/or non-responders in placebo or active arm), that fixed effect model was used in a setting where heterogeneity is expected but could not be tested, and that there is a potential impact of bias due to pooling of doses, we are uncertain if ustekinumab improves clinical response, clinical remission, and mucosal healing after a 1-year treatment regimen when compared with golimumab and adalimumab, as well as clinical response when compared with infliximab and tofacitinib.

Table 4.20. NMA results for clinical response, clinical remission and mucosal healing for pooled doses for NBF patients at 1 year in the base case mimicking a treat-through approach

Treatment	Median OR (CrI) for UST 6 mg/kg → UST 90 mg pooled		
	Clinical response	Clinical remission	Mucosal healing
Placebo – Placebo	8.70 (5.03–15.40); Pr = 100%	5.11 (2.83–9.52); Pr = 100%	5.57 (3.19–9.92); Pr = 100%
Infliximab pooled – Infliximab pooled	2.62 (1.22–5.60); Pr = 99.31%	1.89 (0.83–4.29); Pr = 93.59%	1.43 (0.66–3.09); Pr = 81.59%
Adalimumab 160/80/40mg – Adalimumab 40mg EOW	4.76 (2.25–10.16); Pr = 100%	2.43 (1.10–5.42); Pr = 98.59%	2.91 (1.33–6.39); Pr = 99.62%
Golimumab 200/100mg – Golimumab pooled	3.76 (1.90–7.57); Pr = 99.99%	2.40 (1.13–5.22); Pr = 98.84%	2.79 (1.39–5.69); Pr = 99.81%
Tofacitinib 10mg - Tofacitinib pooled	2.27 (1.06–4.86); Pr = 98.21%	1.51 (0.64–3.51); Pr = 82.97%	1.94 (0.88–4.25); Pr = 95.11%
Vedolizumab 300mg – Vedolizumab 300mg q8w	1.93 (0.75–4.82); Pr = 91.45%	1.47 (0.65–3.33); Pr = 82.38%	1.60 (0.69–3.77); Pr = 86.24%

Abbreviations: CrI = credible interval; EOW = every other week; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR > 1 indicates higher odds of a response with ustekinumab than with the comparator, suggesting that ustekinumab performs better); q8w = every 8 weeks.

Table 4.21. Ranking of treatments from most to least favourable from 1-year NMA for NBF patients, with corresponding SUCRA values

Clinical response		Clinical remission		Mucosal healing	
UST 6-UST pooled	98%	UST 6-UST pooled	89%	UST 6-UST pooled	91%
VDZ 300-VDZ 300 pooled	73%	VDZ 300-VDZ 300 pooled	66%	IFX pooled-IFX pooled	69%
TOC 10-TOC pooled	66%	TOC 10-TOC pooled	62%	VDZ 300-VDZ 300 q8w	63%
IFX pooled-IFX pooled	57%	IFX pooled-IFX pooled	46%	TOC 10-TOC pooled	49%
GOL 200/100-GOL pooled	34%	GOL 200/100-GOL pooled	29%	GOL 200/100-GOL pooled	26%
ADA 160/80/40-ADA 40 EOW	22%	ADA 160/80/40-ADA 40 EOW	27%	ADA 160/80/40-ADA 40 EOW	23%

Abbreviations: ADA = adalimumab; EOW = every other week; GOL = golimumab; IFX = infliximab; SUCRA = surface under the cumulative ranking curve; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Networks of evidence and data inputs

The details, with total number of patients, number and proportion of responders per treatment arm at the end of maintenance and at the end of 1 year, are reported in Appendix 7 Data inputs for 1-year NMA in the MAH NMA report [46].

Clinical response

The network of evidence for clinical response among NBF patients is shown in Figure 4.16. Six studies were included in the analysis.

Figure 4.16. Network of evidence for clinical response for NBF patients in the 1-year base case mimicking a treat-through approach

Abbreviations: ADA = adalimumab; GOL = golimumab; IFX = infliximab; IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

The data inputs from the six studies included in the network of evidence for clinical response among NBF patients are shown in Figure 4.17.

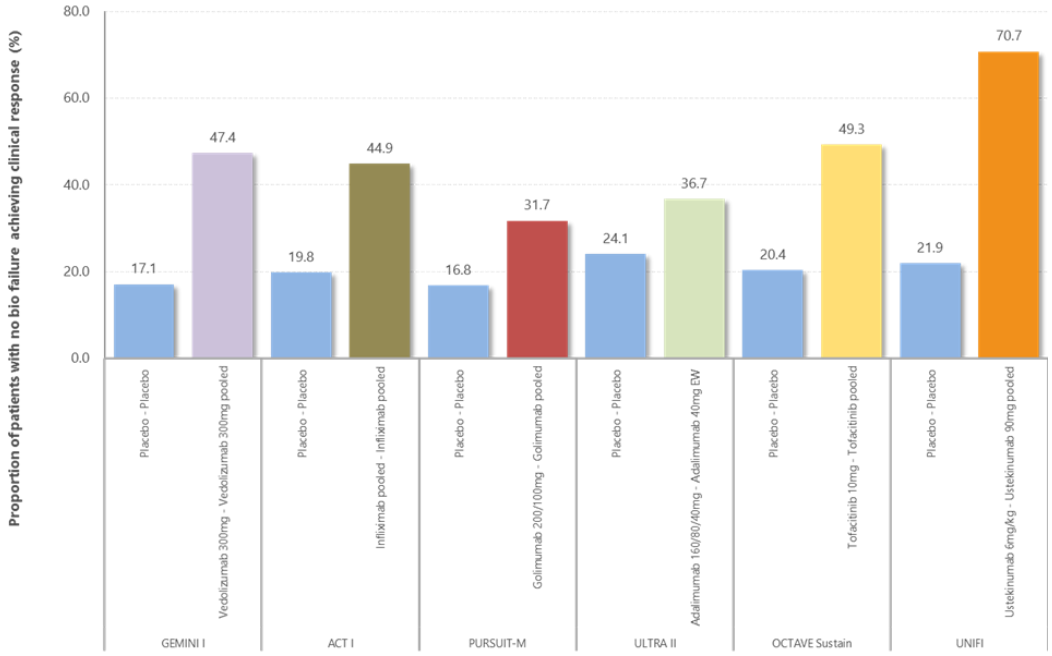


Figure 4.17. Data inputs for clinical response for NBF patients in the 1-year (pooled doses) base case mimicking a treat-through approach

Clinical remission

The network of evidence for clinical remission among NBF patients is shown in Figure 4.18. A total of seven studies were included in the analysis.

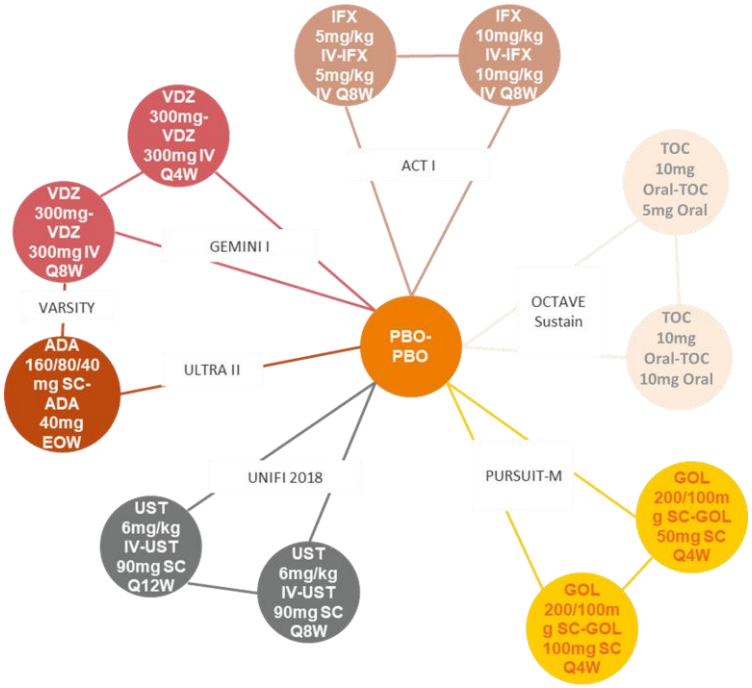


Figure 4.18. Network of evidence for clinical remission for NBF patients in the 1-year base case mimicking a treat-through approach

Abbreviations: ADA = adalimumab; GOL = golimumab; IFX = infliximab; IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

The data inputs from the seven studies included in the network of evidence for clinical remission among NBF patients are shown in Figure 4.19.

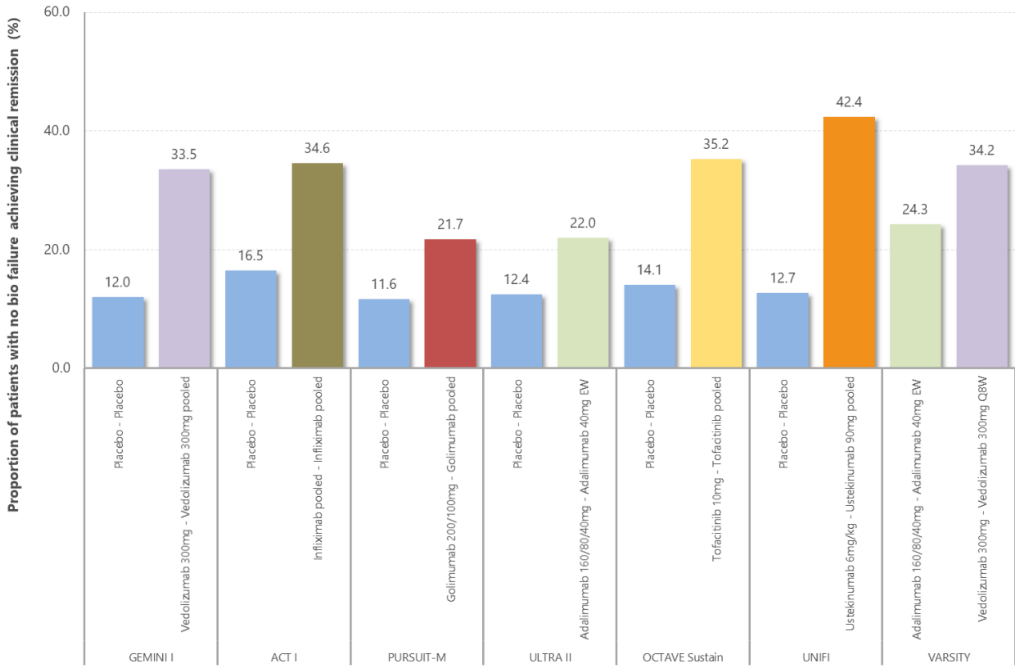


Figure 4.19. Data inputs for clinical remission for NBF patients for the 1-year (pooled doses) base case mimicking a treat-through approach

Mucosal healing

The network of evidence for mucosal healing among NBF patients is shown in Figure 4.20. Seven studies were included in the analysis.

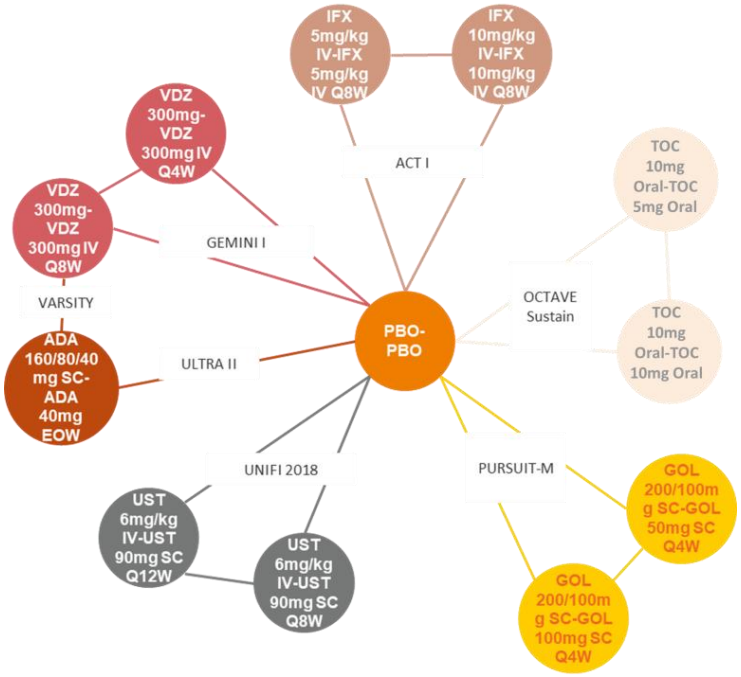


Figure 4.20. Network of evidence for mucosal healing for NBF patients in the 1-year base case mimicking a treat-through approach

Abbreviations: ADA = adalimumab; GOL = golimumab; IFX = infliximab; IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

The data inputs from the seven studies included in the network of evidence for mucosal healing among NBF patients are shown in Figure 4.21.

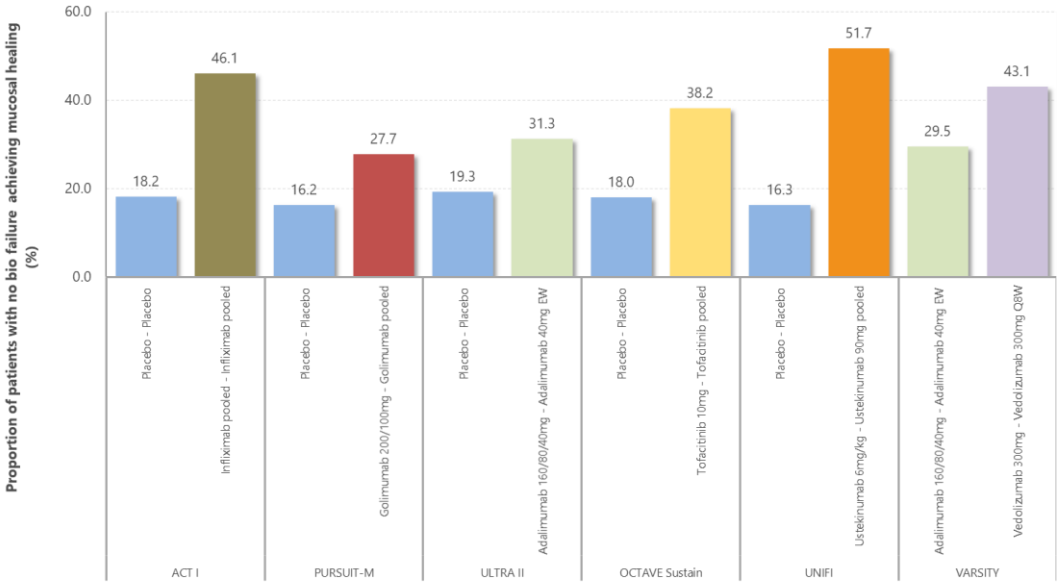


Figure 4.21. Data inputs for mucosal healing for NBF patients in the 1-year (pooled doses) base case mimicking a treat-through approach

BF patients

Results are not available for mucosal healing in the BF population, as no data from GEMINI 1 were available to inform the imputation of induction nonresponder data (required for UNIFI, PURSUIT and OCTAVE).

On the basis of results from the individual trials, a potential dose–response relationship was observed for treatments in 1-year NMA networks for BF patients. Specifically, a dose–response relationship was observed for clinical response to tofacitinib and for clinical remission to tofacitinib and ustekinumab. Therefore, NMAs were built for unpooled doses only.

The 1-year NMA of efficacy for the BF population showed significantly higher odds of achieving a clinical response and clinical remission with both ustekinumab regimes (UST 6 mg/kg → UST 90 mg q8w for induction responders and nonresponders; and UST 6 mg/kg → UST 90 mg q12w for induction responders + UST 90 mg q8w for induction nonresponders(in comparison to placebo (Table 4.23 and Table 4.24). While the point estimates for these comparisons were high, with ORs of 5.3–6.9 for clinical remission and 4.8 for clinical response, the reported CrIs ranged from ~2 to ≥9. Nevertheless, the lower CrI bound was ~2, indicating that ustekinumab may increase the odds of clinical response and remission twofold when compared to placebo. All other comparisons for efficacy in terms of clinical response or clinical remission between ustekinumab and any active treatments were found nonsignificant with wide CrIs, demonstrating that little is known about the effect and that further information is needed. For that reason, similar to what has been discussed for the maintenance NMAs for NBF patients, the fact that all the point estimates for all the comparisons to active treatments, although not significant, point towards higher odds of achieving a clinical response or clinical remission for patients treated with ustekinumab cannot be used to support the absolute ustekinumab efficacy.

For SUCRA rankings, the ustekinumab treatment(s) seemed to be ranked at the top for both clinical response and clinical remission (Table 4.23). Specifically, both ustekinumab doses were ranked first and second for clinical response (SUCRA 79–80%) although their SUCRA values were similar to those for the treatment sequence TOC 10 mg → TOC 10 mg early + TOC 10 mg delayed. Regarding clinical remission, only the UST 6 mg → UST 90 mg SC q8w early + UST 90 mg SC q8w delayed regime was ranked high (rank first, SUCRA 72%). While these rankings seem to show consistency for ustekinumab ranking first, they can also be misleading and the ranking results cannot be taken at face

value. As previously stated, SUCRA does not consider the magnitude of differences in effects between treatments, which for these networks are hindered by very wide CrIs including OR of 1.

Overall, given that the NMA for mucosal healing in the BF group was not feasible, and taking into account various methodological issues for maintenance models for BF patients (data more scarce compared with data in the NMAs considering NBF patients, indirectness and the fact that inconsistency between different sources of evidence was not estimated (including the NMA for clinical remission where inconsistency could be estimated due to a closed loop but the data of GEMINI I trial within the loop were imputed and the loop was apparently used to cross-check the imputation approach by comparing the re-calculated efficacy from GEMINI-1 with the results from VARSITY and ULTRA-II)), the fact that data were recalculated and imputed in majority of trials and that fixed effect model was used in a setting where heterogeneity is expected but could not be tested, we are uncertain about these results.

Table 4.22. NMA results for clinical response (unpooled doses) for BF patients in the 1-year base case mimicking a treat-through approach

Treatment sequence	Median OR (CrI)	
	UST 6mg/kg → UST 90mg q8w for IRs and INRs vs.	**UST 6 mg/kg → UST 90 mg q12w for IRs + UST 90 mg q8w for INRs vs.**
Placebo → placebo	4.83 (2.56–9.25); Pr = 100%	4.82 (2.28–10.30); Pr = 100%
VDZ 300 mg → VDZ 300 mg q8w for IRs + VDZ 300 mg q4w for INRs	1.76 (0.51–6.00); Pr = 81.45%	1.75 (0.48–6.35); Pr = 80.04%
VDZ 300 mg → VDZ 300 mg q4w for IRs and INRs	1.89 (0.53–6.69); Pr = 83.94%	1.88 (0.50–7.06); Pr = 82.54%
ADA 160/80/40 mg → ADA 40 mg EOW	2.03 (0.70–5.72); Pr = 90.52%	2.02 (0.65–6.14); Pr = 88.85%
TOC 10 mg → TOC 5mg for IRs + TOC 10 mg for INRs	1.66 (0.69–3.94); Pr = 87.24%	1.65 (0.63–4.28); Pr = 84.72%
TOC 10 mg → TOC 10 mg for IRs and INRs	1.21 (0.51–2.83); Pr = 66.49%	1.20 (0.46–3.08); Pr = 64.70%
UST 6 mg/kg → UST 90 mg q12w for IRs + UST 90 mg q8w for INRs	1.00 (0.45–2.25); Pr = 50.38%	–
UST 6 mg/kg → UST 90 mg q8w for IRs and INRs	–	1.00 (0.44– 2.23); Pr = 49.62%

Abbreviations: CrI = credible interval; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR >1 indicates a higher likelihood of response with ustekinumab than with the comparator, suggesting that ustekinumab performs better); Pr = Bayesian probability of ustekinumab performing better than the comparator; INR = induction nonresponder; IR = induction responder; ADA = adalimumab; EOW = every other week; GOL = golimumab; IFX = infliximab; IV = intravenous; qXw = every X weeks; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Table 4.23. NMA results for clinical remission (unpooled doses) for BF patients in the 1-year base case mimicking a treat-through approach

Treatment sequence	Median OR (CrI)	
	UST 6 mg/kg → UST 90 mg q8w for IRs and INRs vs.	UST 6 mg/kg → UST 90 mg q12w for IRs + UST 90 mg q8w for INRs vs.
Placebo → placebo	6.89 (2.98–16.90);Pr = 100%	5.34 (1.97–14.62);Pr = 99.94%
VDZ 300 mg → VDZ 300 mg q8w for IRs + VDZ 300 mg q4w for INRs	1.26 (0.31–4.91);Pr = 62.87%	0.97 (0.22–4.11);Pr = 48.53%
VDZ 300 mg → VDZ 300 mg q4w for IRs and INRs	1.32 (0.26–6.63);Pr = 63.48%	1.02 (0.19–5.48);Pr = 51.07%
ADA 160/80/40 mg → ADA 40 mg EOW	1.71 (0.42–6.55);Pr = 77.63%	1.32 (0.29–5.48);Pr = 64.31%
TOC 10 mg → TOC 5 mg for IRs + TOC 10 mg for INRs	1.57 (0.44–5.36);Pr = 76.05%	1.21 (0.31–4.52);Pr = 60.94%
TOC 10 mg → TOC 10 mg for IRs and INRs	1.08 (0.31–3.61);Pr = 54.80%	0.83 (0.21–3.05);Pr = 39.18%
UST 6 mg/kg → UST 90 mg q12w for IRs + UST 90 mg q8w for INRs	1.29 (0.53–3.32);Pr = 70.88%	–
UST 6mg/kg → UST 90 mg q8w for IRs and INRs	–	0.77 (0.30–1.90);Pr = 29.12%

Abbreviations: CrI = credible interval; Pr = Bayesian probability of ustekinumab being better than the comparator; OR = odds ratio (OR >1 indicates a higher likelihood of response with ustekinumab than with the comparator, suggesting that ustekinumab performs better); Pr = Bayesian probability of ustekinumab performing better than the comparator; INR = induction nonresponder; IR = induction responder; ADA = adalimumab; EOW = every other week; GOL = golimumab; IFX = infliximab; IV = intravenous; qXw = every X weeks; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Table 4.24. Ranking of treatments from most to least favourable from the 1-year NMA for BF patients, with corresponding SUCRA values

Clinical response		Clinical remission		Mucosal healing
UST 6-UST 90 SC q8w early + UST 90 SC q8w delayed	80%	UST 6-UST 90 SC q8w early + UST 90 mg SC q8w delayed	72%	NA
UST 6-UST 90 SC q12w early + UST 90 SC q8w delayed	79%	TOC 10-TOC 10 early + TOC 10 delayed	70%	NA
TOC 10-TOC 10 early + TOC 10 delayed	71%	VDZ 300-VDZ 300 q8w early + VDZ 300 q4w delayed	60%	NA
VDZ 300-VDZ 300 q8w early + VDZ 300 q4w delayed	46%	VDZ 300-VDZ 300 q4w early + VDZ 300 q4w delayed	56%	NA
TOC 10-TOC 5 early + TOC 10 delayed	45%	UST 6-UST 90 SC q12w early + UST 90 SC q8w delayed	56%	NA
VDZ 300-VDZ 300 q4w early + VDZ 300 q4w delayed	42%	TOC 10-TOC 5 early + TOC 10 delayed	44%	NA
ADA 160/80/40-ADA 40 EOW	37%	ADA 160/80/40-ADA 40 EOW	41%	NA
PBO-PBO	1%	PBO-PBO	0%	NA

Abbreviations: ADA = adalimumab; EOW = every other week; GOL = golimumab; IFX = infliximab; IV = intravenous; PBO = placebo; qXw = every X weeks; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Networks of evidence and data inputs

Details including the total number of patients and number and proportion of responders per treatment arm at the end of maintenance and at the end of 1 year are reported in Appendix 7 Data inputs for 1-year NMA in the MAH NMA report [46].

Clinical response

The network of evidence for clinical response among BF patients is shown in Figure 4.22. A total of four studies were included in the analysis.

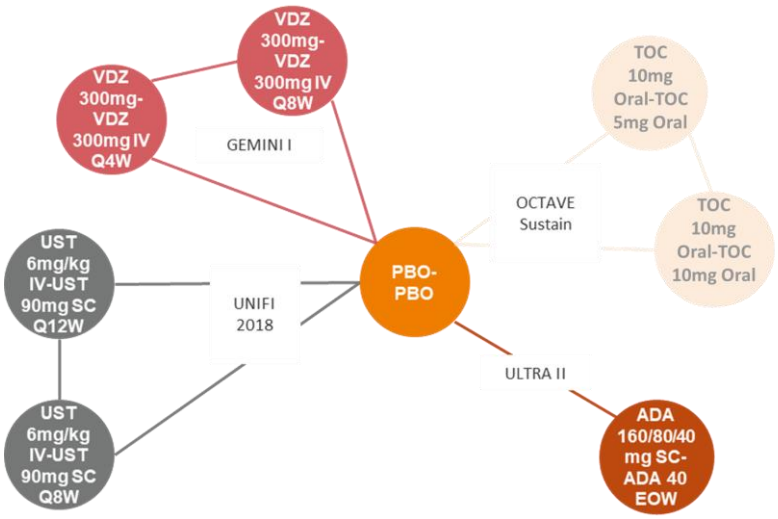


Figure 4.22. Network of evidence for clinical response for BF patients in the 1-year base case mimicking a treat-through approach

Abbreviations: ADA = adalimumab; IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

The data inputs from the four studies included in the network of evidence for clinical response among BF patients are shown in Figure 4.23.

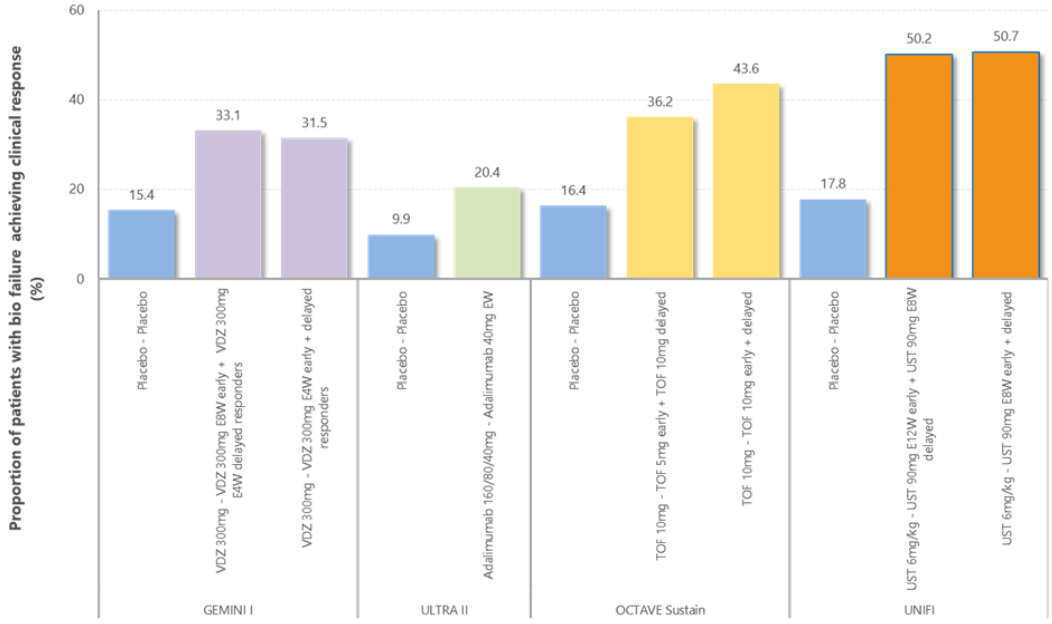


Figure 4.23. Data inputs for clinical response for BF patients in the 1-year base case mimicking a treat-through approach

Clinical remission

The network of evidence for clinical remission among BF patients is shown in Figure 4.24. A total of five studies were included in the analysis.

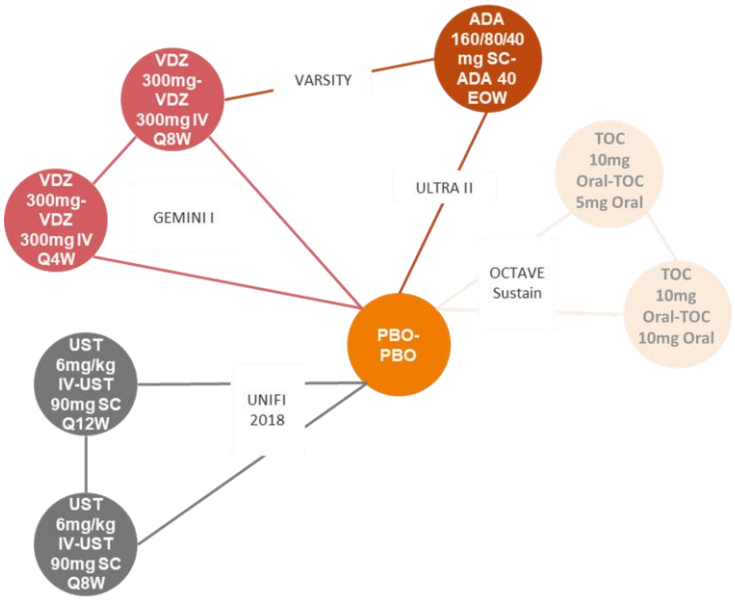


Figure 4.24. Network of evidence for clinical remission for BF patients in the 1-year base case mimicking a treat-through approach

Abbreviations: ADA = adalimumab; IV = intravenous; SC = subcutaneous; TOC = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

The data inputs from the five studies included in the network of evidence for clinical remission among BF patients is shown in Figure 4.25.

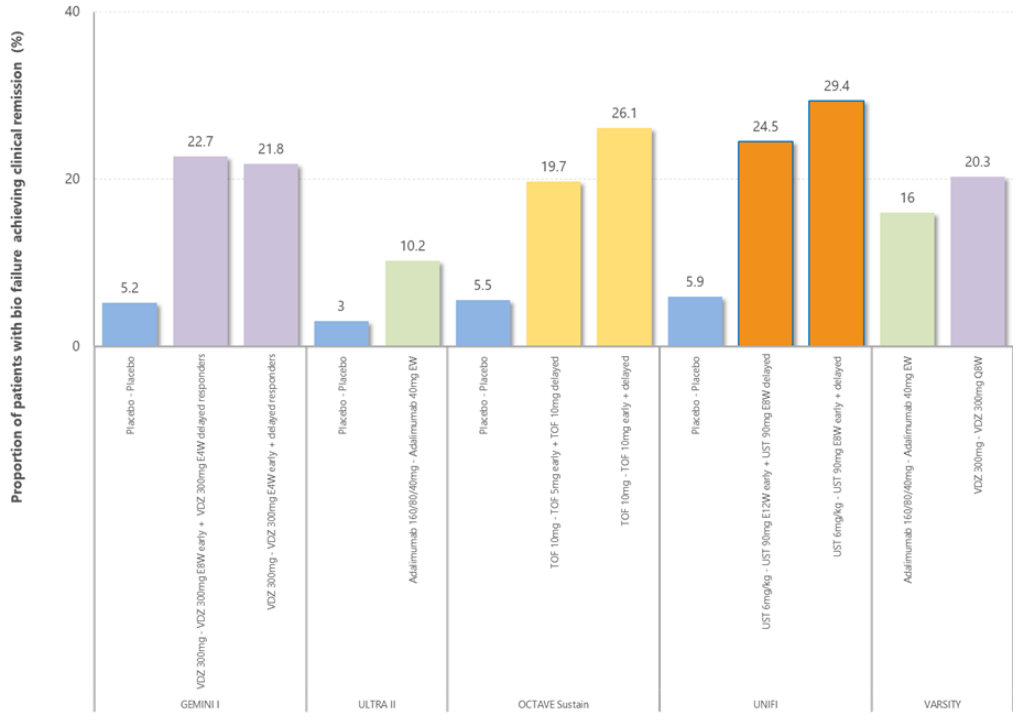


Figure 4.25. Data inputs for clinical remission for BF patients in the 1-year base case mimicking a treat-through approach

Mucosal healing

Results are not available for mucosal healing in the BF population, as no data from GEMINI 1 were available to inform the imputation of induction nonresponder data (required for UNIFI, PURSUIT and OCTAVE).

Patient QoL

Ustekinumab improved IBD-specific and general HRQoL outcomes evaluated using the IBDQ, SF-36 and EQ-5D at week 8, with statistically significant results for both induction doses.

Through to week 44, patients in the ustekinumab q8w and q12w groups generally reported maintained improvements in patient-reported HRQoL assessed using the IBDQ, SF-36 and EQ-5D instruments when compared with patients in the placebo group.

Details can be found in Appendix 2, A2.8. Results related to Quality of Life: for ustekinumab (**Error! Reference source not found.**–Table A20) and for infliximab, adalimumab, golimumab, vedolizumab and tofacitinib trials (**Error! Reference source not found.**–Table A37).

No NMA was conducted for the QoL outcome.

Safety

The safety data for ustekinumab in UNIFI trials include the number of AEs, SAEs, serious infections, malignancies, opportunistic infections, tuberculosis and death at week 8 (induction) and week 44 (maintenance) [60,61]. A summary of the results is listed in Table 4.25.

Safety endpoints analysed in the SLR and NMA and detailed safety results from the infliximab, adalimumab, golimumab, vedolizumab and tofacitinib trials could be found in Appendix 2, A2.9. Safety enpoints analysed in the SLR and NMA with results according the infliximab, adalimumab, golimumab, vedolizumab and tofacitinib trials (Table A38–Table A43).

Table 4.25. Summary of adverse events in the UNIFI trial programme for the safety analysis set

Safety endpoint	**UNIFI full patient population**					
	Induction (week 8)			**Maintenance (week 44)**		
	PBO (*N* = 319)	**UST 130 mg (*N* = 321)**	**UST 6 mg/kg (*N* = 320)**	**PBO SC (*N* = 175)**	**UST 90 mg q12w (*N* = 172)**	**UST 90 mg q8w (*N* = 175)**
AE, *n* (%)	153 (48.0)	133 (41.4)	160 (50.0)	138 (78.9)	119 (69.2)	136 (77.3)
Serious AEs, *n* (%)	22 (6.6)	12 (3.7)	10 (3.1)	17 (9.7)	13 (7.6)	15 (8.5)
Most frequent AEs, *n* (%)						
Worsening UC, *n* (%)	18 (5.6)	9 (2.8)	7 (2.2)	50 (28.6)	19 (11.0)	18 (10.2)
Nasopharyngitis, *n* (%)	NR	NR	NR	28 (16.0)	31 (18)	26 (14.8)
Headache, *n* (%)	14 (4.4)	22 (6.9)	13 (4.1)	7 (4.0)	11 (6.4)	18 (10.2)
Arthralgia, *n* (%)	2 (0.6)	3 (0.9)	6 (1.9)	15 (8.6)	15 (8.7)	8 (4.5)
Any infection, *n* (%)	48 (15.0)	51 (15.9)	49 (15.3)	81 (46.3)	58 (33.7)	86 (48.9)
Serious infection, *n* (%)	4 (1.3)	2 (0.6)	1 (0.3)	4 (2.3)	6 (3.5)	3 (1.7)
AEs of special interest, *n* (%)						
Malignancies (excluding NMSC), *n* (%)	0	0	0	0	1 (0.6)	1 (0.6)
Possible AR and possible DHS, *n* (%)	1 (0.3)	0	0	0	0	0
Cardiovascular events, *n* (%)	1 (0.3)	0	0	0	0	0

Death, *n* (%)	0	0	1 (0.3)	0	0	0
AEs leading to discontinuation, *n* (%)	NR	NR	NR	20 (11.4)	9 (5.2)	5 (2.8)
Abnormal laboratory results, *n* (%)	NA	NA	NA	1	0	0

Abbreviations: AE = adverse event; AR = anaphylactic reaction; DHS = delayed hypersensitivity; NA = not applicable; NMSC = nonmelanoma skin caner; NR = not reported; PBO = placebo; qXw = every X weeks; SC = subcutaneous; UC = ulcerative colitis; UST = ustekinumab.

Comparison of safety

The comparison of safety is derived from a number of sources in order to give a complete picture of the comparative safety of ustekinumab versus its comparators, including data across indications. There is extensive long-term experience with ustekinumab in clinical trials and long-term registries and clinical use for other indications.

First, results from an integrated safety assessment for the use of ustekinumab across all its indications were considered.

Data used to compare safety across products included:

- Safety data from large registries;
- A comparison of safety data from regulatory data assessed by the EMA; and
- An NMA for the induction phase of the trials (understanding the limitations of a safety comparison only for induction).

The efficacy and safety of ustekinumab for the UC indication were demonstrated in the UNIFI clinical trial programme. Ustekinumab has also been extensively studied in clinical trials for its other indications (CD, PsA and PsO for adults and adolescents) and there is considerable experience of real-world use of ustekinumab. An integrated safety analysis incorporating phase 2 and 3 trials across CD, PsO and PsA, which included approximately 6,000 patients treated with ustekinumab, revealed that ustekinumab demonstrated a consistent and favourable safety profile across registrational trials for all the approved indications. This analysis is consistent with the safety data from the UNIFI clinical trial programme.

Robust long-term RWE supports the safety of ustekinumab in the treatment of PsO according to the PSOLAR study. This showed the safety of ustekinumab compared with all three comparators (infliximab, adalimumab and etanercept) studied, and no increase in the risk of malignancy, major adverse cardiovascular events, serious infections or all-cause mortality with ustekinumab was identified.

To demonstrate the in-depth comparisons of safety profiles between the technology and the comparator treatments, including their clinical implications, information from the SmPCs was outlined in the MAH submission file [46]. This is supplemented by tables that detail the safety data reviewed by the EMA in their regulatory assessments.

NMAs for the induction phase only were also conducted for safety endpoints; however, these were considered to be limited for a number of reasons:

- The length of the induction phases in the studies is only 6–8 weeks, which is not sufficiently long to assess the safety profiles of treatments for continued use.
- The sample sizes are small and event counts are low for some of the endpoints, increasing the uncertainty for the results.
- The full trial populations (NBF and BF patients) were included in the NMAs as it was not possible to stratify all study outcomes by the subgroups, which introduces heterogeneity.
- The inclusion/exclusion criteria relevant to infections differed (as further discussed in the MAH submission file [46].

Results for the safety NMA conducted for the induction phase are provided in the MAH submission file [46].

An NMA of the 1-year safety outcomes was not considered appropriate for the following reasons (further details are included in the MAH submission file):

- Different definitions of the placebo safety population, comprising nonhomogeneous placebo arms with different efficacy and exposure levels, can result in spurious conclusions about safety for both SAEs and infections.
- Differences in inclusion criteria exist that may influence results for infections.
- There is insufficient information available for all the comparators for any attempt to correct for these factors

As described previously, other data sources based on an integrated analysis of safety data across indications for ustekinumab, large registry databases and a comparison of the regulatory data for each treatment provides more robust and relevant evidence to compare the safety profile of ustekinumab to other treatments in UC.

The safety profile of ustekinumab demonstrated in the UNIFI clinical programme is consistent with that of other clinical studies in CD, PsO and PsA.

The IV ustekinumab doses of 130 mg and 6 mg/kg were generally well tolerated. The proportions of patients with AEs and SAEs were comparable across treatment groups, with no evidence of an ustekinumab dose effect. Similarly, the proportions of patients who discontinued due to AEs were comparable (ustekinumab no higher than placebo) across treatment groups, with no evidence of an ustekinumab dose effect.

SC ustekinumab at a dose of 90 mg q12w or q8w was generally well tolerated. As observed in induction studies, the proportions of patients with AEs and SAEs were comparable across treatment groups, with no evidence of an ustekinumab dose effect. Similarly, the proportions of patients who discontinued due to AEs were comparable (ustekinumab no higher than placebo) across treatment groups, with no evidence of an ustekinumab dose effect.

In an integrated safety analysis incorporating phase 2 and 3 trials across CD (two phase 2 and three phase 3), PsO (one phase 2 and three phase 3) and PsA (one phase 2 and two phase 3), Ghosh et al. [88] compared the safety of ustekinumab across indications. The analysis includes 3,636 patients treated with ustekinumab (1,582 PsO, 692 PsA and 1,362 CD). The authors reported that AE, SAE and infection/serious infections rates were similar between the ustekinumab and placebo groups across the pooled indications. Overall, ustekinumab demonstrated a favourable safety profile in CD consistent with psoriatic diseases.

PSOLAR [79] is an ongoing, disease-based observational study in which patients eligible for or who are receiving either nonbiologic systemic or biologic agents for the treatment of PsO are followed. This registry is designed to capture AEs of special interest including serious infection data across all therapies used in the treatment of PsO.

An overview of AEs of special interest (AEoSI) was reported by Papp et al. [79] in which cumulative AEoSI rates were reported for ustekinumab, infliximab, other biologics (mostly adalimumab and etanercept) and nonbiologic therapy. The prespecified analyses used attribution rules biased against ustekinumab, with safety events attributed to ustekinumab if patients switched to a different therapy and subsequently experienced an AE. The study included a total of 12,093 patients accounting for 40,388 patient-years. The authors reported that exposure to the combined group of biologics other than ustekinumab (mostly infliximab, etanercept and adalimumab) on registry was associated with a higher risk of serious infection. In addition, the analyses did not identify any higher risk of malignancy, major adverse cardiac event, serious infection or mortality with ustekinumab.

In a separate study focused on the risk of serious infections, Kalb et al. [89] analysed data for 11,466 patients representing 22,311 patient years. The cumulative incidence rate of serious infections was 1.45 per 100-patient years across treatment cohorts, with rates of 0.83, 1.47, 1.97 and 2.49 per 100 patient-years in the ustekinumab, etanercept, adalimumab and infliximab cohorts, respectively. The authors concluded that results from PSOLAR suggest a higher risk of serious infections with adalimumab and infliximab compared with non-methotrexate and nonbiologic therapies, with no higher risk observed with ustekinumab (Figure 4.26).

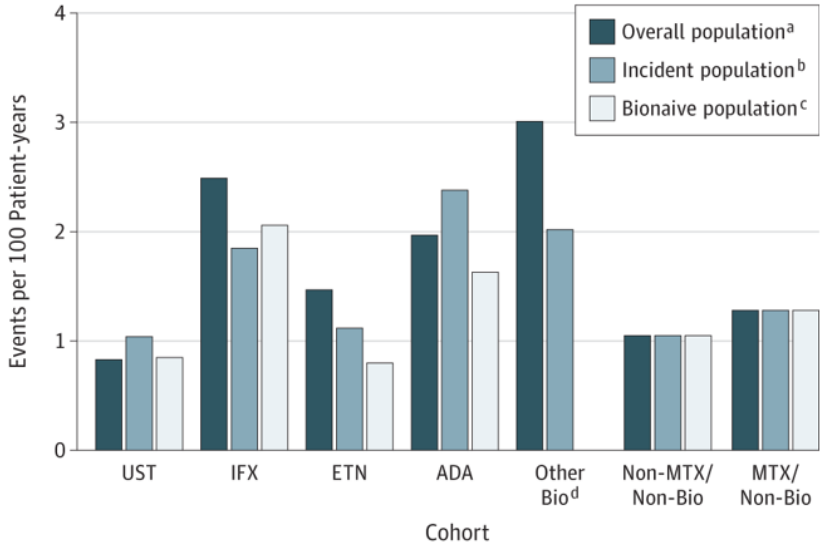


Figure 4.26. Cumulative incidence rates of serious infections of interest per 100 patient-years

Notes: [a] The overall population includes patients who received a biologic agent before (prevalent population) or after (incident population) enrolling in the registry. [b] The incident population, a subset of the overall population, includes patients who received a biologic agent at or after enrolling in the registry but may have been exposed to a different biologic previously. [c] The bionaïve population is a further subset of the incident population and includes patients who received their first biologic agent at or after enrolling in the registry and had never received prior biologic exposure. [d] The other biologics cohort includes patients who received alefacept, efalizumab, golimumab and other investigational biologic agents.

Another analysis of this study investigated patients with concomitant IBD (at the time of analysis there were 12,093 patients and 40,388 total patient-years of follow-up) (Table 4.26), it was demonstrated that patients treated with ustekinumab had numerically lower rates of serious infections compared with patients receiving other biologic and systemic therapies for PsO in the IBD subset as well as overall [90].

Table 4.26. Cumulative incidence rates of serious infections per 100 patient-years within 91 days of biologic administration

	UST (*N* = 4,363)	**IFX (*N* = 1,394)**	**Non-sponsor biologic (*N* = 4,251)**	**Nonbiologic (*N* = 2,083)**	**Total (*N* = 12,093)**
Total IBD population	72 (1.7%)	71 (5.1%)	81 (1.9%)	52 (2.5%)	276 (2.3%)
CD	15 (20.8%)	21 (29.6%)	18 (22.2%)	6 (11.5%)	60 (21.7%)
UC	22 (30.6%)	30 (42.3%)	28 (34.6%)	13 (25.0%)	93 (33.7%)
IC	27 (37.5%)	19 (26.8%)	28 (34.6%)	32 (61.5%)	106 (38.4%)
Pt Yrs	218	226	301	173	918
Rate of SIs/100 PYs	1.38	5.75	4.32	3.47	3.81
Full PSOLAR population					
PYs	7,944	3,301	12,823	16,322	40,389
Rate of SIs/100 PYs	0.93	2.91	1.91	1.43	1.60

Abbreviations: CD = Crohn's disease; IBD = irritable bowel disease; IC = indeterminate colitis; IFX = infliximab; PYs = patient-years; SI = serious infection; UC = ulcerative colitis; UST = ustekinumab.

The EMA safety assessments are summarised in the SmPCs [29]. The relative amount and content of safety information in the SmPCs for the technology and comparator treatments are an important aspect of the safety evidence for the technology and its comparator treatments. This is summarised in the MAH submission file [46] owing to size limitations, and these tables provide an overview of AEs associated with ustekinumab and the comparators as assessed by the EMA.

Comparison of the SmPCs obtained after assessment of detailed data from regulators suggests that Stelara (ustekinumab) has an acceptable safety profile.

An NMA of 1-year safety outcomes was not considered to be feasible because of differences in the definitions of placebo safety populations and in inclusion criteria with regard to safety outcomes and the lack of information for all comparators to attempt to correct for these factors.

An integrated safety analysis, incorporating phase 2 and 3 trials across CD, PsO and PsA, which included approximately 6,000 patients treated with ustekinumab, revealed that ustekinumab demonstrated a favourable and consistent safety profile across registrational trials for the approved indications.

There is robust long-term RWE supporting the safety of ustekinumab in the treatment of PsO from the PSOLAR study. This assessed the safety of ustekinumab versus three comparators (infliximab, adalimumab and etanercept) and showed no increase in the risk of malignancy, major adverse cardiovascular events, serious infections and all-cause mortality with ustekinumab [79,89,91] Another analysis of this study examined patients with concomitant IBD (at the time of analysis there were 12,093 patients and 40,388 total patient-years of follow-up). In line with the overall PSOLAR analysis, it was demonstrated that patients treated with ustekinumab had lower rates of serious infections (*n* = 22; 30.6%) than patients receiving other biologic and systemic therapies (infliximab: *n* = 30, 42.3%; non-sponsor biologics: *n* = 28, 34.6%) for PsO in the IBD subset as well as overall.

To demonstrate the in-depth comparisons of safety profiles between the technology and the comparator treatments, information from the SmPCs was outlined in the MAH submission file [46]. While a qualitative cross-comparison of different safety profiles from the SmPCs is inherently difficult and subject to weighting of different safety attributes, the comparison demonstrates that ustekinumab has an acceptable safety profile.

NMAs were conducted for safety outcomes in the induction phases of clinical trials. These demonstrated that overall, ustekinumab 6mg/kg was associated with similar likelihoods of overall AEs, SAEs and infections when compared to other active comparators. Owing to the low event counts, the results from the NMA for serious infections were associated with high uncertainty. The relevance of the results from these NMAs is limited because of the short time period for assessment and inherent uncertainty given low event counts for some endpoints and small sample sizes.

Results from the integrated analysis, registry databases and SmPCs provide more robust and relevant evidence to describe the safety profile of ustekinumab in comparison to other treatments in UC. This collective information demonstrated that ustekinumab is an innovative, safe and effective treatment for moderately to severely active UC, providing rapid improvement in disease activity and symptoms and long-term maintenance of response and remission.

4.8 Patient involvement

Predefined patient questions related to the impact of condition; experience with currently available treatments; experiences with and expectations for the health technology being assessed; and additional information that patients believe would be helpful to the HTA researchers were used to help in the assessment of the value of the health technology.

A summary of patient involvement is given below.

Impact of condition: moderately to severely active UC

How does moderately to severely active UC affect patient QoL?	Moderately to severely active UC affects a patient's QoL in many ways. The patient may not be able to work or have a social life outside of his/her home and may not be able to participate in family life. Any physical activities may be very difficult. Leaving home for any reason (e.g., work, bringing children to kindergarten/school, grocery shopping, social events) may be impossible because of symptoms. Patients would most like to see a treatment that allows them to have a "normal" life without UC symptoms limiting it. Inability to work, social isolation, depression, seeking information, looking for help.
How does moderately to severely active UC affect carers/unpaid caregivers?	One family member with active UC may put a lot of physical and emotional pressure on other family members or friends. They may need to support the patient by running errands, which can cause stress and fatigue, but also have to deal with the feelings of not being able to help the person with UC.

Experiences with currently available health interventions

How well are adult patients managing moderately to severely active UC with currently available therapies?	While there are many medications and therapies to try, experiencing loss of response or medical contraindications to therapies, constantly having to start and try new therapies, etc. can be very tiring for the patient both physically and emotionally.

Experiences with and expectations for ustekinumab

For those WITH experience in using the new medicine, what difference did it make to their lives?	Visiting hospital to get medicine, better patient adherence (under physician control), better QoL, travel costs, time off work.
For those WITHOUT experience in using the new medicine but who are aware of the results of clinical trials about the new medicine, what are the expectations/limit?	Patients need information; educational seminars about new drugs, side effects, benefits, comparing studies.

Additional information

No additional information (e.g., ethical or social issues) that patients believe would be helpful to the HTA researchers was identified.

Key messages from patients

Moderately to severely active UC affects a patient's QoL in many ways. The greatest challenges in living with moderately to severely active UC are to remain in a good mood and to believe in better days. Current therapies are inadequate because many patients still require surgery. New therapies are needed to ensure that there are several options to choose from in case of intolerance or loss of response so that every patient with UC can work and have a social and family life despite their illness. This new medicine will be important for patients because it will provide new hope for better QoL.

5 DISCUSSION

The included endpoints in this report are the ones that are recommended by guidelines.

The evidence base for ustekinumab for the treatment of UC is supported mainly by the UNIFI clinical trial programme and the NMAs.

The UNIFI trials included patients reflective of a contemporary population of patients with moderately to severely active UC with previous failure or intolerance of conventional and/or biologic therapies, including TNFi agents and/or vedolizumab. Specifically, UNIFI is the first trial to include vedolizumab failures in the study population. This reflects the current relevant patient population (NBF, TNFi failure and vedolizumab failure) in the real world. Compared to placebo, ustekinumab demonstrates a favourable efficacy compared to placebo for remission, response and mucosal healing for the whole patient population, independent of conventional therapy failure or BF. Mucosal healing could not be assessed at 1 year for BF patients and the model of clinical remission for BF at induction.

The list of available endpoints for the NMA could be drawn from the table reporting on characteristics of studies included in the submission file [46], but judging from the data presented for the UNIFI trial, not all the endpoints are listed. Specifically, the major secondary endpoint of corticosteroid-free clinical remission is not reported for the UNIFI trial and was not considered for NMAs, which increases the risk of bias in selection of reported results for the indirect comparison analysis (as discordant pieces of evidence possibly drawn from a relevant secondary endpoint may not be able to inform decision making if available but excluded from analysis).

The comparators, study design and outcomes for this trial were aligned with previous comparator trials in this disease area. For the endpoints of clinical remission and clinical response, the endpoint definitions were consistent across all of the trials, with only a minor difference for OCTAVE for clinical remission (one patient difference between definitions).

However, a potential source of heterogeneity due to differences in outcome measurements was observed for the mucosal healing endpoint, as the endoscopic score for the efficacy analyses used in the NMA was assessed by a local endoscopist in six out of eight studies (VARSITY did not report who performed the readings, while the OCTAVE readings were performed centrally), and no information on whether standardisation of local reading was performed in these studies was presented by the MAH. The EMA guideline on UC [80] recommends that adjudication of endoscopic evidence of activity should be preferably be done via central reading of the examinations, and if local reading of examinations is performed, standardisation of reading should be demonstrated. The NMA also included all relevant comparators used for the treatment of UC, including infliximab 5 and 10mg/kg, which is not the recommended regimen for UC, but the agent is recommended for Crohn's disease and a dose escalation with infliximab to 10mg/kg is an apparently common practice in Europe.

The main efficacy analyses in the submission file are based on results from Bayesian NMAs. The evidence drawn from all the networks was based on indirect data in a setting in which heterogeneity or inconsistency of a network could not be statistically assessed owing to the paucity of the data, and assessment of statistical heterogeneity between trials informing each comparison was limited and probably underpowered. Even in the networks that included a head-to-head comparison, because of data imputation in one of the trials informing a closed loop, inconsistency assessment was not performed. As a network is most justifiable under an assumption of consistency between different sources of evidence, the lack of these data largely increases uncertainty regarding the NMA results. In addition, given the reasons stressed below, it is not certain if the key assumption for NMA—that the different sets of studies included in the analysis should be on average similar in all important factors that may affect the relative effects—is met.

Overall, trials comprising the body of evidence for NMAs were judged as comparable at baseline in terms of patient characteristics (an outlier trial [2] that was not part of NMA was not considered for this conclusion). However, owing to the paucity of the data, statistical assessment of heterogeneity in the distribution of potential effect modifiers across comparisons was not possible. Instead, heterogeneity between individual trials for several baseline demographic and clinical characteristics was assessed by comparing descriptive statistics and reporting data for mixed populations (including BF and NBF patients) and different outcome measures (mean and median). In addition, the data were not stated for all trials and for both subgroups, which increases uncertainty regarding the conclusion. The data for BF patients were particularly under-represented in these analyses, so it was not possible to draw valid conclusions on comparability in baseline characteristics across trials for this subgroup. The

same was true for the maintenance trials, as only two studies (OCTAVE-S and UNIFI) reported baseline characteristics at the onset of the maintenance phase. Therefore, evidence of homogeneity among the trials at baseline is limited and is applicable only to the NBF population in induction trials. Specific findings for this analysis are as follows.

On the basis of data available for mean disease duration, 15 induction trials were similar across an NBF subgroup (range 4–8 years). However only one study reported a corresponding value for the BF subgroup (GEMINI 1, 6–8 years), while the rest of the studies reported overall mean disease duration for a mixed population (NBF and BF patients) of between 6 and 11 years. A study that potentially stood out was NCT00787202, with a mean disease duration for a mixed population of 11 years in one of the intervention groups. Two studies reported baseline characteristics at the onset of maintenance, and the median disease duration was comparable (6–7 years).

The range reported for mean CRP in pivotal phase 3 trials in UC was 11.5–15.7 mg/l [8]. In the MAH submission the mean CRP level in the mixed population at baseline was somewhat lower (9–19 mg/l); placebo arms in the PURSUIT-SC, UNIFI and NCT00787202 trials and the UST 130-mg intervention arm in the UNIFI trial had mean CRP levels below 10 mg/l. Nevertheless, when median values were compared, CRP levels were stable throughout mixed patients population, with median CRP in the range 2.2–4.9 mg/ml. By contrast, median CRP levels were hand more variable for the NBF subgroup: the ACT 1 (8–10 mg/ml) and ACT 2 (6–8 mg/ml) trials reported vales that were twice as high as median CRP reported in other trials for the same subgroup (2.2–5 mg/ml). Such difference suggests possible differences between the studies.

The mean Mayo score reported by Ghosh et al. [8] for pivotal phase 3 trials ranged between 8.4 and 8.9, with a similar range in the submission file [46] for mixed patients (8.0–9.1) and the NBF subgroup (8.2–9.0). For the BF subgroup in GEMINI there was inconsistency between values plotted in Figure 44 in the submission file [46] (mean 8.6–8.7) and values reported in the text (8.6–9.7). However, for the UNIFI trial, only the medians were reported for mixed patients, indicating that the distribution of scores is asymmetric and that mean values would have been higher.

Time points differed across trials for induction and maintenance phases. As partial Mayo scores were consistent within these time ranges in several trials, it was concluded that these time points are equivalent in terms of clinical efficacy. This was shown for the maintenance phase in UNIFI, PURSUIT and GEMINI trials, and for CD patients in the UNIFI (ustekinumab) and CHARM (adalimumab) trials, as well as for the UNIFI induction trial. However, deviation from this pattern was observed in an induction phase trial for CD patients treated with adalimumab, as the largest effect on clinical remission for patients treated with adalimumab was observed at 6 weeks for the dose of 40 mg EOW, and at 8 weeks for the dose of 40 mg weekly. As adalimumab was also one of few drugs that demonstrated significantly worse efficacy than ustekinumab in induction-phase NMAs for both subgroups, there is a possibility that the dose and/or the length of the induction period (6–8 weeks) may be suboptimal for adalimumab for assessment of its true clinical remission in the induction phase.

The MAH identified prior TNFi exposure as the source of heterogeneity and performed analyses separately for NBF and BF patients to minimise the. However, increasingly selective criteria for RCTs of new biologics has been changing the definition of target patient populations, which may affect the adequacy of a common comparator in UC trials. Patients have been enrolled in these trials as biologic-naïve or with failure to a particular TNFi, any TNFi, or any TNFi and/or vedolizumab. In the MAH submission file, while the inclusion criteria for BF for the 2019 UNIFI trial correspond to any TNFi and/or vedolizumab failure [60], in the 2013 GEMINI trial, patients assigned to this category were enrolled because of intolerance or failure to the specific TNFi infliximab (Supplementary Appendix in [57]), and in the 2017 OCTAVE trial because of intolerance or failure to several TNFi agents [51]. In other words, more recent trials enrolled more heterogeneous patients (those with failure or intolerance to a particular TNFi drug, any TNFi, or any TNF and/or vedolizumab) with increasing probability of patients failing more than one biological treatment, which implies that BF patient populations might not be comparable. Moreover, considering the observation that TNFi agents have lower clinical efficacy and mucosal healing rates with increasing disease duration [5,6,7], several theories have been postulated that corroborate this view. 1) It is likely that continuous blockade of key cytokines such as TNF can induce pathomorphisms that are characterised by compensatory activation of other pathways (as evidenced by the occurrence of autoimmune phenomena including PsO and alopecia areata under effective TNFi therapy) and that render patients particularly responsive to other agents affecting such pathways (e.g., blockade of IL-23). 2) Patients with early disease may have a different immunologic composition compared to patients with late disease. This again implies heterogeneity within the BF population.

Moreover, the issue of heterogeneity of patient populations among the UC trials and the representativness of samples has recently become even more critical, as many trials investigating new biologics in UC fail to study representative patient populations. Specifically, in a market in which several drugs are available for prescription, assignment to the placebo arm of a long-term study is not very appealing to patients. Consequently, the proportion of patients available in centres who are recruited to newer clinical trials is so low that the results can no longer be regarded as representative. Ghosh et al. [8] illustrated this difference in populations of patients with CD, a closely related disease, by comparing the time needed to enrol and/or complete a given IBD trial, which has changed dramatically: “In 1999, the 54-week ACCENT 1 trial enrolled 573 anti–tumor necrosis factor (TNF)-therapy naïve patients with CD from 55 centres in 12 countries and was completed in approximately 2 years. In contrast, by 2008 the 52-week GEMINI 2 trial required 1115 patients with CD (62% prior TNFi exposed) from 285 centres in 39 countries and took almost 3.5 years to complete”. A similar difference in patient populations probably applies to UC. Nevertheless, the data from earlier clinical trials on biologics are not likely to be comparable with more recently conducted RCTs as they include different populations of patients.

The induction-phase study designs were all consistent across the trials, so a standard NMA approach could be taken. In the maintenance phase, however, study designs were identified as a substantial source of heterogeneity. To make the maintenance endpoint data between the designs comparable, treatment effects in re-randomised trials were recalculated using a treat-through approach. This approach allows randomisation to be maintained, as treatment arms are based on induction therapy received by the randomised patients. However, recalculation of the total number of patients in the re-randomised responder trials reduced the sample sizes. Apart from recalculation, some data inputs needed for recalculation were missing, so the data in four of seven trials comprising the body of evidence were imputed. The imputation was performed in situation in which data were scarce (only 1 trial per comparison), indirect (there was no head-to-head comparison; only for the clinical remission maintenance networks for both NBF and BF patients was there a closed loop containing a head-to-head comparison that would allow comparison of direct and indirect evidence, but the statistical agreement between direct and indirect treatment effects was not tested as the efficacy data for one of the trials were partly imputed, and thus could not be considered as direct evidence) and heterogeneity could not be assessed statistically but it was reasonable to assume its existence. The MAH performed sensitivity analyses to control the introduction of a bias in assessments, including an ITT approach conditional on the induction response mimicking a response-based design and a multiple imputation approach, but these could not be considered as adequate controls. The imputation for induction responders for NBF patients was based on IPD data from UNIFI, ACT 1 and PURSUIT, for which the MAH provided simulations investigating the change in effects when variation in the estimates used for imputing missing placebo response rates is taken into account. The underlying assumption was that trial data are drawn from the same underlying distribution, with the central tendency value corresponding to a weighted average of the data actually used in imputation and, as expected, in the majority of runs values at or close to the imputed value were generated. As for sensitivity analysis mimicking a response-based approach, the analysis targets only induction responders, so it does not tackle the same underlying question as the base-case approach mimicking a treat-through design and thus cannot be considered as a sensitivity analysis. Finally, the MAH also stated that for the clinical remission network for NBF patients it compared the corresponding arms of the VARSITY head-to-head trial of vedolizumab versus adalimumab with the results from ULTRA 2 for adalimumab and imputed GEMINI 1 results for vedolizumab (data for induction nonresponders in both arms were imputed), which apparently corroborated the imputation, but the imputed data from GEMINI could not be considered as direct evidence.

Nevertheless, the data from earlier clinical trials on biologics are not likely to be comparable with more recently conducted RCTs as they include different populations of patients. Since a valid NMA relies on the assumption that the different sets of studies included in the analysis are on average similar in all important factors that may affect the relative effects, for these NMAs it is very uncertain if the results are valid or are due to bias/artefacts of analysis.

The NMA of 1-year regimens provides ambiguous results of very low evidence certainty for the treatment effects based on treat-through comparisons. Adding to concerns about meeting the key assumption of a valid meta-analysis, the NMA models for both subgroups were built on scarce, indirect and heavily imputed data in a setting in which heterogeneity is expected but could not be tested because of the paucity of the data and use of fixed-effects models as the only available model option. On the basis of the differences observed among registrational clinical trials of biological therapies in adults with IBDs and in meta-analysis of placebo remission rates in UC patients with active disease

showing significant heterogeneity (ranging from 0% to 40%) [8,9], it is reasonable to assume some level of heterogeneity between the trials included in the NMAs. While a random-effects model would explicitly model these differences and capture the uncertainty in the true treatment effect, fixed-effects models are likely to underestimate the uncertainty and possibly come to the wrong conclusion. In addition, having the data drive selection of the models via DIC is considered a questionable approach. The MAH should have stated in their protocol whether fixed-effects or random-effects models were to be used in their primary analyses, depending on their assumptions (not depending on the data). If it is really realistic to assume that there is a single true underlying relative effect of the interventions, a fixed approach can be chosen. As this assumption is typically not tenable, a random approach is preferable, for which a distribution of treatment effects is assumed rather than a single effect. It is acknowledged that a random approach was not feasible because of the data paucity, but the MAH should nevertheless clearly indicate which method was planned. Finally, there is an inconsistency in the submission: while the Methods and Results sections state that DIC values were used to select models, the section on Data limitations states that random-effects models are preferred over fixed-effects models.

The 1-year NMA estimates for both subgroups of patients had very wide CrIs, which hindered interpretation of the results. In addition, owing to high uncertainty and the inability to precisely estimate differences in effect sizes between the treatments, the probability of ustekinumab being better and the SUCRA rankings could not be used as reliable indicators of efficacy of treatments. In particular, it is known that the probability of a treatment being better underperforms when the treatment effect of an intervention is largely uncertain, as an intervention with the most uncertainty can have the greatest probability. The evidence on which the SUCRA rankings is based is insufficient to make an inference. In addition, as SUCRA values by definition are based on any differences among the treatment effects, no matter how small, in this setting it is possible that a treatment with a higher rank is better than another treatment by only a small amount that is not clinically meaningful.

For the NMAs for NBF population in the maintenance phase, additional limitation was identified. As apparently no dose–response relationship was observed, the data for interventions with different doses of the same drug were pooled for treatment arms. The claim that no dose–response relationship exists in these particular networks (whereas the opposite was observed in all the others, including BF patients at 1 year, suggesting an underlying biological phenomenon), which is supported only by observation that higher dose/shorter interval between doses led to higher clinical rates (no statistical testing was performed), does not represent convincing evidence. It is important to note that while pooling of doses might increase the analysis power, as stressed by the MAH, it will also result in overprecise results (as data from two active treatment arms of a trial are pooled together and the sample size of then active arm is increased twice) which could affect comparison with a study having just one active arm (ULTRA 2) and the fixed-effect NMA results [10].

For the NBF population in the maintenance phase, all the point estimates for all the comparisons, irrespective of their significance, point to higher odds of achieving an endpoint for patients treated with ustekinumab compared to those treated with other active treatments or placebo; however, all that can be concluded about the effect sizes from the width of the nonsignificant CrIs (ranging from 0.6 to 3.5) is that not much is known about the effect and that more data are needed. Even in the case of significant findings in this population, the interpretation is hindered. For the NBF population, patients who received ustekinumab 6 mg/kg in induction followed by ustekinumab 90 mg q8w or q12w had significantly higher odds of achieving a clinical response, remission and mucosal healing after a 1-year-long treatment regimen than those receiving placebo or treatment sequences with adalimumab or golimumab. The point estimates for these comparisons were high, with ORs ranging from 2.4 to 8.7, but for some of the comparisons the lower CrI bound was such that the possibility that the true effect is not clinically important cannot be excluded. For example, the NMA for clinical remission suggested an increase of only 10–13% for ustekinumab compared to adalimumab (OR 2.43, CrI 1.10–5.42) or golimumab (OR 2.40, CrI 1.13–5.22). In addition to treatment sequences showing significant results for all three endpoints, ustekinumab increased the odds of achieving a clinical response compared to infliximab or tofacitinib. Again, the interpretation of these findings was not unambiguous, as the CrI for the OR for ustekinumab versus tofacitinib allows for the possibility that the true effect size is below the minimal clinically important difference (OR 2.27, CrI 1.06–4.86), while for ustekinumab versus tofacitinib the CrI range was 1.22–5.6.

For BF patients, in addition to the general limitations of maintenance NMAs already stated, 1-year NMA models were based on even more scarce data with smaller sample sizes and smaller numbers of events in the study arms compared to the NBF NMAs. At the same time, in line with what has been discussed regarding sources of heterogeneity, it was more likely that this population was heterogeneous in trials. The eligibility criteria for BF included failure to TNFi agents or vedolizumab for UNIFI, but for the other trials included TNFi drugs only. However, the trials for this body of evidence were not assessed for initial imbalances in BF patients since they mainly presented baseline data for a mixed population. In addition, placebo imputations for the 1-year NMAs were based on less robust data compared to the data used for the NBF group (only the UNIFI trial was used to impute induction responders to placebo in this population), while results for mucosal healing in the BF population were not available as no data from GEMINI 1 were available to inform the imputation of induction non-responder data (required for UNIFI, PURSUIT and OCTAVE). For the other two endpoints the 1-year NMA of efficacy showed significantly higher odds that both ustekinumab regimes (UST 6 mg/kg → UST 90 mg q8w for induction responders and nonresponders; and UST 6 mg/kg → UST 90 mg q12w for induction responders + UST 90 mg q8w for induction nonresponders) would achieve a clinical response and clinical remission in comparison to placebo. The point estimates were high, with ORs ranging from 5.3 to 6.9 for clinical remission and 4.8 for clinical response, but the certainty of the estimated effect sizes was low as the CrIs ranged from ~2 to ≥9. Nevertheless, the lower CrI bound was ~2, indicating that at a low level of evidence ustekinumab increases the clinical response and remission at least twofold over placebo. No active treatments demonstrated a significant difference to ustekinumab. Although, similar to NBF patients, all point estimates for BF patients pointed towards higher odds of achieving the endpoint with ustekinumab, the main conclusion is that the evidence base is insufficient to make inferences.

Some data limitations for the 1-year NMA biased against ustekinumab; OCTAVE trial re-randomises both placebo and tofacitinib induction responders and estimated clinical remission rate from their combined induction arms; delayed responders had to respond at week 14 in PRSUIT and 16 in UNIFI which potentially underestimates the number of delayed responders included in the golimumab and ustekinumab trial arms; OCTAVE included an open-label treatment for induction non-responders. However the fact that only UST data (UNIFI trial) with the highest placebo response rate among non-biologic failure patients at 1-year were used to impute the data for biologic failure population might have also driven the results the other way.

5.1 Strengths & limitations of the clinical evidence base for ustekinumab in UC

The UNIFI trial demonstrated the efficacy and safety of ustekinumab in moderately to severely active UC patients.

All primary and major secondary endpoints for ustekinumab at induction and maintenance were met and results for remission (Mayo score ≤2 with no subcomponent >1), response (decrease in Mayo score by ≥30% and ≥3 points), endoscopic healing, mucosal healing and corticosteroid reduction were statistically significant.

The UNIFI trial of ustekinumab is stratified by NBF and BF and is the only trial to date that includes patients who were previously treated with vedolizumab, representing a BF treatment group that truly reflects current practice in UC treatment.

Ustekinumab induced strong remission and response rates at induction that were up to three and two times higher, respectively, in the ustekinumab group compared to the placebo group; the remission and response were maintained through to week 44 in the ustekinumab q8w and q12w groups.

In addition to remission as a primary endpoint (defined as a Mayo score ≤2 with no subcomponent >1), ustekinumab is the first biologic treatment in UC that demonstrates evidence of mucosal healing, defined as a composite endpoint of endoscopic and histologic healing in registrational trials, which is considered an important marker of treatment efficacy.

As with many other clinical trials in UC, the UNIFI studies lack a direct comparison versus active comparators (i.e., other nonconventional therapies). Moreover, the trial contained patients refractory (or intolerant) to both TNFi drugs and vedolizumab, for which inclusion of an active comparator is questionable. This limitation was addressed by conducting an NMA.

Strengths and limitations of the NMA

Strengths of the analysis

Clear interpretation: Treat-through trials can be clearly interpreted, as continued treatment is given for approximately 1 full year. Response-based re-randomised trials complicate the interpretation of a full year of treatment given that results are not always directly reported for patients continuing treatment. Therefore, mimicking an ITT approach and recalculating the outcomes of response-based re-randomised trials provides a clear interpretation of the treatment effects for full 1-year regimens. The approach also allows estimation of the true relative treatment effect of active therapies compared to placebo for full 1-year regimens.

Satisfies NMA assumptions: An NMA of maintenance-phase data alone violates assumptions required for an NMA for comparable common arms to connect the studies. The inclusion of induction recalculated treat-through placebo arms allowed a common "true" placebo arm to be used to connect the studies, thus adhering to the assumptions required for NMA.

Robust data for TNFi agents: Although a certain degree of imputation was required to attain treat-through treatment arms across the re-randomised response-based studies, the comparisons to TNFi therapies required minimal or no imputations. No imputation of active-arm efficacy was needed for the TNFi treatments and fewer imputations were required compared to tofacitinib and vedolizumab for the placebo arms.

Maintains randomisation: By considering induction therapy, the NMA conducted maintains the initial randomisation to active drug or placebo. An analysis of maintenance data alone would be subject to selection bias, as not all patients continued in the maintenance phases of the trials reported for the re-randomised response-based trials.

Consistency with head-to-head data: Results from the VARSITY trial [54], the only head-to-head study in UC, could be included in the base-case approach, which strengthened the comparisons made to vedolizumab and adalimumab. Moreover, the efficacy results from the VARSITY trial were similar to those constructed from the individual studies for vedolizumab and adalimumab, validating the base-case approach of combining treat-through data between the trials.

Sensitivity analyses show consistent results: As a sensitivity analysis, an alternative approach was taken to mimic response-based trial designs. This required fewer imputations for the placebo arms compared to the base-case approach, and no imputation for the active arms. No imputations were required for the placebo arms of UNIFI, PURSUIT, ULTRA 2 or ACT 1 [49,55,56,60,61], as efficacy data were reported that could be included in the analysis. The conclusions drawn regarding the treatment effect of ustekinumab were directionally similar to those from the base-case analysis, further validating the base-case approach.

Additional limitations

The UNIFI trial grouped patients according to BF and NBF, whereas other trials grouped patients either according to TNFi naïve/experienced or biologic naïve/experienced. However, only a small proportion of patients in the UNIFI trial were NBF with previous exposure to a biologic therapy (27 out of 961 patients [2.8%] across the induction study arms in the primary efficacy analysis).

A number of limitations are associated with analysis of the maintenance data for the 1-year NMA specifically.

GEMINI 1 data limitations: For GEMINI 1, data published for the vedolizumab arms are limited as efficacy data for induction nonresponders were available from a G-BA document [67] only for the overall population. This required estimation of results by populations for the active and placebo arms. Furthermore, Cohort 1 patients in GEMINI 1 were included in the induction efficacy data, comprising patients who had been blinded and randomised to treatment. The maintenance data were based on Cohort 1 and Cohort 2, and Cohort 2 included patients who had received open-label vedolizumab at induction. Data from GEMINI 1 stratified by cohort in the maintenance phase were not provided. The MAH stated that despite these limitations in the data available, the recalculated 1-year efficacy data from GEMINI 1 corresponded closely to the efficacy data from VARSITY [54]. However, since VARSITY reported data on a head-to-head comparison and GEMINI 1 reported imputed data for comparison of one of the drugs from VARSITY to placebo, it is not clear what exactly was compared. It is understood that inconsistency in networks was not assessed.

Placebo imputations for the BF group: The placebo imputations (where required) for the BF population relied on more limited data and assumptions than for the NBF population. For the BF population, induction responder data relied on a single study (UNIFI) whereas the corresponding data for the NBF population used data from three studies (UNIFI, PURSUIT and ACT 1) [49,55,56,60,61] to strengthen the estimates. In addition, nonresponder data used for the imputation were based on GEMINI (which itself required estimation of the population-specific values). The fact that UNIFI data with high placebo response rates compared to other trials were used for placebo imputation might introduce a bias and reduce the observed clinical effect of other drugs.

Calculation of the total number of re-randomised responders: The total number of patients in the re-randomised responder arms in maintenance were re-calculated to reflect an ITT population. This applied to active arms of GEMINI 1, UNIFI and PURSUIT-M and OCTAVE [51,55,57,61]. This was considered necessary to reflect the more limited sample sizes in the re-randomised arms of these studies and to ensure that this is not overestimated. The recalculation of the total number of patients resulted in smaller patient counts than for the ITT population, leading to a loss of power for detecting treatment differences.

Limitations biasing against ustekinumab

Some of the limitations faced in the analysis were considered to bias results against ustekinumab and therefore can be viewed as conservative.

OCTAVE re-randomises both placebo and tofacitinib induction responders [51]: Data published for the maintenance phase of OCTAVE included patients who responded to induction therapy on placebo or tofacitinib and were subsequently re-randomised. Subgroup analyses have been published for clinical remission at 52 weeks stratified by therapy received at induction for the re-randomised responders. The results showed that the clinical remission rates for the combined induction arms were similar to those for the tofacitinib induction arms. Use of the combined data could potentially overestimate the overall tofacitinib maintenance responses, as the placebo responders from induction in this group would be receiving tofacitinib for the first time. However, on the basis of data from the subgroup analysis and given the small proportion of re-randomised responders who received placebo at induction, this is not expected to have a major impact.

Differences in delayed-response assessment times: In PURSUIT and UNIFI only [55,56,60,61], delayed responders had to respond at week 14 and 16, respectively. This potentially underestimates the number of delayed responders included in the golimumab and ustekinumab trial arms.

OCTAVE included an open-label treatment for induction nonresponders: Induction nonresponders in the OCTAVE trial [51] continued to receive open-label tofacitinib to 60 weeks, which may introduce bias in favour of tofacitinib.

6 CONCLUSION

6.1 Direct evidence

Conclusions from trials of ustekinumab versus active treatment

No head-to-head studies of ustekinumab versus other active therapies in UC have been conducted.

Conclusions from trials of ustekinumab versus placebo

Clinical efficacy

The UNIFI randomised-control trials (induction and maintenance) provided statistically significant results that ustekinumab at both doses was effective at inducing and maintaining a clinical response, clinical remission, endoscopic healing and mucosal healing (a combination of endoscopic and histologic healing), reducing the inflammatory burden and improving health-related quality of life in a population of subjects with moderately to severely active UC who had previously experienced failure or were intolerant of conventional and/or biologic therapy. This included a large proportion of patients who responded between week 8 and week 16. Overall, 56% of patients without a clinical response at week 8 had a clinical response at week 16 after receiving ustekinumab 90 mg SC 8 weeks after induction.

We are uncertain in the long-term efficacy since no efficacy data is available beyond 1 year.

Treatment of subjects with prior failure to biological therapy represents a key goal and an unmet need. Maintenance of efficacy has been demonstrated for subjects who have experienced failure of either anti-TNF or vedolizumab but there were very few patients enrolled in the pivotal trial with failure of both. For patients with failure to anti-TNF and vedolizumab the clinical remission was achieved in 22.7%, 33.3% and 14.8% of those on ustekinumab 90 mg q12w, ustekinumab 90 mg q8w and placebo respectively, but the results did not reach statistical significance. The study was not powered to detect a difference in this subgroup and less than 10 patients were included in each group.

Clinical safety

Adverse event rates in the UNIFI studies were similar in the ustekinumab and placebo arms over the induction and maintenance study phases in the population of adult subjects with moderately to severely active ulcerative colitis. The MAH identified a new non-serious adverse drug reaction of sinusitis (at a frequency of common [≥1/100 and <1/10]). The rate of serious adverse event did not relevantly differ between treatment groups in the induction and maintenance studies. AEs of special interest were similar across treatment groups, including infections, malignancy and cardiovascular events.

It should be mentioned that the sample was too small to conclude that these frequencies were similar, and follow-up was too short to draw conclusion on the risk on malignancy.

The integrated safety summary for ustekinumab revealed a similar safety profile across indications (in patients with psoriasis, psoriatic arthritis, Crohn's disease and ulcerative colitis). Incidences of major cardiovascular AE, malignancies and deaths through to 1 year were ≤0.5/100 patient-years and a comparable safety profile to placebo across indications. A large registry in psoriasis (PSOLAR study) has demonstrated that ustekinumab was not associated with a higher risk of serious infections and that the cumulative risk of serious infections for patients who received ustekinumab was lower than for patients who received etanercept, adalimumab or infliximab. Findings were similar for the subpopulation with concomitant IBD.

6.2 Indirect evidence: NMAs

The Grading of Recommendations Assessment, Development and Evaluation (GRADE)-method was not applied, by either the MAH or the authors of this joint assessment, for rating the quality or certainty of the evidence (as high, moderate, low and very low certainty) from a network meta-analysis.

Conclusions on clinical effectiveness from 1-year NMA

Non-biologic failure (NBF) population

Ustekinumab versus placebo, golimumab and adalimumab

The NMAs in the NBF population suggested that patients who received ustekinumab ~6 mg/kg for induction followed by ustekinumab 90 mg had statistically significant higher odds of achieving a clinical response, clinical remission, and mucosal healing after a 1-year long treatment regimen when compared to placebo, golimumab and adalimumab.

While the point estimates for these comparisons were high, ranging from OR of 2.4 to 8.7, the findings need to be viewed with caution. Because the evidence base was very limited and we could not exclude that effects on clinical remission and mucosal healing for comparisons to golimumab and adalimumab are trivial, we are uncertain if ustekinumab is more effective than these two active treatments mentioned above.

Ustekinumab versus infliximumab and tofacitinib

In addition, NMAs suggested that a statistically significant difference was observed in a clinical response, but not in endpoints of clinical remission and mucosal healing, in patients treated with ustekinumab than for those treated with infliximab or tofacitinib. However, it should be stressed that trivial effects could not be excluded and that evidence based was very limited, which make us uncertain if ustekinumab is more effective in improving clinical response than these two drugs.

Ustekinumab versus vedolizumab

NMAs suggested that no statistically significant difference was observed in a clinical response, clinical remission, and mucosal healing in patients treated with ustekinumab when compared with vedolizumab.

Overall, given that various methodological issues related to maintenance models for non-biological failure patients were raised and discussed, such as the paucity of the data, indirectness and the fact that inconsistency between different sources of evidence was not estimated, that data were recalculated and imputed in majority of trials (either for induction responders and/or non-responders in placebo or active arm), that fixed effect model was used in a setting where heterogeneity is expected but could not be tested, and that there is a potential impact of bias due to pooling of doses, we are uncertain if ustekinumab improves clinical response, clinical remission, and mucosal healing after a 1-year treatment regimen when compared with golimumab and adalimumab, as well as clinical response when compared with infliximab and tofacitinib.

Biological failure (BF) population

Ustekinumab versus placebo

The 1-year NMA of effectiveness for the BF group suggests significantly higher odds of achieving a clinical response and clinical remission in comparison to placebo with both ustekinumab regimes. While the point estimates for these comparisons were high, with ORs of 5.3-6.9 for clinical remission and 4.8 for clinical response endpoint, and wide CrIs (which ranged from OR ~2 to ≥9) suggested that at ustekinumab may have moderate or higher effect, the findings need to be viewed in light of very limited evidence base.

Ustekinumab versus golimumab, adalimumab, infliximab, vedolizumab and tofacitinib

All other comparisons for effectiveness in terms of clinical response or clinical remission between ustekinumab and any active treatments were nonsignificant with wide CrIs, demonstrating that the evidence base is limited. Although SUCRA values suggested that ustekinumab ~6 mg/kg-ustekinumab 90 mg SC q8W early + ustekinumab 90 mg SC q8W delayed is more effective than other active treatments, we are uncertain in this result due to various methodological issues already raised and discussed.

Overall, given that NMA of mucosal healing in the BF group was not feasible as imputation data needed for placebo were not available in this population, and taking into account various methodological issues for the maintenance models for BF patients (even more scarce data than in NMAs in non-biologic failure patients, indirectness and the fact that inconsistency between different sources of evidence was not estimated, the fact that data were recalculated and imputed in majority of trials, that fixed effect model was used in a setting where heterogeneity is expected but could not be tested), we are uncertain in these results.

Conclusions on clinical safety from NMA

Induction NMA

NMAs were conducted for safety outcomes in the induction phases of clinical trials. These suggested that ustekinumab ~6 mg/kg was associated with similar likelihoods of overall adverse events, serious adverse events and overall infections as for other active comparators. Owing to the low event counts, we are uncertain in the NMA results for serious infections.

The relevance of the results from these NMAs is limited by the short time period for assessment and the low event counts for some endpoints and small sample sizes.

1-year NMA

No data on long-term relative safety after 1 year are available because various methodological challenges limit the ability to conduct an NMA on safety after 1 year of treatment.

For the long-term safety (> 1 year) of ustekinumab in UC, further evidence is needed.

7 REFERENCES

[1] Suzuki Y, Motoya S, Hanai H, et al. Efficacy and safety of adalimumab in Japanese patients with moderately to severely active ulcerative colitis. *J Gastroenterol* 2014;49(2):283–94.

[2] Jiang XL, Cui HF, Gao J, et al. Low-dose infliximab for induction and maintenance treatment in Chinese patients with moderate to severe active ulcerative colitis. *J Clin Gastroenterol* 2015;49(7):582–8.

[3] Probert C, Hearing SD, Schreiber S, et al. Infliximab in moderately severe glucocorticoid resistant ulcerative colitis: a randomised controlled trial. *Gut* 2003 Jul;52(7):998-1002.

[4] Sandborn WJ, Ghosh S, Panes J, et al. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. *N Engl J Med* 2012;367(7):616–24.

[5] Schreiber S, Colombel JF, Bloomfield R, et al. Increased response and remission rates in short-duration Crohn's disease with subcutaneous certolizumab pegol: an analysis of PRECiSE 2 randomized maintenance trial data. *Am J Gastroenterol* 2010;105:1574–82.

[6] Schreiber S, Reinisch W, Colombel JF, et al. Subgroup analysis of the placebo-controlled CHARM trial: increased remission rates through 3 years for adalimumab-treated patients with early Crohn's disease. *J Crohns Colitis* 2013;7:213–21.

[7] Panés J, López-Sanromán A, Bermejo F, et al. Early azathioprine therapy is no more effective than placebo for newly diagnosed Crohn's disease. *Gastroenterology* 2013;145:766–74.

[8] Ghosh S, Sandborn WJ, Colombel JF, et al. Interpreting registrational clinical trials of biological therapies in adults with inflammatory bowel diseases. *Inflamm Bowel Dis* 2016;22(11):2711–23.

[9] Su C, Lewis JD, Goldberg B, et al. A meta-analysis of the placebo rates of remission and response in clinical trials of active ulcerative colitis. *Gastroenterology* 2007;132(2):516–26.

[10] Altman DG, Deeks JJ. Meta-analysis, Simpson's paradox, and the number needed to treat. *BMC Med Res Methodol* 2002;2:3.

[11] Ungaro R, Colombel J-F, Lissoos T, et al. A treat-to-target update in ulcerative colitis: a systematic review. *Am J Gastroenterol* 2019;114(6):874–83.

[12] Moller FT, Andersen V, Wohlfahrt J, et al. Familial risk of inflammatory bowel disease: a population-based cohort study 1977–2011. *Am J Gastroenterol* 2015;110(4):564–71.

[13] Ungaro R, Mehandru S, Allen PB, et al. Ulcerative colitis. *Lancet* 2017;389(10080):1756–70.

[14] Fumery M, Singh S, Dulai PS, et al. Natural history of adult ulcerative colitis in population-based cohorts: a systematic review. *Clin Gastroenterol Hepatol* 2018;16(3):343–56.e3.

[15] Koutroubakis IE. Recent advances in the management of distal ulcerative colitis. *World J Gastrointest Pharmacol Ther* 2010;1(2):43.

[16] Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* 2017;390(10114):2769–78.

[17] Ghosh S, Mitchell R. Impact of inflammatory bowel disease on quality of life: results of the European Federation of Crohn's and Ulcerative Colitis Associations (EFCCA) patient survey. *J Crohns Colitis* 2007;1(1):10–20.

[18] Wanderås MH, Moum BA, Høivik ML, et al. Predictive factors for a severe clinical course in ulcerative colitis: results from population-based studies. *World J Gastrointest Pharmacol Ther* 2016;7(2):235.

[19] Ford AC, Moayyedi P, Hanauer SB, et al. Ulcerative colitis. *BMJ* 2013;346:f432.

[20] Alper A, Zhang L, Pashankar DS. Correlation of erythrocyte sedimentation rate and C-reactive protein with pediatric inflammatory bowel disease activity. *J Pediatr Gastroenterol Nutr* 2017;65(2):e25–7.

[21] Molodecky NA, Soon IS, Rabi DM, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012;142(1):46–54.e42.

[22] Jakubowski A, Zagórowicz E, Kraszewska E, et al. Rising hospitalization rates for inflammatory bowel disease in Poland. *Polish Arch Intern Med* 2014;124(4):180–90.

[23] M'Koma AE. Inflammatory bowel disease: an expanding global health problem. *Clin Med Insights Gastroenterol* 2013;6:33–47.

[24] Selinger CP, Leong RW. Mortality from inflammatory bowel diseases. *Inflamm Bowel Dis* 2012;18(8):1566–72.

[25] Bernstein CN, Ng SC, Lakatos PL, et al. A review of mortality and surgery in ulcerative colitis. *Inflamm Bowel Dis* 2013;19(9):2001–10.

[26] Winther KV, Jess T, Langholz E, et al. Survival and cause-specific mortality in ulcerative colitis: follow-up of a population-based cohort in Copenhagen County. *Gastroenterology* 2003;125(6):1576–82.

[27] Hoivik ML, Moum B, Solberg IC, et al. Work disability in inflammatory bowel disease patients 10 years after disease onset: results from the IBSEN study. *Gut* 2013;62(3):368–75.

[28] European Crohn's and Colitis Organisation. *e-Guideline algorithms: ulcerative colitis 2017*. http://www.e-guide.ecco-ibd.eu

[29] European Medicines Agency. *Summary of Product Characteristics: Stelara*. https://www.ema.europa.eu/en/documents/product-information/stelara-epar-product-information_en.pdf

[30] European Medicines Agency. *Summary of Product Characteristics: Humira*. https://www.ema.europa.eu/en/documents/product-information/humira-epar-product-information_en.pdf

[31] European Medicines Agency. *Summary of Product Characteristics: Remicade*. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000240/WC500050888.pdf

[32] European Medicines Agency. *Summary of Product Characteristics: Simponi*. https://www.ema.europa.eu/en/documents/product-information/simponi-epar-product-information_en.pdf

[33] European Medicines Agency. *Summary of Product Characteristics: Entyvio*. https://www.ema.europa.eu/en/documents/product-information/entyvio-epar-product-information_en.pdf

[34] European Medicines Agency. *Summary of Product Characteristics: Xeljanz*. https://www.ema.europa.eu/en/documents/product-information/xeljanz-epar-product-information_en.pdf

[35] European Medicines Agency. *Summary of Product Characteristics: Amgevita*. https://www.ema.europa.eu/en/documents/product-information/amgevita-epar-product-information_en.pdf

[36] European Medicines Agency. *Summary of Product Characteristics: Imraldi*. https://www.ema.europa.eu/en/documents/product-information/imraldi-epar-product-information_en.pdf

[37] European Medicines Agency. *Summary of Product Characteristics: Kromeya*. https://www.ema.europa.eu/en/documents/product-information/kromeya-epar-product-information_en.pdf

[38] European Medicines Agency. *Summary of Product Characteristics: Halimatoz*. https://www.ema.europa.eu/en/documents/product-information/halimatoz-epar-product-information_en.pdf

[39] European Medicines Agency. *Summary of Product Characteristics: Idacio*. https://www.ema.europa.eu/en/documents/product-information/idacio-epar-product-information_en.pdf

[40] European Medicines Agency. *Summary of Product Characteristics: Hefiya*. https://www.ema.europa.eu/en/documents/product-information/hefiya-epar-product-information_en.pdf

[41] European Medicines Agency. *Summary of Product Characteristics: Hulio*. https://www.ema.europa.eu/en/documents/product-information/hulio-epar-product-information_en.pdf

[42] European Medicines Agency. *Summary of Product Characteristics: Remsina*. https://www.ema.europa.eu/en/documents/product-information/remsima-epar-product-information_en.pdf

[43] European Medicines Agency. *Summary of Product Characteristics: Zessly*. https://www.ema.europa.eu/en/documents/product-information/zessly-epar-product-information_en.pdf

[44] European Medicines Agency. *Summary of Product Characteristics: Flixabi*. https://www.ema.europa.eu/en/documents/product-information/flixabi-epar-product-information_en.pdf

[45] European Medicines Agency. *Summary of Product Characteristics: InflectRa*. https://www.ema.europa.eu/en/documents/product-information/inflectra-epar-product-information_en.pdf

[46] Janssen. *Ulcerative colitis_Ustekinumab STC Final*. 2019-07-01. Pdf (EUnetHTA PTJA07 webpage)

[47] European Network for Health Technology Assessment. Joint action 3. Ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (UC) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies. *Project ID: PTJA07. Project Description and Planning*. EUnetHTA; 2019.

[48] Kobayashi T, Suzuki Y, Motoya S, et al. First trough level of infliximab at week 2 predicts future outcomes of induction therapy in ulcerative colitis-results from a multicenter prospective randomized controlled trial and its post hoc analysis. *J Gastroenterol* 2016;51(3):241–51.

[49] Rutgeerts P, Sandborn WJ, Feagan BG, et al. Infliximab for induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2005;353:2462–76.

[50] Silva R, Bissoli G, Farinelli E, et al. Infliximab versus adalimumab: clinical and endoscopy response in ulcerative colitis patients. A prospective study. *J Crohns Colitis* 2017;11(Suppl 1):S496.

[51] Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2017;376(18):1723–36.

[52] Reinisch W, Sandborn WJ, Panaccione R, et al. 52-week efficacy of adalimumab in patients with moderately to severely active ulcerative colitis who failed corticosteroids and/or immunosuppressants. *Inflamm Bowel Dis* 2013;19(8):1700–9.

[53] Sandborn WJ, Van Assche G, Reinisch W, et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. *Gastroenterology* 2012;142(2):257–65.e3.

[54] Schreiber S, Peyrin-Biroulet L, Loftus EV, et al. VARSITY: a double-blind, double-dummy, randomised, controlled trial of vedolizumab versus adalimumab in patients with active ulcerative colitis [ECCO abstract OP34]. *J Crohns Colitis* 2019;13(suppl 1).

[55] Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab maintains clinical response in patients with moderate-to-severe ulcerative colitis. *Gastroenterology* 2014;146(1):96–109.e1.

[56] Hibi T, Imai Y, Senoo A, et al. Efficacy and safety of golimumab 52-week maintenance therapy in Japanese patients with moderate to severely active ulcerative colitis: a phase 3, double-blind, randomized, placebo-controlled study—(PURSUIT-J study). *J Gastroenterol* 2017;52(10):1101–11.

[57] Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2013;369(8):699–710.

[58] Kobayashi T, Suzuki Y, Watanabe K, et al. A phase 1, multiple-dose study of vedolizumab in Japanese patients with ulcerative colitis. *J Clin Pharmacol* 2019;59(2):271–9.

[59] Motoya S, Watanabe K, Ogata H, et al. Vedolizumab in Japanese patients with ulcerative colitis: a phase 3, randomized, double-blind, placebo-controlled study. *PLoS One* 2019;14(2):e0212989.

[60] Janssen Research and Development LLC. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab induction and maintenance therapy in subjects with moderately to severely active ulcerative colitis. *Clinical Study Report CNTO1275UCO3001 Induction*. Janssen; 2018.

[61] Janssen Research and Development LLC. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab induction and maintenance therapy in subjects with moderately to severely active ulcerative colitis. *Clinical Study Report CNTO1275UCO3001 Maintenance*. Janssen; 2018.

[62] Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. The Cochrane Collaboration; 2011. http://handbook.cochrane.org

[63] European Network For Health Technology Assessment. *Comparators & Comparisons: Direct and Indirect Comparisons*. EuNetHTA; 2015. https://www.eunethta.eu/wp-content/uploads/2018/03/Direct_comparators_comparisons.pdf

[64] Gemayel NC, Rizzello E, Atanasov P, et al. Dose escalation and switching of biologics in ulcerative colitis: a systematic literature review in real world evidence. *Curr Med Res Opin*. In press. https://doi.org/10.1080/03007995.2019.1631058

[65] Dias S, Sutton AJ, Ades A, et al. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decision Making* 2013;33(5):607–17.

[66] Dias S, Welton NJ, Sutton AJ, et al. *NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials*. NICE; 2011. http://nicedsu.org.uk/wp-content/uploads/2017/05/TSD2-General-meta-analysis-corrected-2Sep2016v2.pdf

[67] Federal Joint Committee (G-BA) document – Vedolizumab. 2015. https://www.g-ba.de/downloads/40-268-6001/2015-01-08_AM-RL-XII_Vedolizumab_2014-07-15-D-122_ZD.pdf

[68] Sands BE, Abreu MT, Ferry GD, et al. Design issues and outcomes in IBD clinical trials. *Inflamm Bowel Dis* 2005;11(Suppl 1):S22–8.

[69] Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162:777–84.

[70] Janssen. *UC_NMA Study Report*. 2019.

[71] Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6(7):e1000100.

[72] European Network For Health Technology Assessment. *Process of Information Retrieval for Systematic Reviews and Health Technology Assessments on Clinical Effectiveness.* EuNetHTA; 2017. https://www.eunethta.eu/wp-content/uploads/2018/01/Guideline_Information_Retrieval_V1-2_2017.pdf

[73] Brignardello-Petersen R, Bonner A, Alexander PE, Siemieniuk RA, Furukawa TA, Rochwerg B, Hazlewood GS, Alhazzani W, Mustafa RA, Murad MH, Puhan MA, Schünemann HJ, Guyatt GH, For the GRADE Working Group, Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. *Journal of Clinical Epidemiology* (2017), doi: 10.1016/j.jclinepi.2017.10.005.

[74] Panes J, Vermeire S, Lindsay JO, Sands BE, Su C, Friedman G, et al. Tofacitinib in Patients with Ulcerative Colitis: Health-Related Quality of Life in Phase 3 Randomised Controlled Induction and Maintenance Studies. *J Crohns Colitis*. 2018;12(2):145-56

[75] Panés J, Su C, Bushmakin AG, et al. Direct and indirect effects of tofacitinib on treatment satisfaction in patients with ulcerative colitis. *J Crohns Colitis* 2016;10(11):1310–5.

[76] Singh S, Fumery M, Sandborn W, et al. Systematic review with network meta-analysis: first-and second-line pharmacotherapy for moderate-severe ulcerative colitis. *Aliment Pharmacol Ther* 2018;47(2):162–75.

[77] National Institute for Health and Care Excellence. *Tofacitinib for Moderately to Severely Active Ulcerative Colitis*. Technology Appraisal Guidance TA547 https://www.nice.org.uk/guidance/ta547/history

[78] National Institute for Health and Care Excellence. *Ulcerative Colitis: Management*. Clinical Guidance 166. https://www.nice.org.uk/guidance/cg166

[79] Papp K, Gottlieb AB, Naldi L, et al. Safety surveillance for ustekinumab and other psoriasis treatments from the Psoriasis Longitudinal Assessment and Registry (PSOLAR). *J Drugs Dermatol* 2015;14(7):706–14.

[80] European Medicines Agency. *Guideline on the Development of New Medicinal Products for the Treatment of Ulcerative Colitis*. EMA; 2019. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-development-new-medicinal-products-treatment-ulcerative-colitis-revision-1_en.pdf

[81] Trigo-Vicente C, Gimeno-Ballester V, García-López S, et al. Systematic review and network meta-analysis of treatment for moderate-to-severe ulcerative colitis. *Int J Clin Pharm* 2018;40(6):1411–19.

[82] Reinisch W, Sandborn WJ, Hommes DW, et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. *Gut* 2011;60(6):780–7.

[83] European Network For Health Technology Assessment. *Internal Validity of Randomised Controlled Trials*. EUnetHTA; 2015. https://www.eunethta.eu/wp-content/uploads/2018/01/16_WP7-SG3-GL-int_val_RCTs_amend2015.pdf

[84] Cope S, Donohue JF, Jansen JP, et al. Comparative efficacy of long-acting bronchodilators for COPD: a network meta-analysis. *Respir Res* 2013;14:100.

[85] Schünemann HJ, Oxman AD, Higgins JPT, et al. Presenting results and 'Summary of findings' tables. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. The Cochrane Collaboration; 2011. Chapter 11. http://handbook.cochrane.org

[86] Chaimani A, Salanti G, Leucht S, et al. Common pitfalls and mistakes in the set-up, analysis and interpretation of results in network meta-analysis: what clinicians should look for in a published article. *Evidence Based Mental Health* 2017;20(3):88.

[87] Mills EJ, Ioannidis JP, Thorlund K, et al. How to use an article reporting a multiple treatment comparison meta-analysis. *JAMA* 2012;308(12):1246–53.

[88] Ghosh S, Sands BE, De Villiers W, et al. A pooled safety analysis from the ustekinumab Crohn's disease and psoriatic diseases phase 2 and 3 clinical trial programs. *United Eur Gastroenterol J* 2016;4(5S):A157.

[89] Kalb RE, Fiorentino DF, Lebwohl MG, et al. Risk of serious infection with biologic and systemic treatment of psoriasis: results from the Psoriasis Longitudinal Assessment and Registry (PSOLAR). *JAMA Dermatol* 2015;151(9):961–9.

[90] Loftus E Jr, Augustin M, Bissonnette R, et al. Prevalence of inflammatory bowel disease amongst patients with psoriasis and incidence of serious infections in this subset: results from the PSOLAR registry. *J Crohns Colitis* 2016;10(Suppl 1):S418.

[91] Strober BE, Bissonnette R, Fiorentino D, et al. Comparative effectiveness of biologic agents for the treatment of psoriasis in a real-world setting: results from a large, prospective, observational study (Psoriasis Longitudinal Assessment and Registry [PSOLAR]). J Am Acad Dermatol 2016;74(5):851–61.e4.

[92] Feagan BG, Reinisch W, Rutgeerts P, Sandborn WJ, Yan S, Eisenberg D, et al. The effects of infliximab therapy on health-related quality of life in ulcerative colitis patients. *Am J Gastroenterol*. 2007;102(4):794-802

[93] Sandborn WJ, Van Assche GA, Thakkar R, Lazar A, Kron M, Yang M, et al. Adalimumab Improves Health-Related Quality of Life for 52 Weeks in Patients With Ulcerative Colitis. *Gastroenterology.* 2011;140(5).

[94] Probert CS, Sebastian S, Gaya DR, Hamlin PJ, Gillespie G, Rose A, et al. Golimumab induction and maintenance for moderate to severe ulcerative colitis: results from GO-COLITIS (Golimumab: a Phase 4, UK, open label, single arm study on its utilization and impact in ulcerative Colitis). BMJ Open Gastroenterol. 2018;5(1):e000212-e.

[95] 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(2):229-39 e5.

APPENDIX 1: GUIDELINES FOR DIAGNOSIS AND MANAGEMENT

The table below takes into account only guidelines which were published in English and refers to the stage of treatment at which ustekinumab is to be used (after failure of conventional therapy and/or a biologic).

Table A1 Overview of guidelines used for this assessment [46]

Name of society/organisation issuing guidance	Date of issue	Country/ies to which applicable	Summary of recommendation	Level of evidence (A,B,C)/ class of recommendation (I, IIa, IIb, III)
NICE (NICE 2019)	2019	United Kingdom	**Infliximab** is recommended as an option for the treatment of acute exacerbations of severely active ulcerative colitis only in patients in whom ciclosporin is contraindicated or clinically inappropriate, based on a careful assessment of the risks and benefits of treatment in the individual patient. [2008] In people who do not meet the criterion described above, infliximab should only be used for the treatment of acute exacerbations of severely active ulcerative colitis in clinical trials. [2008][1] For guidance on biologics and Janus kinase inhibitors see the NICE technology appraisal guidances. **Infliximab, adalimumab and golimumab** are recommended, within their marketing authorisations, as options for treating moderately to severely active ulcerative colitis 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. Golimumab is recommended only if the company provides the 100 mg dose of golimumab at the same cost as the 50 mg dose, as agreed in the patient access scheme. The choice of treatment between infliximab, adalimumab or golimumab should be made on an individual basis after discussion between the responsible clinician and the patient about the advantages and disadvantages of the treatments available. This should take into consideration therapeutic need and whether or not the patient is likely to adhere to	Not specified. Recommendation based on the consultation with experts, people using services, carers and the public.

[1] These recommendations are from 'Infliximab for acute exacerbations of ulcerative colitis' (NICE technology appraisal guidance 217).

Name of society/organisation issuing guidance	Date of issue	Country/ies to which applicable	Summary of recommendation	Level of evidence (A,B,C)/ class of recommendation (I, IIa, IIb, III)
			treatment. If more than 1 treatment is suitable, the least expensive should be chosen (taking into account administration costs, dosage and price per dose)[2]. **Vedolizumab** is recommended, within its marketing authorisation, as an option for treating moderately to severely active ulcerative colitis in adults only if the company provides vedolizumab with the discount agreed in the patient access scheme[3]. **Tofacitinib** is recommended, within its marketing authorisation, as an option for treating moderately to severely active ulcerative colitis in adults when conventional therapy or a biological agent cannot be tolerated or the disease has responded inadequately or lost response to treatment. It is recommended only if the company provides tofacitinib with the discount agreed in the commercial arrangement.[4]	
ECCO (Harbord 2017)	2017	Europe	ECCO statement 11K Patients with moderate colitis refractory to thiopurines should be treated with **anti-TNF** [EL:1], preferably combined with thiopurines, at least for infliximab [EL:2], or vedolizumab [EL:2]. In case of treatment failure, **a different anti-TNF** [E:L4] **or vedolizumab** [EL:2] should be considered, and colectomy recommended if further medical therapy does not achieve a clear clinical benefit [EL:5]. ECCO statement 12I **Anti-TNF** or **vedolizumab** may be used as first-line biological therapy. Vedolizumab is effective in patients failing anti-TNF [EL2]. In patients responding to vedolizumab, maintenance therapy with vedolizumab is appropriate [EL2].	EL based on the Oxford Centre for Evidence-Based Medicine. Recommendation based on the consensus statements (≥80% agreement of participants).
WGO (Bernstein 2015)	2015	World	Treatment of ulcerative colitis - medium resources available If remission is not maintained with 5-aminosalicylic acid, then azathioprine or 6-mercaptopurine/azathioprine should be	Not specified.

[2] This recommendation is from 'Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy' (NICE technology appraisal guidance 329).

[3] This recommendation is from 'Vedolizumab for treating moderately to severely active ulcerative colitis' (NICE technology appraisal guidance 342).

[4] This recommendation is from 'Tofacitinib for moderately to severely active ulcerative colitis' (NICE technology appraisal guidance 547).

Name of society/organisation issuing guidance	Date of issue	Country/ies to which applicable	Summary of recommendation	Level of evidence (A,B,C)/ class of recommendation (I, IIa, IIb, III)
			considered; if azathioprine fails, **anti-TNF or vedolizumab** should be considered. If biological agents are available, then depending on the severity of the illness their use may be indicated instead of trials of immunomodulator monotherapy. - extensive resources available **Infliximab or vedolizumab** intravenously, or Humira (**adalimumab) or golimumab** subcutaneously, are options for ambulatory patients with moderate to severe disease. In case of azathioprine failure, **anti-TNF or vedolizumab** should be considered.	
GETECCU (Gomollon 2012)	2013	Spain	For patients with moderate ulcerative colitis flare after failure to corticosteroids, treatment should be as for severe flare. In treatment of severe ulcerative colitis flare after failure to corticosteroids the following are recommended: cyclosporine (QE: moderate, R: strong), infliximab (QE: moderate, R: strong) or surgery is suggested (QE: low, R: weak). After failure of treatment with **cyclosporine or infliximab** the other non-used medical agent may be considered. When remission on infliximab is not maintained, **surgery** or **adalimumab** can be used.	QE based on the GRADE methodology. Recommendation divided into 4 categories (implication for the clinician): strong – to *do it*; weak – to *probably do it*; weak against an intervention – to *probably do not do it*; strong against an intervention –*not to do it*.
SE (Manz 2011)	2011	Switzerland	If a patient was a prior azathioprine or mercaptopurine failure, **infliximab** at 5 mg/kg can be started. In case of lack of response after two infusions or no remission after ten weeks, the dose should be doubled to 10 mg/kg. In a patient still not responding, third-line **immunosuppressive therapy or surgery** must be considered at a specialised centre.	Not specified. Recommendation based on consensus of experts.
ISGE (Orlando 2011)	2011	Italy	**Infliximab** induction regimen can be used in patients with moderate-to-severe ulcerative colitis who are refractory to systemic corticosteroids [EL: 1b, RG: A] and in corticosteroid-dependent patients who are intolerant/refractory to thiopurines [EL: 2b, RG: C]. A **colectomy** is recommended if there is no improvement after infliximab [EL: 5, RG: D].	EL and RG based on the Oxford Centre for Evidence Based Medicine.

Abbreviations: NICE – National Institute for Health and Care Excellence, ECCO – European Crohn's and Colitis Organisation, WGO – World Gastroenterology Organisation, GETECCU – Spanish Group of Ulcerative Colitis and Crohn's disease, SE – Swiss experts, ISGE – Italian Society of Gastroenterology, EL – level of evidence, QE – quality of evidence, RG – grades of recommendation, R - recommendation.
Sources: NICE 2019, Harbord 2017, Bernstein 2015, Gomollon 2012, Manz 2011, Orlando 2011

APPENDIX 2: DETAILS METHOD AND RESULTS SECTIONS OF MAIN TEXT

A2.1 Excluded studies with reasons

Table A2. Publications excluded based on screening of full text documents by Reasons of Exclusion from Submission Dossier

#	Reference	Rationale for exclusion	Specific reason for 'Not study type of interest'
1	Parikh A.;Fox I.;Leach T.;Xu J.;Scholz C.;Patella M.;Feagan B. G. Long-term clinical experience with vedolizumab in patients with inflammatory bowel disease. Inflammatory bowel diseases. 2013;19(8):1691-1699	Not study type of interest	Observational study
2	Parikh A.;Leach T.;Wyant T.;Scholz C.;Sankoh S.;Mould D. R.;Ponich T.;Fox I.;Feagan B. G. Vedolizumab for the treatment of active ulcerative colitis: a randomized controlled phase 2 dose-ranging study. Inflammatory bowel diseases. 2012;18(8):1470-1479	Not population of interest	
3	Perks B. Randomized non-inferiority trial fails to find inferiority switching from infliximab originator to CT-P13 biosimilar. GaBI Journal. 2017;6(4):	Not intervention of interest	
4	Peyrin-Biroulet L.;Danese S. Tofacitinib: janus bifrons in ulcerative colitis treatment. Gastroenterology. 2013;144(5):1136-1138	Duplicates	
5	Pierik M. J.;Van Der Meulen-De Jong A. E.;Bloemsaat-Minekus J. P. J.;Van Megen Y. J. B.;Dijkstra G. Switching from the originator infliximab to biosimilar CT-P13 did not change the quality of life and clinical efficacy for IBD patients in stable remission in daily clinical practice (interim analysis). Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S339-S340	Not study type of interest	Observational study
6	Pouillon L.;Bossuyt P.;Peyrin-Biroulet L. Tofacitinib Is the Right OCTAVE for Ulcerative Colitis. Gastroenterology. 2017;153(3):862-864	Duplicates	
7	Reinisch W.;Feagan B.;Yan S.;Sandborn W. J.;Rutgeerts P.;Eisenberg D. Improvement in health related quality of life in infliximab-treated moderate-to-severe active ulcerative colitis patients: improvement overall and by baseline disease activity. UEGW abstract database. 2006;:Abstract MON?G?298	Not study type of interest	Observational study
8	Reinisch W.;Gibson P.;Sandborn W. J.;Feagan B.;Marano C.;Strauss R.;Johanns J.;Zhang H.;Padgett L.;Adedokun O. J.;Colombel J. F.;Collins J.;Rutgeerts P.;Tarabar D. Safety, efficacy, and pharmacokinetics of golimumab in patients with moderately to severely Active ulcerative colitis: PURSUIT-SC long-term extension. Journal of Crohn's and Colitis. 2016;10:S248-S249	Not study type of interest	Long-term follow-up study of an RCT population
9	Reinisch W.;Gibson P. R.;Sandborn W. J.;Feagan B. G.;Strauss R.;Johanns J.;Padgett L.;Adedokun O. J.;Colombel J. F.;Collins J.;Rutgeerts P.;Tarabar D.;Marano C. Long-term Benefit of Golimumab	Not study type of interest	Long-term follow-up study of an RCT population

	for Patients with Moderately-to-Severely Active Ulcerative Colitis: Results from the PURSUIT-Maintenance Extension. J Crohns Colitis. 2018;:		
10	Reinisch W.;Lewis J. D.;Dassopoulos T.;Ginsburg P.;Sands B. E.;Feagan B.;Danese S.;Roseth A. G.;Rosario M.;Yang H.;et al. Clinical response and remission with vedolizumab across a range of baseline fecal calprotectin levels in ulcerative colitis: results from GEMINI 1. Journal of crohn's and colitis.. 2015;9:S47	Not outcome of interest	
11	Reinisch W.;Sandborn W. J.;Kumar A.;Pollack P. F.;Lazar A.;Thakkar R. B. 52-week clinical efficacy with adalimumab in patients with moderately to severely active ulcerative colitis who failed corticosteroids and/or immunosuppressants. Gut. 2011;60:A139-A140	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
12	Reinisch W.;Sandborn W. J.;Panaccione R.;Huang B.;Pollack P. F.;Lazar A. 52-week efficacy of adalimumab in patients with moderately to severely active ulcerative colitis who failed corticosteroids and/or immunosuppressants. . 2013;19:	Duplicates	
13	Roblin X.;Osterman M. T.;Glover S.;Navaneethan U.;Popa M. A.;Wyant T.;James A.;Lasch K.;Rosario M. The relationship between vedolizumab drug concentrations at or before week 6 and remission at week 14 in ulcerative colitis patients from GEMINI 1. Journal of Crohn's and Colitis. 2016;10:S72-S73	Not outcome of interest	
14	Roblin X.;Paul S.;Boschetti G.;Phelip J. M.;Del Tedesco E.;Berger A.;Nancey S.;Williet N.;Flourie B. Interest in the addition of azathioprine (AZA) to the switch of anti-TNF in IBD patients in loss of response with undetectable anti-TNF trough levels and anti-drug antibodies: a prospective randomised trial. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S414-S415	Not intervention of interest	
15	Rosario M.;Abhyankar B.;Sankoh S.;Dirks N.;Lasch K.;Sandborn W. Relationship between vedolizumab pharmacokinetics and endoscopic outcomes in patients with Ulcerative Colitis. Journal of crohn's and colitis.. 2015;9:S46	Not outcome of interest	
16	Rosario M.;Wyant T.;Leach T.;Sankoh S.;Scholz C.;Parikh A.;Fox I.;Feagan B. G. Vedolizumab Pharmacokinetics, Pharmacodynamics, Safety, and Tolerability Following Administration of a Single, Ascending, Intravenous Dose to Healthy Volunteers. Clinical Drug Investigation. 2016;36(11):913-923	Not population of interest	
17	Rubin D. T.;Tudor D.;Khalid J. M.;Patel H. Improvements in subcomponents of the inflammatory bowel disease questionnaire in patients treated with vedolizumab: results from GEMINI trial data. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S395	Not outcome of interest	
18	Rutgeerts P.;Feagan B.;Marano C.;Strauss R.;Johanns J.;Zhang H. Phase 2/3 randomized, placebo-controlled, double-blind study of SC golimumab induction in moderate to severe UC. Journal of gastroenterology and hepatology. 2013;28:591	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main

			publication for the RCT)
19	Rutgeerts P.;Feagan B. G.;Marano C.;Strauss R.;Johanns J.;Zhang H.;Guzzo C.;Colombel J. F.;Reinisch W.;Gibson P.;Collins J.;Jarnerot G.;Sandborn W. Phase 3 study to evaluate the SC golimumab maintenance therapy in moderate to severe UC: Pursuit-maintenance. Journal of Gastroenterology and Hepatology. 2013;28:593	Not study type of interest	Observational study
20	Rutgeerts P.;Feagan B. G.;Marano C. W.;Padgett L.;Strauss R.;Johanns J.;Adedokun O. J.;Guzzo C.;Zhang H.;Colombel J. F.;et al. Randomised clinical trial: a placebo-controlled study of intravenous golimumab induction therapy for ulcerative colitis. Alimentary pharmacology & therapeutics. 2015;42(5):504-514	Not outcome of interest	
21	Sandborn W.;Colombel J. F.;Panaccione R.;Lasch K.;Mody R.;Green A.;Abhyankar B. Deep remission as a predictor of clinical outcomes in vedolizumab-treated patients with ulcerative colitis. Journal of crohn's and colitis.. 2015;9:S299-S300	Not outcome of interest	
22	Sandborn W.;Colombel J. F.;Panaccione R.;Lasch K.;Mody R.;Green A.;Abhyankar B. Deep remission as a predictor of clinical outcomes in vedolizumab-treated patients with ulcerative colitis. Gastroenterology. 2015;148(4 SUPPL. 1):S256	Not outcome of interest	
23	Sandborn W.;Colombel J. F.;Panaccione R.;Lasch K.;Sankoh S.;Abhyankar B. Deep clinical remission in patients with Ulcerative Colitis: evaluating the effects of vedolizumab on various combinations of endoscopic and patient-reported outcomes. Journal of crohn's and colitis.. 2015;9:S237-S238	Not outcome of interest	
24	Sandborn W.;Feagan B.;Colombel J. F.;Reinisch W.;Gibson P.;Rutgeerts P.;Weng H.;Yao R.;Marano C.;Zhang H.;Strauss R. Relationship between clinical outcomes and disease duration, extent, and severity in patients with ulcerative colitis who received 6 weeks of treatment with golimumab. United European Gastroenterology Journal. 2014;2(1):A538-A539	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
25	Sandborn W. J. Mucosal healing in inflammatory bowel disease. Reviews in Gastroenterological Disorders. 2008;8(4):271-272	Not population of interest	
26	Sandborn W. J. Mucosal healing with infliximab: Results from the active ulcerative colitis trials. Gastroenterology and Hepatology. 2012;8(2):117-119	Not study type of interest	Literature review
27	Sandborn W. J.;Danese S.;Panes J.;Zhang H.;Woodworth D.;Marren A.;Su C. Onset of efficacy of tofacitinib for induction therapy in patients with active ulcerative colitis in two multinational, phase 3 clinical trials. United european gastroenterology journal. Conference: 24th united european gastroenterology week, UEG 2016. Austria. 2016;4(5 Supplement 1):A441-A442	Not outcome of interest	
28	Sandborn W. J.;Danese S.;Panés J.;Zhang H.;Woodworth D.;Marren A.;Su C. Onset of efficacy of tofacitinib for induction therapy in patients with active ulcerative colitis in two multinational, phase 3 clinical trials. American Journal of Gastroenterology. 2016;111:S260-S261	Not outcome of interest	

29	Sandborn W. J.;Panes J.;D'Haens G. R.;Sands B. E.;Su C.;Moscariello M.;Jones T. V.;Pedersen R. D.;Friedman G. S.;Lawendy N.;et al. Tofacitinib for the treatment of ulcerative colitis: up to 4.4 years of safety data from global clinical trials. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S45-S46	Not outcome of interest	
30	Sandborn W. J.;Panés J.;Zhang H.;Yu D.;Niezychowski W.;Su C. Correlation Between Concentrations of Fecal Calprotectin and Outcomes of Patients With Ulcerative Colitis in a Phase 2 Trial. Gastroenterology. 2016;150(1):96-102	Not outcome of interest	
31	Sandborn W. J.;Rutgeerts P.;Feagan B. G.;Reinisch W.;Olson A.;Johanns J.;Lu J.;Horgan K.;Rachmilewitz D.;Hanauer S. B.;et al. Colectomy rate comparison after treatment of ulcerative colitis with placebo or infliximab. Gastroenterology. 2009;137(4):1250-60	Not outcome of interest	
32	Sandborn W. J.;Rutgeerts P.;Zhang H.;Adedokun O. J.;Xu S.;Shraim R.;Marano C. Characterization of ulcerative colitis patients in the Golimumab PURSUIT-Maintenance study: post-hoc analyses of patients who maintained and did not maintain clinical response through week 54. Journal of crohn's and colitis. Conference: 12th congress of the european crohn's and colitis organisation, ECCO 2017. Spain. 2017;11(Supplement 1):S304-S305	Not outcome of interest	
33	Sandborn W. J.;Rutgeerts P.;Zhang H.;Adedokun O. J.;Xu S.;Shraim R.;Marano C. Characterization of ulcerative colitis patients in the golimumab pursuit-maintenance study: post-hoc analyses of patients who maintained and did not maintain clinical response through week 54. Gastroenterology. Conference: digestive disease week 2017, DDW 2017. United states. 2017;152(5 Supplement 1):S598	Not outcome of interest	
34	Sandborn W. J.;Van Assche G.;Thakkar R. B.;Lazar A.;Kron M.;Yang M. Adalimumab improves health-related quality of life for 52 weeks in patients with ulcerative colitis. Gut. 2012;61:A400-1	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
35	Sandborn W. J. Colombel J. F. D'Haens G. Van Assche G. Wolf D. Kron M. Lazar A. Robinson A. M. Yang M. Chao J. D.;Thakkar R. One-year maintenance outcomes among patients with moderately-to-severely active ulcerative colitis who responded to induction therapy with adalimumab: subgroup analyses from ULTRA 2. Alimentary pharmacology & therapeutics. 2013;37(2):204-213	Not population of interest	
36	Sands B.;D'Haens G.;Sandborn W. J.;Hibi T.;Su C.;Niezychowski W.;Ghosh S.;Zhang H.;Yu D.;Woodworth D.;Healey P.;Marren A.;Panés J. Tofacitinib has induction efficacy in moderately to severely active ulcerative colitis, regardless of prior TNF inhibitor therapy. American Journal of Gastroenterology. 2016;111:S261	Not outcome of interest	
37	Sands B.;Feagan B.;Yan S.;Sandborn W.;Rutgeerts P.;Eisenberg D. Improvement in health related quality of life in infliximab-treated moderate-to-severe active ulcerative colitis patients: improvement overall and by baseline disease activity. Inflammatory bowel diseases. 2007;13(Supp 5):648	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)

38	Sands B.;Hanauer S.;Colombel J. F.;Danese S.;Abreu M. T.;Ahuja V.;Ponich T.;Hilmi I.;Sankoh S.;Smyth M.;Abhyankar B.;Fox I.;Feagan B. Reductions in corticosteroid use in patients with ulcerative colitis or Crohn's disease treated with vedolizumab. United European Gastroenterology Journal. 2013;1(1):A370	Not outcome of interest	
39	Sands B. E.;Cohen R.;Isaacs K.;Fedorak R. N.;Abhyankar B.;Sankoh S.;Smyth M. Infusion-related reactions with vedolizumab treatment in patients with UC or CD during the GEMINI 1 and GEMINI 2 clinical trials. Journal of Crohn's and Colitis. 2015;9:S392-S393	Not outcome of interest	
40	Sands B. E.;Cohen R. D.;Isaacs K. L.;Fedorak R. N.;Abhyankar B.;Sankoh S.;Smyth M. D. Infusion-related reactions with vedolizumab treatment in patients with UC or CD during the GEMINI 1 and GEMINI 2 clinical trials. Gastroenterology. 2015;148(4):S232-S233	Not outcome of interest	
41	Sands B. E.;Dubinsky M. C.;Vermeire S.;Sankoh S.;Rosario M.;Milch C. Effects of increased vedolizumab dosing frequency on disease activity in ulcerative colitis and crohn's disease. United European Gastroenterology Journal. 2014;2(1):A67	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
42	Sands B. E.;Shafran I.;Farraye F. A.;Cheifetz A. S.;Abhyankar B.;Sankoh S.;Smyth M. Efficacy and safety of retreatment with vedolizumab in patients with ulcerative colitis. Journal of crohn's and colitis.. 2015;9:S37	Not intervention of interest	
43	Sands B. E.;Shafran I.;Farraye F. A.;Cheifetz A. S.;Abhyankar B.;Sankoh S.;Smyth M. D. Efficacy and safety of retreatment with vedolizumab in patients with ulcerative colitis. Gastroenterology. 2015;148(4):S278	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
44	Sands B. E.;Taub P. R.;Feagan B. G.;Armuzzi A.;Friedman G. S.;Moscariello M.;Lawendy N.;Pedersen R. D.;Chan G.;Nduaka C. I.;et al. The effect of tofacitinib on serum lipids and cardiovascular safety in patients with ulcerative colitis: results from the tofacitinib ulcerative colitis clinical programme. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S023	Not outcome of interest	
45	Seagrove A.;Rapport F.;Williams J. Infliximab or ciclosporin: patients' treatment preferences and the impact of ulcerative colitis (UC) on their lives. Gut.. 2014;63:A65	Not study type of interest	Observational study
46	Seagrove A. C.;Alam M. F.;Alrubaiy L.;Cheung W. Y.;Clement C.;Cohen D.;Grey M.;Hilton M.;Hutchings H.;Morgan J.;et al. Randomised controlled trial. Comparison Of iNfliximab and ciclosporin in STeroid Resistant Ulcerative Colitis: trial design and protocol (CONSTRUCT). BMJ open. 2014;4(4):e005091	Not intervention of interest	
47	Sheridan J.;Coe C. A.;Doran P.;Egan L.;Cullen G.;Kevans D.;Leyden J.;Galligan M.;O'Toole A.;McCarthy J.;et al. Protocol for a multicentred randomised controlled trial investigating the use of personalised golimumab dosing tailored to inflammatory load in ulcerative colitis: the GOAL-ARC study (GLM dose Optimisation to Adequate Levels to Achieve Response in Colitis) led by the	Not study type of interest	Trial protocol

	INITIAtive group (NCT 0268772). BMJ open gastroenterology. 2018;5(1) (no pagination):		
48	Simmons J.;Jewell D. P. Infliximab for ulcerative colitis. Digestive and Liver Disease. 2002;34(9):616-618	Not study type of interest	Review article
49	Singh S.;Loftus Jr E. V. Management of severe steroid-refractory ulcerative colitis: Cyclosporine or infliximab?. Gastroenterology. 2013;144(5):1138-1140	Duplicates	
50	Smith K. Golimumab shows promise in treatment of active ulcerative colitis. Nature Reviews Gastroenterology and Hepatology. 2013;10(7):386	Not study type of interest	Review article
51	Stallmach A.;Bokemeyer B.;Axler J.;Curtis R.;Ehehalt R.;Feagan B.;Geransar P.;James A.;Kaviya A.;Khalid J. M.;et al. Sustained remission with vedolizumab in patients with moderately to severely active ulcerative colitis: a GEMINI 1 post hoc analysis of week 14 remitters. Journal of crohn's and colitis. Conference: 12th congress of the european crohn's and colitis organisation, ECCO 2017. Spain. 2017;11(Supplement 1):S42-S43	Not outcome of interest	
52	Strik A.;Van De Vrie W.;Bloemsaat-Minekus J.;Nurmohamed M.;Bossuyt P.;Bodelier A.;Rispens T.;Van Megen Y.;D'Haens G. Unchanged infliximab serum concentrations after switching from the originator infliximab to the biosimilar CT-P13 in patients with quiescent inflammatory bowel disease: Results from a prospective study (SECURE). Journal of Crohn's and Colitis. 2018;12:S477-S478	Not population of interest	
53	Strik A. S.;van de Vrie W.;Bloemsaat-Minekus J. P. J.;Nurmohamed M.;Bossuyt P. J. J.;Bodelier A.;Rispens T.;van Megen Y. J. B.;D'Haens G. R. Serum concentrations after switching from originator infliximab to the biosimilar CT-P13 in patients with quiescent inflammatory bowel disease (SECURE): an open-label, multicentre, phase 4 non-inferiority trial. The Lancet Gastroenterology and Hepatology. 2018;3(6):404-412	Not population of interest	
54	Suzuki Y.;Motoya S.;Hanai H.;Hibi T.;Nakamura S.;Lazar A.;Robinson A. M.;Skup M.;Mostafa N. M.;Huang B.;Thakkar R.;Watanabe M. Four-year maintenance treatment with adalimumab in Japanese patients with moderately to severely active ulcerative colitis. Journal of Gastroenterology. 2017;:43374	Not study type of interest	Observational study
55	Travis S. Does it all ADA up? Adalimumab for ulcerative colitis. Gut. 2011;60(6):741-742	Not study type of interest	Review article
56	Vande Casteele N.;Ferrante M.;Van Assche G.;Ballet V.;Compernolle G.;Van Steen K.;Simoens S.;Rutgeerts P.;Gils A.;Vermeire S. Trough concentrations of infliximab guide dosing for patients with inflammatory bowel disease. Gastroenterology. 2015;148(7):1320-9	Not population of interest	
57	Vande Casteele N.;Gils A.;Ballet V.;Compernolle G.;Peeters M.;Van Steen K.;Simoens S.;Ferrante M.;Van Assche G.;Vermeire S.;et al. Randomised controlled trial of drug level versus clinically based dosing of infliximab maintenance therapy in IBD: final results of the taxit study. United european gastroenterology journal.. 2013;1(1 SUPPL. 1):A1	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
58	Vermeire S.;Loftus E. V.;Khalid J. M.;Tudor D.;Akbari M.;Demuth D.;Peyrin-Biroulet L. Vedolizumab	Not study type of interest	Post hoc analysis (also

	treatment persistence up to 3 years: post hoc analysis in vedolizumab-naive patients from the GEMINI LTS study. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S30-S31		outcomes of interest already captured in the main publication for the RCT)
59	Vermeire S.;Loftus E. V.;Khalid J. M.;Tudor D.;Akbari M.;Demuth D.;Peyrin-Biroulet L. Vedolizumab treatment persistence up to 3 years: Post hoc analysis in vedolizumab-naïve patients from the GEMINI LTS study. Journal of Crohn's and Colitis. 2018;12:S30-S31	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
60	Volkers A. G.;Jansen J. M. Similar trial-efficacy of infliximab-biosimilar compared to infliximab-biological in patients with inflammatory bowel disease in remission-a randomized, controlled, double blind, phase 4 noninferiority trial. United european gastroenterology journal. Conference: 25th united european gastroenterology week, UEG 2017. Spain. 2017;5(5 Supplement 1):A307	Not population of interest	
61	Williams J. G.;Alam M. F.;Alrubaiy L.;Arnott I.;Clement C.;Cohen D.;Gordon J. N.;Hawthorne A. B.;Hilton M.;Hutchings H. A.;et al. Infliximab versus ciclosporin for steroid-resistant acute severe ulcerative colitis (CONSTRUCT): a mixed methods, open-label, pragmatic randomised trial. The lancet. Gastroenterology & hepatology. 2016;1(1):15-24	Not intervention of interest	
62	Williams J. G.;Alam M. F.;Alrubaiy L.;Clement C.;Cohen D.;Croft G. P.;Grey M.;Hutchings H. A.;Morgan J. M.;Rapport F.;Russell I. T.;Seagrove A. C.;Watkins A. Comparative clinical effectiveness of infliximab and ciclosporin for acute severe ulcerative colitis: Early results from the construct trial. United European Gastroenterology Journal. 2014;2(1):A28-A29	Not study type of interest	Observational study
63	Williams J. G.;Alam M. F.;Alrubaiy L.;Clement C.;Cohen D.;Grey M.;Hilton M.;Hutchings H. A.;Longo M.;Morgan J. M.;et al. Comparison Of iNfliximab and ciclosporin in STeroid Resistant Ulcerative Colitis: pragmatic randomised Trial and economic evaluation (CONSTRUCT). Health technology assessment (winchester, england). 2016;20(44):1-320	Not intervention of interest	
64	Yajnik V.;Khan N.;Dubinsky M.;Axler J.;Green A.;Abhyankar B.;Lasch K. Efficacy and safety of vedolizumab with advancing age in patients with ulcerative colitis: results from the GEMINI 1 study. Journal of crohn's and colitis.. 2015;9:S363-S364	Not outcome of interest	
65	Yajnik V.;Khan N.;Dubinsky M.;Axler J.;James A.;Abhyankar B.;Lasch K. Efficacy and Safety of Vedolizumab in Ulcerative Colitis and Crohn's Disease Patients Stratified by Age. Advances in therapy. 2017;34(2):542-559	Not outcome of interest	
66	Yajnik V.;Khan N.;Dubinsky M.;Axler J. L.;Green A.;Abhyankar B.;Lasch K. Efficacy and safety of vedolizumab with advancing age in patients with ulcerative colitis: results from the GEMINI 1 study. Gastroenterology.. 2015;148(4 SUPPL. 1):S278-S279	Not outcome of interest	
67	Reinisch W.;Colombel J. F.;Feagan B. G.;Han C.;Marano C.;Strauss R.;Gibson P.;Sandborn W. J.;Huyck S.;Cornillie F.;Rutgeerts P. Patient-reported outcomes can be used to monitor continuous clinical response in patients with moderately to severely active ulcerative colitis treated with golimumab: Results from the pursuit maintenance study. United European Gastroenterology Journal.	Not outcome of interest	

	2015;3(5):A616		
68	3 P.;Reinisch W.;Feagan B.;Sandborn W.;Tarabar D.;Habzda Z.;Weng H.;Yao R.;Zhang H.;Adedokun O.;Marano C.;Strauss R. Weight-based maintenance dosing of golimumab in patients with moderate to severe ulcerative colitis. United European Gastroenterology Journal. 2014;2(1):A529-A530	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
69	Halfvarson J.;Järnerot G. Treatment of choice for acute severe steroid-refractory ulcerative colitis is remicade. Inflammatory Bowel Diseases. 2009;15(1):143-145	Not study type of interest	Review article
70	Hanauer S.;Panaccione R.;Danese S.;Cheifetz A.;Reinisch W.;Higgins P. D. R.;Woodworth D. A.;Zhang H.;Friedman G. S.;Lawendy N.;Quirk D.;Nduaka C. I.;Su C. Tofacitinib Induction Therapy Reduces Symptoms Within 3 Days for Patients with Ulcerative Colitis. Clin Gastroenterol Hepatol. 2018;:	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
71	Hyams Jeffrey;Damaraju Lakshmi;Blank Marion;Johanns Jewel;Guzzo Cynthia;Winter Harland S. Induction and Maintenance Therapy With Infliximab for Children With Moderate to Severe Ulcerative Colitis. Clinical gastroenterology and hepatology. 2012;10(4):391-399	Not population of interest	
72	Hyams J. S.;Griffiths A. M.;Veereman G.;Turner D.;Chan D.;Adedokun O. J.;Padgett L.;Strauss R. A multi-center open-label study assessing pharmacokinetics, efficacy, and safety of subcutaneous golimumab in pediatric patients with moderately-severely active ulcerative colitis. Gastroenterology. 2016;150(4):S132	Not population of interest	
73	Järnerot G.;Hertervig E.;Friis-Liby I.;Blomquist L.;Karlén P.;Grännö C.;Vilien M.;Ström M.;Danielsson A.;Verbaan H.;et al. Infliximab as rescue therapy in severe to moderately severe ulcerative colitis: a randomized, placebo-controlled study. Gastroenterology. 2005;128(7):1805-1811	Not population of interest	
74	Joergensen K. K.;Olsen I. C.;Goll G. L.;Lorentzen M.;Bolstad N.;Berset I. P.;Haavardsholm E. A.;Lundin K. E.;Mork C.;Kvien T. K.;et al. Biosimilar infliximab (CT-P13) is not inferior to originator infliximab: explorative ibd subgroup-analyses in Crohn's disease and ulcerative colitis from the nor-switch trial. Gastroenterology. Conference: digestive disease week 2017, DDW 2017. United states. 2017;152(5 Supplement 1):S65-S66	Not outcome of interest	
75	Jorgensen K. K.;Goll G. L.;Sexton J.;Olsen I. C.;Bolstad N.;Lundin K. E.;Berset I. P.;Haavardsholm E. A.;Mork C.;Kvien T. K.;et al. Long-term efficacy and safety of biosimilar infliximab (CT-P13) after switching from originator infliximab: explorative subgroup analyses in IBD from the NOR-SWITCH EXTENSION trial. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S348-S349	Not outcome of interest	
76	Jorgensen K. K.;Olsen I. C.;Goll G. L.;Lorentzen M.;Bolstad N.;Berset I. P.;Haavardsholm E. A.;Lundin K. E.;Mork C.;Kvien T. K.;et al. Biosimilar infliximab (CT-P13) is not inferior to originator infliximab: explorative IBD subgroup-analyses in Crohn's disease and ulcerative colitis from the NOR-SWITCH trial. Journal of crohn's and colitis. Conference: 12th congress of the european crohn's and colitis organisation, ECCO 2017. Spain. 2017;11(Supplement 1):S62-S63	Not outcome of interest	

77	Jorgensen K. K.;Olsen I. C.;Goll G. L.;Lorentzen M.;Bolstad N.;Haavardsholm E. A.;Lundin K. E. A.;Mork C.;Jahnsen Jo;Kvien T. K. Switching from originator infliximab to biosimilar CT-P13 compared with maintained treatment with originator infliximab (NOR-SWITCH): a 52-week, randomised, double-blind, non-inferiority trial. Lancet. 2017;(no pagination):	Not population of interest	
78	Jørgensen K. K.;Olsen I. C.;Goll G. L.;Lorentzen M.;Bolstad N.;Haavardsholm E. A.;Lundin K. E. A.;Mørk C.;Jahnsen J.;Kvien T. K. Switching from originator infliximab to biosimilar CT-P13 compared with maintained treatment with originator infliximab (NOR-SWITCH): a 52-week, randomised, double-blind, non-inferiority trial. Lancet (london, england). 2017;389(10086):2304-2316	Not population of interest	
79	Kaasen Jørgensen K. K.;Goll G. L.;Sexton J.;Olsen I. C.;Bolstad N.;Lundin K. E. A.;Berset I. P.;Haavardsholm E. A.;Mørk C.;Kvien T. K.;Jahnsen J. Biosimilar infliximab (CT-P13) is not inferior to originator infliximab: Explorative subgroupanalyses in Crohn's disease and ulcerative colitis in the nor-switch extension trial. United European Gastroenterology Journal. 2017;5(8):1140-1141	Not outcome of interest	
80	Kim J. H.;Cheon J. H. Understanding the role of adalimumab in the treatment of moderately to severely active ulcerative colitis. Gut and Liver. 2016;10(2):162-163	Not study type of interest	Review article
81	Kobayashi T.;Suzuki Y.;Motoya S.;Hirai F.;Ogata H.;Ito H.;Sato N.;Ozaki K.;Watanabe M.;Hibi T. First trough level of infliximab at week 2 predicts future outcomes of induction therapy in ulcerative colitis: a post-hoc analysis of a multicenter prospective randomized controlled trial. American journal of gastroenterology.. 2015;110:S787	Not outcome of interest	
82	Kobayashi T.;Suzuki Y.;Motoya S.;Hirai F.;Ogata H.;Ito H.;Sato N.;Ozaki K.;Watanabe M.;Hibi T. Factors associated with the first trough level of infliximab at week 2 that predicts short-and long-term outcomes in ulcerative colitis. Gastroenterology. 2016;150(4):S412	Not outcome of interest	
83	Laharie D.;Bourreille A.;Branche J.;Allez M.;Bouhnik Y.;Filippi J.;Zerbib F.;Savoye G.;Nachury M.;Moreau J.;et al. Ciclosporin versus infliximab in patients with severe ulcerative colitis refractory to intravenous steroids: a parallel, open-label randomised controlled trial. Lancet (london, england). 2012;380(9857):1909-1915	Not intervention of interest	
84	Laharie D.;Bourreille A.;Branche J.;Allez M.;Bouhnik Y.;Filippi J.;Zerbib F.;Savoye G.;Vuitton L.;Moreau J.;et al. Long-term outcomes in a cohort of patientswith acute severe ulcerative colitis refractory tointravenous steroids treated with cyclosporine orinfliximab. Journal of crohn's and colitis.. 2015;9:S10-S11	Not intervention of interest	
85	Laharie D.;Bourreille A.;Branche J.;Allez M.;Bouhnik Y.;Filippi J.;Zerbib F.;Savoye G.;Vuitton L.;Moreau J.;et al. Long-term outcome of patients with steroidrefractory acute severe UC treated with ciclosporin or infliximab. Gut. (no pagination), 2017. 2017;Date of Publication: January 04:	Not study type of interest	Observational study
86	Laharie D.;Bourreille A.;Branche J.;Allez M.;Bouhnik Y.;Filippi J.;Zerbib F.;Savoye G.;Vuitton L.;Moreau J.;et al. Long-term outcome of patients with steroid-refractory acute severe UC treated with ciclosporin or infliximab. Gut. 2018;67(2):237-243	Not intervention of interest	

87	Lewis J.;Reinisch W.;Bressler B.;Parikh A.;Yang H.;Rosario M.;Røseth A.;Danese S.;Feagan B. G.;Sands B. E.;Ginsburg P.;Dassopoulos T.;Xu J.;Wyant T. Fecal calprotectin reductions in patients with mucosal healing during vedolizumab induction therapy in GEMINI 1. Gastroenterology. 2016;150(4):S989	Not outcome of interest	
88	Lichtenstein G. R.;Ciorba M. A.;Rogler G.;Quirk D.;Nduaka C. I.;Pedersen R. D.;Lawendy N.;Chan G.;Su C.;Panes J. Tofacitinib for the treatment of ulcerative colitis: analysis of malignancy rates from the OCTAVE clinical programme. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S48-S49	Not outcome of interest	
89	Lichtenstein G. R.;Diamond R. H.;Wagner C. L.;Fasanmade A. A.;Olson A. D.;Marano C. W.;Johanns J.;Lang Y.;Sandborn W. J. Clinical trial: benefits and risks of immunomodulators and maintenance infliximab for IBD-subgroup analyses across four randomized trials. Alimentary pharmacology & therapeutics. 2009;30(3):210-226	Not outcome of interest	
90	Lichtenstein G. R.;Loftus E. V.;Bloom S.;Lawendy N.;Friedman G.;Zhang H.;Wang W.;Thorpe A. J.;Nduaka C.;Su C. Tofacitinib, an oral Janus Kinase inhibitor, in the treatment of ulcerative colitis: open-label, long-term extension study. United european gastroenterology journal. Conference: 25th united european gastroenterology week, UEG 2017. Spain. 2017;5(5 Supplement 1):A39-A40	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
91	Loftus E. V.;Colombel J. F.;Feagan B.;Vermeire S.;Sandborn W.;Sands B.;Danese S.;D'Haens G.;Kaser A.;Panaccione R.;Rubin D.;Shafran I.;O'Byrne S.;Geransar P.;James A.;Kaviya A.;Khalid J. M. Long-term effectiveness and safety of vedolizumab in patients with ulcerative colitis: 5-year cumulative exposure of GEMINI 1 completers rolling into the GEMINI open-label extension study. Journal of Crohn's and Colitis. 2017;11:S182-S183	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
92	Loftus E. V.;Colombel J. F.;Feagan B. G.;Vermeire S.;Sandborn W. J.;Sands B. E.;Danese S.;D'Haens G. R.;Kaser A.;Panaccione R.;et al. Long-term efficacy of vedolizumab for ulcerative colitis. Journal of crohn's & colitis. 2017;11(4):400-411	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
93	Loftus E. V.;Colombel J. F.;Feagan B. G.;Vermeire S.;Sandborn W. J.;Sands B. E.;Danese S.;D'Haens G. R.;Kaser A.;Panaccione R.;Rubin D. T.;Shafran I.;O'Byrne S.;Geransar P.;James A.;Kaviya A.;Khalid J. M. Long-term effectiveness and safety of vedolizumab in patients with ulcerative colitis: 5-year cumulative exposure of GEMINI 1 completers rolling into the GEMINI open-label extension study. Gastroenterology. 2017;152(5):S602	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
94	Loftus E. V.;Colombel J. F.;Previtali A.;Smyth M. Response and remission rates with up to 3 years of vedolizumab treatment in patients with ulcerative colitis. Journal of crohn's and colitis. Conference: 11th congress of the european crohn's and colitis organisation, ECCO 2016. Netherlands. Conference start: 20160316. Conference end: 20160319. 2016;10:S58-S59	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
95	Loftus E. V.;Colombel J. F.;Previtali A.;Smyth M. D. Response and remission rates with up to 3 years of vedolizumab treatment in patients with ulcerative colitis. Gastroenterology.	Not study type of interest	Post hoc analysis (also outcomes of interest already

	2016;150(4):S805		captured in the main publication for the RCT)
96	Loftus E. V.;Sloan S.;Ramachandran P.;Yang Z.;Guo C. Y.;Gasink C. Comparison of rates of active tuberculosis infection in the phase 2 and 3 clinical trial programs for anti-IL12/23 and anti-TNFs. Gastroenterology. Conference: digestive disease week 2017, DDW 2017. United states. 2017;152(5 Supplement 1):S596	Not population of interest	
97	Loftus Jr E. V.;Siegel C. A.;Panaccione R.;Sandborn W. J.;Smyth M.;Green A.;Xu J.;Abhyankar B. Corticosteroid dose reduction in ulcerative colitis patients treated with vedolizumab. United European Gastroenterology Journal. 2015;3(5):A255-A256	Not outcome of interest	
98	Loftus Jr E. V.;Siegel C. A.;Panaccione R.;Sandborn W. J.;Smyth M. D. L.;Green A.;Xu J.;Abhyankar B. Corticosteroid dose reduction in ulcerative colitis patients treated with vedolizumab during the GEMINI 1 trial. American journal of gastroenterology.. 2015;110:S790	Not outcome of interest	
99	Michetti P.;Braegger F.;Kempf C.;Allez M. Efficacy of vedolizumab (VDZ) by disease extension in ulcerative colitis. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S402-S403	Not outcome of interest	
100	Motoya S.;Watanabe M.;Kim H. J.;Kim Y. H.;Han D. S.;Yuasa H.;Tabira J.;Isogawa N.;Arai S.;Kawaguchi I.;Hibi T. Tofacitinib induction and maintenance therapy in East Asian patients with active ulcerative colitis: subgroup analyses from three phase 3 multinational studies. Intest Res. 2018;16(2):233-245	Not outcome of interest	
101	Mshimesh B. A. R. Efficacy and safety of adalimumab versus infliximab in patients suffered from moderate to severe active ulcerative colitis. Asian journal of pharmaceutical and clinical research. 2017;10(3):300-307	Not study type of interest	Observational study
102	Nct A Study of Adalimumab in Japanese Subjects With Moderately to Severely Active Ulcerative Colitis. Https://clinicaltrials.gov/show/nct00853099. 2009;:	Not outcome of interest	
103	Nct Conventional Step-Up Versus Infliximab Monotherapy in Patients With Ulcerative Colitis (P05553). Https://clinicaltrials.gov/show/nct00984568. 2009;:	Not intervention of interest	
104	Nct A Study to Evaluate the Effectiveness and Safety of Infliximab in Chinese Patients With Active Ulcerative Colitis. Https://clinicaltrials.gov/show/nct01551290. 2012;:	Not outcome of interest	
105	Nct A Safety and Effectiveness Study of Golimumab in Japanese Patients With Moderately to Severely Active Ulcerative Colitis. Https://clinicaltrials.gov/show/nct01863771. 2013;:	Not outcome of interest	
106	Nct Phase III Study of MLN0002 (300 mg) in the Treatment of Ulcerative Colitis. Https://clinicaltrials.gov/show/nct02039505. 2014;:	Not outcome of interest	
107	Nct Efficacy and Safety of Infliximab-biosimilar (Inflectra) Compared to Infliximab-innovator (Remicade) in Patients With Inflammatory Bowel Disease in Remission: the SIMILAR Trial.	Not outcome of interest	

	Https://clinicaltrials.gov/show/nct02452151. 2015;:		
108	Nct Efficacy and Safety of Vedolizumab Subcutaneously (SC) as Maintenance Therapy in Ulcerative Colitis. Https://clinicaltrials.gov/show/nct02611830. 2015;:	Not outcome of interest	
109	Nct An Efficacy and Safety Study of Vedolizumab Intravenous (IV) Compared to Adalimumab Subcutaneous (SC) in Participants With Ulcerative Colitis. Https://clinicaltrials.gov/show/nct02497469. 2015;:	Not outcome of interest	
110	Nct A Study to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Participants With Moderately to Severely Active Ulcerative Colitis. Https://clinicaltrials.gov/show/nct02407236. 2015;:	Not outcome of interest	
111	Nct Vedolizumab Subcutaneous (SC) Versus Intravenous (IV) in Ulcerative Colitis or Crohn's Disease. Https://clinicaltrials.gov/show/nct02913508. 2016;:	Not outcome of interest	
112	Nct Optimising Infliximab Induction Therapy for Acute Severe Ulcerative Colitis. Https://clinicaltrials.gov/show/nct02770040. 2016;:	Not outcome of interest	
113	Nct Efficacy and Safety of Vedolizumab Intravenous (IV) in the Treatment of Primary Sclerosing Cholangitis in Subjects With Underlying Inflammatory Bowel Disease. Https://clinicaltrials.gov/show/nct03035058. 2017;:	Not outcome of interest	
114	Nct A Study to Evaluate the Efficacy, Safety, and Pharmacokinetics of UTTR1147A Compared With Placebo and With Vedolizumab in Participants With Moderate to Severe Ulcerative Colitis (UC). Https://clinicaltrials.gov/show/nct03558152. 2018;:	Not outcome of interest	
115	Ng S. C.;Palo W.;Blake A.;Rana-Khan Q.;Bhayat F. Vedolizumab clinical and post-marketing safety experience of opportunistic infections. Journal of crohn's and colitis. Conference: 12th congress of the european crohn's and colitis organisation, ECCO 2017. Spain. 2017;11(Supplement 1):S41	Not outcome of interest	
116	Ng S. C.;Palo W.;Blake A.;Rana-Khan Q.;Bhayat F. Vedolizumab clinical and post-marketing safety experience of opportunistic infections. Gastroenterology. 2017;152(5):S575-S576	Not outcome of interest	
117	Ochsenkühn T.;Sackmann M.;Göke B. Infliximab for acute, not steroid-refractory ulcerative colitis: a randomized pilot study. European journal of gastroenterology & hepatology. 2004;16(11):1167-1171	Not population of interest	
118	Osterman M. T.;Roblin X.;Glover S. C.;Navaneethan U.;Popa M. A.;Wyant T.;James A.;Lasch K.;Rosario M. Association of vedolizumab drug concentrations at or before week 6 with remission at week 14 in moderately to severely active ulcerative colitis patients from gemini 1. Gastroenterology. 2016;150(4):S105	Not outcome of interest	
119	Panaccione R.;Ghosh S.;Middleton S.;Márquez J. R.;Flint L.;Van Hoogstraten H. J. F.;Chen A. C.;Zheng H.;Danese S.;Rutgeerts P. Improvement in patient quality of life during treatment with infliximab, azathioprine, or combination infliximab+azathioprine for moderate to severe ulcerative colitis. United European Gastroenterology Journal. 2013;1(1):A223	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)

120	Panaccione R.;Ghosh S.;Middleton S.;Marquez J. R.;Scott B. B.;Flint L.;van Hoogstraten H. J.;Chen A. C.;Zheng H.;Danese S.;et al. Combination therapy with infliximab and azathioprine is superior to monotherapy with either agent in ulcerative colitis. Gastroenterology. 2014;146(2):392-400	Not intervention of interest	
121	Panes J.;Bressler B.;Colombel J. F.;Lawendy N.;Maller E.;Zhang H.;Woodworth D. A.;Chan G.;Salese L.;Su C. Efficacy and safety of tofacitinib retreatment for ulcerative colitis after treatment interruption: results from the OCTAVE clinical trials. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S366-S367	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
122	Panes J.;Rubin D. T.;Vermeire S.;Lindsay J. O.;Sands B. E.;Su C.;Friedman G.;Zhang H.;Kayhan C.;Manuchehri A.;et al. Maintenance of quality of life improvement in a phase 3 study of tofacitinib for patients with moderately to severely active ulcerative colitis. Journal of crohn's and colitis. Conference: 12th congress of the european crohn's and colitis organisation, ECCO 2017. Spain. 2017;11(Supplement 1):S315-S317	Not outcome of interest	
123	Panes J.;Su C.;Bushmakin A.;Cappelleri J.;Healey P. Direct and indirect effects of tofacitinib on treatment satisfaction in patients with ulcerative colitis. Journal of crohn's and colitis.. 2015;9:S380-S381	Not outcome of interest	
124	Panes J.;Su C.;Bushmakin A. G.;Cappelleri J. C.;Healey P. J. Direct and indirect effects of tofacitinib on treatment satisfaction in patients with ulcerative colitis. Gastroenterology.. 2015;148(4 SUPPL. 1):S266	Not outcome of interest	
125	Panés J.;Su C.;Bushmakin A. G.;Cappelleri J. C.;Mamolo C.;Healey P. Randomized trial of tofacitinib in active ulcerative colitis: analysis of efficacy based on patient-reported outcomes. BMC gastroenterology. 2015;15:14	Duplicates	
126	Panés J.;Su C.;Marren A.;Yu D.;Woodworth D. A.;Zhang H.;Healey P. J. Improvement in patient-reported outcomes in two phase 3 induction studies of tofacitinib in patients with moderately to severely active ulcerative colitis. Gastroenterology. 2016;150(4):S1003	Not outcome of interest	
127	IBD: Induction and maintenance therapy for ulcerative colitis-vedolizumab more effective than placebo. Nature Reviews Gastroenterology and Hepatology. 2013;:	Not study type of interest	Review article
128	Vedolizumab (Entyvio) for inflammatory bowel disease. Medical Letter on Drugs and Therapeutics. 2014;56(1451):86-88	Not outcome of interest	
129	Golimumab (Simponi) for ulcerative colitis. Med Lett Drugs Ther. 2014;56(1439):43276	Not study type of interest	Review article
130	Randomised clinical study: discrepancies between patient-reported outcomes and endoscopic appearance in moderate to severe ulcerative colitis. Alimentary pharmacology and therapeutics. 42 (9) (pp 1082-1092), 2015. Date of publication: 01 nov 2015.. 2015;:	Not outcome of interest	
131	Abstracts of the 10th Congress of ECCO. Journal of crohn's & colitis. 2015;Conference: 10th	Not study type of interest	Review article

	Congress of the European Crohn's and Colitis Organisation, ECCO 2015 Barcelona Spain. Conference Start: 20150218 Conference End: 20150221. Conference Publication: (var.pagings). 9:		
132	A SPECIAL MEETING REVIEW EDITION: CCFA/Advances in Inflammatory Bowel Diseases 2015: Highlights in Ulcerative Colitis and Crohn's Disease: A Review of Selected Presentations From the CCFA/Advances in Inflammatory Bowel Diseases 2015 Clinical and Research Conference * December 10-12, 2015 * Orlando, FloridaSpecial Reporting on:* New and Future Adhesion Molecule-Based Therapies in IBD* Efficacy and Safety of Vedolizumab for Inflammatory Bowel Disease in Clinical Practice* A Multicenter, Double-Blind, Placebo-Controlled Phase 3 Study of Ustekinumab, a Human IL-12/23p40 Monoclonal Antibody, in Moderate-Severe Crohn's Disease Refractory to Anti-TNFalpha: UNITI-1* Intravenous Iron Sucrose for Treatment of Iron Deficiency Anemia in Pediatric Inflammatory Bowel Disease* Does Vedolizumab Affect Postoperative Outcomes in Patients Undergoing Abdominal Operations for Inflammatory Bowel Disease?PLUS Meeting Abstract Summaries With Expert Commentary by: Gary R. Lichtenstein, MDProfessor of MedicineDirector, Center for Inflammatory Bowel DiseaseUniversity of Pennsylvania Health SystemHospital of the University of PennsylvaniaPhiladelphia, Pennsylvania. Gastroenterol Hepatol (N Y). 2016;12(2 Suppl 1):43831	Not study type of interest	Review article
133	Sustained remission with vedolizumab in patients with moderately to severely active ulcerative colitis: a GEMINI 1 post hoc analysis of week 14 remitters. Gastroenterology. 2017;Conference: Digestive Disease Week 2017, DDW 2017. United States. 152(5 Supplement 1):S604	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
134	Akobeng A. K. Infliximab for induction and maintenance therapy for ulcerative colitis. Journal of Pediatric Gastroenterology and Nutrition. 2006;42(5):589-590	Not study type of interest	Review article
135	Ananthakrishnan A. N. Infliximab or ciclosporin for acute severe ulcerative colitis?. The Lancet Gastroenterology and Hepatology. 2016;1(1):43161	Not study type of interest	Review article
136	Anonymous Late-breaking Abstracts 25th UEG Week 2017. United european gastroenterology journal. Conference: 25th united european gastroenterology week, UEG 2017. Spain. 2017;5(8) (no pagination):	Not study type of interest	Review article
137	Anonymous 13th Congress of ECCO. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1) (no pagination):	Not study type of interest	Review article
138	Arijs I.;De Hertogh G.;Lemmens B.;Van Lommel L.;de Bruyn M.;Vanhove W.;Cleynen I.;Machiels K.;Ferrante M.;Schuit F.;et al. Effect of vedolizumab (anti-?4?7-integrin) therapy on histological healing and mucosal gene expression in patients with UC. Gut. 2018;67(1):43-52	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
139	Armuzzi A.;De Pascalis B.;Lupascu A.;Fedeli P.;Leo D.;Mentella M. C.;Vincenti F.;Melina D.;Gasbarrini G.;Pola P.;et al. Infliximab in the treatment of steroid-dependent ulcerative colitis.	Not study type of interest	Post hoc analysis (also outcomes of interest already

	European review for medical and pharmacological sciences. 2004;8(5):231-233		captured in the main publication for the RCT)
140	Balzola F.;Cullen G.;Ho G. T.;Russell R. K. Subcutaneous golimumab induces clinical response and remission in patients with moderate to severe ulcerative colitis. Inflammatory Bowel Disease Monitor. 2013;14(1):24	Duplicates	
141	Bonovas S.;Peyrin-Biroulet L.;Danese S. Infliximab biosimilar CT-P13 for inflammatory bowel disease. The Lancet Gastroenterology and Hepatology. 2018;3(6):373-375	Not intervention of interest	
142	Bradley C. A. IBD: tofacitinib effective in ulcerative colitis. Nature reviews gastroenterology and hepatology. 2017;14(7):388	Not study type of interest	Observational study
143	Casteele N. V.;Jairath V.;Levesque B.;Feagan B. G.;Sandborn W. J. Severely active ulcerative colitis is associated with high baseline infliximab clearance, reduced serum half-life and worse endoscopic outcomes. Gastroenterology. Conference: digestive disease week 2017, DDW 2017. United states. 2017;152(5 Supplement 1):S379-S380	Not outcome of interest	
144	Colombel J. F.;Han C.;Reinisch W.;Feagan B.;Marano C.;Strauss R.;Johanns J.;Zhang H.;Gibson P.;Collins J.;et al. Predictive value of patient-reported outcomes to mucosal healing in patients with moderately to severely active ulcerative colitis. Value in health.. 2014;17(3):A40	Not study type of interest	Observational study
145	Colombel J. F.;Loftus E. V.;Siegel C. A.;Lewis J.;Smyth M. D.;Sankoh S.;Abhyankar B. Efficacy of vedolizumab with concomitant corticosteroid or immunomodulator use in patients with ulcerative colitis from GEMINI 1. Gastroenterology. 2015;148(4):S277-S278	Not outcome of interest	
146	Colombel J. F.;Loftus Jr E. V.;Siegel C. A.;Lewis J. D.;Smyth M.;Sankoh S.;Abhyankar B. Efficacy of vedolizumab with concomitant corticosteroid or immunomodulator use in patients with ulcerative colitis from GEMINI 1. Journal of crohn's and colitis.. 2015;9:S296-S297	Not outcome of interest	
147	Colombel J. F.;Panaccione R.;Ghosh S.;Sandborn W. J.;Rutgeerts P.;Hanauer S.;Van Assche G.;Reinisch W.;Peyrin-Biroulet L.;Robinson A. M.;Lau W.;Huang B.;Pappalardo B.;Read H. A. Long-term safety of adalimumab in clinical trials in adult patients with Crohn's disease or ulcerative colitis. United European Gastroenterology Journal. 2015;3(5):A428-A429	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
148	Colombel J. F.;Rutgeerts P.;Reinisch W.;Esser D.;Wang Y.;Lang Y.;Marano C. W.;Strauss R.;Oddens B. J.;Feagan B. G.;et al. Early mucosal healing with infliximab is associated with improved long-term clinical outcomes in ulcerative colitis. Gastroenterology. 2011;141(4):1194-1201	Not outcome of interest	
149	Colombel J. F.;Sandborn W.;Ghosh S.;Wolf D. C.;Panaccione R.;Feagan B. G.;Reinisch W.;Robinson A.;Lazar A.;Kron M.;Huang B.;Thakkar R. Adalimumab maintains remission for up to 4 years in patients with ulcerative colitis. Gastroenterology. 2014;146(5):S589-S590	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
150	Colombel J. F.;Sandborn W. J.;Ghosh S.;Wolf D. C.;Panaccione R.;Feagan B.;Reinisch	Duplicates	

	W.;Robinson A. M.;Lazar A.;Kron M.;et al. Four-year maintenance treatment with adalimumab in patients with moderately to severely active ulcerative colitis: data from ULTRA 1, 2, and 3. American journal of gastroenterology. 2014;109(11):1771-1780		
151	Colombel J. F.;Sands B. E.;Rutgeerts P.;Sandborn W.;Danese S.;D'Haens G.;Panaccione R.;Loftus E. V.;Sankoh S.;Fox I.;Parikh A.;Milch C.;Abhyankar B.;Feagan B. G. The safety of vedolizumab for ulcerative colitis and Crohn's disease. Gut. 2017;66(5):839-851	Not study type of interest	Literature review
152	Colombel Md J. F.;Sandborn Md W. J.;Ghosh Md S.;Wolf Md D. C.;Panaccione Md R.;Feagan Md B.;Reinisch M. D. PhD W.;Robinson Pharm D. A. M.;Lazar Md A.;Kron PhD M.;Huang PhD B.;Skup PhD M.;Thakkar Md R. B. Four-Year Maintenance Treatment With Adalimumab in Patients with Moderately to Severely Active Ulcerative Colitis: Data from ULTRA 1, 2, and 3. American Journal of Gastroenterology. 2014;:	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
153	Cross R. K.;Barr E.;Rajan D. Infliximab is not associated with excess weight gain in women: An analysis of clinical trial data from accent I, accent II, act 1, and sonic. Gastroenterology. 2017;152(5):S109	Not outcome of interest	
154	D Laharie;A Bourreille;J Branche Ciclosporine versus infliximab in patients with severe ulcerative colitis refractory to intravenous steroids: a parallel, open-label randomised controlled trial. Lancet. 2012;380:1909-1915	Not population of interest	
155	D'Haens G.;Colombel J. F.;Dubinsky M.;Abhyankar B.;James A.;Lasch K. The efficacy of vedolizumab by disease activity and prior tumour necrosis factor-alpha antagonist failure in patients with ulcerative colitis or Crohn's disease: Post-hoc analyses from the GEMINI 1 and GEMINI 2 studies. Journal of Crohn's and Colitis. 2016;10:S58	Not outcome of interest	
156	D'Haens G. R.;Colombel J. F.;Dubinsky M.;Abhyankar B.;James A.;Lasch K. The efficacy of vedolizumab by disease activity and prior tumor necrosis factor a antagonist failure in patients with ulcerative colitis or Crohn's disease: Post HOC analyses from the gemini 1 and gemini 2 studies. Gastroenterology. 2016;150(4):S804-S805	Not outcome of interest	
157	Feagan B.;Kaser A.;Smyth M.;Panaccione R.;Sankoh S.;Abhyankar B. Long-term efficacy of vedolizumab therapy for ulcerative colitis. United European Gastroenterology Journal. 2014;2(1):A66-A67	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
158	Feagan B.;Lasch K.;Khalid J. M.;Cao C.;Wojtowicz A. M.;James A.;Colombel J. Vedolizumab demonstrates early symptomatic improvement in ulcerative colitis: a Gemini 1 post hoc analysis. United european gastroenterology journal. Conference: 25th united european gastroenterology week, UEG 2017. Spain. 2017;5(5 Supplement 1):A40-A41	Not outcome of interest	
159	Feagan B.;Sandborn W. J.;Reinisch W.;Ghosh S.;Robinson A. M.;Lazar A.;Zhou Q.;Skup M.;Thakkar R. B. Predictors of hospitalization in patients with moderately to severely active ulcerative colitis from ultra 1 and ultra 2. United European Gastroenterology Journal.	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main

	2014;2(1):A220-A221		publication for the RCT)
160	Feagan B.;Siegel C. A.;Melmed G.;Isaacs K.;Lasch K.;Rosario M.;Green A.;Abhyankar B. Efficacy of vedolizumab maintenance therapy with and without continued immunosuppressant use in GEMINI 1 and GEMINI 2. American journal of gastroenterology.. 2015;110:S791	Not outcome of interest	
161	Feagan B. G.;Dubinsky M. C.;Lukas M.;Quirk D.;Nduaka C. I.;Maller E.;Lawendy N.;Kayhan C.;Wang W.;Chan G.;et al. Efficacy and safety of an additional 8 weeks of tofacitinib induction therapy: results of the OCTAVE open study for tofacitinib 8-week induction non-responders. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S50	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
162	Feagan B. G.;Greenberg G. R.;Wild G.;Fedorak R. N.;Paré P.;McDonald J. W.;Dubé R.;Cohen A.;Steinhart A. H.;Landau S.;et al. Treatment of ulcerative colitis with a humanized antibody to the alpha4beta7 integrin. New England journal of medicine. 2005;352(24):2499-2507	Not population of interest	
163	Feagan B. G.;Greenberg G. R.;Wild G.;Fedorak R. N.;Paré P.;McDonald J. W. D.;Dubé R.;Cohen A.;Steinhart A. H.;Landau S.;Aguzzi R. A.;Fox I. H.;Vandervoort M. K. Treatment of ulcerative colitis with a humanized antibody to the ?4?7 integrin. New England Journal of Medicine. 2005;352(24):2499-2507	Duplicates	
164	Feagan B. G.;Ha C. Y.;Taub P. R.;Quirk D.;Nduaka C. I.;Salese L.;Chan G.;Friedman G. S.;Wang W.;Su C.;et al. Correlation of lipid levels with reduction in inflammation in patients with ulcerative colitis: data from the tofacitinib OCTAVE clinical trials. Journal of crohn's and colitis. Conference: 13th congress of european crohn's and colitis organisation, ECCO 2018. Austria. 2018;12(Supplement 1):S453-S454	Not outcome of interest	
165	Feagan B. G.;Lasch K.;Lissoos T.;Cao C.;Wojtowicz A. M.;Khalid J. M.;Colombel J. F. Rapid response to vedolizumab therapy in biologic-naive patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol. 2018;:	Not outcome of interest	
166	Feagan B. G.;Sandborn W. J.;Lazar A.;Thakkar R. B.;Huang B.;Reilly N.;Chen N.;Yang M.;Skup M.;Mulani P.;et al. Adalimumab therapy is associated with reduced risk of hospitalization in patients with ulcerative colitis. Gastroenterology. 2014;146(1):110-118	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
167	Feagan B. G.;Siegel C. A.;Melmed G. Y.;Isaacs K.;Lasch K.;Rosario M.;Green A.;Abhyankar B. Efficacy of vedolizumab with and without continued immunosuppressant use in GEMINI 1 and GEMINI 2. United European Gastroenterology Journal. 2015;3(5):A17-A18	Not outcome of interest	
168	Gao J.;Jiang X. L. Low-dose infliximab for corticosteroid-refractory ulcerative colitis: impact of number of infusions on efficacy and safety. World chinese journal of digestology. 2013;21(15):1453-1457	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)

169	Ghosh S.;Sandborn W.;Panés J.;D'Haens G.;Reilly N.;Lazar A.;Huang B.;Robinson A. M.;Thakkar R. Different methodologies for determining mayo score in clinical studies can influence disease activity assessments in ulcerative colitis. United European Gastroenterology Journal. 2014;2(1):A532-A533	Not study type of interest	Observational study
170	Gibson P. R.;Feagan B. G.;Sandborn W. J.;Marano C.;Strauss R.;Johanns J.;Padgett L.;Collins J.;Tarabar D.;Hebzda Z.;Rutgeerts P.;Reinisch W. Maintenance of Efficacy and Continuing Safety of Golimumab for Active Ulcerative Colitis: PURSUIT-SC Maintenance Study Extension Through 1 Year. Clin Transl Gastroenterol. 2016;7:e168	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
171	Jean-Frederic C.;Walter R.;Osterman Mark T.;Torpe Andrew J.;Nduaka Chudy I.;Haiying Z.;Nervin L.;Friedman Gary S.;Chinyu S. Maintenance of remission with tofacitinib in patients with ulcerative colitis: subpopulation analysis from an open-label, long-term extension study. American journal of gastroenterology. Conference: 2018 annual scientific meeting of the american college of gastroenterology, ACG 2018. United states. 2018;113(Supplement 1):S4	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
172	Lichtenstein G. R.;Loftus E. V.;Bloom S.;Lawendy N.;Friedman G. S.;Zhang H.;Wang W.;Thorpe A. J.;Nduaka C. I.;Su C. Tofacitinib, an oral Janus Kinase inhibitor, in the treatment of ulcerative colitis: open-label, long-term extension study. American journal of gastroenterology. Conference: 82nd annual scientific meeting of the american college of gastroenterology. United states. 2017;112(Supplement 1):S395	Not study type of interest	Post hoc analysis (also outcomes of interest already captured in the main publication for the RCT)
173	William S.;Jean-Frederic C.;Remo P.;Parambir D.;Maria R.;Charlie C.;Morris B.;Karen L. Relationship between vedolizumab concentrations and deep remission in patients with moderately-to-severely active ulcerative colitis: a GEMINI 1 post hoc analysis. American journal of gastroenterology. Conference: 2018 annual scientific meeting of the american college of gastroenterology, ACG 2018. United states. 2018;113(Supplement 1):S2	Not outcome of interest	
174	Brian F.;Karen L.;Mona K. J.;Charlie C.;Abigail W.;Alexandra J.;Jean-Frederic C. Vedolizumab demonstrates early symptomatic improvement in ulcerative colitis: a GEMINI 1 post hoc analysis. American journal of gastroenterology. Conference: 2018 annual scientific meeting of the american college of gastroenterology, ACG 2018. United states. 2018;113(Supplement 1):S24	Not outcome of interest	
175	Feagan B. G.;Lasch K.;Khalid J. M.;Cao C.;Wojtowicz A. M.;James A.;Colombel J. F. Vedolizumab demonstrates early symptomatic improvement in ulcerative colitis: a gemini 1 post hoc analysis. American journal of gastroenterology. Conference: 82nd annual scientific meeting of the american college of gastroenterology. United states. 2017;112(Supplement 1):S371-S373	Not outcome of interest	
176	Feagan B. G.;Vermeire S.;Sandborn W. J.;Reinisch W.;Panes J.;Tarabar D.;Su C.;Niezychowski W.;Zhang H.;Friedman G. S.;et al. Tofacitinib for maintenance therapy in patients with active ulcerative colitis in the phase 3 octave sustain trial: results by local and central endoscopic assessments. American journal of gastroenterology. Conference: 82nd annual scientific meeting of the american college of gastroenterology. United states. 2017;112(Supplement 1):S329	Not outcome of interest	

177	Kobayashi K.;Suzuki Y.;Watanabe K.;Oda K.;Mukae M.;Yamada A., et al. A Phase 1, Multiple-Dose Study of Vedolizumab in Japanese Patients With Ulcerative Colitis. Journal of Clinical Pharmacology. 2019;2(271-279):	Not outcome of interest	
178	Feagan B. G.;Lasch K.;Lissoos T.;Cao C.;Wojtowicz A. M.;Khalid J. M., et al. Rapid Response to Vedolizumab Therapy in Biologic-Naive Patients With Inflammatory Bowel Diseases. Clinical Gastroenterology and Hepatology. 2019;1(130-138.e7):	Not outcome of interest	
179	Hanauer S.;Panaccione R.;Danese S.;Cheifetz A.;Reinisch W.;Higgins P. D. R., et al. Tofacitinib Induction Therapy Reduces Symptoms Within 3 Days for Patients With Ulcerative Colitis. Clinical Gastroenterology and Hepatology. 2019;1(139-147):	Not outcome of interest	
180	Osterman M. T.;Rosario M.;Lasch K.;Barocas M.;Wilbur J. D.;Dirks N. L., et al. Vedolizumab exposure levels and clinical outcomes in ulcerative colitis: determining the potential for dose optimisation. Alimentary Pharmacology and Therapeutics. 2019;4(408-418):	Not outcome of interest	
181	Reinisch W.;Colombel J. F.;Gibson P. R.;Rutgeerts P.;Sandborn W. J.;Tarabar D., et al. Continuous clinical response is associated with a change of disease course in patients with moderate to severe ulcerative colitis treated with golimumab. Inflammatory Bowel Diseases. 2019;1(163-171):	Not outcome of interest	
182	Sandborn W. J.;Colombel J. F.;Panaccione R.;Dulai P. S.;Rosario M.;Cao C., et al. Deep Remission with Vedolizumab in Patients with Moderately to Severely Active Ulcerative Colitis: A GEMINI 1 post hoc Analysis. Journal of Crohn's and Colitis. 2019;2(172-181):	Not outcome of interest	

A2.2 The Methods of NMA

A2.2.1. The selection criteria for inclusion of clinical studies in NMA

Selection criteria

Based on the results of an SLR of randomised controlled trials, studies were included in the NMA if they met the following inclusion criteria:

- **Outcomes** - reported one of the following efficacy outcomes:
  - clinical response,
  - clinical remission,
  - mucosal healing
- **Timepoints** - reported the efficacy outcome for one of the timepoints of assessment:
  - end of induction: 6-8 weeks
  - end of maintenance: 44-54 weeks of maintenance treatment (corresponding to approximately a yearlong treatment regimen: induction followed by maintenance)
- **Comparators** – included any of the following comparators (with doses and regimens corresponding to the EMA licences):
  - adalimumab,
  - infliximab,
  - golimumab,
  - tofacitinib,
  - vedolizumab
- **Population** - included subjects with moderate to severe UC who have either:
  - Not failed on a previous biologic therapy (non-biologic failure), or
  - Failed on a previous biologic therapy (biologic failure)

A2.2.2. Outputs generated by WinBUGS

Outputs generated by WinBUGS are posterior distributions for each parameter of interest. Summary statistics of these distributions are provided as results in this report: median odds ratio (OR), 95% credible interval [81] and probabilities for the treatment of interest to perform better than the comparator. An interpretation of each of these statistics is provided below.

- Odds ratios (OR): Relative treatment effects for ustekinumab versus each comparator (OR > 1 suggests ustekinumab performs better for efficacy endpoints, OR < 1 suggests ustekinumab performs better for safety endpoints).
- 95% credible intervals: Given the data and model specified, there is a 95% chance that the true value lies between the interval.
- Bayesian pairwise probability: Probability for ustekinumab to perform better than the comparators to achieve the endpoints. A study by Cope et al, 2013 [84] suggested that a probability ≥ 0.85 indicates a treatment is likely to be more effective than the comparator. For efficacy endpoints reflecting a positive outcome, it means that a high probability reflects a

better performance of ustekinumab compared to the other treatments. Whereas for safety endpoints reflecting negative outcomes, lower probabilities of ustekinumab versus comparators are in favour of ustekinumab.

A2.2.3. Treat-trough trial design

Figure A1. Treat-through trial design schematic

- Active and placebo arms at induction: The patients randomised to induction phase continue with the same treatment to the maintenance phase.
- This design is conventional and allows for a straightforward interpretation of the effectiveness of a continued 1-year regimen versus placebo.

A2.2.4. Response-based re-randomised trial design

Although the primary analyses of maintenance data may be reported based on the induction responders who enter the maintenance phase, the studies still capture maintenance outcomes for both induction responders and non-responders for the active arms at least. One-year outcomes are captured from these trials and results from these trials can be re-calculated to correspond closely to one-year outcomes from treat-through trial arms.

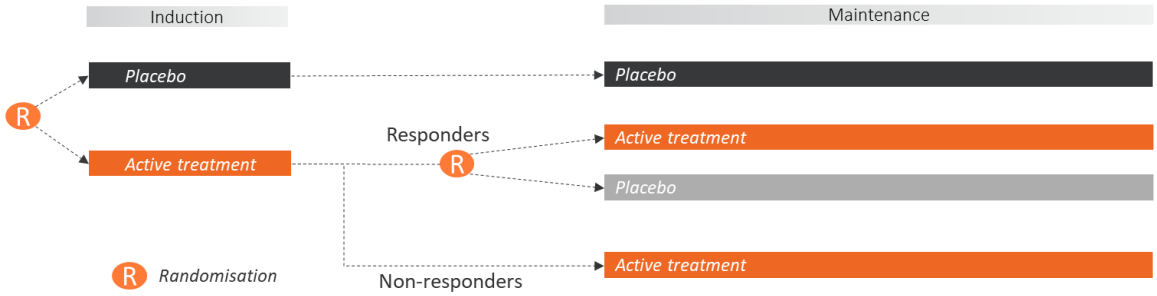


Figure A2. Response-based re-randomised trial design schematic

- **Active arm at induction**: Responders to active treatment during the induction phase are re-randomised to the treatment or placebo arm for the maintenance phase; non-responders are treated-through for the maintenance phase up to 1 year.
- **Placebo arm at induction**: Patients either remain on placebo from induction to maintenance regardless of response (as depicted above for GEMINI I) or continued to receive placebo maintenance treatment based on response at the end of induction (PURSUIT and UNIFI).

- In the OCTAVE, PURSUIT and UNIFI trial designs, exposure to long-term ineffective placebo is minimised as only responders continue in the trial. Non-responders cross over to active treatment or are included in an open-label extension phase of the study.
- This design complicates the interpretation of continued one-year treatment regimens, as typically the efficacy of a regimen over the course of a full year is not directly reported (though these can be re-calculated as further discussed)

A2.2.5. Carry-over effect

In the UNIFI trial [60,61], there is evidence of a carry-over effect of induction therapy with ustekinumab affecting maintenance outcomes for patients who receive placebo. Evidence of carry-over effects are also found in the trials for other biologic treatments for UC however the magnitude of the carry-over effects differ across studies. This is illustrated in Table A3 showing the evolution of the partial mayo score in the maintenance phase of UNIFI, with partial mayo scores in patients on placebo post-active ustekinumab induction therapy being notably lower (i.e. showing a better response) compared to the partial mayo scores for patients on placebo post-active golimumab induction therapy in PURSUIT. A similar difference is observed in Crohn's Disease between ustekinumab and adalimumab in terms of clinical remission.

This is further illustrated by substantial differences ($p < 0.001$) in response rates at the end of maintenance for the re-randomised placebo arms of UNIFI, PURSUIT, OCTAVE and GEMINI I, where patients received active induction therapy. The response rates across the placebo arms ranged between 26.6% to 50.6% in the non-biologic failure group and 15.8% to 38.6% in the biologic failure. For both patient groups the highest response rate was observed in the UNIFI trial for patients who received ustekinumab induction therapy.

For further details on the carry-over effect differences, see the MAH submission file [46].

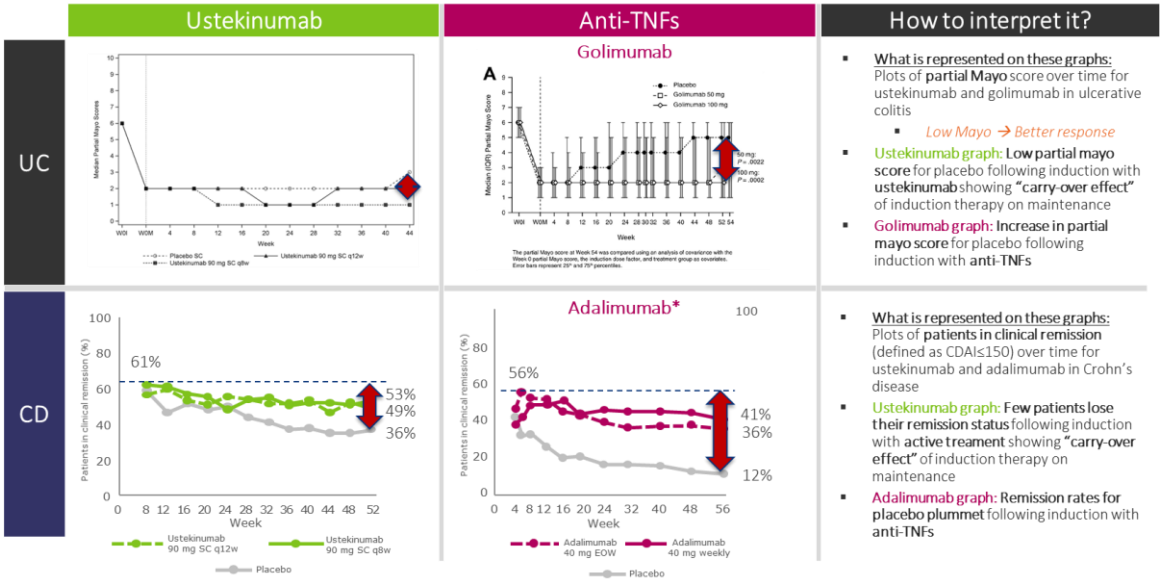


Figure A3. Evidence of carry-over of the induction with ustekinumab vs. anti-TNFs in UC and Crohn's Disease

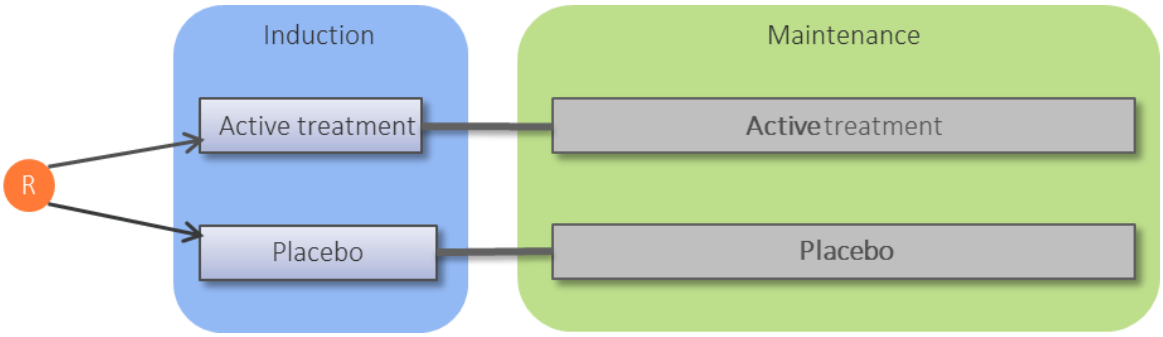
Abbreviation: CD: Crohn's disease; UC: Ulcerative Colitis; * Adalimumab trial patients all received active treatment in induction

A2.2.5.1. Approaches for comparisons at 1-year and choice of base case

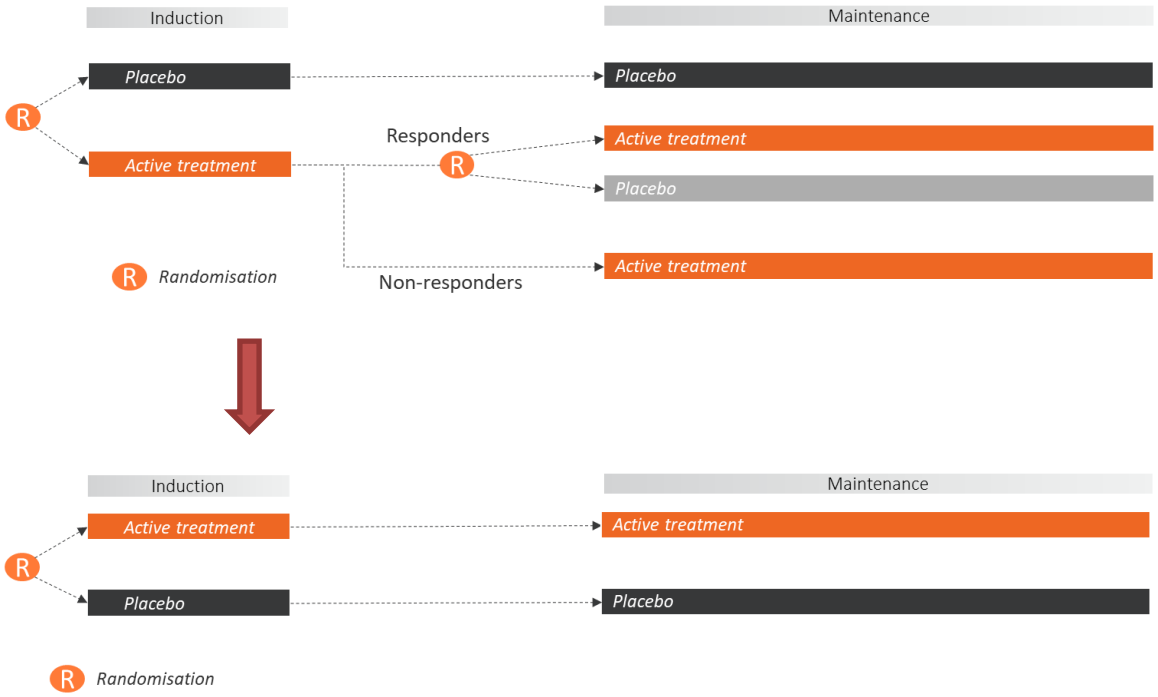
Approaches for comparisons at 1-year and choice of base case

Based on the study designs and data available, there are two approaches that can be taken to assess 1-year treatment effects:

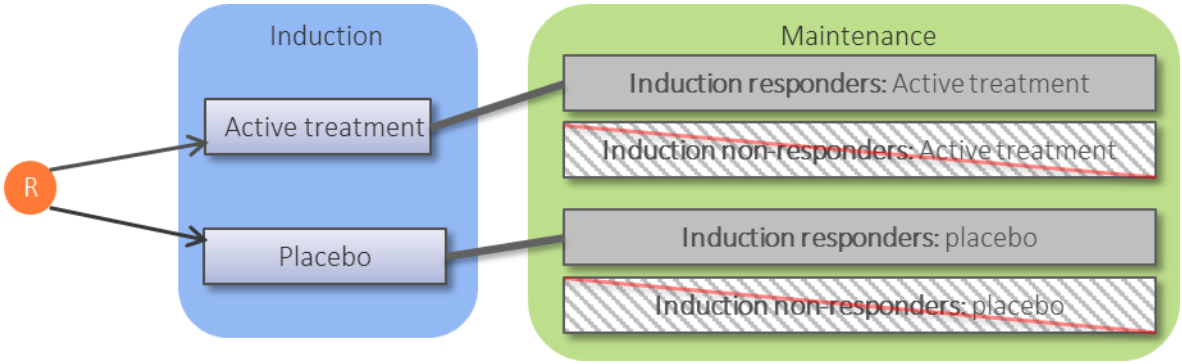
1) **Intention-to-treat (ITT) approach mimicking a treat-through trial design**



- NMA of treat-through arms (induction to maintenance)
- Includes treat-through trial data directly
- Involves re-calculating data from response-based trials to correspond to a treat-through design, maintaining the initial randomisation:



2) ITT approach conditional on induction response



- NMA using an ITT approach with induction to maintenance conditional on responding to induction therapy
- Delayed responders are considered as non-responders post-induction
- Includes response-based trial data directly (responders to induction therapy)
- Involves re-calculating data from treat-through trials to correspond to a response-based design:

Both approaches required re-calculating the data from trials to correspond to a common set of inputs for the network meta analysis. The second approach was considered to be limited as assumes that patients who do not respond at the end of induction also do not respond at the end of maintenance (i.e. ignoring delayed responders).

The approach to mimic an ITT treat-through design was considered to be the most appropriate and hence chosen as the base case. This approach more closely reflected clinical practice, allows for a clearer interpretation of the treatment effects and enables including of head-to-head trial data. The approach mimicking the response-based trial design was conducted as a sensitivity analysis. The second approach was assessed in a sensitivity analysis. Full details on the advantages and disadvantages of each approach can be found in the NMA report of the MAH submission file [46].

A2.2.5.2. Calculations to mimic a treat-through approach

For the re-randomised response-based trials, the calculations required to attain the inputs for the 1-year NMA are as follows (Table A3):

Table A3. Calculations to mimic a treat-through approach

Period	Endpoint data	Calculation
Induction	% Response end of induction of the ITT population	A
	% No response end of induction of the ITT population	B
Maintenance	% Response end of maintenance of the induction responders	C
	% Response end of maintenance of the induction non-responders	D
1-year	% Response end of 1-year of the induction responders	A X C
	% Response end of 1-year of the induction non-responders	B X D
	% Response end of 1-year of the ITT population	(A X C) + (B X D)

For the two approaches assessed, this leads to the following estimates for the % at the end of each sequence:

- Base case approach mimicking a treat-through design: **Induction responders and induction non-responders = (A X C) + (B X D)**
- Sensitivity analysis mimicking a response-based design: **induction responders = A X C**

Where trials only report short-term induction data for patients remaining on the same treatment, imputation was required to estimate the long-term (end of maintenance i.e. 1-year) sequences for both the base case and sensitivity approaches. The methods for including data for both active and placebo arms are described in the following sections. Additionally, the approach requires re-calculation of the total number of patients in the re-randomised response-based trials, further described in the NMA report of the MAH submission file [46].

Full details of the calculations undertaken for each study are provided in the NMA report [46] for the example of clinical response in non-biologic failure patients.

A2.3. Study pool – list of relevant studies used for the assessment

Table A4. Study pool – list of relevant studies used for the assessment

Study reference/ID	Study category			
	Study for marketing authorization of the technology under assessment (yes/no)[a]	**Study included in assessment (yes/no)**	**Sponsored or third-party study[b]**	**Available documentation[c]**
Ustekinumab				
UNIFI NCT02407236	Yes	Yes	Janssen Research & Development, LLC	https://clinicaltrials.gov/ct2/show/record/NCT02407236?term = NCT02407236&rank = 1 CSRs Janssen. A Phase 3, Randomized, Double-blind, Placebo-controlled, Parallel-group, Multicenter Study to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Subjects with Moderately to Severely Active Ulcerative Colitis. Clinical Study Report CNTO1275UCO3001 Induction. 2018. [60] Janssen. A Phase 3, Randomized, Double-blind, Placebo-controlled, Parallel-group, Multicenter Study to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Subjects with Moderately to Severely Active Ulcerative Colitis. Clinical Study Report CNTO1275UCO3001 Maintenance. 2018. [61]
Adalimumab				
ULTRA 1 NCT00385736	N.A	Yes	Abbott	https://clinicaltrials.gov/ct2/show/record/NCT00385736 Reinisch W, Sandborn WJ, Hommes DW, et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. Gut. 2011;60(6):780-787. [52] Reinisch W, Sandborn WJ, Panaccione R, et al. 52-week efficacy of adalimumab in patients with moderately to severely active ulcerative colitis who failed corticosteroids and/or immunosuppressants. Inflamm Bowel Dis. 2013;19(8):1700-1709. [82]

ULTRA 2 NCT00408629	N.A	Yes	Abbott	https://clinicaltrials.gov/ct2/show/record/NCT00408629 Sandborn WJ, van Assche G, Reinisch W, et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2012;142(2):257-65 e3. [53]
A Study of Adalimumab in Japanese Subjects With Moderately to Severely Active Ulcerative Colitis NCT00853099	N.A	Yes	AbbVie (prior sponsor, Abbott)	https://clinicaltrials.gov/ct2/show/record/NCT00853099 Suzuki Y, Motoya S, Hanai H, et al. Efficacy and safety of adalimumab in Japanese patients with moderately to severely active ulcerative colitis. J Gastroenterol. 2014;49(2):283-294. [1]
Infliximab				
Jiang 2015 N.A	N.A	Yes	?	Jiang X CH, et al. Low-dose Infliximab for Induction and Maintenance Treatment in Chinese Patients With Moderate to Severe Active Ulcerative Colitis. J Clin Gastroenterol. 2015;49(582-588). [2]
Probert 2003 N.A	N.A	Yes	"Institutional grant support was received from Schering-Plough and from the BMBF Competence Network "Chronic Inflammatory Bowel Disease" (Germany). We are indebted to Hazel Taylor for performing the statistical analysis. We wish to thank Schering Plough Ltd (UK) and Essex Pharma (Germany)forprovidingthedrugsuse dinthisInvestigatorInstigated Study."	Probert CS, Hearing SD, Schreiber S, Kühbacher T, Ghosh S, Arnott ID, Forbes A. Infliximab in moderately severe glucocorticoid resistant ulcerative colitis: a randomised controlled trial. Gut. 2003 Jul;52(7):998-1002. doi: 10.1136/gut.52.7.998. PMID: 1280 [3]
Japic CTI-060298 N.A	N.A	Yes	?	Kobayashi T, Suzuki Y, Motoya S, et al. First trough level of infliximab at week 2 predicts future outcomes of induction therapy in ulcerative colitis-results from a multicenter prospective randomized controlled trial and its post hoc analysis. J Gastroenterol. 2016;51(3):241-251. [48]
ACT 1 NCT00036439	N.A	Yes	Centocor, Inc.	https://clinicaltrials.gov/ct2/show/record/NCT00036439 Rutgeerts P. SW, Feagan b., et al. Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2005;353:2462-2476. [49]

ACT 2 NCT00096655	N.A	Yes	Centocor, Inc.	https://clinicaltrials.gov/ct2/show/record/NCT00096655 Rutgeerts P. SW, Feagan b., et al. Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2005;353:2462-2476 [49]
Golimumab				
PURSUIT-M NCT00488631	N.A	Yes	Janssen Research & Development, LLC	https://clinicaltrials.gov/ct2/show/record/NCT00488631 Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab maintains clinical response in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2014;146(1):96-109 e1. [55]
PURSUIT-J NCT01863771	N.A	Yes, but not in NMA	Janssen Pharmaceutical K.K.	https://clinicaltrials.gov/ct2/show/record/NCT01863771 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(10):1101-1111. [56]
PURSUIT-SC NCT00487539	N.A	Yes	Janssen Research & Development, LLC	https://clinicaltrials.gov/ct2/show/record/NCT00487539 Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab maintains clinical response in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2014;146(1):96-109 e1. [55]
Tofacitinib				
A Study To Investigate The Safety And Efficacy Of CP-690,550 In Patients With Moderate And Severe Ulcerative Colitis. NCT00787202	N.A	Yes	Pfizer	https://clinicaltrials.gov/ct2/show/record/NCT00787202 Sandborn WJ, Ghosh S, Panes J, et al. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. N Engl J Med. 2012;367(7):616-624. [4]
OCTAVE - Induction 1 (OCTAVE-I1)	N.A	Yes	Pfizer	https://clinicaltrials.gov/ct2/show/record/NCT01465763 Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as Induction and

NCT01465763				Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2017;376(18):1723-1736. [51]
OCTAVE - Induction 2 (OCTAVE-I2) NCT01458951	N.A	Yes	Pfizer	https://clinicaltrials.gov/ct2/show/record/NCT01458951 Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2017;376(18):1723-1736. [51]
OCTAVE – Sustain NCT01458574	N.A	Yes	Pfizer	https://clinicaltrials.gov/ct2/show/record/NCT01458574 Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med. 2017;376(18):1723-1736. [51]
Vedolizumab				
GEMINI 1 NCT00783718	N.A	Yes	Millennium Pharmaceuticals, Inc.	https://clinicaltrials.gov/ct2/show/record/NCT00783718 Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2013;369(8):699-710. [57]
Phase III Study of MLN0002 (300 mg) in the Treatment of Ulcerative Colitis NCT02039505 MLN0002/CCT-101 U1111-1151-6762 (Other Identifier: WHO) JapicCTI-142403 (Registry Identifier: JapicCTI)	N.A	Yes	Takeda	https://clinicaltrials.gov/ct2/show/record/NCT02039505 Motoya S., Watanabe K., Ogata H., Kanai T., Matsui T., Suzuki Y., et al. Vedolizumab in Japanese patients with ulcerative colitis: A Phase 3, randomized, double-blind, placebo-controlled study. PLoS ONE. 2019;14(2). [59]
VARSITY	N.A	Yes	Takeda	https://clinicaltrials.gov/ct2/show/record/NCT02497469

NCT02497469 EudraCT 2015-000939-33				Schreiber LP-B, Edward. V. Loftus J, et al. VARSITY: A double-blind, double-dummy, randomised, controlled trial of vedolizumab versus adalimumab in patients with active ulcerative colitis. ECCO Conference Abstract 2019. [54]

A2.4. Characteristics of the studies included, Induction UNIFI-I and Maintance UNIFI-M studies

Table A5. Characteristics of the studies included, Induction UNIFI-I and Maintance UNIFI-M studies

Study	Design	Population	Dose Regimen of intervention and comparator	Key Endpoints and follow up
UNIFI induction study (NCT02407236) **(CNTO1275UCO3001)** **CSR** [60]	Phase III, randomised, double-blind, placebo-controlled, parallel-group, multicentre study	Adult patients aged 18 years or older with moderately to severely active ulcerative colitis (N = 961) (Mayo score: 6-12; endoscopy score: ≥2) Have failed biologic therapy, that is, have received treatment with 1 or more tumour necrosis factor (TNF) antagonists or vedolizumab at a dose approved for the treatment of UC, and have a documented history of failure to respond to or tolerate such treatment; OR Be naïve to biologic therapy (TNF antagonists or vedolizumab) or have received biologic therapy but have not demonstrated a history of failure to respond to, or tolerate, a biologic therapy and have a prior or current UC medication history that includes at least 1 of the following: a. Inadequate response to or failure to tolerate current treatment with oral corticosteroids or immunomodulators (6-mercaptopurine [6-MP] or azathioprine [AZA]) OR b. History of failure to respond to, or tolerate, at least 1 of the following therapies: oral or IV corticosteroids or immunomodulators (6-MP or AZA) OR c. History of corticosteroid dependence (that is, an inability to successfully taper corticosteroids without a	1:1:1 ratio of placebo IV (n = 319), ustekinumab 130 mg IV (n = 320), and ustekinumab ~6 mg/kg IV (n = 322)	**Primary** Clinical remission at Week 8 (Mayo score ≤2 with no individual subscore > 1), based on centrally read endoscopic subscores **Secondary (Major)** Clinical response at week 8 Endoscopic healing at Week 8 Mucosal healing at week 8 Change from induction baseline in total score of the IBDQ at Week 8 Safety evaluations include an assessment of AEs and routine laboratory analyses.

		return of the symptoms of UC) Before the first administration of study agent, the following conditions must be met: vedolizumab must have been discontinued for at least 4 months and anti-tumor necrosis factors (TNFs) for at least 8 weeks Exclusion: Presence of a stoma; severe extensive colitis and is at imminent risk of colectomy Has UC limited to the rectum only or to < 20 centimeters (cm) of the colon history of a fistula Participants with history of extensive colonic resection (for example, less than 30 cm of colon remaining) that would prevent adequate evaluation of the effect of study agent on clinical disease activity Participants with history of colonic mucosal dysplasia. Asia 13.8%; Eastern Europe 38.2%, Rest of world 48% (North America, Western Europe, Israel, Australia, New Zeland)		
UNIFI maintenance study (NCT02407236) **(CNTO1275UCO3001)** **CSR** [61]	See above Patients in response to ustekinumab at end of 8-week induction phase were eligible to be re-randomised in 44-week maintenance phase	Adult patients aged 18 years or older with moderately to severely active ulcerative colitis who responded to ustekinumab treatment at Week 8 of the induction study (n = 523)	1:1:1 ratio of placebo IV (n = 175), ustekinumab 90 mg SC q12w (n = 172), and ustekinumab 90 mg SC q8w (n = 176)	**Primary** Clinical remission at Week 44 (Mayo score ≤2 with no individual subscore > 1), based on centrally read endoscopic subscores **Secondary (Major)** Maintenance of clinical response through Week 44; Endoscopic healing at Week 44; Corticosteroid

				free clinical remission at Week 44; Maintenance of clinical remission through Week 44 among the subjects who had achieved clinical remission at maintenance baseline The full Mayo score (including an endoscopy) were assessed at the final efficacy visit. Selected patient-reported outcomes (PRO) and health economics data were also collected. Safety evaluations include an assessment of AEs and routine laboratory analyses.

Abbreviations: 5-ASA = 5-aminosalicylic acid; 6-MP = 6-mercaptopurine; IBDQ = inflammatory bowel disease questionnaire; MTX = methotrexate; SF = steroid-free; TNF = tumour necrosis factor; UC = ulcerative colitis

A2.5. Characteristics of the additional studies included in NMA related to adalimumab, infliximab, golimumab, tofacitinib and vedolizumab

Table A6. Characteristics of the additional studies included in NMA related to adalimumab, infliximab, golimumab, tofacitinib and vedolizumab

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary outcome measures and follow-up time points
ULTRA 1 (NCT00385736) [52,82]	To assess the efficacy and safety of **adalimumab** for the induction and maintenance of clinical remission in anti-TNF naive patients with moderately to severely active ulcerative colitis	Phase III/IV, multinational, double-blind (induction) and open-label (maintenance) trial Multicenter, Randomized	Inclusion: Adult patients with moderately to severely active UC (Mayo score: 6-12; endoscopy score: 2-3); patients; patients may be concurrently treated with oral corticosteroids were receiving a stable dose prior to baseline. Patients treated with immunomodulators received at least a consecutive 90-day course prior to baseline of azathioprine (or 6-MP). Exclusion: Previous receipt of any anti-TNF agent or any biological agent, including adalimumab; receipt of intravenous corticosteroids within 14 days prior to screening and during screening; receipt of cyclosporine, tacro-	Induction: • Adalimumab 80/40mg (n = 130) • Adalimumab 160/80/40mg (n = 130) • Placebo (n = 130) Maintenance: • Adalimumab 40mg (n = 390) The study enrolled 576 participants, including 186 participants under the original protocol and protocol Amendments 1 and 2, and 390 participants under protocol Amendments 3 and	Induction: Proportion of participants with Clinical remission Per Mayo Score at week 8 Maintenance: Proportion of participants With Clinical remission, response, and mucosal healing at week 52 as a Secondary Outcome Measures	Induction: Proportion of patients with clinical response per Mayo score; proportion of patients with mucosal healing; proportion of patients with subscores indicative of mild disease at week 8 Maintenance: Proportion of patients who increased to weekly dosing and their rate of clinical remission at week 52, steroid tapering and SF remission at week 52, and remission per partial Mayo score (partial Mayo score ≤2 with no subscore > 1) over time

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary outcome measures and follow-up time points
			limus, mycophenolate mofetil, or MTX within 60 days or cyclosporine, tacrolimus, or mycophenolate mofetil within 30 days prior to baseline See ULTRA 1 trial for more details	4.		Safety: - Summary of Treatment-Emergent Adverse Events in Patients Who Received Any Dose of Adalimumab Through Week 52
ULTRA 2 (NCT00408629) [53]	To assess the efficacy and safety of adalimumab for the induction and maintenance of clinical remission in anti-TNF naive patients with moderately to severely active ulcerative colitis	Phase II, multinational, double-blind trial Multicenter, Randomized https://els-jbs-prod-cdn.literatumonline.com/cms/attachment/2011272685/2033816830/gre1.jpg	Inclusion: Adult patients with moderately to severely active UC (Mayo score: 6-12; endoscopy score: ≥2); patients may be treated concurrently with corticosteroids and/or azathioprine or 6-mercaptopurine Exclusion: Previous receipt of any anti-TNF agent or any biological agent, including adalimumab See ULTRA 2 trial for more details	Induction: • Adalimumab 160mg/80mg (n = 258) • Placebo (n = 260) Maintenance: • Adalimumab 40mg every other week starting at Week 4 (n = 258) • Placebo (n = 260)	Clinical remission at week 8 and 52 Primary outcome:Proportion of Participants Who Achieved Clinical Remission Per Mayo Score at Week 52 Secondary outcome:Proportion of Participants Who Achieved Clinical Response Per Mayo Score at Week 8	• Proportion of patients who achieved clinical remission at both weeks 8 and 52 (sustained) • Clinical response at week 8, week 52, and both weeks 8 and 52 (sustained) • Mucosal healing at week 8, week 52, and both weeks 8 and 52 (sustained) Safety profile See ULTRA 2 for

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
						more details
NCT00853099 [1]	To evaluate ada-limumab for induc-tion and mainte-nance treatment in anti-TNF–naive Japanese patients with UC who were refractory to cortico-steroids, immuno-modulators, or both	Phase II/III, double-blind, trial in Japan Phase III Multi-Center, Ran-domized https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3925299/figure/Fig1/	Inclusion: Patients with mod-erately to severely active UC (Mayo score: 6-12; endos-copy score: ≥2) who are ≥15 years; pa-tients may have received corticoster-oids or immuno-modulators concur-rently at baseline; patients may be refractory to cortico-steroids, immuno-modulators, or both 15yrs and older Exclusion: Patients who have received anti-TNF therapies or other biologic agents, discontinued oral corticosteroids with-in 2 weeks before baseline, received corticosteroid injec-tion, cyclosporine, tacrolimus, or myco-	Induction: • Adalimumab 80/40mg (n = 87) • Adalimumab 160/80mg (n = 90) • Placebo (n = 96) Maintenance: • Adalimumab 40mg (n = 177) • Placebo (n = 96) 274 participants: 96-placebo 87(adalimumab 80/40 mg) 90(adalimumab 160/80 mm Hg)	Clinical response, clinical remission, and mucosal healing at weeks 8, 32, and 52 Primary Outcome Measures : Percentage of Par-ticipants With Clini-cal Remission at 8 Weeks and at 52 weeks Secondary Outcome Measures: Percent-age of Participants With Clinical Remis-sion, With a Clinical Response, With Mucosal Healing, With Rectal Bleed-ing Subscore, …. at 8, 32, and 52 Weeks	Rectal bleeding subscore, physi-cians global as-sessment, and stool frequency indicative of mild disease (score ≤1) and IBDQ response (C16-point increase from base-line in IBDQ score) at weeks 8, 32, and 52 See NCT00853099 for more details

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			phenolate mofetil within 4 weeks be-fore baseline See NCT00853099 for more details			
Jiang 2015 [2]	To study efficacy, safety, or dosage of **infliximab** in Chi-nese patients with moderate to severe ulcerative colitis	Double-blind trial in China Multi-Center, Ran-domized, placebo controlled	Inclusion: Adult patients with moderately to se-verely active UC (Mayo score: 6-12; endoscopy score: ≥2); patients may be treated concurrently with corticosteroids only or in combina-tion with azathio-prine and drugs containing 5-aminosalicylates Exclusion: Patients who re-ceived corticoster-oids or drugs con-taining 5-aminosalicylates rectally within 2 weeks before screening; patients who were previously exposed to inflixi-mab or any other anti–tumour necro-sis factor See trial for more	Induction and Maintenance: • Infliximab 3.5mg (n = 41) • Infliximab 5mg (n = 41) Placebo (n = 41) weeks 0, 2, and 6 and then every 8 weeks through week 22. Patients were followed up through week 30. 123 participants, 18-65 yrs.	Clinical response	Clinical response or clinical remission with discontinuation of corticosteroids at week 30; clinical remission and mu-cosal healing at weeks 8 and 30; clinical response at week 8 in patients with a medical histo-ry of disease refrac-tory to corticoster-oids See trial for more details

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			details			
Probert 2003 [3]	To conduct a double blind, randomised, placebo controlled trial of infliximab in the treatment of moderately severe glucocorticoid re-sistant ulcerative colitis	Double-blind, Ran-domized, placebo controlled trial in United Kingdom and Germany	Inclusion: Adult patients with moderately to se-verely active UC (Mayo score: 6-12; endoscopy score: ≥2); patients may be treated concurrently with a stable dose of 6-mercaptopurine or azathioprine for more than three months prior to trial; patients may have failed to respond to conventional treat-ment with glucocor-ticoids Exclusion: Patients who had received cyclospor-in, any therapeutic agent used to direct-ly reduce TNF, or any investigational drug within three months of enrol-ment, as well as those who had re-cently commenced treatment (within the last three months)	Induction: • Infliximab 5mg (n = 23) Placebo 2.2 mg monobasic sodium phosphate mono-hydrate, 6.1 mg dibasic so-dium phosphate dihydrate, 500 mg sucrose, and 0.5 mg poly-sorbate-80 (n = 20) All patients were randomised to re-ceive a blinded infusion of infliximab 5 mg/kg body weight or pla-cebo at week 0 and a second identical infusion at week 2. At week 6, all pa-tients were reassessed and those who con-tinued to have active ulcerative colitis	Clinical remission six weeks after the first infu-sion using the UCSS	Secondary end points: Changes in UCSS, Baron score, quality of life, and serum C reactive protein levels, and change in daily glucocorticoid dose See trial for further details

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			with 6- MP or AZA See trial for further details	were offered open label treatment with 10 mg/kg infliximab.		
Japic CTI060298 [48]	As part of a phase 3 randomised con-trolled trial of inflixi-mab in UC, to as-sess the predictive value of the first TL at week 2 for short- and long-term re-sponse	Double-blind trial in Japan https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4766223/figure/Fig1/	Inclusion: 16 and older Patients with mod-erately to severely active UC (Mayo score: 6-12; endos-copy score: ≥2); patients may be treated concurrently with corticosteroids, oral aminosalicylates, AZA and 6-MP Exclusion: Patients treated with other biologics, MTX, calcineurin inhibitors, or cy-tapheresis within the previous 18 months See trial for more details	Induction and Maintenance: • Infliximab 5mg (n = 104) • Placebo (n = 104) at weeks 0, 2, and 6. Patients with evi-dence of a response by week 8 continued treatment at weeks 14 and 22.	Clinical response The primary end point was a clinical response at week 8. Secondary end points were clinical remission or MH response at week 8 as well as clinical response or clinical remission at week 30. The Mayo score was determined at weeks 0, 8, and 30.	Clinical remission and mucosal healing Safety -incidence of adverse events See trial for more details
ACT 1 (NCT00036439) [49]	To conduct 54-week and 30-week studies of infliximab in pa-tients with moder-ate-to-severe ulcera-tive colitis	Phase III, multi-national, Random-ized, double-blind trial	Inclusion: Patients with mod-erately to severely active UC (Mayo score: 6-12; endos-copy score: ≥2); patients may be	Induction and Maintenance: • Infliximab 5mg/kg (n = 121) • Infliximab 10mg/kg (n =	Primary outcome: Clinical response at week 8 defined as a de-crease from base-line in the Mayo score by = 30% and	Clinical response or clinical remission with discontinuation of corticosteroids at weeks 30 and 54, clinical remission and mucosal healing

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary outcome measures and follow-up time points
			treated concurrently with corticosteroids alone or in combination with AZA or mercaptopurine Exclusion: Patients treated with infliximab or any other anti-TNF agent; rectally administered corticosteroids or medications containing 5-ASA two weeks before screening See ACT 1 trial for more details	122) • Placebo (n = 121)	= 3 points, with a decrease in the rectal bleeding subscore of = 1 or a rectal bleeding subscore of 0 or 1, at week 8. Secondary outcome: Clinical remission Mayo score of = 2 points, with no individual subscore > 1 at week 8	at weeks 8, 30 and 54, clinical response at week 8 in patients with a history of disease refractory to corticosteroids See ACT 1 trial for more details Changes in Inflammatory Bowel Disease Questionnaire (IBDQ) and Medical Outcomes Study 36-Item Short Form (SF-36)
ACT 2 (NCT00096655) [49]	To conduct 54-week and 30-week studies of infliximab in patients with moderate-to-severe ulcerative colitis	Phase III, multinational, Randomized, double-blind trial	Inclusion: Patients with moderately to severely active UC (Mayo score: 6-12; endoscopy score: ≥2); patients may be treated concurrently with corticosteroids alone or in combination with AZA or MP and medications contain-	Induction and Maintenance: • Infliximab 5mg/kg (n = 121) • Infliximab 10mg/kg (n = 120) • Placebo (n = 123)	Clinical response at week 8	Clinical response or clinical remission with discontinuation of corticosteroids at weeks 30, clinical remission and mucosal healing at weeks 8 and 30, clinical response at week 8 in patients with a history of disease refractory to corticosteroids

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			ing 5- ASA Exclusion: Patients treated with infliximab or any other anti-TNF agent, rectally administered corticosteroids or medications containing 5-ASA two weeks before screening See ACT 2 trial for more details			
PURSUIT-J (NCT01863771) [56]	To evaluate the efficacy and safety of **golimumab** as maintenance therapy in the Japanese population	Phase III, Randomized, open-label (induction) and double-blind (maintenance) trial in Japan	Inclusion: Adult patients with moderately to severely active UC (Mayo score: 6-12; endoscopy score: ≥2); patients may have had inadequate response to or had failed to tolerate one or more conventional therapies (oral 5-ASA, oral corticosteroids, AZA and/or MP) or had demonstrated corticosteroid dependence Exclusion:	Induction: • Golimumab 200mg (n = 144) Maintenance: • Golimumab 100mg (n = 32) • Golimumab 100mg (n = 60) • Placebo (n = 31) https://link.springer.com/article/10.1007%2Fs00535-017-1326-1	Clinical response at week 54	Clinical remission and mucosal healing at week 30 and week 54; proportion of patients who maintained clinical remission at both week 30 and week 54 among patients induced into clinical remission with SC golimumab See PURSUIT-J trial for more details

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			Patients previously treated with TNFi therapy See PURSUIT-J trial for more details	https://clinicaltrials.gov/ct2/show/results/NCT00488631		
PURSUIT-SC (NCT00487539) [55]	To evaluate the safety and efficacy of the selected SC golimumab induction regimens	Phase II/III, multi-national, Random-ized, double-blind trial	Inclusion: Adult patients with moderately to se-verely active UC (Mayo score: 6-12; endoscopy score: ≥2); patients concur-rently treated with oral 5-aminosalicylates or corticosteroids were to receive a stable dose for at least 2 weeks before base-line, and patients receiving AZA and/or 6-MP were to receive a stable dose for at least 4 weeks before base-line. Exclusion: Patients previously treated with biologic anti-TNF agent(s)	Induction: • Golimumab 100/50mg (n = 42) • Golimumab 200/100mg (n = 42) • Golimumab 400/200mg (n = 43) • Placebo (n = 42) Maintenance: • Golimumab 200/100mg (n = 258) • Golimumab 400/200mg (n = 258) • Placebo (n = 258) https://marlin-prod.literatumonline.com/cms/attachment	Primary outcome: Clinical response at week 6	Secondary outcome: Clinical remission at week 6, mucosal healing, IBDQ score change See PURSUIT-SC trial for more details

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			natalizumab or other agents targeting the α-4 integrin, B-cell depleting agents (rituximab), or T-cell depleting agents (alemtuzumab, visilizumab) within 12 months of the first study-agent injection (or contin-ued B- or T-cell depletion > 12 See PURSUIT-SC trial for more details	/2011260507/20337 28763/gr1_lrg.jpg https://www.gastrojo ur-nal.org/article/S0016 -5085(13)00846-9/pdf		
PURSUIT-M (NCT00488631) [55]	To evaluate SC golimumab mainte-nance therapy ad-ministered every 4 weeks through week 52	Phase III, multi-national, double-blind trial	**Inclusion Criteria:** Participants who received all study agent administra-tions and completed the Week 6 Mayo score evaluation in induction study NCT00488774 or NCT00487539 Participants who completed the Week 0 visit for this maintenance study NCT00488631 on the same day as the Week 6 visit of the induction study	Maintenance: • Golimumab 50mg (n = 154) • Golimumab 100mg (n = 154) Placebo (n = 156)	Maintenance of clinical response through week 54	Clinical remission at both weeks 30 and 54, mucosal healing at both weeks 30 and 54, clinical re-mission at both weeks 30 and 54, corticosteroid-free clinical remission at week 54

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			NCT00488774 or NCT00487539 **Exclusion Criteria:** Participants who increased the dose of their concomitant (given at the same time) UC medica-tions since Week 0 of induction study NCT00488774 or NCT00487539 Participants who initiated a concomi-tant UC medication since Week 0 of an induction study NCT00488774 or NCT00487539 Participants who had a partial or total colectomy (surgery to remove part or all of the colon) or an ostomy (surgical construction of an artificial opening (stoma) for external fistulization of a duct or vessel by inser-tion of a tube with or without a supportive stent) since Week 0 of an induction study			

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			NCT00488774 or NCT00487539 Participants with signs or symptoms of latent or active granulomatous in-fection (including TB); a nontubercu-lous mycobacterial infection or oppor-tunistic infection; or infection with HIV (Human Immunode-ficiency Virus), hep-atitis B, or hepatitis C Participants with signs and symptoms of any malignancy or suggestive of a possible lymphopro-liferative disease (disorders character-ized by proliferation of lymphoid tissue, general or unspeci-fied)			
NCT00787202 [4]	A Randomized, Placebo Controlled, Double Blind, Paral-lel Group Multi-Center Study In Order To Investigate Safety And Efficacy	Phase II, multi-national, random-ized, double-blind trial	Inclusion: Adult patients with moderately to se-verely active UC (Mayo score: 6-12; endoscopy score: ≥2)	Induction: • Tofacitinib 0.5mg (n = 31) • Tofacitinib 3mg (n = 33) • Tofacitinib 10mg (n = 33)	Primary end point: Clinical response at 8 weeks	Clinical remission at, endoscopic re-sponse, and endo-scopic remission at 8 weeks See trial for more

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
	Of CP- 690 550 In Subjects With Mod-erate To Severe Ulcerative Colitis		Exclusion: Patients receiving AZA or 6-MP or MTX within 7 days, or cyclosporine, mycophenolate, or tacrolimus within 4 weeks, or anti-TNF agents within 8 weeks See trial for more details	• Tofacitinib 15mg (n = 49) • Placebo (n = 48) https://www.nejm.org/doi/full/10.1056/NEJMoa1112168		details
OCTAVE - Induction 1 (OCTAVE-I1; NCT01465763) [51]	To further evaluate the efficacy and safety of **tofacitinib** in patients with moderate to severe ulcerative colitis who have failed or be intolerant to one of following treatments for ulcerative colitis: oral steroids, azathi-opurine/6-mercaptopurine, or anti-TNF-alpha therapy. as induction and maintenance thera-py	Phase III, multi-national, random-ized, double-blind trial	Inclusion: Adult patients with moderately to se-verely active UC documented diag-nosis of UC at least 4 months prior to entry into the study (Mayo score: 6-12; endoscopy score: 2-3); Subjects must have failed or be intoler-ant of at least one of the following treat-ments for UC: Corticosteroids (oral or intravenous); Azathioprine or 6 mercaptopurine (6 MP);	Induction: • Tofacitinib 10mg (n = 476) • Placebo (n = 122) https://www.nejm.org/doi/full/10.1056/NEJMoa1606910	Primary efficacy end point Clinical remission at week 8	Key secondary end point Mucosal heal-ing, at week 8 See OCTAVE-1 trial for more details

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			Anti TNF-alpha therapy. patients may con-currently be treated with oral aminosalic-ylates and oral glu-cocorticoids Exclusion: Patients who had inadequate washout for the following medications prior to baseline: AZA, 6-MP, or MTX within 2 weeks, TNF antago-nist or interferon therapy within 8 weeks Presence of inde-terminate colitis, microscopic colitis, ischemic colitis, infectious colitis, or clinical findings suggestive of Crohn's disease. Subjects with dis-ease limited to distal 15 cm. Subjects without previous treatment for UC (ie, treatment naïve).			

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			Subjects displaying clinical signs of fulminant colitis or toxic megacolon. See OCTAVE-1 trial for more details			
OCTAVE - Induction 2 (OCTAVE-I2; NCT01458951) [51]	To further evaluated the efficacy and safety of tofacitinib in patients with moderate to severe ulcerative colitis who have failed or be intolerant to one of following treatments for ulcerative colitis: oral steroids, azathi-opurine/6-mercaptopurine, or anti-TNF-alpha therapy. as induction and maintenance thera-py	Phase III, multi-national, random-ized, double-blind trial	See inclusion and exclusion criteria for OCTAVE-1 trial	Induction: • Tofacitinib 10mg (n = 429) • Placebo (n = 112) https://www.nejm.org/doi/full/10.1056/NEJMoa1606910	Primary Outcome: Clinical remission at week 8	Secondary Out-come: Mucosal healing, at week 8 See OCTAVE-1 trial for more details
OCTAVE – Sustain (NCT01458574) [51]	To further evaluated the efficacy of tofa-citinib as induction and maintenance therapy The study proposes to assess whether compared to place-bo, tofacitinib is effective, safe, and	Phase III, multi-national, random-ized, double-blind trial	**Inclusion Criteria:** Subjects who met study entry criteria and completed 8-week induction treatment from Study A3921094 or A3921095; who achieved clinical response in Study A3921094 or	Maintenance: • Tofacitinib 5mg (n = 198) • Tofacitinib 10mg (n = 197) • Placebo (n = 198) https://www.nejm.org/doi/full/10.1056/NEJMoa1606910	Clinical remission at week 52	Mucosal healing at week 52, remission that was sustained at weeks 24 and 52, and glucocorticoid-free among patients who were in remis-sion at mainte-nance-trial entry See OCTAVE-

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
	tolerable mainte-nance therapy in UC; does mainte-nance therapy more effectively achieves mucosal healing and improves quality of life; to assess phar-macokinetic expo-sure during mainte-nance therapy		A3921095; negative test for pregnancy prior to study enrollment; Subjects who are willing and able to comply with sched-uled visits, treat-ment plan, laborato-ry tests, and other study procedures; personally signed and dated informed consent docu-ment(s) **Exclusion Criteria**: Subjects who had major protocol viola-tion (as determined by the Sponsor) in Study A3921094 or A3921095; Pres-ence of indetermi-nate colitis, micro-scopic colitis, is-chemic colitis, infec-tious colitis, or clini-cal findings sugges-tive of Crohn's dis-ease; have had surgery for UC or in the opinion of the inves-tigator, are likely to require surgery for			Sustain trial for more details

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary outcome measures and follow-up time points
			UC during the study period.			
GEMINI 1 (NCT00783718) [57]	To evaluate the efficacy and safety of **vedolizuma**b in patients with ulcerative colitis The primary purpose of this study was to determine the effect of vedolizumab induction treatment on clinical response at 6 weeks and to determine the effect of vedolizumab maintenance treatment on clinical remission at 52 weeks.	Phase III, randomized, multi-national, double-blind and open-label (induction) and double-blind (maintenance) trial	Inclusion: Adult patients with moderately to severely active UC (Mayo score: 6-12; endoscopy score: ≥2); patients demonstrated, over the previous 5 year period, an inadequate response to, loss of response to, or intolerance at least 1 of the following agents: may have been treated with one or more glucocorticoids, immunosuppressive medications (i.e., AZA and 6-MP), or TNF antagonists Exclusion: Patients previously treated with TNF antagonists within 60 days before enrolment or cyclosporine, thalidomide, or investigational	Induction: • Placebo (n = 149) non-ITT (cohort 1) • Vedolizumab 300mg (all) (n = 225) (cohort 1) • Vedolizumab 300mg (n = 521) (cohort 2) https://clinicaltrials.gov/ct2/show/results/NCT00783718 Maintenance: • Vedolizumab 300mg Q8W (n = 122) • Vedolizumab 300mg Q4W (n = 125) • Placebo (n = 126) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6357899/	Induction: Proportion of patients with clinical response at Week 6 Maintenance: Proportion of patients in clinical remission at Week 52	Induction: Proportion of patients in clinical remission at Week 6; proportion of patients with mucosal healing at Week 6 Maintenance: Proportion of patients with durable clinical response; proportion of patients with mucosal healing at Week 52; proportion of patients with durable clinical remission See GEMINI 1 trial for more details

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			drugs within 30 days before enrolment, or if they had been treated previously with vedolizumab, natalizumab, efali-zumab, or rituximab Evidence of ab-dominal abscess at the initial screening visit Extensive colonic resection, subtotal or total colectomy Ileostomy, colosto-my, or known fixed symptomatic steno-sis of the intestine Chronic hepatitis B or C infection Active or latent tu-berculosis See GEMINI 1 for more details			
NCT02039505 [59]	The purpose of this study is to evaluate efficacy, safety and pharmacokinetics of the vedolizumab (MLN0002) induc-tion and mainte-nance therapy in Japanese partici-pants with moderate	Phase III, double-blind and open-label (induction) and dou-ble-blind (mainte-nance) trial in Japan This Phase 3 study (ClinicalTrials.gov:	**Inclusion Criteria:** Eligible patients were aged 15–80 years with total or left-sided UC diag-nosis 6 months before enrollment into the study, based on the Japa-nese Diagnostic	Induction: • Vedolizumab 300mg (n = 246) Cohort 1: Vedolizumab 300 mg n = 164	The primary end-point for the induc-tion phase (Cohort 1 only) was clinical response at Week 10, defined as a reduction of ≥3 points and ≥30% from baseline in the full Mayo score],	In the maintenance phase (Cohort 1 and 2 responders to vedolizumab induc-tion), the primary endpoint was clinical remission at Week 60; secondary end-points were durable clinical

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
	or severe ulcerative colitis (UC)	NCT02039505) consisted of two sequential and integrated, randomized, double-blind, placebo-controlled, parallel-group, multicenter trials plus an open-label cohort, to investigate the efficacy, safety, and PK of vedolizumab as induction and maintenance therapy in Japanese patients with moderate-to-severe UC.	Criteria for UC (2012 revision). Patients had moderate-to-severely-active UC, defined as baseline full Mayo [24] score of 6–12 with an endoscopic subscore 2. Participants meeting the following treatment failure criteria with at least one of the following agents within 5 years before signing on the informed consent: corticosteroids, immunomodulators (azathioprine [AZA] or 6-mercaptopurine), or TNFα antagonist	Placebo N = 82 • Placebo (n = 83) Cohort 2: Vedolizumab 300 mg n = 46 Maintenance: • Vedolizumab 300mg (n = 246) • Placebo (n = 83) Maintenance Phase: Placebo Continuation Vedolizumab 300 mg n = 26 https://clinicaltrials.gov/ct2/show/results/NCT02039505?view = results	and ≥1 point decrease on the rectal bleeding subscore or an absolute rectal bleeding subscore ≤1. Secondary endpoints were clinical remission (defined as a full Mayo score ≤2 and no subscore > 1), and mucosal healing (defined as an endoscopic subscore ≤1) at Week 10.	response (defined as a clinical response at both Weeks 10 and 60), mucosal healing, durable remission (defined as a clinical remission at both Weeks 10 and 60), and corticosteroid-free remission at Week 60 (defined as a clinical remission at Weeks 60 without corticosteroid in patients receiving concomitant corticosteroid therapy at Week 0).
VARSITY NCT02497469 [54]	The purpose of this study is to evaluate the efficacy and safety of vedolizumab intravenous	Phase III, multinational, double-blind trial Randomized, Double-Dummy	Inclusion Criteria: 1.Has a diagnosis of ulcerative colitis established at least	Maintenance: • Vedolizumab 300mg (n = 111) • Adalimumab 160/80/40 mg	Primary Outcome Measures : Percentage of Participants Achieving	Secondary Outcome Measures : Percentage of Participants Achieving

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary outcome measures and follow-up time points
	(IV) treatment compared to adalimumab subcutaneous (SC) treatment over a 52-week treatment period	Head to head trial	3 months prior to screening by clinical and endoscopic evidence and corroborated by a histopathology report. 2.Has moderately to severely active ulcerative colitis as determined by a Mayo score of 6 to 12 with an endoscopic subscore greater than or equal to > = 2 within 14 days prior to the randomization. 3.Has evidence of ulcerative colitis proximal to the rectum (> = 15 centimeter [cm] of involved colon). 4.a.Has had previous treatment with tumor necrosis factor–alpha (TNF-α) antagonists without documented clinical response to treatment (eg, due to lack of response [primary nonresponders], loss of response, or	(n = 120) A total of 769 patients were randomised to VDZ (n = 383) or ADA (n = 386) Experimental: Vedolizumab IV Vedolizumab 300 milligram (mg), infusion, intravenously over 30 minutes on Day 1 and Weeks 2, 6, 14, 22, 30, 38, and 46. Adalimumab placebo-matching injection, subcutaneously on Day 1, Week 2, and every 2 weeks thereafter up to Week 50. Active Comparator: Adalimumab SC Adalimumab 160 mg, injection, subcutaneously on Day 1, adalimumab 80 mg, injection, subcutaneously at Week 2, then adalimumab 40 mg, injection, subcutaneously every 2	Clinical Remission [Time Frame: Week 52]	Mucosal Healing [Time Frame: Week 52] Percentage of Participants Using Oral Corticosteroids at Baseline who Discontinue Corticosteroids and are in Clinical Remission [Time Frame: Week 52]

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			intolerance [sec-ondary nonre-sponder], or b.Has previously used a TNF-α an-tagonist (except adalimumab), and discontinued its use due to reasons oth-er than safety 5. Is naïve to TNF-α antagonist therapy but is failing current treatment (ie, corticosteroids, 5-aminosalicylate, or immunomodula-tors). Exclusion Criteria: 1.The subject has had an extensive colonic resection, subtotal or total colectomy. 2.The subject has any evidence of an active infection dur-ing Screening. 3.The subject has a positive progressive multifocal leu-koencephalopathy (PML) subjective symptom checklist	weeks thereafter up to Week 50. Vedoli-zumab placebo-matching infusion, intravenously on Day 1 and Weeks 2, 6, 14, 22, 30, 38, and 46. https://clinicaltrials.gov/ct2/show/record/NCT02497469		

Study reference/ID	Objective	Study design	Eligibility criteria	Intervention and Comparator (N enrolled)	Primary outcome measure and folow-up time point	Secondary out-come measures and follow-up time points
			at Screening or before the administration of study drug at Day 1. 4. The subject has received any inves-tigational or ap-proved biologic or biosimilar agent (other than those listed below) within 60 days or 5 half lives prior to screen-ing (whichever is longer). 5. The subject has had prior exposure to vedolizumab, natalizumab, efalizumab, ada-limumab, etroli-zumab, AMG-181, anti–mucosal ad-dressin cell adhesion mole-cule-1 (MAdCAM-1)-antibodies, or rituximab.			

Abbreviations: 5-ASA: 5-aminosalicylic acid; 6-MP: 6-mercaptopurine; IBDQ: inflammatory bowel disease questionnaire; MTX: Methotrexate; SF: steroid-free; TNF: tumour necrosis factor; UC: ulcerative colitis

A2.6. Baseline patient characteristics of studies used in NMA

Table A7. Baseline patient characteristics of studies used in NMA [46]

Trial	Population	Phase	Arms	Age (Mean)	Males %	Weight – Kg (Mean)	CRP level - mg/L (Mean)	Disease duration (Mean)	Mayo score (mean)
NCT00853099	**Non-biologic failure**	Induction	PBO (n = 96)	41.3	72.9	60.8	3.4*	7.8	8.5
			ADA 80/40mg (n = 87)	44.4	57.5	58.7	3.1*	8.3	8.6
			ADA 160/80mg (n = 90)	42.5	67.8	60.1	2.2*	7.8	8.6
GEMINI	**Full patient population**	Induction	Cohort 1: PBO (n = 149)	41.2	61.7	72.4	NR	7.1	8.6
			Cohort 1: VDZ 300mg (n = 225)	40.1	58.7	72.4	NR	6.1	8.5
			Cohort 2: VDZ 300mg (n = 521)	40.1	57.8	74.2	NR	7.2	8.6
		Maintenance†	PBO (n = 126)	40.3	55	74.7	NR	7.8	8.4
			VDZ 300mg q8w (n = 122)	41	57	78.2	NR	6.2	8.4
			VDZ 300mg q4w (n = 125)	38.6	54	71.8	NR	7.6	8.3
	Non-biologic failure	Induction	Cohort 1: PBO (n = 76)	40.5	62	70	NR	6.1	8.5
			Cohort 1: VDZ 300mg (n = 130)	39.7	53	69.2	NR	5.8	8.4
			Cohort 2: VDZ 300mg (n = 258)	40.6	59	72.7	NR	6.4	8.5
		Maintenance†	PBO (n = 79)	39.5	57	71.3	NR	6.4	8.4
			VDZ 300mg q8w (n = 72)	41	54	76.1	NR	5.8	8.3
			VDZ 300mg q4w (n = 73)	38.3	53	70	NR	7	8.2
	Biologic Failure	Induction	Cohort 1: PBO (n = 63)	41.8	56	74.2	NR	8	8.6
			Cohort 1: VDZ 300mg (n = 82)	39.7	61	74.9	NR	6.4	8.7
			Cohort 2: VDZ 300mg (n = 222)	40.2	55	75.3	NR	8	8.6
		Maintenance†	PBO (n = 38)	41.6	55	81.2	NR	9.8	8.2
			VDZ 300mg q8w (n = 43)	41.3	56	79.1	NR	6.8	8.5

			VDZ 300mg q4w (n = 40)	39.9	53	72.7	NR	8.1	8.4
ULTRA 1	**Non-biologic failure**	Induction	PBO (n = 130)	37*	63.1	78.7	3.2*	5.35*	8.7
			ADA 80/40mg (n = 130)	40*	60	76.8	6.4*	6.91*	9
			ADA 160/80mg (n = 130)	36.5*	63.8	75.5	3.3*	6.06*	8.8
ULTRA 2	**Full patient populatoin**	Induction	PBO (n = 260)	41.3	61.8	77.1	13.1	8.5	8.9
			ADA 160/80/40mg (n = 248)	39.6	57.3	75.3	14.5	8.1	8.9
		Maintenance - responder†	ADA 40mg EOW (n = 19)	39.6	36.8	78.5	3.9*	7.23*	8
			ADA 40mg weekly (n = 20)	39.8	80	78.4	1.4*	7.1*	8.8
		Maintenance – non responder†	ADA 40mg EOW (n = 19)	41.2	58.6	73.8	8.3*	4.96*	9.1
			ADA 40mg weekly (n = 48)	38.1	60.4	78.3	3.7*	6.79*	9.3
ACT 1	**Non-biologic failure**	Induction	PBO (n = 121)	41.4	59.5	76.8	17	6.2	8.4
			IFX 5mg (n = 121)	42.4	64.5	80	14	5.9	8.5
			IFX 10mg (n = 122)	41.8	59	76.9	16	8.4	8.4
ACT 2	**Non-biologic failure**	Induction	PBO (n = 123)	39.3	57.7	76.1	16	6.5	8.3
			IFX 5mg (n = 121)	40.5	62.8	78.4	13	6.7	8.3
			IFX 10mg (n = 120)	40.3	56.7	79.6	14	6.5	8.5
OCTAVE-I1	**Full patient population**	Induction	PBO (n = 123)	41.8	63.1	72.7	4.7*	6*	9.1
			TFB 10mg (n = 476)	41.3	58.2	72.9	4.4*	6.5*	9
OCTAVE-I2	**Full patient population**	Induction	PBO (n = 112)	40.4	49.1	73.2	5*	6.2*	8.9
			TFB 10mg (n = 429)	41.1	60.4	74.4	4.6*	6*	9
OCTAVE-I1+I2	**Full patient population**	Induction	PBO (n = 234)	41.1	56.4	NR	NR	8.1	9
			TFB 10mg (n = 905)	41.2	59.2	NR	NR	8.1	9
OCTAVE-S	**Full patient population**	Maintenance‡	PBO (n = 198)	43.4	58.6	76.2	1*	7.2*	3.3
			TFB 5mg (n = 198)	41.9	52	73.4	0.7*	6.5*	3.3
			TFB 10mg (n = 197)	42.9	55.8	74.6	0.9*	6.8*	3.4

PURSUIT-SC	**Full patient populatoin**	Induction	PBO (n = 258)	39.7	50.4	NR	9.6	6.4	8.3
			GOL 200/100mg (n = 258)	39.7	54.3	NR	11.5	6.4	8.7
			GOL 400/200mg (n = 258)	40.9	59.7	NR	12	6.5	8.6
	Non-biologic failure	Induction	PBO (n = 331)	39	52.9	NR	10.7	6	8.3
			GOL 100/50mg (n = 72)	40.9	55.6	NR	8.2	6.6	8.2
			GOL 200/100mg (n = 331)	40	54.4	NR	11.3	6.4	8.6
			GOL 400/200mg (n = 331)	40.7	60.7	NR	13.2	6.4	8.5
PURSUIT-M	**Non-biologic failure**	Maintenance - Patients who failed conventional therapy + non responders in induction phases†	PBO (n = 129)	38	47.3	NR	9.5	6.3	8.2
			GOL 100mg (n = 230)	40.3	57	NR	9.6	6.2	8.2
			GOL 100mg (n = 405)	41.2	65.9	NR	13.2	6.1	8.6
		Maintenance - Patients who failed conventional therapy + responders in induction phases†	PBO (n = 156)	40.2	48.1	NR	9.6	6.9	8.3
			GOL 50mg (n = 154)	41.4	50	NR	8.5	6.8	8.1
			GOL 100mg (n = 154)	39.1	57.8	NR	8.9	7.2	8.5
PURSUIT-J	**Non-biologic failure**	Induction	I: GOL 200mg (n = 144)	42.4	68	61.51	4.9	5.08*	8*
		Maintenance†	M: DB: PBO (n = 31)	42.9	61	59.48	4.06	5.74*	8*
			M: DB: GOL 100mg (n = 32)	39.3	59	64.59	5.31	5.35*	8*
			M: OL: GOL 100mg (n = 60)	42.1	70	60.97	4.68	4.57*	8*
NCT00787202	**Full patient population**	Induction	PBO (n = 48)	42.5	48	74.6	9.7	8.8	8.2
			TFB 0.5mg (n = 31)	43.8	55	75.6	18.8	8.8	8.6
			TFB 3mg (n = 33)	42.5	58	73.8	12.6	8.9	8.3
			TFB 10mg (n = 33)	43.2	64	75.9	11.3	10.9	8

			TFB 15mg (n = 49)	41.2	53	74.1	17.1	7.6	8
Jiang 2015	**Non-biologic failure**	Induction	PBO (n = 41)	34.5	60.9	61.2	35.1	4.4	NR
			IFX 3.5mg (n = 41)	34.1	58.5	63.1	35.7	4.3	NR
			IFX 5mg (n = 41)	34.3	63.4	62.8	35.8	4.4	NR
Probert 2003	**Non-biologic failure**	Induction	PBO (n = 20)	NR	NR	72*	12	4.92*	NR
			IFX 5mg (n = 23)	NR	NR	66*	9	6.25*	NR
Japic CTI060297	**Non-biologic failure**	Induction	PBO (n = 104)	37.8	64.4	60.3	7	7.1	8.5
			IFX 5mg (n = 104)	40	63.5	57.6	10	8.1	8.6
UNIFI	**Full patient population**	Induction	PBO (n = 319)	40*	61.8	70*	9.8	5.97*	9*
			UST 130mg (n = 320)	42*	59.4	72*	9.6	5.9*	9*
			UST 6mg/kg (n = 322)	41*	60.6	71.8*	12.1	6.03*	9*
		Maintenance‡	PBO (n = 175)	42*	61.9	71*	3.73	5.56*	4*
			UST 90mg q12w (n = 172)	39*	55.8	70*	3.91	5.95*	4*
			UST 90mg q8w (n = 176)	39*	53.4	70*	4.95	6.36*	4*

* Median † The baseline values were obtained at the beginning of the induction phase for patients entering the maintenance phase ‡ The baseline values were obtained at the beginning of the maintenance phase

Abbreviations: ADA=adalimumab; GOL=golimumab; IFX=infliximab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo; CrI=credible interval; DB=Double blind; EOW=every other week; OL=open label; Pr=Bayesian probability for ustekinumab to be better than its comparator

A2.7. Effectiveness results based on full patient population and subgroup analyses based on biological failure status

Results for clinical remission, clinical response and mucosal healing based on full patient population are presented in Table A8 till Table A10 [46].

Table A8. Results summary for clinical remission in the full patient population

Study reference/ID	Induction/Maintenance	Outcome intervention n/N (%)	Outcome Comparator n/N (%)
Ustekinumab trials (technology)			
UNIFI	Induction	UST 130mg 57/320 (17.8)	PBO 22/319 (6.9)
	Induction	UST 6mg/kg 61/322 (18.9)	PBO 22/319 (6.9)
	Maintenance	UST 90mg Q12W 66/172 (38.4)	PBO 42/175 (24.0)
	Maintenance	UST 6mg/kg Q8W 77/176 (43.8)	PBO 42/175 (24.0)
Infliximab trials			
ACT 1	Induction	IFX 5mg/kg 47/121 (38.8)	PBO 18/121 (14.9)
	Induction	IFX 10mg/kg 39/122 (32.0)	PBO 18/121 (14.9)
	Maintenance	IFX 5mg/kg 42/121 (34.7)	PBO 20/121 (16.5)
	Maintenance	IFX 10mg/kg 42/122 (34.4)	PBO 20/121 16.5
ACT 2	Induction	IFX 5mg/kg 41/121 (33.9)	PBO 7/123 (5.7)
	Induction	IFX 10mg/kg 33/120 (27.5)	PBO 7/123 (5.7)
Japic CTI060298	Induction	IFX 5mg 21/104 (20.2)	PBO 11/104 (10.6)
Jiang 2015	Induction	IFX 3.5mg/kg 21/41 (51.2)	PBO 9/41 (21.9)
	Induction	IFX 5mg/kg 22/41 (53.7)	PBO 9/41 (21.9)
Adalimumab trials			
ULTRA 1	Induction	ADA 80/40mg NA/130 (10.0)	PBO NA/130 (9.2)
	Induction	ADA 160/80mg NA/130 (18.5)	PBO NA/130 (9.2)
	Maintenance	ADA 80/40mg, 160/80mg, placebo pooled 115/390 (29.5)	NA
ULTRA 2	Induction	ADA 40mg EOW NA/248 (16.5)	PBO NA/246 (9.3)

	Maintenance	ADA 40mg EOW NA (17.3)	PBO NA (8.5)
NCT00853099	Induction	ADA 80/40mg NA/87 (14)	PBO NA/96 (11)
	Induction	ADA 160/80mg NA/90 (10)	PBO NA/96 (11)
	Maintenance	ADA 40mg EOW NA/96 (23)	PBO NA/177 (7)
VARSITY	Maintenance	VDZ 300 mg 120/383 (31.3)	ADA 160/80/40mg 87/386 (22.5)
Golimumab trials			
PURSUIT-M	Maintenance	GOL 50mg (NA/151); (23.2)	PBO (NA/154); (15.6)
	Maintenance	GOL 100mg (NA/151); (27.8)	PBO (NA/154); (15.6)
PURSUIT-J	Induction	GOL 200 mg (27/144); (18.8)	NA
	Maintenance	GOL 100mg (16/32); (50.0)	PBO (2/31); (6.5)
PURSUIT-SC	Induction	GOL 200/100mg (45/253); (17.8)	PBO (16/251); (6.4)
	Induction	GOL 400/200mg (46/257); (17.9)	PBO (16/251); (6.4)
Vedolizumab trials			
GEMINI	Induction	VDZ 300 mg 38/225 (16.9)	PBO 8/149 (5.4)
	Maintenance	VDZ 300 mg q4w 56/125 (44.8)	PBO 20/126 (15.9)
	Maintenance	VDZ 300 mg q8w 51/122 (41.8)	PBO 20/126 (15.9)
NCT02039505	Induction	VDZ 300 mg 30/164 (18.3)	PBO 10/82 (12.2)
NCT00787202	Induction	TOC 10mg 16/33 (48)	PBO 5/48 (10)
	Induction	TOC 15mg 20/49 (41)	PBO 5/48 (10)
VARSITY	Maintenance	VDZ 300 mg 120/383 (31.3)	ADA 160/80/40mg 87/386 (22.5)
Tofacitinib trials			
OCTAVE-I1	Induction	TOC 10mg 88/476 (18.5)	PBO 10/122 (8.2)
OCTAVE-I2	Induction	TOC 10mg 71/429 (16.6)	PBO 4/112 (3.6)
OCTAVE – Sustain	Maintenance	TOC 5mg 68/198 (34.3)	PBO 22/198 (11.1)
	Maintenance	TOC 10mg 80/197 (40.6)	PBO 22/198 (11.1)

Abbreviations: ADA=adalimumab; GOL=golimumab; IFX=infliximab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo; NA=Not Applicable

Table A9. Results summary for clinical response in the full patient population

Study reference/ID	Induction/ Maintenance	Outcome intervention n/N (%)	Outcome comparator n/N (%)
Ustekinumab trials (technology)			
UNIFI	Induction	UST 130mg 164/320 (51.3)	PBO 100/319 (31.3)
	Induction	UST 6mg/kg 199/322 (61.8)	PBO 100/319 (31.3)
	Maintenance	UST 90mg SC q12w 117/172 (68.0)	PBO 78/175 (44.6)
	Maintenance	UST 90mg SC q8w 125/176 (71.0)	PBO 78/175 (44.6)
Infliximab trials			
Silva 2017	NR	IFX 12/21 (57.14)	ADA 6/10 (60)
Jiang 2015	Induction	IFX 3.5mg/kg 30/41 (73.1)	PBO 15/41 (36.6)
	Induction	IFX 5mg/kg 32/41 (78.1)	PBO 15/41 (36.6)
Japic CTI060298	Induction	IFX 5mg/kg 57/104 (54.8)	PBO 37/104 (35.6)
ACT 1	Induction	IFX 5mg/kg 84/121 (69.4)	PBO 45/121 (61.5)
	Induction	IFX 10mg/kg 75/122 (61.5)	PBO 45/121 (61.5)
	Maintenance	IFX 5mg/kg 55/121 (45.5)	PBO 24/121 (19.8)
	Maintenance	IFX 10mg/kg 54/121 (44.3)	PBO 24/121 (19.8)
ACT 2	Induction	IFX 5mg/kg 78/121 (64.5)	PBO 36/123 (29.3)
	Induction	IFX 10mg/kg 83/120 (69.2)	PBO 36/123 (29.3)
Tofacitinib trials			
OCTAVE-I1	Induction	TOC 10mg 285/476 (59.9)	PBO 40/122 (32.8)
OCTAVE-I2	Induction	TOC 10mg 236/429 (55.0)	PBO 32/112 (28.6)

OCTAVE-Sustain	Maintenance	TOC 5mg 102/198 (51.5)	PBO 40/198 (20.2)
	Maintenance	TOC 10mg 122/197 (61.9)	PBO 40/198 (20.2)
NCT00787202	Induction	TOC 0.5mg 10/31 (32)	PBO 20/48 (42)
	Induction	TOC 3mg 16/33 (48)	PBO 20/48 (42)
	Induction	TOC 10mg 20/33 (61)	PBO 20/48 (42)
	Induction	TOC 15mg 38/49 (78)	PBO 20/48 (42)
Vedolizumab trials			
GEMINI 1	Induction	VDZ 300mg 106/225 (47.1)	PBO 38/149 (25.5)
	Maintenance	VDZ 300mg q8w 69/122 (56.6)	PBO 30/126 (23.8)
	Maintenance	VDZ 300mg q4w 65/125 (52.0)	PBO 30/126 (23.8)
NCT02039505	Induction	VDZ 300mg 65/164 (39.6)	PBO 27/82 (32.9)
	Maintenance	VDZ 300mg 27/41 (65.9)	PBO 15/42 (35.7)
Adalimumab trials			
ULTRA 1	Induction	ADA 80/40mg NA (51.5)	PBO NA (44.6)
	Induction	ADA 160/80mg NA (54.6)	PBO NA (44.6)
	Maintenance	ADA 166/390 (42.6)	NA
ULTRA 2	Induction	ADA NA (50.4)	PBO NA (34.6)
	Maintenance	ADA NA (30.2)	PBO NA (18.3)
NCT00853099	Induction	ADA 80/40mg NA (43)	PBO NA (35)
	Induction	ADA 160/80mg NA (50)	PBO NA (35)
	Maintenance	ADA 40mg EOW NA (31)	PBO NA (18)

Golimumab trials			
PURSUIT-SC	Induction	GOL 200/100mg NA (51.0)	PBO NA (30.3)
	Induction	GOL 400/200mg NA (54.9)	PBO NA (30.3)
PURSUIT-M	Maintenance	GOL 50mg NA (47.0)	PBO NA (31.2)
	Maintenance	GOL 100mg NA (49.7)	PBO NA (31.2)
PURSUIT-J	Maintenance	GOL 100mg NA (56.3)	PBO NA (19.4)

Abbreviations: ADA=adalimumab; GOL=golimumab; IFX=infliximab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo; NA=Not Applicable

Table A10. Results summary for mucosal healing in the full patient population

Study reference/ID	Induction/ Maintenance	Outcome intervention n/N (%)	Outcome comparator n/N (%)
Ustekinumab trials (technology)			
UNIFI	Induction	UST 130mg 84/320 (26.3)	PBO 44/319 (13.8)
	Induction	UST 6mg/kg 87/322 (27.0)	PBO 44/319 (13.8)
	Maintenance	UST 90mg SC q12w 75/172 (43.6)	PBO 50/175 (28.6)
	Maintenance	UST 90mg SC q8w 90/176 (51.1)	PBO 50/175 (28.6)
Infliximab trials			
Jiang 2015	Induction	IFX 3.5mg/kg 23/41 (56.1)	PBO 10/41 (24.4)
	Induction	IFX 5 mg/kg 24/41 (58.5)	PBO 10/41 (24.4)
Japic CTI060298	Induction	IFX 5 mg/kg 48/104 (46.2)	PBO 29/104 (27.9)
ACT 1	Induction	IFX 5 mg/kg 75/121 (62.0)	PBO 41/121 (33.9)
	Induction	IFX 10 mg/kg 72/122 (59.0)	PBO 41/121 (33.9)
	Maintenance	IFX 5mg/kg 55/121 (45.5)	PBO 22/121 (18.2)
	Maintenance	IFX 10 mg/kg 57/122 (46.7)	PBO 22/121 (18.2)
ACT 2	Induction	IFX 5mg/kg 73/121 (60.3)	PBO 38/123 (30.9)
	Induction	IFX 10 mg/kg 74/120 (61.7)	PBO 38/123 (30.9)
Tofacitinib trials			
OCTAVE-I1	Induction	TOC 10mg 149/476 (31.3)	PBO 19/122 (15.6)
OCTAVE-I2	Induction	TOC 10mg 122/429 (28.4)	PBO 13/112 (11.6)
OCTAVE-Sustain	Maintenance	TOC 5mg 74/198 (37.4)	PBO 26/198 (13.1)
	Maintenance	TOC 10mg 90/197 (45.7)	PBO 26/198 (13.1)
Vedolizumab trials			
GEMINI 1	Induction	VDZ 300mg 92/225 (40.9)	PBO 37/149 (24.8)
	Maintenance	VDZ 300mg q8w 63/122 (51.6)	PBO 25/126 (19.8)

	Maintenance	VDZ 300mg q4w 70/125 (56.0)	PBO 25/126 (19.8)
NCT02039505	Induction	VDZ 300mg 60/164 (36.6)	PBO 25/82 (30.5)
	Maintenance	VDZ 300mg 26/41 (63.4)	PBO 14/42 (33.3)
VARSITY	Maintenance	VDZ 152/383 (39.7)	ADA 107/386 (27.7)
Adalimumab trials			
ULTRA 1	Induction	ADA 80/40mg NA (37.7)	PBO NA (41.5)
	Induction	ADA 160/80mg NA (46.9)	PBO NA (41.5)
	Maintenance	ADA 148/390 (37.9)	NA
ULTRA 2	Induction	ADA NA (41.1)	PBO NA (31.7)
	Maintenance	ADA NA (31.7)	PBO NA (15.4)
NCT00853099	Induction	ADA 80/40mg NA (39)	PBO NA (30)
	Induction	ADA 160/80 NA (44)	PBO NA (30)
	Maintenance	ADA 40mg EOW NA (29)	PBO NA (16)
Golimumab trials			
PURSUIT-SC	Induction	GOL 200/100mg 107/253 (42.3)	PBO 72/251 (28.7)
	Induction	GOL 400/200mg 116/257 (45.1)	PBO 72/251 (28.7)
PURSUIT-M	Maintenance	GOL 50mg NA (41.7)	PBO NA (26.6)
	Maintenance	GOL 100mg NA (42.4)	PBO NA (26.6)
PURSUIT-J	Maintenance	GOL 100mg NA (59.4)	PBO NA (16.1)

Abbreviations: ADA=adalimumab; GOL=golimumab; IFX=infliximab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo; NA=Not Applicable

Subgroup analyses results for clinical remission, clinical response and mucosal healing based on biological failure status are presented in Table A11-Table A16.

Table A11. Results summary for clinical response in the non-biologic failure population

Study reference/ID	Induction/Maintenance	Outcome intervention (n/N); %	Outcome Comparator (n/N); %
Ustekinumab trials (technology)			
UNIFI	Induction	UST 130mg (n = 90/156); 57.7	PBO (n = 158); 35.4
	Induction	UST 6mg/kg (n = 104/156); 66.7	PBO (n = 158); 35.4
	Maintenance	UST 90mg Q12W (n = 102); 76.5	PBO (n = 87); 50.6
	Maintenance	UST 6mg/kg Q8W (n = 85); 77.6	PBO (n = 87); 50.6
Infliximab trials			
ACT 1	Induction	IFX 5mg/kg (n = 84/121); 69.4	PBO (n = 45/121); 37.2
	Induction	IFX 10mg/kg (n = /75122); 61.5	PBO (n = 45/121); 37.2
	Maintenance	IFX 5mg/kg (n = 55/121); 45.5	PBO (n = 24/121); 19.8
	Maintenance	IFX 10mg/kg (n = 54/122); 44.3	PBO (n = 24/121); 19.8
ACT 2	Induction	IFX 5mg/kg (n = 78/121); 64.5	PBO (n = 36/123); 29.3
	Induction	IFX 10mg/kg (n = 83/120); 69.2	PBO (n = 36/123); 29.3
Japic CTI060298	Induction	IFX 5mg (n = 57/104); 54.8	PBO (n = 37/104); 35.6
Jiang 2015	Induction	IFX 3.5mg/kg (n = 30/41); 73.1	PBO (n = 15/41); 36.6
	Induction	IFX 5mg/kg (n = 32/41); 78.1	PBO (n = 15/41); 36.6
Adalimumab trials			
ULTRA 1	Induction	ADA 80/40mg (n = 130); 51.5	PBO (n = 130); 44.6
	Induction	ADA 160/80mg (n = 130); 54.6	PBO (n = 130); 44.6
	Maintenance	ADA 40mg EOW (n = 390); 42.6	NA
ULTRA 2	Maintenance	ADA 40mg EOW (n = 150); 36.7	PBO (n = 145); 24.1
NCT00853099	Maintenance	ADA 40mg EOW (n = NR); 31.0	PBO (n = NR); 18.0

Golimumab trials			
PURSUIT-M	Maintenance	GOL 50mg (n = 151); 47.0	PBO (n = 154); 31.2
	Maintenance	GOL 100mg (n = 151); 49.7	PBO (n = 154); 31.2
PURSUIT-J	Induction	GOL 200 mg (n = 63/144); 43.8	NA
	Maintenance	GOL 100mg (n = 18/32); 56.3	PBO (n = 6/31); 19.4
PURSUIT-SC	Induction	GOL 200/100mg (n = 253); 51.0	PBO (n = 251); 30.3
	Induction	GOL 400/200mg (n = 257); 54.9	PBO (n = 251); 30.3
Vedolizumab trials			
GEMINI	Induction	VDZ 300 mg (n = 69/130); 53.1	PBO (n = 20/76); 26.3
	Maintenance	VDZ 300 mg (n = 88/145); 60.7	PBO (n = 21/79); 26.6
NCT02039505	Induction	VDZ 300 mg (n = 79); 53.2	PBO (n = 41); 36.6
Tofacitinib trials			
NCT00787202	Induction	TOC 10mg (n = 23); 60.9	PBO (n = 33); 45.5
	Induction	TOC 15mg (n = 34); 82.4	PBO (n = 33); 45.5
OCTAVE I1+I2	Induction	TOC 10mg (n = 440); 64.5	PBO (n = 110); 39.1
OCTAVE – Sustain	Maintenance	TOC 5mg (n = 115); 56.5	PBO (n = 109); 24.8
	Maintenance	TOC 10mg (n = 104); 64.4	PBO (n = 109); 24.8

Abbreviations: ADA=adalimumab; GOL=golimumab; IFX=infliximab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo; NA=Not Applicable

Table A12. Results summary for clinical response in the biologic failure population

Study reference/ID	Induction/Maintenance	Outcome intervention (n/N); %	Outcome Comparator (n/N); %
Ustekinumab trials (technology)			
UNIFI	Induction	UST 130mg (n = 74/164); 45.1	PBO (n = 44/161); 27.3
	Induction	UST 6mg/kg (n = 95/166); 57.2	PBO (n = 44/161); 27.3
	Maintenance	UST 90mg Q12W (n = 70); 55.7	PBO (n = 88); 38.6
	Maintenance	UST 6mg/kg Q8W (n = 91); 64.8	PBO (n = 88); 38.6
Adalimumab trials			
ULTRA 2	Induction	ADA 160/80/40mg (n = 98); 36.7	PBO (n = 101); 28.7
	Maintenance	ADA 40mg EOW (n = 98); 20.4	PBO (n = 101); 9.9
Vedolizumab trials			
GEMINI	Induction	VDZ 300 mg (n = 32/82); 39.0	PBO (n = 13/63); 20.6
	Maintenance	VDZ 300 mg (n = 37/83); 44.6	PBO (n = 6/38); 15.8
Tofacitinib trials			
NCT00787202	Induction	TOC 10mg (n = 10); 60.9	PBO (n = 15); 33.3
	Induction	TOC 15mg (n = 15); 66.7	PBO (n = 15); 33.3
OCTAVE I1+I2	Induction	TOC 10mg (n = 465); 51.0	PBO (n = 124); 23.4
OCTAVE – Sustain	Maintenance	TOC 5mg (n = 83); 44.6	PBO (n = 89); 14.6
	Maintenance	TOC 10mg (n = 93); 59.1	PBO (n = 89); 14.6

Abbreviations: ADA=adalimumab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo

Table A13. Results summary for clinical remission for the non-biologic failure population

Study reference/ID	Induction/Maintenance	Outcome intervention (n/N); %	Outcome Comparator (n/N); %
Ustekinumab trials (technology)			
UNIFI	Induction	UST 130mg (n = 31/156); 19.9	PBO (n = 15/159); 9.4
	Induction	UST 6mg/kg (n = 29/156); 18.6	PBO (n = 15/159); 9.4
	Maintenance	UST 90mg sc Q12W (n = 102); 49.0	PBO (n = 87); 31.0
	Maintenance	UST 90mg sc Q8W (n = 85); 48.2	PBO (n = 87); 31.0
Infliximab trials			
ACT 1	Induction	IFX 5mg/kg (n = 47/121); 38.8	PBO (n = 18/121); 14.9
	Induction	IFX 10mg/kg (n = 39/122); 32.0	PBO (n = 18/121); 14.9
	Maintenance	IFX 5mg/kg (n = 42/121); 34.7	PBO (n = 20/121); 16.5
	Maintenance	IFX 10mg/kg (n = 42/122); 34.4	PBO (n = 20/121); 16.5
ACT 2	Induction	IFX 5mg/kg (n = 41/121); 33.9	PBO (n = 7/123); 5.7
	Induction	IFX 10mg/kg (n = 33/120); 27.5	PBO (n = 7/123); 5.7
Japic CTI060298	Induction	IFX 5mg (n = 21/104); 20.2	PBO (n = 11/104); 10.6
Jiang 2015	Induction	IFX 3.5mg/kg (n = 21/41); 51.2	PBO (n = 9/41); 21.9
	Induction	IFX 5mg/kg (n = 22/41); 53.7	PBO (n = 9/41); 21.9
Probert 2003	Induction	IFX 5mg (n = 9/23); 39.0	PBO (n = 6/20); 30.0
Adalimumab trials			
ULTRA 1	Induction	ADA 80/40mg (n = 130); 10.0	PBO (n = 130); 9.2
	Induction	ADA 160/80mg (n = 130); 18.5	PBO (n = 130); 9.2
	Maintenance	ADA 40mg EOW (n = 390); 25.6	NA
ULTRA 2	Induction	ADA 160/80/40mg (n = 150); 21.3	PBO (n = 145); 11.0
	Maintenance	ADA 40mg EOW (n = 150); 22.0	PBO (n = 145); 12.4

NCT00853099	Induction	ADA 80/40mg (n = 87); 14.0	PBO (n = 96); 11.0
	Induction	ADA 160/80mg (n = 90); 10.0	PBO (n = 96); 11.0
	Maintenance	ADA 40mg EOW (n = 177); 23.0	PBO (n = 96); 7.0
VARSITY	Maintenance	VDZ 300mg (n = 304); 34.2	NA
	Maintenance	ADA 160/80/40mg (n = 305); 24.3	NA
Golimumab trials			
PURSUIT-M	Maintenance	GOL 50mg (n = 151); 33.1	PBO (n = 154); 22.1
	Maintenance	GOL 100mg (n = 151); 33.8	PBO (n = 154); 22.1
PURSUIT-J	Induction	GOL 200 mg (n = 27/144); 18.8	NA
	Maintenance	GOL 100mg (n = 16(32); 50.0	PBO (n = 2/31); 6.5
PURSUIT-SC	Induction	GOL 200/100mg (n = 45/253); 17.8	PBO (n = 16/251); 6.4
	Induction	GOL 400/200mg (n = 46/257); 17.9	PBO (n = 16/251); 6.4
	Maintenance	GOL 200/100mg	PBO
	Maintenance	GOL 400/200mg	PBO
Vedolizumab trials			
GEMINI	Induction	VDZ 300 mg (n = 30/130); 23.1	PBO (n = 5/76); 6.6
	Maintenance	VDZ 300 mg (n = 68/145); 46.9	PBO (n = 15/79); 19.0
NCT02039505	Induction	VDZ 300 mg (n = 79); 27.8	PBO (n = 41); 14.6
	Maintenance	VDZ 300 mg (n = 24); 54.2	PBO (n = 28); 35.7
Kobayashi 2018	Induction	VDZ 150mg	PBO
	Induction	VDZ 300mg	PBO
VARSITY	Maintenance	VDZ 300mg (n = 304); 34.2	NA
	Maintenance	ADA 160/80/40mg (n = 305); 24.3	NA
Tofacitinib trials			
NCT00787202	Induction	TOC 10mg ;	PBO ;
	Induction	TOC 15mg	PBO ;

OCTAVE I1+I2	Induction	TOC 10mg (n = 440); 24.1	PBO (n = 110); 11.8
OCTAVE I1	Induction	TOC 10mg (n = 222); 25.2	PBO (n = 57); 15.8
OCTAVE I2	Induction	TOC 10mg (n = 195); 22.1	PBO (n = 47); 8.5
OCTAVE – Sustain	Maintenance	TOC 5mg (n = 115); 41.7	PBO (n = 79); 11.0
	Maintenance	TOC 10mg (n = 104); 44.2	PBO (n = 79); 11.0

Abbreviations: ADA=adalimumab; GOL=golimumab; IFX=infliximab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo; NA=Not Applicable

Table A14. Results summary for clinical remission for the biologic failure population

Study reference/ID	Induction/Maintenance	Outcome intervention (n/N); %	Outcome Comparator (n/N); %
Ustekinumab trials (technology)			
UNIFI	Induction	UST 130mg (n = 19/164); 11.6	PBO (n = 2/161); 1.2
	Induction	UST 6mg/kg (n = 22/166); 12.7	PBO (n = 2/161); 1.2
	Maintenance	UST 90mg sc Q12W (n = 70); 22.9	PBO (n = 88); 17.0
	Maintenance	UST 90mg sc Q8W (n = 91); 39.6	PBO (n = 88); 17.0
Adalimumab trials			
ULTRA 2	Induction	ADA 160/80/40mg (n = 98); 9.2	PBO (n = 101); 6.9
	Maintenance	ADA 40mg EOW (n = 98); 10.2	PBO (n = 101); 3.0
VARSITY	Maintenance	VDZ 300mg (n = 79); 20.3	NA
	Maintenance	ADA 160/80/40mg (n = 81); 16.0	NA
Vedolizumab trials			
GEMINI	Induction	VDZ 300 mg (n = 28/82); 9.8	PBO (n = 2/63); 3.2
	Maintenance	VDZ 300 mg (n = 30/83); 36.1	PBO (n = 2/38); 5.3
NCT02039505	Induction	VDZ 300 mg (n = 85); 9.4	PBO (n = 41); 9.8
	Maintenance	VDZ 300 mg (n = 17); 58.8	PBO (n = 14); 21.4
VARSITY	Maintenance	VDZ 300mg (n = 79); 20.3	NA
	Maintenance	ADA 160/80/40mg (n = 81); 16.0	NA
Tofacitinib trials			
OCTAVE I1+I2	Induction	TOC 10mg (n = 465); 11.4	PBO (n = 124); 0.8
OCTAVE – Sustain	Maintenance	TOC 5mg (n = 83); 24.1	PBO (n = 89); 11.2
	Maintenance	TOC 10mg (n = 93); 36.6	PBO (n = 89); 11.2

Abbreviations: ADA=adalimumab; GOL=golimumab; IFX=infliximab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo; NA=Not Applicable

Table A15. Results summary for mucosal healing for the non-biologic failure population

Study reference/ID	Induction/Maintenance	Outcome intervention (n/N); %	Outcome Comparator (n/N); %
Ustekinumab trials (technology)			
UNIFI	Induction	UST 130mg (n = 156); 34.6	PBO (n = 158); 20.8
	Induction	UST 6mg/kg (n = 156); 33.3	PBO (n = 158); 20.8
	Maintenance	UST 90mg Q12W (n = 102); 55.9	PBO (n = 87); 34.5
	Maintenance	UST 6mg/kg Q8W (n = 85); 57.6	PBO (n = 87); 34.5
Infliximab trials			
ACT 1	Induction	IFX 5mg/kg (n = 75/121); 62.0	PBO (n = 41/123); 33.9
	Induction	IFX 10mg/kg (n = 120); 59.0	PBO (n = 41/123); 33.9
	Maintenance	IFX 5mg/kg (n = 55/121); 45.5	PBO (n = 22/121); 18.2
	Maintenance	IFX 10mg/kg (n = 57/122); 46.7	PBO (n = 22/121); 18.2
ACT 2	Induction	IFX 5mg/kg (n = 73/121); 60.3	PBO (n = 38/123); 30.9
	Induction	IFX 10mg/kg (n = 74/120); 61.7	PBO (n = 38/123); 30.9
Japic CTI060298	Induction	IFX 5mg (n = 48/104); 46.2	PBO (n = 29/104); 27.9
Jiang 2015	Induction	IFX 3.5mg/kg (n = 23/41); 56.1	PBO (n = 10/41); 24.4
	Induction	IFX 5mg/kg (n = 24/41); 58.5	PBO (n = 10/41); 24.4
Adalimumab trials			
ULTRA 1	Induction	ADA 80/40mg (n = 130); 37.7	PBO (n = 130); 41.5
	Induction	ADA 160/80mg (n = 130); 46.9	PBO (n = 130); 41.5
	Maintenance	ADA 40mg EOW (n = 309); 54.0	NA
ULTRA 2	Induction	ADA 160/80/40mg (n = 150); 49.3	PBO (n = 145); 35.2
	Maintenance	ADA 40mg EOW (n = 150); 31.3	PBO (n = 145); 19.3
NCT00853099	Induction	ADA 80/40mg (n = 87); 39.0	PBO (n = 96); 30.0
	Induction	ADA 160/80mg (n = 90); 44.0	PBO (n = 96); 30.0

	Maintenance	ADA 40mg EOW (n = 177); 29.0	PBO (n = 96); 16.0
VARSITY	Maintenance	VDZ 300mg (n = 304); 43.1	NA
	Maintenance	ADA 160/80/40mg (n = 305); 29.5	NA
Golimumab trials			
PURSUIT-M	Maintenance	GOL 50mg (n = 151); 41.7	PBO (n = 154); 53.2
	Maintenance	GOL 100mg (n = 151); 42.4	PBO (n = 154); 53.2
PURSUIT-J	Induction	GOL 200 mg (n = 53/144); 36.8	NA
	Maintenance	GOL 100mg (n = 19/32); 59.4	PBO (n = 5/31); 16.1
PURSUIT-SC	Induction	GOL 200/100mg (n = 107/253); 42.3	PBO (n = 72/251); 28.7
	Induction	GOL 400/200mg (n = 116/257); 45.1	PBO (n = 16/251); 28.7
Vedolizumab			
GEMINI	Induction	VDZ 300 mg (n = 64/130); 49.2	PBO (n = 19/76); 25.0
	Maintenance	VDZ 300 mg (n = 87/145); 60.0	PBO (n = 19/79); 24.1
NCT02039505	Induction	VDZ 300 mg (n = 79); 48.1	PBO (n = 41); 31.7
	Maintenance	VDZ 300 mg (n = 24); 62.5	PBO (n = 28); 35.7
VARSITY	Maintenance	VDZ 300mg (n = 304); 43.1	NA
	Maintenance	ADA 160/80/40mg (n = 305); 29.5	NA
Tofacitinib trials			
OCTAVE I1+I2	Induction	TOC 10mg (n = 440); 38.2	PBO (n = 110); 21.8
OCTAVE I1	Induction	TOC 10mg (n = 222); 39.6	PBO (n = 57); 25.3
OCTAVE I2	Induction	TOC 10mg (n = 195); 36.4	PBO (n = 47); 19.1
OCTAVE – Sustain	Maintenance	TOC 5mg (n = 115); 42.6	PBO (n = 109); 13.8
	Maintenance	TOC 10mg (n = 104); 51.0	PBO (n = 109); 13.8

Abbreviations: ADA=adalimumab; GOL=golimumab; IFX=infliximab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo; NA=Not Applicable

Table A16. Results summary for mucosal healing in the biologic failure population

Study reference/ID	Induction/Maintenance	Outcome intervention (n/N); %	Outcome Comparator (n/N); %
Ustekinumab trials (technology)			
UNIFI	Induction	UST 130mg (n = 164); 18.3	PBO (n = 160); 6.9
	Induction	UST 6mg/kg (n = 166); 21.1	PBO (n = 160); 6.9
	Maintenance	UST 90mg Q12W (n = 70); 25.7	PBO (n = 88); 22.7
	Maintenance	UST 6mg/kg Q8W (n = 91); 45.1	PBO (n = 88); 22.7
Adalimumab trials			
ULTRA 2	Induction	ADA 160/80/40mg (n = 98); 28.6	PBO (n = 101); 26.7
	Maintenance	ADA 40mg EOW (n = 98); 15.3	PBO (n = 101); 9.9
VARSITY	Maintenance	VDZ 300mg (n = 79); 26.6	NA
	Maintenance	ADA 160/80/40mg (n = 81); 21.0	NA
Vedolizumab trials			
GEMINI	Induction	VDZ 300 mg (n = 25/82); 30.6	PBO (n = 13/63); 20.6
	Maintenance	VDZ 300 mg (n = 37/83); 44.6	PBO (n = 3/38); 7.9
NCT02039505	Induction	VDZ 300 mg (n = 85); 27.1	PBO (n = 41); 29.3
	Maintenance	VDZ 300 mg (n = 17); 64.7	PBO (n = 14); 28.6
VARSITY	Maintenance	VDZ 300mg (n = 79); 26.6	NA
	Maintenance	ADA 160/80/40mg (n = 81); 21.0	NA
Tofacitinib trials			
OCTAVE I1+I2	Induction	TOC 10mg (n = 465); 22.2	PBO (n = 124); 6.5
OCTAVE – Sustain	Maintenance	TOC 5mg (n = 83); 30.1	PBO (n = 89); 12.4
	Maintenance	TOC 10mg (n = 93); 39.8	PBO (n = 89); 12.4

Abbreviations: ADA=adalimumab; GOL=golimumab; IFX=infliximab; TOC=tofacitinib; UST=ustekinumab; VDZ=vedolizumab; PBO=placebo; NA=Not Applicable

A2.8. Results related to Quality of Life

All studies below are extracted by MAH.

Ustekinumab trials - Total IBDQ Score

UNIFI-Induction (MAH data) [60]

Clinically significant improvement from Baseline in Total IBDQ Score at Week 8

Clinically significant improvements in IBDQ from baseline in total IBDQ score at Week 8 was reached for both the ~6 mg/kg and 130 mg ustekinumab groups.

When considering a > 20-point improvement from baseline in total IBDQ score at Week 8, significantly greater proportions of subjects in the ~6 mg/kg and 130 mg ustekinumab groups had improvements (62.1% and 61.3%, respectively) compared with subjects in the placebo group (37.0%; $p < 0.001$ for both comparisons). This was important as > 20 point improvement is considered a clinically important objective end point for clinical improvement for UC patients.

When considering a ≥16-point improvement from baseline in total IBDQ score at Week 8, significantly greater proportions of subjects in the ~6 mg/kg and 130 mg ustekinumab groups had improvements (68.6% and 66.6%, respectively) compared with subjects in the placebo group (44.2%; $p < 0.001$ for both comparisons).

UNIFI- Maintenance (MAH data) [61]

Clinically significant improvement from Baseline in Total IBDQ Score at Week 44

Clinically significant improvements in IBDQ from induction baseline in total IBDQ score at Week 44 was reached for both the ustekinumab q8w and q12w groups. When considering a > 20-point improvement from baseline in total IBDQ score at Week 44, significantly greater proportions of subjects in the q8w and q12w groups had improvements (69.9% and 66.3%, respectively) compared with subjects in the placebo group (42.9%; $p < 0.001$ for both comparisons). When considering a ≥16-point improvement from induction baseline in total IBDQ score at Week 44, significantly greater proportions of subjects in the q8w and q12w groups had improvements (73.3% and 68.6%, respectively) compared with subjects in the placebo group (47.4%; $p < 0.001$ for both comparisons).

Table A17. IBDQ scores in UNIFI Induction trial (UST 130 mg, UST 6mg/kg, PBO) (colorless columns) and UNIFI Maintenance trial (UST 90 mg q12w, UST 90 mg q8w, PBO) (columns in grey)

IBDQ																					
Baseline score						Change from baseline score						Total score in induction/maintenance						≥20 point improvement		≥16 point improvement	
Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR	n	%	n	%
126	129	33,14	NR	NR	'05;149	33,4	31,5	32,53	NR	NR	7.5;53.5	159,2	164	37,16	NR	NR	47;220	196	61,3	213	66,6
127	126	33,27	NR	NR	104.0;152	35	31	31,86	NR	NR	11.0;56.0	161,9	166	35,64	NR	NR	34;224	200	62,1	221	68,6
127,4	126	3445	NR	NR	0.0155	16,1	10	31,39	NR	NR	-2.0;34.0	143,5	147	39,96	NR	NR	48;214	118	37	141	44,2
NR	NR	NR	NR	NR	NR	NR	1,5	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	66,3	NR	68,6
NR	NR	NR	NR	NR	NR	NR	5	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	69,9	NR	73,3
NR	NR	NR	NR	NR	NR	NR	-7	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	42,9	NR	47,4

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Table A18. SF-36 scores in UNIFI Induction trial (UST 130 mg, UST 6mg/kg, PBO) (colorless columns) and UNIFI Maintenance trial (UST 90 mg q12w, UST 90 mg q8w, PBO) (columns in grey)

SF-36																	
Baseline score (PCS)						Change from baseline score(PCS)						Baseline score (MCS)					
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	50	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR
43,1	43,7	7,85	NR	NR	37.7;48.5	4,7	4,4	6,49	NR	NR	O.2;8.9	40,1	40	10,85	NR	NR	32.4;49.2
43,1	43,6	7,73	NR	NR	38.1;49.0	5,2	4,1	6,16	NR	NR	1.0;8.9	40,5	40,7	10,59	NR	NR	33.0;48.3
43,6	43,5	7,96	NR	NR	20.0;60.0	2,1	1,6	6,39	NR	NR	-1.4;5.4	40,5	41,1	11,43	NR	NR	32.6;49.3
NR	NR	NR	NR	NR	NR	NR	0	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	1,4	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	-1	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

SF-36 (continued)													
Change from baseline score (MCS)						Total score in induction/maintenance						≥5 point improvement	
Mean	Mean	Mean	Mean	Mean	Mean	Mean	Median	SD	SE	95% CI	IQR	n	%
5,3	3,8	9,63	NR	NR	-0.9;10.9	NR	NR	NR	NR	NR	NR	154	48,3
5,1	3,6	9,72	NR	NR	-1.0;10.4	NR	NR	NR	NR	NR	NR	146	45,3
2,2	1	10,2	NR	NR	-4.0;7.2	NR	NR	NR	NR	NR	NR	83	26
NR	0,4	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	50
NR	0,8	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	53,4
NR	-1,2	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	30,3

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Ustekinumab trials - SF-36 Physical and Mental Component Scores

UNIFI- Induction (MAH data) *[60]*

Change from Baseline in the SF-36 Physical and Mental Component Scores at Week 8

At Week 8, median increases from baseline in the Physical Component Summary (PCS) scores were significantly greater for subjects in the ~ 6mg/kg and 130 mg ustekinumab groups (4.1 and 4.4, respectively compared with subjects in the placebo group (1.6; $p < 0.001$ for both comparisons).

At Week 8, median increases from baseline in the Mental Component Summary (MCS) scores were significantly greater for subjects in the ~6 mg/kg and 130 mg ustekinumab groups (3.6 and 3.8, respectively) compared with subjects in the placebo group (1.6; $p < 0.001$ for both comparisons).

UNIFI- Maintenance (MAH data) *[61]*

Improvement of ≥5-points from induction baseline in the SF-36 Physical and Mental Component Scores at Week 44

Among subjects with a ≥5-point improvement (from induction baseline) in the SF-36 PCS score at maintenance baseline, significantly greater proportions of the ustekinumab q8w and q12w groups maintained their ≥5-point improvement through maintenance Week 44 (62.4% and 59.5%, respectively) compared with subjects in the placebo group (38.3%, $p = 0.002$ and $p = 0.004$, respectively). In addition, significantly greater proportions of subjects in the ustekinumab q8w and q12w groups had a ≥5-point improvement from baseline in the SF-36 PCS score at Week 44 (53.4% and 50.0%, respectively) compared with subjects in the placebo group (30.3%; $p < 0.001$ for both comparisons).

Among subjects with a ≥5-point improvement (from induction baseline) in the SF-36 MCS score at maintenance baseline, significantly greater proportions of the ustekinumab q8w and q12w groups maintained their ≥5-point improvement through maintenance Week 44 (59.8% and 58.3%, respectively) compared with subjects in the placebo group (36.1%, $p = 0.001$ and $p = 0.002$, respectively). In addition, significantly greater proportions of subjects in the ustekinumab q8w and q12w groups had a ≥5-point improvement from baseline in the SF-36 MCS score at Week 44 (54.0% and 47.1%, respectively) compared with subjects in the placebo group (28.6%; $p < 0.001$ for both comparisons).

Ustekinumab trials - EQ-5D Index, EQ-5D Dimensions, and Health State VAS Scores

UNIFI- Induction (MAH data) *[60]*

Change from baseline in EQ-5D Index, EQ-5D Dimensions, and Health State VAS Scores at Week 8

At baseline, the median EQ-5D index and health state VAS scores were similar across all treatment groups. At Week 8, the median changes from baseline in EQ-5D and health state EQ-VAS were significantly greater for subjects in the ~ 6mg/kg and 130 mg ustekinumab groups compared with those in the placebo group ($p < 0.001$ for both comparisons, for both the EQ-5D index and health state VAS).

At baseline, the distributions for each of the five dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression were generally consistent across treatment groups. At Week 8, significantly greater proportions of subjects had improvement in the dimensions of usual activities, pain/discomfort and anxiety/depression for each ustekinumab group compared to placebo ($p \leq 0.002$). An improvement in the self-care dimension was also noted in the ~6 mg/kg group ($p = 0.004$) compared with the placebo group.

UNIFI- Maintenance (MAH data) *[61]*

Change from baseline in EQ-5D Index, EQ-5D Dimensions, and health state VAS scores through Week 44

At maintenance baseline, the median EQ-5D index and EQ-5D health state VAS scores were similar across all treatment groups. Over time through Week 44, the EQ-5D index and EQ-5D health state VAS scores were maintained for subjects in the ustekinumab q8w and q12w groups and decreased (worsened) for subjects in the placebo group.

At maintenance baseline, the distribution for each of the five dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression were generally consistent across treatment groups. Overall, at Week 44, greater proportions of subjects in the ustekinumab q8w and q12w groups reported EQ-5D dimension scores that improved or were maintained from maintenance baseline compared with subjects in the placebo group.

Table A19. EQ-5D scores in UNIFI Induction trial (UST 130 mg, UST 6mg/kg, PBO) (colorless columns) and UNIFI Maintenance trial (UST 90 mg q12w, UST 90 mg q8w, PBO) (columns in grey)

EQ-5D																	
Baseline score						Change from baseline score						Total score in induction/ maintenance					
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR
067	071	0204	NR	NR	0.60:0.77	009	006	0182	NR	NR	0.0:0.19	NR	NR	NR	NR	NR	NR
067	071	0195	NR	NR	0.64:0.77	011	006	0172	NR	NR	0.0:0.21	NR	NR	NR	NR	NR	NR
066	071	0208	NR	NR	0.56:0.77	004	001	0182	NR	NR	-0.02:0.13	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	-0019	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Table A20. EQ-5D VAS scores in UNIFI Induction trial(UST 130 mg, UST 6mg/kg, PBO) (colorless columns) and UNIFI Maintenance trial (UST 90 mg q12w, UST 90 mg q8w, PBO) (columns in grey)

EQ-5D VAS																	
Baseline score						Change from baseline score						Total score in induction/maintenance					
Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	1QR	Mean	Median	SD	SE	95% CI	IQR
54,14	55	20,545	NR	NR	40.0;70.0	13,64	10	20,394	NR	NR	0.0;20.5	NR	NR	NR	NR	NR	NR
55,76	55	19,333	NR	NR	40.0;70.0	13.51	10	18,447	NR	NR	0.0;25.0	NR	NR	NR	NR	NR	NR
55,11	60	20,815	NR	NR	40.0;70.0	5,71	5	19,584	NR	NR	-5.0;20.0	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	-5	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Infliximab trials

There is no QoL data in Japic CTI060298 and Jiang studies [2,48].

QoL data in ACT 1 and ACT2 **[**49,92]

Baseline scores for the pooled patient population indicated substantial impairment in HRQL. Improvement at week 8 in the total IBDQ score was significantly greater in the infliximab 5-mg/kg (40, *P* < 0.001) and 10-mg/kg (36, *P* < 0.001) groups compared with the placebo group. Improvement at week 8 was also significantly greater in the infliximab 5- and 10-mg/kg groups for the PCS (6.8 and 5.9, respectively) and MCS (5.9 and 6.4, respectively) compared with placebo (PCS = 3.7, MCS = 3.0, *P* < 0.01 for all comparisons). Continued benefit was seen at weeks 30 and 54 with infliximab maintenance therapy (*P* < 0.001 for all comparisons). Improvement in total IBDQ score correlated significantly (*P* < 0.001) with improvement in both PCS and MCS scores, and Mayo score.

Probert et al [3] showed that improvement in the IBDQ and EuroQol was not significantly different between the groups (p = 0.22 and 0.3, respectively, Mann-Whitney U test). Twenty eligible patients were given open labelled infusions. Remission was achieved in 3/11 (27%) patients initially treated with infliximab and in 1/9 (11%) patients treated with placebo.

Table A21. IBDQ scores in Infliximab trial (Probert, 2003; IFX 5mg/kg vs PBO) extracted by MAH for Measurement of Quality of life in UC [46]

IBDQ																					
Baseline score						Change from baseline score						Total score in induction/maintenance						≥20 point improvement		≥ 16 point improvement	
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR	n	%	n	%
127	NR	40	NR	NR	NR	36	NR	49	NR	NR	NR	163	NR	40	NR	NR	NR	NR	NR	NR	NR
114	NR	29	NR	NR	NR	25	NR	28	NR	NR	NR	139	NR	43	NR	NR	NR	NR	NR	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Table A22. EQ-5D scores in Infliximab trial (Probert, 2003; IFX 5mg/kg vs PBO) extracted by MAH for Measurement of Quality of life in UC [46]

EQ-5D																	
Baseline score						Change from baseline score						Total score in induction/maintenance					
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR
52	NR	16	NR	NR	NR	7	NR	17	NR	NR	NR	59	NR	19	NR	NR	NR
49	NR	17	NR	NR	NR	4	NR	16	NR	NR	NR	54	NR	23	NR	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Adalimumab trials

There is no QoL data in ULTRA 1 and in the VARSITY trial.

Significantly more adalimumab-treated patients were IBDQ responders (increase in IBDQ score ≥16 points from baseline) throughout Weeks 8, 32, and 52 compared with placebo. Mean changes from baseline IBDQ scores were consistently greater for adalimumab- vs. placebo-treated patients. In anti-TNF-naïve patients, the improvements from baseline for IBDQ were 34±38 and 22±37 at Week 8 (p = 0.002), 33±43 and 24±43 at Week 32 (p = 0.03), and 33±44 and 23±42 at Week 52 (p = 0.02) for ADA and placebo, respectively. The IBDQ responder rates were 68% and 52% at Week 8 (p = 0.004), 42% and 27% at Week 32 (p = 0.006), 32% and 21% at Week 52 (p = 0.040) for ADA and placebo, respectively, among anti-TNF-naïve patients [93].

Table A23. IBDQ scores in ULTRA 2 trial (ADA 40 mg EOW vs PBO, induction and maintenance)

IBDQ			
≥20 point improvement		**≥ 16 point improvement**	
n	**%**	**n**	**%**
NR	NR	112	45,5
NR	NR	144	58,1
NR	NR	40	16,3
NR	NR	65	26,2

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Table A24. IBDQ scores in NCT00853099 trial (ADA 80/40, ADA 160/80 and PBO in Induction; ADA 40 mgEOW and PBO in Maintenance)

IBDQ									
Baseline score						**≥20 point improvement**		**≥ 16 point improvement**	
Mean	**Median**	**SD**	**SE**	**95%CI**	**IQR**	**n**	**%**	**n**	**%**
144.9	NR	28.7	NR	NR	NR	NR	NR	42	48,3
146	NR	31.7	NR	NR	NR	NR	NR	38	42.2
148.2	NR	28.9	NR	NR	NR	NR	NR	38	39.6
NR	NR	NR	NR	NR	NR	NR	NR	45	25.4
NR	NR	NR	NR	NR	NR	NR	NR	12	12.5

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Golimumab trials

There is no QoL data in PURSUIT-J study.

Statistically significant improvements from baseline to week 6 were observed for the IBDQ total score and for each IBDQ domain score (bowel symptoms, emotional function, systemic symptoms and social function), as well as the EQ-5D index score and associated visual analogue scale score ($p < 0.0001$). Improvement of HRQoL was sustained through week 54 [94].

Table A25. IBDQ scores in PURSUIT-M trial (Maintenance GOL 50 mg, GOL 100 mg, PBO)

IBDQ									
Baseline score						≥20 point improvement		≥ 16 point improvement	
Mean	Median	SD	SE	95%CI	IQR	n	%	n	%
NR	NR	NR	NR	NR	NR	6	22.2	NR	NR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Table A26. IBDQ scores in PURSUIT-SC trial (Induction GOL 400/200 mg, PBO)

IBDQ					
Change from baseline score					
Mean	Median	SD	SE	95%CI	IQR
27.0	22.5	33.72	NR	NR	0.5;48.5
26.9	21.0	34.28	NR	NR	0;50
14.8	11.0	31.25	NR	NR	(-)3;29

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Vedolizumab trials

There is no QoL data in NCT02039505 and in the VARSITY studies.

Compared with placebo-treated patients, vedolizumab-treated patients had greater improvements (152-201%) in IBDQ, EQ-5D visual analogue scale (VAS), and EQ-5D utility scores. Greater proportions (6.9-19.9%) of vedolizumab-treated patients than placebo-treated patients met MCID thresholds for all the instruments. Vedolizumab-treated patients with lower baseline disease activity and those without prior tumour necrosis factor (TNF) antagonist failure had greater HRQL improvements. Among 127 patients with clinical remission based on complete Mayo Clinic scores, > 80% also had IBDQ remission; > 70% of the 150 patients with IBDQ remission demonstrated clinical remission. Vedolizumab therapy was associated with significant improvements in HRQL measures compared with placebo. Benefits were greater in patients with lower disease activity and no prior TNF antagonist failure [95].

Table A27. IBDQ scores in GEMINI trial (Induction VDZ 300 mg, PBO; Maintenance VDZ 300 mg q4w, VDZ 300 mg q8w, PBO)

IBDQ											
Baseline score						Change from baseline score					
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR
125	NR	35	NR	NR	NR	NR	NR	NR	NR	NR	NR
126	NR	34	NR	NR	NR	NR	NR	NR	NR	NR	NR
123,7	NR	34	NR	NR	NR	NR	NR	NR	3,3	NR	NR
124,5	NR	34	NR	NR	NR	NR	NR	NR	3,4	NR	NR
122,2	NR	34							3,3	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Table A28. SF-36 scores in GEMINI trial (Induction VDZ 300 mg, PBO; Maintenance VDZ 300 mg q4w, VDZ 300 mg q8w, PBO)

SF-36																	
Baseline score (PCS)						Change from baseline score (PCS)						Baseline score (MCS)					
Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR
NR	NR	NR	NR	NR	NR	29	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	11	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
41,1	NR	7,7	3,1	NR	NR	49	NR	NR	NR	NR	NR	37,9	NR	11,2	NR	NR	NR
40	NR	8,5	3,1	NR	NR	48,4	NR	NR	NR	NR	NR	39,2	NR	11,6	NR	NR	NR
39,7	NR	8,4	3,0	NR	NR	27,3	NR	NR	NR	NR	NR	38,5	NR	12	NR	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Table A29. ED-5Q scores in GEMINI trial (Induction VDZ 300 mg, PBO; Maintenance VDZ 300 mg q4w, VDZ 300 mg q8w, PBO)

EQ-5D																	
Baseline score						Change from baseline score						Total score in induction/maintenance					
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
0,674	NR	0,227	0,02	NR	NR	0,141	NR	NR	0,019	NR	NR	NR	NR	NR	NR	NR	NR
0,673	NR	0,235	0,021	NR	NR	0,131	NR	NR	0,019	NR	NR	NR	NR	NR	NR	NR	NR
0,677	NR	0,213	0,019	NR	NR	0,083	NR	NR	0,019	NR	NR	NR	NR	NR	NR	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Table A30. ED-5Q VAS scores in GEMINI trial (Induction VDZ 300 mg, PBO; Maintenance VDZ 300 mg q4w, VDZ 300 mg q8w, PBO)

EQ-5D VAS																	
Baseline score						Change from baseline score						Total score in induction/maintenance					
Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	1QR	Mean	Median	SD	SE	95% CI	IQR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
53,6	NR	20,3	1,8	NR	NR	19,4	NR	NR	1,7	NR	NR	NR	NR	NR	NR	NR	NR
56,6	NR	20,9	1,9	NR	NR	19	NR	NR	1,7	NR	NR	NR	NR	NR	NR	NR	NR
54,6	NR	20,2	1,8	NR	NR	9,7	NR	NR	1,7	NR	NR	NR	NR	NR	NR	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Tofacitinib trials

In OCTAVE Induction 1 and 2, mean changes from baseline IBDQ were greater with tofacitinib 10 mg BID at Week 8 [28.9 and 31.5] versus placebo [15.4 and 17.2; p < 0.0001]; mean changes from baseline SF-36v2 Physical and Mental Component Summaries [PCS/MCS] were also greater with 10 mg BID [PCS: 6.8 and 6.8; MCS: 6.8 and 7.6] versus placebo [PCS: 2.5 and 4.6; MCS: 3.5 and 4.4; p < 0.01]. In OCTAVE Sustain atWeek 52, changes in IBDQ were maintained with tofacitinib 5 mg [-1.3] and 10 mg BID [0.6], and larger with placebo [-20.2; p < 0.0001]. Changes in SF-36v2 PCS/MCS were also maintained with 5 mg [PCS: 0.0; MCS: -1.0] and 10 mg BID [PCS: 0.3; MCS: 0.1] versus placebo [PCS: -5.2; MCS: -6.7; p < 0.0001] at Week 52 in OCTAVE Sustain.Tofacitinib 10 mg BID induction therapy significantly improved HRQoL versus placebo at Week 8. Improvements were maintained through 52 weeks' maintenance therapy with tofacitinib 5 mg and 10 mg BID [74,75].

Table A31. IBDQ scores in OCTAVE-I1 trial (Induction TOC 10 mg, PBO)

IBDQ											
Baseline score						Change from baseline score					
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR
126	NR	27.1	NR	NR	NR	28.9	NR	NR	1.2	NR	NR
127	NR	27.9	NR	NR	NR	15.4	NR	NR	2.2	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Table A32. SF-36 scores in OCTAVE-I1 trial (Induction TOC 10 mg, PBO)

SF-36																							
Baseline score (PCS)						Change from baseline score (PCS)						Baseline score (MCS)						Change from baseline score (MCS)					
Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR
41.2	NR	8.3	NR	NR	NR	6.8	NR	NR	0.3	NR	NR	39	NR	12	NR	NR	NR	6.8	NR	NR	0.5	NR	NR
41.5	NR	8	NR	NR	NR	2.5	NR	NR	0.6	NR	NR	38.7	NR	12	NR	NR	NR	3.5	NR	NR	0.9	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Table A33. IBDQ scores in OCTAVE-I2 trial (Induction TOC 10 mg, PBO)

IBDQ											
Baseline score						Change from baseline score					
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR
123,7	NR	24.4	NR	NR	NR	31.5	NR	NR	1.4	NR	NR
120,1	NR	23.2	NR	NR	NR	17.2	NR	NR	2.5	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Table A34. SF-36 scores in OCTAVE-I2 trial (Induction TOC 10 mg, PBO)

SF-36																							
Baseline score (PCS)						Change from baseline score (PCS)						Baseline score (MCS)						Change from baseline score (MCS)					
Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR
40.5	NR	8.2	NR	NR	NR	6.8	NR	NR	0.4	NR	NR	37.8	NR	11.2	NR	NR	NR	7.6	NR	NR	0.5	NR	NR
40.2	NR	7.6	NR	NR	NR	4.6	NR	NR	0.7	NR	NR	38.3	NR	11.2	NR	NR	NR	4.4	NR	NR	1	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Table A35. IBDQ scores in OCTAVE-Sustain trial (Maintenance TOC 5 mg, TOC 10 mg, PBO)

IBDQ											
Baseline score						Change from baseline score					
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR
167.4	NR	22.8	NR	NR	NR	-1.3	NR	NR	2.3	NR	NR
167.7	NR	21.3	NR	NR	NR	0.6	NR	NR	2.3	NR	NR
166.7	NR	21.5	NR	NR	NR	-20.2	NR	NR	2.9	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

Table A36. SF-36 scores in OCTAVE-Sustain trial (Maintenance TOC 5 mg, TOC 10 mg, PBO)

SF-36																							
Baseline score (PCS)						Change from baseline score (PCS)						Baseline score (MCS)						Change from baseline score (MCS)					
Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR	Mean	Median	SD	SE	95% CI	IQR
50.5	NR	6.8	NR	NR	NR	0	NR	0.8	NR	NR	NR	49	NR	9.3	NR	NR	NR	-1	NR	NR	1	NR	NR
49.3	NR	7.1	NR	NR	NR	0.3	NR	0.7	NR	NR	NR	48.9	NR	9.6	NR	NR	NR	0.1	NR	NR	1	NR	NR
50	NR	7.2	NR	NR	NR	-5.2	NR	0.9	NR	NR	NR	47.8	NR	10.6	NR	NR	NR	-6.7	NR	NR	1.2	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range

Table A37. IBDQ scores in NCT00787202 trial (Induction TOC 10 mg, TOC 15 mg, PBO)

IBDQ											
Baseline score						Change from baseline score					
Mean	Median	SD	SE	95%CI	IQR	Mean	Median	SD	SE	95%CI	IQR
134.5	NR	32.5	NR	NR	NR	30.4	NR	NR	39.8	NR	NR
124	NR	34.9	NR	NR	NR	50.7	NR	NR	35.6	NR	NR
123.2	NR	29.5	NR	NR	NR	27.8	NR	NR	29.8	NR	NR

Abbreviations: SD=standard deviation; SE=standard error; CI=confidence interval; IQR=interquartile range; IBDQ=Inflammatory Bowel Disease Questionnaire

A2.9. Safety enpoints analysed in the SLR and NMA with results according the infliximab, adalimumab, golimumab, vedolizumab and tofacitinib trials

Table A38. Methods of data collection and analysis of overall adverse events

Study reference /ID	Endpoint definition	Method of analysis
All trials	All safety endpoints event	Adverse events were either voluntarily reported by the subject or were obtained by means of interviewing subjects in a nondirected manner at study visits.

Table A39. Overview of adverse events: Infliximab trials [2,3,48,49]

Safety endpoint	ACT 1 Non-biologic failure Maintenance (Week 54)			ACT 2 Non-biologic failure Maintenance (Week 30)			Jiang 2015 Non-biologic failure Maintenance (Week 8)			Japic CTI060298 Non-biologic failure			
										Induction (Week 8)		Maintenance (Week 8)	
	PBO (N = 121)	IFX 5mg/kg (N = 121)	IFX 10mg/kg (N = 122)	PBO (N = 123)	IFX 5mg/kg (N = 121)	IFX 10mg/kg (N = 120)	PBO (N = 41)	IFX 3.5mg/kg (N = 41)	IFX 5mg/kg (N = 41)	PBO (N = 104)	IFX 5mg/kg (N = 104)	PBO (N = 104)	IFX 5mg/kg (N = 104)
Adverse events, n (%)	103 (85.1)	106 (87.6)	111 (91)	90 (73.2)	99 (81.8)	96 (80)	16 (39)	16 (39)	17 (41.5)	86 (82.7)	85 (81.7)	94 (90.4)	100 (96.2)
Serious adverse events, n (%)	31 (25.8)	26 (21.5)	29 (23.8)	24 (19.5)	13 (10.7)	11 (9.2)	NR	NR	NR	13 (12.5)	9 (8.7)	19 (18.3)	18 (17.3)
Any infection, n (%)	47 (38.8)	53 (43.8)	60 (49.2)	29 (23.6)	33 (27.3)	34 (28.3)	5 (12.2)	5 (12.2)	6 (14.6)	35 (33.7)	33 (31.7)	51 (49)	62 (59.6)
Serious infection, n (%)	5 (4.1)	3 (2.5)	8 (6.6)	1 (0.8)	2 (1.7)	3 (2.5)	0 (0)	0 (0)	1 (2.4)	2 (1.9)	1 (1)	2 (1.9)	1 (1.1)
Adverse events leading to discontinuation, n (%)	11 (9.1)	10 (8.3)	11 (9)	12 (9.8)	2 (1.7)	5 (4.2)	2 (4.9)	0 (0)	1 (2.4)	8 (7.7)	5 (4.8)	8 (7.7)	7 (6.7)

Abbreviations: IFX=infliximab; PBO=placebo; NR=not reported

Table A40. Overview of adverse events: Adalimumab trials [1,52,82]

Safety endpoint	ULTRA 1 Non-biologic failure Maintenance (Week 52)	NCT00853099 Full patient population Induction (Week 8)		
	Any ADA (N = 557)	PBO (N = 96)	ADA 80/40mg (N = 87)	ADA 160/80mg (N = 90)
Adverse events, n (%)	NR	NR	NR	NR
Serious adverse events, n (%)	NR	NR	NR	NR
Any infection, n (%)	NR	15 (15.6)	11 (12.6)	17 (18.9)
Serious infection, n (%)	NR	0 (0)	0 (0)	3 (3.3)
Adverse events leading to discontinuation, n (%)	78 (14.0)	2	1	3

Abbreviations: ADA=adalimumab; PBO=placebo; NR=not reported

Table A41. Overview of adverse events: Golimumab trials [55]

Safety endpoint	PURSUIT-SC Non-biologic failure Induction (Week 6)				PURSUIT-J Non biologic failure				PURSUIT-M Non biologic failure Maintenance (Week 54)					
					Maintenance (double-blind; Week 68)		Maintenance (open-label ; Week 68)	Induction (Week 68)						
	PBO (N = 126)	GOL 100/50mg (N = 71)	GOL 200/100mg (N = 331)	GOL 400/200 mg (N = 332)	PBO (N = 31)	GOL 100mg (N = 32)	GOL 100mg (N = 60)	GOL 200mg (N = 144)	PBO (N = 156)	GOL 50mg (N = 154)	GOL 100mg (N = 154)	PBO/GOL NR/100mg (N = 76)	GOL 50mg/100mg (N = 25)	GOL 100mg/200mg (N = 14)
Adverse events, n (%)	126 (38.2)	34 (47.9)	37.5 (124)	38.9 (129)	22 (71)	31 (96.9)	47 (78.3)	65 (45.1)	103 (66)	112 (72.7)	113 (73.4)	54 (71.1)	16 (64)	9 (64.3)
Serious adverse events, n (%)	20 (6.1)	2(2.8)	9(2.7)	11(3.3)	4 (12.9)	1 (3.1)	6 (10)	5 (3.5)	12 (7.7)	13 (8.4)	22 (14.3)	8 (10.5)	5 (20)	1 (7.1)
Any infection, n (%)	40 (12.1)	8 (11.3)	39 (11.8)	41 (12.3)	11 (35.5)	21 (65.6)	26 (43.3)	NR	44 (28.2)	60 (39)	60 (39)	26 (34.2)	10 (40)	4 (0)
Serious infection, n (%)	6 (1.8)	0 (0)	1 (0.3)	3 (0.9)	NR	NR	NR	NR	3 (1.9)	5 (3.2)	5 (3.2)	1 (1.3)	1 (4)	0 (0)
Adverse events leading to discontinuation, n (%)	3 (0.9)	2 (2.8)	1 (0.3)	1 (0.3)	NR	NR	NR	10 (6.9)	10 (6.4)	8 (5.2)	14 (9.1)	8 (10.5)	4 (16)	0 (0)

Abbreviations: NR=Not Reported; GOL=golimumab; PBO=placebo

Table A42. Overview of adverse events: Vedolizumab trials [57,58,59]

Safety endpoint	GEMINI 1 Full patient population Induction (Week 52)			GEMINI 1 Full patient population Maintenance (Week 8)			NCT02039505 Full patient population Induction (Week 6)			NCT02039505 Full patient population Maintenance (Week 52)		Kobayashi Full patient popula-tion Maintenance (Week 34; 239 days)	
	PBO (N = 149)	VDZ 300mg cohort 1 patients (N = 225)	VDZ 300mg cohort 2 patients (N = 521)	PBO (N = 126)	VDZ 300mg Q8w (N = 122)	VDZ 300mg Q4w (N = 125)	PBO (N = 82)	VDZ cohort 1 (N = 164)	VDZ cohort 2 (N = 46)	PBO (N = 42)	VDZ (N = 41)	VDZ 150mg (n = 4)	VDZ 300mg (n = 6)
Adverse events, n (%)	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Serious adverse events, n (%)	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Any infection, n (%)	22 (15)	31 (14)	71 (14)	89 (71)	87 (71)	90 (72)	NR	NR	NR	NR	NR	0	4 (66.7)
Serious infection, n (%)	3 (2)	1 (< 1)	3 (< 1)	4 (3)	3 (2)	2 (2)	NR	NR	NR	NR	NR	NR	NR
Adverse events leading to discontinuation, n (%)	4	0	7	15	7	6	8 (4.9)	2 (2.4)	6 (13.0)	2 (4.9)	6 (14.3)	0	0

Abbreviations: NR=Not Reported; VDZ=vedolizumab; PBO=placebo

Table A43. Overview of adverse events: Tofacitinib trials [4,51]

Safety endpoint	OCTAVE-I1 Full patient population Induction (Week 8)		OCTAVE-I2 Full patient population Induction (Week 8)		OCTAVE-Sustain Full patient population Maintenance (Week 52)			NCT00787202 Full patient population Induction (Week 12)				
	PBO (N = 122)	TOC 10mg (N = 476)	PBO (N = 112)	TOC 10mg (N = 429)	PBO (N = 198)	TOC 5mg (N = 198)	TOC 10mg (N = 196)	PBO (N = 48)	TOC 0.5mg (n = 31)	TOC 3mg (n = 33)	TOC 10mg (n = 33)	TOC 15mg (n = 49)
Adverse events, n (%)	73 (59.8)	269 (56.5)	59 (52.7)	232 (54.1)	149 (75.3)	143 (72.2)	156 (79.6)	NR	NR	NR	NR	NR
Serious adverse events, n (%)	5 (4.1)	16 (3.4)	9 (8)	18 (4.2)	13 (6.6)	10 (5.1)	11 (5.6)	NR	NR	NR	NR	NR
Any infection, n (%)	19 (15.6)	111 (23.3)	17 (15.2)	78 (18.2)	48 (24.2)	71 (35.9)	78 (39.8)	7 (15)	8 (26)	3 (9)	9 (27)	3 (6)
Serious infection, n (%)	0 (0)	6 (13)	0 (0)	1 (0.2)	2 (1)	2 (1)	1 (0.5)	NR	NR	NR	NR	NR
Adverse events leading to discontinuation, n (%)	2 (1.6)	18 (3.8)	8 (7.1)	17 (4)	37 (18.7)	18 (9.1)	19 (9.7)	4 (8)	2 (6)	0	1 (3)	2 (4)

Abbreviations: NR=Not Reported; TOC=tofacitinib; PBO=placebo

APPENDIX 3: EVIDENCE GAPS

Table A44. Recommendations for research

Research question: What is the relative clinical effectiveness and safety of ustekinumab in head-to-head comparisons, in patients with moderately to severely active ulcerative colitis? **Rationale:** No direct comparison of ustekinumab and any biological therapy for UC is available.	
Evidence	Direct evidence is needed
Population	Patients with moderately to severely active ulcerative colitis (subgroups according to biological failure status)
Intervention	Ustekinumab
Comparator	Adalimumab, golimumab, infliximab, vedolizumab, tofacitinib
Outcome(s)	Clinical response, clinical remission, endoscoping healing, QoL, hospitalizations and UC related surgeries, AEs and SAEs
Time stamp	Beyond 1 year
Burden of disease	High
Study design	Pragmatic RCTs and prospective cohor studies