CADTH

CADTH COMMON DRUG REVIEW

Clinical Review Report

USTEKINUMAB (STELARA/STELARA I.V.)

Janssen Inc.

Indication: Treatment of adult patients with moderately to severely active ulcerative colitis who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a biologic or have medical contraindications to such therapies.

Service Line:	CADTH Common Drug Review
Version:	Final (with redactions)
Publication Date:	September 2020
Report Length:	121 Pages

CADTH

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

CADTH

Table of Contents

Tables

Figures

Abbreviations

AE	adverse event
CDR	CADTH Common Drug Review
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CrI	credible interval
CRP	C-reactive protein
EQ-5D	EuroQol 5-Dimensions
EQ-5D-3L	EuroQol 5-Dimensions 3-Levels
HRQoL	health-related quality of life
IBD	inflammatory bowel disease
IBDQ	Inflammatory Bowel Disease Questionnaire
ICC	intraclass correlation
ITC	indirect treatment comparison
ITT	intention to treat
MCID	minimal clinically important difference
MCS	Mental Component Summary
NMA	network meta-analysis
OR	odds ratio
PCS	Physical Component Summary
PGA	Physician's Global Assessment
RCT	randomized controlled trial
SC	subcutaneous
SF-36	Short Form (36) Health Survey
TNF	tumour necrosis factor
UC	ulcerative colitis
VAS	Visual Analogue Scale
WPAI-GH	Work Productivity and Activity Questionnaire – General Health

Drug	Ustekinumab (Stelara/Stelara I.V.)
Indication	Treatment of adult patients with moderately to severely active ulcerative colitis who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a biologic or have medical contraindications to such therapies.
Reimbursement request	As per indication.
Dosage form(s) and route of administration)/strength(s)	Solution for intravenous infusion, 130 mg/26 mL (5 mg/mL) and solution for subcutaneous injection, 90 mg/1.0 mL.
NOC date	January 23, 2020
Sponsor	Janssen, Inc.

Executive Summary

Introduction

Ulcerative colitis (UC) is a chronic inflammatory bowel disease (IBD) that involves inflammation of the intestinal mucosae affecting the rectum and variable levels of proximal extension into the colon. Age of onset of signs or symptoms is typically less than 30 years. It has a worldwide distribution, with a global incidence of 1.2 to 20.3 cases per 100,000 persons per year, and a prevalence of 7.6 to 246.0 cases per 100,000 per year. Canada is among the countries with the highest incidence and prevalence of IBD, with approximately 270,000 Canadians living with UC or Crohn disease. The incidence of UC ranges in different Canadian provinces from 8.4 to 21.4 per 100,000 people.

Ustekinumab is a human monoclonal antibody affecting the interleukin pathways in the pathogenesis of IBD and other immune-modulated conditions. It is approved by Health Canada for the treatment of adults with chronic moderate-to-severe plaque psoriasis and for the treatment of adult patients with active psoriatic arthritis, and Crohn disease. The current indication under review is for the treatment of adult patients with moderately to severely active UC who have had an inadequate response to, lost response to, or were intolerant to either conventional therapy or a biologic, or have medical contraindications to such therapies. The recommended dosage for ustekinumab in the treatment of UC is a single weight-based IV infusion (approximating 6 mg/mL) followed by a 90 mg subcutaneous (SC) dose eight weeks later, then 90 mg SC every eight weeks thereafter. Consideration should be given to discontinuing treatment in patients who show no evidence of therapeutic benefit 16 weeks after the IV induction dose. This drug has been previously reviewed by CADTH through the CADTH Common Drug Review (CDR) process for each of the Health Canada–approved indications.

This review aims to evaluate the beneficial and harmful effects of ustekinumab IV infusion (induction phase) and SC injection (maintenance phase) for the treatment of adult patients with moderately to severely active UC who have failed or were intolerant to treatment with immunomodulators or corticosteroids — but never failed treatment with a biologic — or have failed or were intolerant to treatment with a biologic.

Stakeholder Engagement

Patient Input

- Two patient groups answered CADTH's call for patient input: Crohn's and Colitis Canada and the Gastrointestinal Society. Both entities aim to support research for IBDs and improve the lives of adults and children with UC by providing support and information about treatments, research, and quality-of-life issues, while working closely with health care professionals, the government, and other patient groups. They use different channels (newsletters, group meetings, lectures, and websites in both English and French) to inform those who have been recently diagnosed or have been living with UC or another gastrointestinal-related condition for years.
- Both groups described the circumstances of living with an IBD and what patients have to endure: how UC represents a disabling, lifelong gastrointestinal condition that primarily affects working-age individuals. Symptoms associated with UC, such as bloody diarrhea, bloating, abdominal pain, cramping, and fatigue, affect their day-to-day lives, sometimes causing them to experience isolation, anxiety, and debilitating, frequent, and urgent bowel movements. Their quality of life is deeply affected during periods of active disease, with patients spending a lot of time in the bathroom; even in periods of remission, patients have to stay near a bathroom. UC forces them to limit their activities, sometimes because of the stigma associated with an IBD. Both patient groups described the concerns expressed by patients about future flares, which sometimes worsen and are unpredictable.
- Patients often seek treatment options that can reduce or eliminate their symptoms and regularly long for treatments that could protect their ability to work, attend school and social events, and perform basic day-to-day activities. The patient groups reported that many current treatments can have undesirable effects because they must be used long term (e.g., glucocorticoids) and that individuals with UC are continuously struggling for a normal life. They require new and effective options to achieve mucosal healing and decrease debilitating symptoms. According to Crohn's and Colitis Canada, patients preferred drugs that are convenient and easy to use. For instance, most patients were pleased with not having to travel to a clinic to administer ustekinumab, providing them with the opportunity for a normal life. Given that all individuals respond differently to therapies, it was considered imperative that patients have a variety of options for treatment.

Clinician Input

- According to advice obtained from a clinical expert, the goal of medical treatment for moderate-to-severe UC should include endoscopic and histologic healing of colonic inflammation. Patients who do not achieve both clinical and endoscopic remission will be at increased risk for symptom relapse as well as the need for surgery, and at possible increased long-term risk for the development of colorectal cancer. Specific goals of UC treatment should include improving symptoms and quality of life, achieving mucosal and histologic remission, reducing the risk for future symptomatic relapse, avoiding the need for colectomy and end ileostomy or ileal pouch-anal anastomosis, preventing the development of colorectal cancer, allowing women of childbearing age to achieve pregnancy if desired, avoiding the need for short- or long-term steroid use, allowing optimal male fertility, and achieving durable clinical response with minimal development of anti-drug antibodies or primary or secondary loss of response and minimal adverse

events (AEs) (infections, malignancy, neurological events, thrombosis, or other cardiovascular events).

- Among the unmet needs are the large proportion of patients undergoing induction therapy with immunomodulators (e.g., azathioprine) who fail to achieve remission (up to 50%). This includes patients undergoing induction therapy with infliximab, vedolizumab, adalimumab, or tofacitinib who fail to achieve clinical remission during induction (primary nonresponse). Secondary loss of response can occur with all of these therapies and may be related to the development of anti-drug antibodies or breakthrough of the inflammatory response beyond the targeted mechanism of action.
- The population specified by the approved indication will be the target for treatment with ustekinumab. Ustekinumab would be most commonly used for induction of remission and maintenance therapy as monotherapy.
- According to expert input, failure of response to IV induction at eight weeks (a primary nonresponse) and secondary loss of response during maintenance therapy could be considered reasons for reassessing disease activity. If drug antibodies are detected or if adequate drug levels are identified in the presence of active inflammation on endoscopy, the drug is proven to be ineffective and should be discontinued.
- IV ustekinumab induction is usually given at an infusion clinic. Close follow-up by specialists familiar with UC is required, both to monitor clinical response and AEs. It is unlikely that patient treatment response will vary among physicians due to the standardization of doses and the administration intervals, which are usually followed as they are laid out in clinical trials.

Clinical Evidence

Pivotal Studies and Protocol Selected Studies

Description of Studies

- One double-blind randomized controlled trial (RCT), the UNIFI trial, was included in the review. The study was composed of two phases: an induction phase and a maintenance phase. The induction phase included 961 patients randomized to one of three arms: placebo IV (n = 319), ustekinumab IV (weight-based dosing of approximating 6 mg/kg; n = 322), or ustekinumab 130 mg (n = 320). All patients received a single administration of the treatment they were randomized to. Patients were evaluated at week 8 post-randomization for clinical remission, defined using the Mayo score. Two definitions of clinical remission were used for all patients, regardless of geographical location, to accommodate US and global regulatory preferences (US versus outside the US). Patients who were not in clinical remission at this stage received an additional single dose of ustekinumab, either 90 mg SC if they initially received ustekinumab (any dose), or approximately 6 mg/kg IV if they were initially allocated to placebo.
- Those in the induction ustekinumab arms (either dose) who responded to induction at week 8 were eligible to continue to the maintenance phase, as were those in the induction placebo arm who did not respond at week 8 but responded at week 16 to ustekinumab 6 mg/kg IV administered at week 8. These groups of patients formed the randomized population of the maintenance phase.
- Patients in the induction ustekinumab arms who did not respond at week 8 but responded at week 16 (delayed responders) were allowed to continue into the maintenance phase and continued to receive ustekinumab 90 mg SC every eight weeks. At the same time,

patients in the placebo arm who were in clinical remission continued to receive placebo during the rest of the maintenance phase (44 weeks). These patients were grouped in to the non-randomized population of the maintenance phase.

- Finally, all patients who did not respond to ustekinumab at both week 8 and week 16 were excluded from the maintenance phase and were followed up for safety through week 44.

Efficacy Results

From the induction phase, the groups who received IV ustekinumab 130 mg or 6 mg/kg had a higher proportion of patients who achieved clinical remission (15.6% and 15.5%, respectively) than those who received placebo (5.3%) ($P < 0.001$ for both comparisons) at week 8 based on the global definition of clinical remission (Mayo score of ≤ 2 points, with no individual subscore > 1). Similar results were reported based on the US definition of clinical remission (an absolute stool number of ≤ 3, a Mayo rectal bleeding subscore of 0, and a Mayo endoscopy subscore of 0 or 1). Sensitivity analyses supported the robustness of the primary analyses for both ustekinumab treatment groups versus placebo. Pre-specified subgroups of interest for this review were: history of conventional therapy for UC, history of biologics for UC, disease severity, disease extent, and risk of progression. Overall, the subgroup analyses were consistent with the primary analysis for the full study population, with a greater percentage of patients achieving clinical remission at week 8 with ustekinumab than with placebo.

Other efficacy outcomes of interest for this review, such as clinical response at week 8, endoscopic healing, health-related quality-of-life measures, and mucosal healing were statistically significantly improved in the ustekinumab groups compared with placebo (Table 1).

Of the 961 patients randomly allocated to ustekinumab or placebo in the induction phase, 783 were eligible to enter the maintenance phase, of which 523 were assigned to the randomized population (due to their response to ustekinumab IV), while 260 patients were allocated to the non-randomized population because they were late responders or responded to placebo only. Those in the randomized population were again assigned to receive SC maintenance injections of ustekinumab 90 mg (either every 12 weeks [n = 172 patients] or every eight weeks [n = 176]), or placebo (n = 175).

In the randomized population of the maintenance phase, the percentage of patients who had clinical remission (global and US definition) at week 44 was statistically significantly higher among patients assigned to 90 mg of SC ustekinumab every 12 weeks (approximately 39%) or every eight weeks (approximately 43%) than among those assigned to placebo (approximately 24.0%) ($P = 0.002$ and $P < 0.001$, respectively). Sensitivity analyses supported the primary analysis. Subgroup analyses were also generally consistent with the primary analysis for the full population. However, it was reported in UNIFI maintenance that for the subgroup by induction treatment received (ustekinumab 6 mg/kg IV [approximately], 130 mg IV, or placebo IV), there may be a lower maintenance-treatment effect on clinical remission (particularly for the every 12 weeks regimen) for patients who had received the 130 mg IV induction treatment or the placebo IV induction treatment. The sample sizes for these analyses were relatively small and estimates were imprecise.

Statistically significantly higher proportions of patients in the ustekinumab groups at week 44 maintained clinical response, corticosteroid-free remission, and endoscopic healing

compared with the placebo group. A significantly greater proportion of patients treated with either dose of ustekinumab compared with placebo also maintained clinical remission to week 44; however, the difference between groups was only numerically larger in the group treated with ustekinumab 90 mg every eight weeks.

The data also suggested greater improvement in health-related quality of life, mucosal healing, and productivity with ustekinumab versus placebo. However, these outcomes were not included in the hierarchical analysis plan and therefore not adjusted for inflated type I error. There were too few events related to colectomies (three patients treated with placebo and two patients in the combined ustekinumab group) upon which to draw conclusions.

Harms Results

There were fewer serious AEs with ustekinumab (combined total of 3.4% and 7.3% for the two ustekinumab groups) versus with placebo (6.6% and 9.7%) in the induction and maintenance phases of the UNIFI study, respectively. The higher frequency in the placebo group was seemingly driven by a larger percentage of patients reporting UC as an AE, likely reflecting a lack of efficacy from placebo. A larger percentage of patients in the placebo group (11.6%) withdrew from the maintenance phase due to an AE compared with those in the ustekinumab groups (5.1%); no patients withdrew from the induction phase due to AEs. Through 52 weeks of exposure, there were two deaths (one each from acute respiratory distress syndrome and hemorrhage from esophageal varices) and seven cases of cancer (one each of prostate, colon, renal papillary, and rectal cancer, and three non-melanoma skin cancers) among 825 patients who received ustekinumab, and no deaths and one case of cancer (testicular) among 319 patients who received placebo.

Table 1: Summary of Key Results

Induction study	Ustekinumab IV			Placebo
	130 mg N = 320	~6 mg/kg N = 322	Combined N = 642	N = 319
Clinical remission at week 8 (ITT)				
Number of patients in clinical remission (global definition),[a] n (%)	50 (15.6)	50 (15.5)	100 (15.6)	17 (5.3)
Percentage difference versus placebo, (95% CI);[b] P value[c]	10.3 (5.7 to 14.9); < 0.001	10.2 (5.6 to 14.8); < 0.001	10.2 (6.6 to 13.9); < 0.001	–
Number of patients in clinical remission (US definition),[d] n (%)	53 (16.6)	61 (18.9)	114 (17.8)	20 (6.3)
Percentage difference against placebo, (95% CI);[b] P value[c]	10.3 (4.8 to 15.8); < 0.001	12.7 (7.0 to 18.4); < 0.001	11.5 (7.0 to 16.0); < 0.001	–
Clinical response at week 8 (ITT)				
Number of patients in clinical response, n (%)	164 (51.3)	199 (61.8)	363 (56.5)	100 (31.3)
Percentage difference against placebo, (95% CI);[b] P value[c]	19.9 (12.5 to 27.3); < 0.001	30.5 (23.2 to 37.8); < 0.001	25.2 (18.9 to 31.5); < 0.001	–
HRQoL: Total IBDQ score at week 8[e]				
Baseline total IBDQ score, mean (SD)	126.0 (33.1)	127.0 (33.3)	126.5 (33.2)	127.4 (34.5)
Change from baseline in total IBDQ score, mean (SD)	33.4 (32.5)	35.0 (31.9)	34.2 (32.2)	16.1 (31.4)
P value[f]	< 0.001	< 0.001	< 0.001	–

Induction study	Ustekinumab IV			Placebo
	130 mg N = 320	~6 mg/kg N = 322	Combined N = 642	N = 319
HRQoL: EQ-5D scores at week 8[g,h]				
Baseline index score, mean (SD)	0.67 (0.204)	0.67 (0.195)	0.67 (0.199)	0.66 (0.208)
Change from baseline in EQ-5D index score, mean (SD)	0.090 (0.182)	0.110 (0.172)	0.100 (0.177)	0.040 (0.182)
Mean difference versus placebo, (95% CI); P value	0.050 (0.021 to 0.078); P < 0.001	0.070 (0.042 to 0.097); P < 0.001	0.060 (0.035 to 0.084); P < 0.001	–
Work productivity at week 8[g]				
Baseline percentage of work time missed due to health, mean (SD)	18.0 (30.22)	17.7 (29.07)	17.8 (29.59)	19.3 (32.32)
Change from baseline in percentage of work time missed due to health, mean (SD)	−5.9 (31.39)	−9.1 (23.84)	−7.6 (27.70)	−3.7 (30.41)
Mean difference versus placebo to (95% CI); P value	−2.20 (−9.02 to 4.62); 0.52	−5.4 (−11.21 to 0.41); 0.06	−3.9 (−9.37 to 1.57); 0.16	–
Harms	**N = 321**	**N = 320**	**N = 641**	**N = 319**
Patients with ≥ 1 adverse event, n (%)	133 (41.4)	160 (50.0)	293 (45.7)	153 (48)
Patients with ≥ 1 serious adverse event	12 (3.7)	10 (3.1)	22 (3.4)	21 (6.6)
Serious infections, n (%)	2 (0.6)	1 (0.3)	3 (0.5)	4 (1.3)

Maintenance study	Ustekinumab SC			Placebo
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 176	Combined N = 348	N = 175
Clinical remission at week 44 (ITT)				
Number of patients in clinical remission (global definition),[a] n (%)	66 (38.4)	77 (43.8)	143 (41.1)	42 (24.0)
Difference against placebo, (95% CI);[b] P value[c]	14.5 (5.5 to 23.6); 0.002	19.7 (10.3 to 29.0); < 0.001	17.1 (9.3 to 24.9); < 0.001	–
Number of patients in clinical remission (US definition),[d] n (%)	68 (39.5)	75 (42.6)	143 (41.1)	43 (24.6)
Percentage difference against placebo, (95% CI);[b] P value[c]	15.1 (6.0 to 24.2); 0.002	17.9 (8.6 to 27.2); < 0.001	16.5 (8.7 to 24.3); < 0.001	–
Corticosteroid-free clinical remission at week 44 (ITT)				
Number of patients in clinical remission (global definition),[a] n (%)	65 (37.8)	74 (42.0)	139 (39.9)	41 (23.4)
Difference against placebo, (95% CI);[b] P value[c]	14.5 (5.5 to 23.6); 0.002	18.5 (9.3 to 27.8); < 0.001	16.5 (8.8 to 24.3); < 0.001	–
Number of patients in clinical remission (US definition),[d] n (%)	65 (37.8)	74 (42.0)	139 (39.9)	41 (23.4)
Percentage difference against placebo, (95% CI);[b] P value[c]	14.5 (5.5 to 23.6); 0.002	18.5 (9.3 to 27.8); < 0.001	16.5 (8.8 24.3); < 0.001	–
Maintenance of clinical response at week 44 (ITT)				

Maintenance study	Ustekinumab SC			Placebo
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 176	Combined N = 348	N = 175
Number of patients in clinical response, n (%)	117 (68.0)	125 (71.0)	242 (69.5)	78 (44.6)
Difference against placebo, (95% CI); P value	23.5 (13.7 to 33.3); < 0.001	26.4 (16.6 to 36.1); < 0.001	25.0 (16.4 to 33.6); < 0.001	–
HRQoL: Total IBDQ score at week 44[g] (ITT)				
Maintenance baseline total IBDQ score: mean (SD) median (IQR)	175.4 (29.75) 180.5 (155.0 to 200.0)	174.1 (26.76) 177.0 (159.0 to 195.0)	174.7 (28.25) 178.0 (156.0 to 198.0)	174.3 (29.15) 181.0 (153.0 to 197.0)
Change from maintenance baseline in total IBDQ score: mean (SD) median (IQR) P value[i]	−3.0 (32.89) 1.5 (−14.0 to 16.5) < 0.001	3.9 (31.54) 5.0 (−7.0 to 20.0) < 0.001	0.5 (32.36) 3.0 (−11.0 to 18.0) < 0.001	−15.1 (35.43) −7.0 (−40.0 to 8.0)
Difference against placebo, (95% CI); P value	23.4 (12.96 to 33.12); < 0.001	27.0 (16.70 to 36.48)	25.2 (16.22 to 33.70); < 0.001	–
HRQoL: EQ-5D index score at week 44[g,h] (ITT)	0.008 (0.1656)	0.025 (0.1674)	0.017 (0.166)	−0.048 (0.158)
Maintenance baseline index score: Mean (SD) median (IQR)	0.810 (0.1563) 0.795 (0.726 to 1.000)	0.801 (0.1588) 0.795 (0.714 to 1.000)	0.806 (0.1574) 0.795 (0.721 to 1.000)	0.820 (0.1516) 0.837 (0.728 to 1.000)
Change from maintenance baseline in EQ-5D index score: mean (SD) median (IQR)	0.008 (0.1656) 0 (−0.062 to 0.107)	0.025 (0.1674) 0 (−0.042 to 0.121)	0.017 (0.1665) 0 (−0.052 to 0.111)	−0.048 (0.1587) −0.019 (−0.163 to 0.031)
Mean difference versus placebo, (95% CI); P value	0.056 (0.021 to 0.090); 0.001	0.065 (0.030 to 0.099); < 0.001	0.065 (0.035 to 0.094); < 0.001	–
Work productivity at week 44[g] (ITT)	−2.0 (22.16)	2.1 (19.07)	0 (20.70)	4.7 (21.83)
Maintenance baseline % of work time missed due to health, mean (SD) median (IQR)	9.4 (25.65) 0 (0.0 to 0.0)	2.8 (8.56) 0 (0.0 to 0.0)	6.1 (19.33) 0 (0.0 to 0.0)	6.5 (17.13) 0 (0.0 to 0.0)
Change from baseline in % of work time missed due to health: mean (SD) median (IQR)	−2.0 (22.16) 0 (0.0 to 0.0)	2.1 (19.07) 0 (0.0 to 0.0)	0.0 (20.70) 0 (0.0 to 0.0)	4.7 (21.83) 0 (0.0 to 0.0)
Mean difference versus placebo, (95% CI); P value	−6.7 (−12.92 to −0.47); 0.03	−2.6 (−8.36 to 3.16); 0.374	−4.7 (−9.83 to 0.43); 0.072	–
Harms	**N = 172**	**N = 333**	**N = 505**	**N = 277**
Patients with ≥ 1 adverse event, n (%)	119 (69.2)	253 (76.0)	372 (73.7)	209 (75.5)

Maintenance study	Ustekinumab SC			Placebo
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 176	Combined N = 348	N = 175
Patients with ≥ 1 serious adverse event, n (%)	13 (7.6)	26 (7.8)	39 (7.7)	24 (8.7)
Serious infections, n (%)	6 (3.5)	5 (1.5)	11 (2.2)	5 (1.8)
Malignancies, n (%)	1 (0.6)	1 (0.3)	2 (0.4)	1 (0.4)

ANCOVA = analysis of covariance; CI = confidence interval; EQ-5D = EuroQol 5-Dimensions; HRQoL = health-related quality of life; IBDQ = Inflammatory Bowel Disease Questionnaire; IQR = interquartile range; ITT = intention to treat; q.8.w. = every eight weeks; q.12.w. = every 12 weeks; SC = subcutaneous; SD = standard deviation; VAS = Visual Analogue Scale.

[a] Mayo score of ≤ 2 points with no individual subscore > 1.

[b] The CIs were based on the Wald statistic with Mantel-Haenszel weight.

[c] The P values were based on the Cochran-Mantel-Haenszel test.

[d] An absolute stool number of ≤ 3, a Mayo rectal bleeding subscore of 0, and a Mayo endoscopy subscore of 0 or 1.

[e] Outcome was outside the statistical testing hierarchy for US-based analyses.

[f] ANCOVA on the van der Waerden normal scores with baseline IBDQ score, biologic failure status, region, and group as covariates.

[g] Outcome was outside the statistical testing hierarchy.

[h] Cochran-Mantel-Haenszel chi-square (row mean scores) test stratified by biologic failure status and region.

[i] ANCOVA on the van der Waerden normal scores with the respective baseline value, clinical remission status at maintenance baseline, induction treatment, and maintenance treatment group as covariates.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Critical Appraisal

- Overall, the risk of bias was low for the included trial, with no limitations in the randomization process, blinding, differences in baseline characteristics, or assessment of outcomes. No major limitations were noted in the attrition rate of patients throughout both phases of the study. In terms of the external validity, one concern was the number of patients (157 out of 233 [67%]) who initially did not respond in the induction study at week 8 and received a second dose of ustekinumab SC 90 mg and responded at week 16. The clinical expert consulted by CADTH indicated the proportion of delayed responders seemed high and, in clinical practice, clinicians may opt to administer a second dose of ustekinumab to induce remission.

Indirect Treatment Comparisons

Description of Studies

One systematic review and network meta-analysis (NMA) of indirect treatment comparisons (ITCs) was included.

Efficacy Results

This synthesis assesses the efficacy of ustekinumab indirectly compared with other interventions, namely, infliximab, adalimumab, vedolizumab, ustekinumab, golimumab, tofacitinib, and placebo. It evaluates three outcomes — clinical remission, clinical response, and mucosal healing — in patients considered biologic and non-biologic failures, and also in the induction and maintenance phases of drug administration. Based on the NMA of the induction phase, ustekinumab had higher odds of clinical response, clinical remission, and mucosal healing against placebo and adalimumab (in biologic and non-biologic failure patients for clinical response, but only in biologic failure patients for clinical remission and mucosal healing). For the rest of the comparisons, ustekinumab either did not increase or decrease the odds of any of these outcomes when compared with infliximab, vedolizumab,

golimumab, and tofacitinib. In the maintenance phase, ustekinumab had higher odds of clinical response in non-biologic failure patients when compared with adalimumab, golimumab, tofacitinib, and placebo but not against vedolizumab, while in the biologic failure patients, it was only better than placebo. For clinical remission, ustekinumab provided higher odds against golimumab, adalimumab, and placebo in the non-biologic failure group (but not against vedolizumab, infliximab, or tofacitinib); while in the biologic failure patients, ustekinumab was only better than placebo. Lastly, ustekinumab had higher odds of mucosal healing in non-biologic failure patients than adalimumab, golimumab, and placebo, but it was no better than infliximab, tofacitinib, and vedolizumab.

Harms Results

The ITC submitted and evaluated did not include an assessment of the AEs.

Critical Appraisal

This systematic review and NMA of ITCs were performed under the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) report checklist. The search strategy was properly conducted based on a protocol and statement of an important clinical question. The review was well performed in terms of an adequate search and the extraction and analysis of data. However, the limitations of the NMA include uncertainty about the effect estimates, particularly for the one-year outcomes, mostly due to concerns of heterogeneity, intransitivity, and uncertainty due to the use of multiple assumptions of the imputation process, and overestimated precision for reported comparisons, although a multiple-imputation sensitivity analysis was performed for clinical response in non-biologic failure patients who, overall, showed the same conclusions. Finally, individual studies had a moderate risk of bias, with concerns from the randomization process, unclear blinding, and unbalanced dropout rates with no intention-to-treat (ITT) analysis.

Conclusions

Based on one trial, ustekinumab is more effective than placebo for inducing and maintaining clinical remission and clinical response, maintaining a corticosteroid-free remission, and inducing and maintaining endoscopic healing in patients who have moderate-to-severe UC despite current or previous treatment with conventional or biologic therapy.

Based on one review of ITCs, although with better odds for all outcomes when compared with placebo, ustekinumab had no clear superiority over other common comparators with the same indication, although there is still uncertainty due to inconsistency in the body of evidence and risk of bias that decreases our confidence in this result.

Although AEs were not different between ustekinumab and placebo, the number of events were low and more long-term studies are needed to assess possible harms.

Introduction

Disease Background

UC is a chronic IBD that involves inflammation of the mucosae of the large intestine, starting distally in the rectum and with variable levels of proximal extension into the colon. Although it may affect any age group, its onset is usually during young adulthood, peaking between 15 and 30 years of age.[3,4]

UC has a worldwide distribution, albeit with a predominance in high-income Western countries, with a global incidence of 1.2 to 20.3 cases per 100,000 persons per year, and a prevalence of 7.6 to 246.0 cases per 100,000 per year.[5] Canada is among the countries with the highest incidence and prevalence of IBD, with approximately 270,000 Canadians living with UC or Crohn disease. The incidence of UC in different Canadian provinces ranges from 8.4 to 21.4 cases per 100,000 people.[6]

The risk of death from UC is increased within the first year after diagnosis but, beyond that point, patients remain at the same risk as the general population.[7] The diagnosis of UC implies a burden for patients, families, and health care systems, as it affects quality of life in different domains, including school, work, and social interactions. Increasing costs within the health care system is also an issue. In Canada, approximately $1.2 billion is spent annually by the health care system in patients with IBD, while there is an estimated indirect cost to society of nearly $1.5 billion in domains such as loss of work and productivity, disability coverage, and premature retirement or death.[8,9]

The etiology of UC is not completely understood, although evidence of the role of genetic and environmental factors, as well as correlations between UC and the microbiota, is accumulating.[5]

Symptoms start gradually in most cases, with following periods of spontaneous remissions and relapses. Bloody diarrhea with or without mucus is the most common initial manifestation. Depending on the extension and severity of disease, symptoms, beside frequent evacuations with blood and mucus, can include urgency or tenesmus, fever, abdominal pain, and weight loss.[5,10] Prognosis is usually good, with the majority of patients not needing a colectomy and remitting within the first decade.[11]

Severity of disease may be defined differently, depending on the index or score used, for example, the Mayo Clinic score or the Montreal classification. The extent of endoscopic disease has been categorized as proctitis (distal to the rectosigmoid junction or within 18 cm of the anal verge), left-sided colitis (extending anywhere from the sigmoid to the splenic flexure), or extensive colitis (extending beyond the splenic flexure).[12]

Standards of Therapy

Current guidelines suggest assessing the level of clinical activity or severity (mild, moderate, severe) as well as the extension (proctitis, left-sided colitis, or pancolitis).[12,13] The goal is to obtain a sustained remission free of steroids and with proper support for managing other domains to increase quality of life, such as psychosocial support, and understanding the patient's own values and preferences, emphasizing the prevention of morbidity due to surgery or hospitalization.[12]

First-line treatments for inducing remission include either orally or rectally administered sulfasalazine and 5-aminosalicylates (mesalamine, olsalazine, and balsalazide). Half of patients are expected to enter remission within two weeks. Rectal administration of 5-aminosalicylates or glucocorticoid are considered only for patients who have distal disease (e.g., proctitis).[5] If mild-to-moderate left-sided or extensive UC is present, a mixture of rectal and oral 5-aminosalicylates can be used, with escalating doses of oral 5-aminosalicylates. Next steps for patients with poor response to rectal therapies and 5-aminosalicylates include oral glucocorticoids or immunosuppressive drugs, such as azathioprine or 6-mercaptopurine, as second-line therapy to induce complete remission. Glucocorticoids can also be considered first-line therapy if patients start with moderate-to-severe active UC.[5,12] Patients who continue to require glucocorticoids at this step are considered to have moderate-to-severe active UC and are candidates to receive vedolizumab or anti–tumour necrosis factor (anti-TNF) therapy to induce complete glucocorticoid-free remission. Vedolizumab (an α4β7 inhibitor), anti-TNF therapies (infliximab, adalimumab, golimumab), and tofacitinib (a selective Janus kinase inhibitor) are part of the group of medications collectively known as biologics and are considered immune-modifying therapies for the induction or maintenance of remission for patients with UC.

Drug

Ustekinumab is a human monoclonal antibody designed to interfere with the interleukin pathways in the pathogenesis of immune-modulated conditions (specifically, interleukin-12 and interleukin-23). It has been approved by Health Canada for the treatment of adults with chronic moderate-to-severe plaque psoriasis who are candidates for systemic therapy or phototherapy and for the treatment of adult patients with active psoriatic arthritis, and for the treatment of adult patients with Crohn disease.[14,15]

The current indication under review is for the treatment of adult patients with moderately to severely active UC who have had an inadequate response to, lost response to, or were intolerant to either conventional therapy or a biologic or have medical contraindications to such therapies. The recommended dosage of ustekinumab for the induction of remission of UC is as a single IV dose based on body weight (approximating 6 mg/kg). Maintenance dosing using 90 mg SC should be administered eight weeks after the IV induction dose, then every eight weeks thereafter. For some patients (e.g., "those with low inflammatory burden," per the product monograph), an alternative maintenance regimen of ustekinumab 90 mg SC every 12 weeks may be administered at the discretion of the treating physician. Patients who respond inadequately to 90 mg SC dosing every 12 weeks may be switched to receive the drug every eight weeks. Immunomodulators and corticosteroids may be continued during treatment with ustekinumab. The product monograph approved by Health Canada recommends that consideration be given to discontinuing treatment in patients who show no evidence of therapeutic benefit 16 weeks after the IV induction dose. It also recommends that, in patients who have responded to treatment with ustekinumab, corticosteroids may be reduced or discontinued in accordance with standard of care. The product monograph notes that ustekinumab should be used only by physicians who have enough knowledge of the indication for which it is being considered (e.g., UC) and who have fully familiarized themselves with the efficacy and safety profile of the drug.

Ustekinumab has been previously reviewed by CADTH through the CADTH CDR process. First, for the treatment of adults with chronic moderate-to-severe plaque psoriasis who are candidates for systemic therapy or phototherapy, and for the treatment of adult patients with active psoriatic arthritis, alone or in combination with methotrexate. The former CADTH

Canadian Expert Drug Advisory Committee (CEDAC) recommended that ustekinumab be reimbursed for patients with severe, debilitating psoriasis with clinical criteria.[16] Later, the CADTH Canadian Drug Expert Committee (CDEC) recommended that ustekinumab not be reimbursed at the submitted price for the treatment of psoriatic arthritis[17] and, more recently, CDEC recommended that ustekinumab be reimbursed for the treatment of adult patients with moderately to severely active Crohn disease who have had an inadequate response to, loss of response to, or were intolerant to either immunomodulators or one or more TNF-alpha antagonists, or who have had an inadequate response to, an intolerance to, or demonstrated dependence on corticosteroids, following clinical criteria.[18]

The key characteristics of the drug and other main comparators are presented in Table 2.

Table 2: Key Characteristics of Ustekinumab and Main Comparators

	Ustekinumab	Infliximab	Vedolizumab	Golimumab	Tofacitinib	Adalimumab
Mechanism of Action	Human IgG1 monoclonal antibody that neutralizes cellular responses mediated by IL-12 and IL-23.	Anti-TNF. IgG1κ monoclonal antibody that neutralizes the biological activity of TNF alpha by specifically binding to its receptors.	IgG1 monoclonal antibody that binds to the human α4β7 integrin, acting as a gut-selective anti-inflammatory biologic.	Anti-TNF. Human monoclonal antibody that binds to human TNF (p55 or p75 receptors).	Selective Janus kinase inhibitor. Blocks several cytokine pathways and lymphocyte activation.	Anti-TNF. Human IgG1 monoclonal antibody that binds specifically to TNF alpha and blocks its interaction with the p55 and p75 cell surface TNF receptors.
Indication[a]	Treatment of adult patients with moderately to severely active UC who have had an inadequate response to, lost response to, or were intolerant to either conventional therapy or a biologic, or have medical contraindications to such therapies.	Induction and maintenance of clinical remission and mucosal healing, and reduction or elimination of corticosteroid use in adult patients with moderately to severely active UC who have had an inadequate response to conventional therapy.	Treatment of adult patients with moderately to severely active UC who have had an inadequate response, loss of response to, or were intolerant to either conventional therapy or infliximab, a TNF alpha antagonist.	To induce and maintain clinical response in adult patients with moderately to severely active UC who have had an inadequate response to or have medical contraindications for conventional therapy, including corticosteroids, amino salicylates, azathioprine, or 6-MP.	For the treatment of adult patients with moderately to severely active UC with an inadequate response, loss of response to, or intolerance to either conventional UC therapy or a TNF alpha inhibitor.	For the treatment of adult patients with moderately to severely active UC who have had an inadequate response to conventional therapy, including corticosteroids and/or azathioprine or 6-MP, or who are intolerant to such therapies.
Route of Administration	Intravenous induction followed by subcutaneous for maintenance	Intravenous	Intravenous	Subcutaneous	Oral	Subcutaneous
Recommended Dose	Induction: IV infusion of single-use weight-based dose (~6 mg/kg): 250 mg for those weighing ≤ 55 kg, 390 mg for those weighing > 55 kg to ≤ 85 kg, or 520 mg for those weighing > 85 kg.	Induction dose of 5 mg/kg at 0, 2, and 6 weeks, followed by 5 mg/kg every eight weeks thereafter.	300 mg administered by IV infusion at 0, 2, and 6 weeks, and then every 8 weeks thereafter.	200 mg initially administered by subcutaneous injection at week 0, followed by 100 mg at week 2 and then 50 mg every 4 weeks thereafter.	Tofacitinib tablets, 10 mg (as tofacitinib citrate) orally twice daily.	160 mg at week 0, followed by 80 mg at week 2 administered by subcutaneous injection.

	Ustekinumab	Infliximab	Vedolizumab	Golimumab	Tofacitinib	Adalimumab
	Maintenance: SC injection of 90 mg every 8 or 12 weeks.					
Serious Adverse Effects or Safety Issues	Immunomodulating drugs have the potential to increase the risk of infections and malignancy. No clinically significant differences have been found in terms of malignancies.	Infections and malignancies have been observed in patients receiving infliximab.	Infections and malignancies are reported in patients taking vedolizumab but no clinically significant differences have been found.	Upper respiratory infections and reactions at the site of injection, but no clinically significant differences with placebo.	Can increase the risk of thromboses (pulmonary and deep vein thrombosis). Increased risk of serious infections, including herpes zoster infections.	Serious infections (pneumonia), malignancies, and neurologic events have been reported more frequently in patients taking adalimumab.
Other					Not recommended in combination with biological UC therapies or with potent immunosuppressants such as azathioprine and cyclosporine.	

6-MP = 6-mercaptopurine; Ig = immunoglobulin; IL = interleukin; JAK = Janus kinase; TNF = tumour necrosis factor; UC = ulcerative colitis.

[a] Health Canada–approved indication.

Source: Product monographs of ustekinumab (Stelara),[19] infliximab (Remicade),[20] vedolizumab (Entyvio),[21] golimumab (Simponi),[22] tofacitinib (Xeljanz),[23,24] and adalimumab (Humira).[25]

Stakeholder Engagement

Patient Group Input

This section was prepared by CADTH staff based on the input provided by patient groups.

Two patient groups responded to CADTH's call for patient input for the ustekinumab (Stelara) submission for UC: Crohn's and Colitis Canada and the Gastrointestinal Society.

Crohn's and Colitis Canada is a national, volunteer-based charity with more than 65,000 supporters. The organization aims to support research for IBDs and improve the lives of affected adults and children by providing support and information on treatments, research, and quality-of-life issues. Since 1974, it has received investments totalling more than $122 million. A medical science liaison from the sponsor of Stelara (Janssen) provided a briefing to Crohn's and Colitis Canada explaining the mechanism of action of ustekinumab. Crohn's and Colitis Canada also solicited help from Canadian gastroenterologists to identify patients who have had experience taking ustekinumab. Over the last two years, the organization has received between $5,000 and $10,000 from Roche, and in excess of $50,000 from Pfizer Canada, Janssen, AbbVie, Merck, and Takeda.

The Gastrointestinal Society is an organization committed to helping individuals with gastrointestinal and liver conditions by supporting research, advocating for patient access to health care, and promoting overall gastrointestinal and liver health. The organization informs Canadians through different channels such as newsletters, lectures, and websites in both English and French. The Gastrointestinal Society also holds support group meetings for those recently diagnosed as well as for individuals who have been living with a gastrointestinal condition for years. Its staff and advisors work closely with health care professionals, other patient groups, and government. The Gastrointestinal Society indicated it did not receive any outside help in preparing this submission. The organization has received more than $50,000 from Janssen over the last two years.

Crohn's and Colitis Canada provided information from its own website and databases, such as its 2018 report, *Impact of Inflammatory Bowel Disease in Canada.* This information also included data from a national online survey conducted in 2011 that included input from 430 respondents living in Canada, plus responses from a series of interviews and surveys of 13 Canadian patients being treated with Stelara. The Gastrointestinal Society used two questionnaires to survey 565 Canadians with IBD. The Gastrointestinal Society also had contact with patients affected by IBD through one-on-one conversations at the BadGut lectures and patient roundtables, and through phone, email, and social media interactions.

Patients from both groups describe UC as a disabling, lifelong gastrointestinal condition that primarily affects working-age Canadians. Symptoms associated with UC include bloody diarrhea, bloating, abdominal pain, cramping, and fatigue. Individuals with UC are at an increased risk of colon cancer. Patient groups often describe experiences of isolation, anxiety, and debilitating, frequent, and urgent bowel movements. Results from Crohn's and Colitis Canada's 2011 survey indicated that 73% of respondents affected by an IBD experience 5 to 20 or more bowel movements a day. During periods of active disease, patients report spending a lot of time in the bathroom and, even in periods of remission, they have to stay near a bathroom. One respondent stated, "When you have to go to the washroom 20 times a day, it impacts everything that you do." Another patient expressed, "When the disease takes control of your body, you feel very tired. When my large bowel is

affected, I get bloody diarrhea quick and practically live in the bathroom. It plays havoc with my head, I can't sleep, and I get headaches and other problems as a result." The patient group added that individuals with UC must limit their activities sometimes because of the stigma associated with an IBD. Declared one patient: "You simply can't lead a normal life of working and going to the office." Overall, the Gastrointestinal Society described this as a chronic disease, one where there is a constant concern regarding future and possibly worse and unpredictable flares, many times disrupting patients' lives.

First-line treatments for UC include anti-inflammatory drugs such as 5-aminosalicylates and corticosteroids to control disease flares. Nonresponders and more severe cases of UC are treated with second-line treatments such as immunomodulators or immunosuppressants. Third-line treatment includes biologics such as anti-TNF drugs. While current treatments are often effective in patients with moderate colitis, they fail to maintain remission for those with severe colitis.

Patients often seek treatment options that can reduce or eliminate their symptoms. Additionally, patients would like a treatment that can protect their ability to work productively, attend school and social events, and perform basic day-to-day activities. The patient groups report that some of these treatments, such as steroids, can have negative impacts associated with long-term use. According to the Gastrointestinal Society, only 28% of patients thought the available medications were adequate, while 54% found them to be somewhat adequate, and 18% said they were not adequate. The patient groups report that patients are still suffering and require new and effective options to achieve mucosal healing and decrease debilitating symptoms.

According to Crohn's and Colitis Canada, patients being treated with Stelara found the drug to be convenient and easy to use. Most patients were pleased with not having to travel to a clinic to administer the medication, providing them with the ability to live a normal life, with some reporting that Stelara "has been the difference between not really living and living."

Patients interviewed by the Gastrointestinal Society noted that ustekinumab has the potential to improve health and quality of life. One patient stated, "It is always good to have hope that there is another option out there for treatment. It is scary when you are running out of options and when whatever you are on is not working." Patients are hoping for treatment options that can mitigate their symptoms and protect their ability to work, attend school and social events, and perform basic day-to-day activities. Many patients interviewed considered frequency and urgency of bowel movements to be the most important symptom to control. "The simple ability to live life," one patient said, is the most important aspect of potential treatment is to achieve remission for the longest period possible. Moreover, patients would like a more convenient form of administrating treatments, such as self-injection rather than IV infusions in the clinic. One patient reported being frightened to self-inject but would "gladly accept it" for better convenience.

Given that all individuals respond differently to treatment, the submissions noted it is important that patients have a variety of treatment options available. Moreover, inadequate access to medication can result in patient suffering and excess usage of health care resources.

Clinician Input

All CADTH review teams include at least one clinical specialist with expertise in the diagnosis and management of the condition for which the drug is indicated. Clinical experts are a critical part of the review team and are involved in all phases of the review process (e.g., providing guidance on the development of the review protocol, assisting in the critical appraisal of clinical evidence, interpreting the clinical relevance of the results, and providing guidance on the potential place in therapy). The following input was provided by one clinical specialist with expertise in the diagnosis and management of IBD, specifically, UC.

Description of the Current Treatment Paradigm for the Disease

The prevalence of UC in Canada is among the highest in the world (0.4%). Over the past 15 years, the incidence has been rising, particularly in children. More than 4,500 new cases of UC are now diagnosed annually in Canada.[26]

UC can present as mild, moderate, or severe disease. Mild symptoms include minor diarrhea and trivial rectal bleeding. Moderate symptoms include significant diarrhea and significant rectal bleeding. Severe colitis can present with tachycardia, fever, volume loss from severe diarrhea, and anemia, often requiring hospitalization. The degree of disease activity can be objectively assessed using validated clinical criteria (e.g., partial Mayo score). Endoscopic criteria can also be used to assess disease activity with the endoscopic Mayo score. Other markers of disease activity include C-reactive protein (CRP) as well as fecal calprotectin levels.

Treatment Goals

Treatment for active UC involves an induction phase as well as long-term maintenance. Active disease is likely to relapse. Mild disease is usually managed with oral or rectally administered 5-aminosalicylate products. Moderate-to-severe disease requires escalation of therapy, including periodic steroid therapy for rapid relief of symptoms as required. Immunomodulatory therapy with azathioprine and methotrexate can be used for moderate disease for steroid sparing and prevention of disease relapse. Patients relapsing and requiring frequent courses of steroids or resistant to immunomodulatory therapy require escalation to biologic therapy (infliximab, golimumab, adalimumab, vedolizumab).

The goal of medical treatment for moderate-to-severe UC is beyond symptom remission. It ideally should include endoscopic and histologic healing of colonic inflammation. Patients who do not achieve both clinical and endoscopic remission are at increased risk for symptom relapse as well as the need for surgery, and possible increased long-term risk for the development of colorectal cancer. Specific goals of UC treatment include:

- improving symptoms
- improving quality of life
- achieving mucosal and histologic remission
- reducing the risk for future symptomatic relapse
- avoiding the need for colectomy and end ileostomy or ileal pouch-anal anastomosis
- preventing the development of colorectal cancer
- allowing women of childbearing age to achieve pregnancy, if desired
- avoiding requirement for short- or long-term steroid use

- allowing optimal male fertility
- achieving durable clinical response with minimal development of anti-drug antibodies or primary or secondary loss of response and minimal AEs (infections, malignancy, neurological events, thrombosis, or other cardiovascular events).

Unmet Needs

Among the unmet needs that clinicians and patients currently face is that a significant proportion of patients undergoing induction therapy with immunomodulators (e.g., azathioprine) fail to achieve remission (up to 50%). In patients undergoing induction therapy with infliximab, vedolizumab, adalimumab, or tofacitinib, failure to achieve clinical remission during induction (primary nonresponse) can occur in up to 50% of patients. Secondary loss of response can occur with all of these therapies and may be related to the development of anti-drug antibodies or breakthrough of the inflammatory response beyond the targeted mechanism of action.

For IV formulations of biologic medications, patients are required to attend infusion clinics. This may not be desirable or possible for patients who either live in remote locations or are required to travel away from their place of residence for work. An SC or oral medication would likely be valuable for these individuals.

There is a lack of data on the safety profile of some biologic medications in relation to pregnancy and lactation and use by elderly patients, those with pre-existing cardiovascular disease (for certain biologic medications), and those who have had previous malignancies now in remission (e.g., breast cancer, skin cancer).

Place in Therapy

Ustekinumab is a treatment for induction of remission in active UC as well as a long-term maintenance drug.

To date, there is scarce evidence regarding dual therapy with ustekinumab and an immunomodulator such as azathioprine for induction of remission or maintenance. Therefore, ustekinumab would be most commonly used for induction of remission and maintenance therapy as monotherapy. Ustekinumab is likely to be used in accordance with the indication. However, where ustekinumab fits in the overall armamentarium of treatments for moderate-to-severe UC in clinical settings remains to be determined. It may be reasonable to use ustekinumab as a second-line therapy after failure of alternative treatments (e.g., infliximab) or when there has been a primary or secondary loss of response.

Patient Population

Patients best suited to ustekinumab therapy are those with UC with moderate-to-severe symptom and endoscopic scores, patients who have failed steroid induction, and patients who have failed immunomodulator therapy or other biologic therapies (e.g., infliximab).

Ustekinumab is best suited for patients with moderate-to-severe UC identified by the Mayo score, the endoscopic Mayo score, and ancillary lab testing (e.g., elevated CRP and fecal calprotectin levels).

Least suitable patients would be those with mild symptoms of UC, those in hospital with severe disease, those with active malignancy or active infection (e.g., tuberculosis), and pregnant women, due to limited safety data.

Assessing Response to Treatment

The parameters used in clinical practice to assess response to treatment include a decrease in the Mayo score (clinical remission defined as a partial Mayo score of < 2), complete mucosal healing identified endoscopically (endoscopic Mayo score of 0), CRP level and fecal calprotectin level returning to normal, and the overall improvement in the patient's quality of life.

It is unlikely that treatment response will vary across physicians due to the standardization of the doses and administration intervals, which are usually followed as they are laid out in clinical trials.

According to expert input, the assessment of response to the induction dose of ustekinumab depends on the treatment and severity of UC but may occur between four and eight weeks after the initiation of induction treatment. For ustekinumab, assessment will occur at the week 8 post-induction dose. The clinical expert noted that, for some patients, there appears to be a delay in achieving remission at week 8 during induction. In practice, treating physicians may opt to wait another four to eight weeks for these patients to achieve induction remission. The decision in practice to wait to see whether a delayed response occurs would depend on many factors (see next paragraph). Following remission on induction, patients should be assessed at least annually. Patients should be seen and evaluated promptly for symptoms that suggest a secondary loss of response.

Discontinuing Treatment

Failure of response to IV induction (at eight weeks), i.e., a primary nonresponse, and secondary loss of response during maintenance therapy would be considered reasons for reassessing disease activity. If the symptom score worsens or if the endoscopic score worsens along with corroborating evidence of inflammation, such as elevated CRP or increased fecal calprotectin, levels of the drug and drug antibodies should be obtained. If drug levels are sub-therapeutic (trough drug levels), without the development of drug antibodies, dose escalation (shortening the injection interval) should be carried out. If drug antibodies are detected or if adequate drug levels are identified in the presence of active inflammation on endoscopy, the drug is proven to be ineffective and should be discontinued.

Prescribing Conditions

IV ustekinumab induction is usually given at an infusion clinic. SC injections can be administered at home and require minimal training. Administration of ustekinumab is complex. Close follow-up by specialists familiar with UC is required, both to monitor clinical response and AEs. These patients should be followed by a physician skilled in the administration of this drug (e.g., gastroenterologist).

Clinical Evidence

The clinical evidence included in this review of ustekinumab is presented in three sections. Section 1, the systematic review, includes pivotal studies provided in the sponsor's submission to CADTH and Health Canada, as well as those studies that were selected according to an a priori protocol. Section 2 includes indirect evidence from the sponsor and indirect evidence selected from the literature that met the selection criteria specified in the review. Section 3 includes additional relevant studies that were considered to address important gaps in the evidence included in the systematic review.

Systematic Review (Pivotal and Protocol Selected Studies)

Objectives

To perform a systematic review of the beneficial and harmful effects of ustekinumab IV infusion (induction phase) and SC injection (maintenance phase) for the treatment of adult patients with moderately to severely active UC who have failed or were intolerant to treatment with immunomodulators or corticosteroids — but never failed treatment with a biologic — or have failed or were intolerant to treatment with a biologic.

Methods

Studies selected for inclusion in the systematic review include pivotal studies provided in the sponsor's submission to CADTH and Health Canada, as well as those meeting the selection criteria presented in Table 3.

The literature search for clinical studies was performed by an information specialist using a peer-reviewed search strategy according to the PRESS (Peer Review of Electronic Search Strategies) checklist (https://www.cadth.ca/resources/finding-evidence/press).[27]

Published literature was identified by searching the following bibliographic databases: MEDLINE All (1946–) through Ovid, Embase (1974–) through Ovid, and PubMed. The search strategy comprised both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were Stelara and UC. Clinical trial registries were searched: the US National Institutes of Health's clinicaltrials.gov and the World Health Organization's International Clinical Trials Registry Platform search portal. No filters were applied to limit the retrieval by study type. Retrieval was not limited by publication date or by language. Conference abstracts were excluded from the search results. See Appendix 2 for the detailed search strategies. The initial search was completed on September 10, 2019. Regular alerts updated the search until the meeting of CDEC on January 15, 2020. Grey literature (literature that is not commercially published) was identified by searching relevant websites from the following sections of the *Grey Matters: A Practical Tool For Searching Health-Related Grey Literature* checklist (https://www.cadth.ca/grey-matters):[28] Health Technology Assessment Agencies, Health Economics, Clinical Practice Guidelines, Drug and Device Regulatory Approvals, Advisories and Warnings, Drug Class Reviews, Clinical Trials Registries, and Databases (Free). Google was used to search for additional internet-based materials. These searches were supplemented by reviewing bibliographies of key papers and through contacts with appropriate experts. In addition, the sponsor of the drug was contacted for information regarding unpublished studies. See Appendix 2 for more information on the grey literature search strategy.

Table 3: Inclusion Criteria for the Systematic Review

Patient population	Adults with moderately to severely active ulcerative colitis who: • have failed or were intolerant to treatment with immunomodulators or corticosteroids but never failed treatment with a biologic or • have failed or were intolerant to treatment with a biologic. Subgroups: • Patients experienced with previous (versus no previous) conventional therapy • Patients experienced with previous (versus no previous) anti-TNF drugs • Disease severity (e.g., moderate versus severe) • Disease extent (extensive versus limited colitis) • Low versus high-risk of progression
Intervention	Ustekinumab. **Induction:** solution for a single intravenous tiered dose based on body weight (approximately 6 mg/kg). **Maintenance:** subcutaneous injection of 90 mg (90 mg/1.0 mL vial) starting 8 weeks after the intravenous dose and then every 8 weeks thereafter.
Comparators	• Adalimumab • Golimumab • Infliximab • Tofacitinib • Vedolizumab • Conventional therapy: any combination of aminosalicylates, corticosteroids, and immunomodulators.
Outcomes	**Efficacy outcomes:** • Clinical remission[a] (global definition[b]), including corticosteroid-free clinical remission • Clinical response[a,c] • Health-related quality of life[a] • Need for colectomy • Mucosal healing determined by histology or endoscopy • Productivity[a] **Harm outcomes:** AEs, SAEs, WDAEs, mortality Notable harms and harms of special interest: thrombosis (any type), hypersensitivity (anaphylaxis and/or angioedema), serious infections (including herpes zoster), malignancy, major cardiovascular event.
Study design	Published and unpublished phase III and IV RCTs

AE = adverse event; HRQoL = health-related quality of life; RCT = randomized controlled trial; SAE = serious adverse event; WDAE = withdrawal due to adverse event.

[a] These outcomes were identified as being of importance to patients in the input received by CADTH from patient groups.

[b] Mayo score of ≤ 2 points with no individual subscore > 1.

[c] A decrease from baseline in the Mayo score of ≥ 30% and ≥ 3 points, with either a decrease from baseline in the rectal bleeding subscore of ≥ 1 point or a rectal bleeding subscore of 0 or 1.

The literature search was performed by an information specialist using a peer-reviewed search strategy.

Two CDR clinical reviewers independently selected studies for inclusion in the review based on titles and abstracts, according to the predetermined protocol. Full-text articles of all citations considered potentially relevant by at least one reviewer were acquired. Reviewers independently made the final selection of studies to be included in the review, and differences were resolved through discussion.

Findings From the Literature

From the literature, we identified one study[29] that was subdivided into two reports, one for each phase of the study,[1,2] for inclusion in the systematic review (Figure 1). The included study and each phase of the study are summarized in Table 4. A list of excluded studies is presented in Appendix 2.

Figure 1: Flow Diagram for Inclusion and Exclusion of Studies

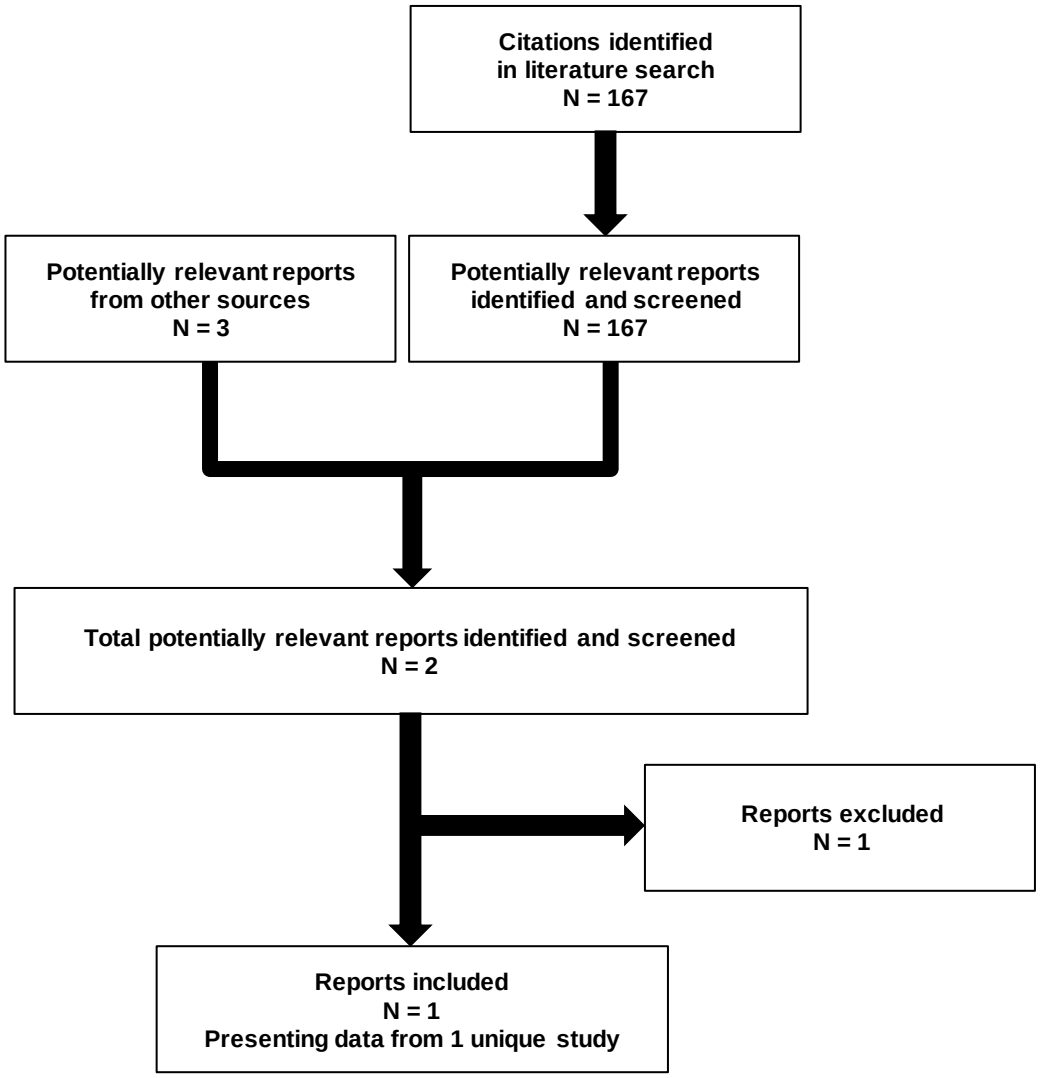


Table 4: Details of Included Studies

		UNIFI induction	UNIFI maintenance
Designs & Populations	**Study Design**	Double-blind, parallel-group, placebo-controlled RCT	Double-blind, parallel-group, placebo-controlled RCT
	Locations	Australia, Canada, Japan, Korea, New Zealand, US, Austria, Belgium, Bulgaria, Czech Republic, Denmark, France, Germany, Hungary, Israel, Italy, Netherlands, Poland, Romania, Russia, Serbia, Slovakia, UK, Ukraine	Australia, Canada, Japan, Korea, New Zealand, US, Austria, Belgium, Bulgaria, Czech Republic, Denmark, France, Germany, Hungary, Israel, Italy, Netherlands, Poland, Romania, Russia, Serbia, Slovakia, UK, Ukraine
	Randomized (N)	961	523
	Inclusion Criteria	• Adults > 18 years of age with moderately to severely active UC (Mayo score of 6 to 12, including an endoscopy subscore of ≥ 2 as assessed during the central review of the video of the endoscopy). Patients may have experienced biologic failures, i.e., received treatment with one or more TNF antagonists or vedolizumab and either did not respond initially, responded initially but then lost response, or were intolerant to the medication. or • Patients who may have been biologic-naive or may have been exposed to biologic therapy but did not demonstrate an inadequate response to or intolerance to treatment with a biologic drug. These patients must have demonstrated an inadequate response to, or have failed to tolerate, at least one of the following conventional UC therapies: oral or IV corticosteroids, or the immunomodulators azathioprine or 6-mercaptopurine. Patients who demonstrated corticosteroid dependence (i.e., an inability to successfully taper corticosteroids without a return of the symptoms of UC) were also eligible for entry into the study.	• Moderately to severely active UC who had an inadequate response or had failed to tolerate conventional therapy or biologic therapy, and who demonstrated a clinical response to the study drug during the induction study. These included: ○ patients who were randomized to receive ustekinumab (130 mg IV or ~6 mg/kg IV) at week 0 of the induction study and were in clinical response at induction week 8 ○ patients who were randomized to receive placebo at week 0 of the induction study and were not in clinical response at induction week 8 but were in clinical response at induction week 16 after receiving a dose of IV ustekinumab (~6 mg/kg) at induction week 8.
	Exclusion Criteria	• Severe extensive colitis. • UC limited to the rectum only or < 20 cm of the colon. • Presence of a stoma, a fistula, a bowel obstruction, or adenomatous colonic polyps that were not removed. • Diagnosis of indeterminate colitis, microscopic colitis, ischemic colitis, or Crohn disease, or clinical findings suggestive of Crohn disease. • A stool culture or other examination that was positive for an enteric pathogen, including *Clostridium difficile* toxin, in the previous 4 months, unless a repeat examination was negative and there were no signs of ongoing infection with that pathogen.	• Patients who did not demonstrate clinical response to the study drug following induction. • Patients who initiated or increased the dose of UC-specific medication (or any prohibited medication) during the induction study.

		UNIFI induction	UNIFI maintenance
Drugs	**Intervention**	Single dose of ustekinumab IV at week 0 as follows: • Low-dose group: 130 mg • High-dose group: ~6 mg/kg IV (weight ≤ 55 kg: 260 mg; weight > 55 and ≤ 85 kg: 390 mg; > 85 kg: ustekinumab 520 mg)	Ustekinumab SC injection at week 0/baseline visit of the maintenance phase, randomized to: • ustekinumab 90 mg SC q.12.w • ustekinumab 90 mg SC q.8.w.
	Comparator(s)	Placebo IV (10 mM L-histidine, 8.5% (w/v) sucrose, 0.04% (w/v) polysorbate 80, 0.4 mg/mL L-methionine, and 20 mcg/mL disodium salt)	Placebo as a sterile liquid for SC injection at a fill volume of 1.0 mL in a single-use dose containing L-histidine, sucrose, and polysorbate 80 at pH 6.0
Duration	**Phase**		
	Run-in	8 weeks of screening	Induction study
	Double blind	8 weeks	44 weeks
	Follow-up	Up to 16 weeks	Up to 220 weeks
Outcomes	**Primary End Point**	Proportion of patients on clinical remission at week 8; two definitions were used: • the global definition (outside the US): a Mayo score of ≤ 2 points, with no individual subscore > 1 • The US definition: an absolute stool number ≤ 3, a Mayo rectal bleeding subscore of 0, and a Mayo endoscopy subscore of 0 or 1	Clinical remission at week 44, with two definitions: • global definition: a Mayo score of ≤ 2 points, with no individual subscore > 1 • US definition: an absolute stool number ≤ 3, a Mayo rectal bleeding subscore of 0, and a Mayo endoscopy subscore of 0 or 1
	Secondary and Exploratory End Points	Secondary: • endoscopic healing at week 8 (endoscopy subscore of 0 or 1) • clinical response at week 8 (decrease from baseline Mayo score of > 30% and > 3 points), with either a decrease from baseline in the rectal bleeding subscore of > 1 or a rectal bleeding subscore of 0 or 1 • change from baseline in IBDQ score at week 8 Exploratory: • modified Mayo score • UCEIS • BSFS	Secondary: • efficacy in maintaining clinical response in patients induced into clinical response • endoscopic healing in patients induced into clinical response • achieving corticosteroid-free clinical remission • maintaining clinical remission in patients induced into clinical remission with ustekinumab • mucosal healing • health-related quality of life • pharmacokinetics and immunogenicity as well as levels of CRP, fecal calprotectin, and lactoferrin Exploratory: • response using the Mayo score without the Physician's Global Assessment subscore
Notes	**Publications**	Sands (2018)[30] Danese (2019)[31] Adedokun (2019)[32] Sands (2019a)[29] Sands (2019b)[33] Li (2019)[34]	Sands (2019a)[29] Sands (2019c)[35] Sandborn (2019)[36] Van Assche (2019)[37]

BSFS = Bristol Stool Form Scale; CRP = C-reactive protein; DB = double blind; IBDQ = Inflammatory Bowel Disease Questionnaire; q.8.w. = every eight weeks; q.12.w. = every 12 weeks; RCT = randomized controlled trial; UCEIS = Ulcerative Colitis Endoscopic Index of Severity; w/v = weight by volume.

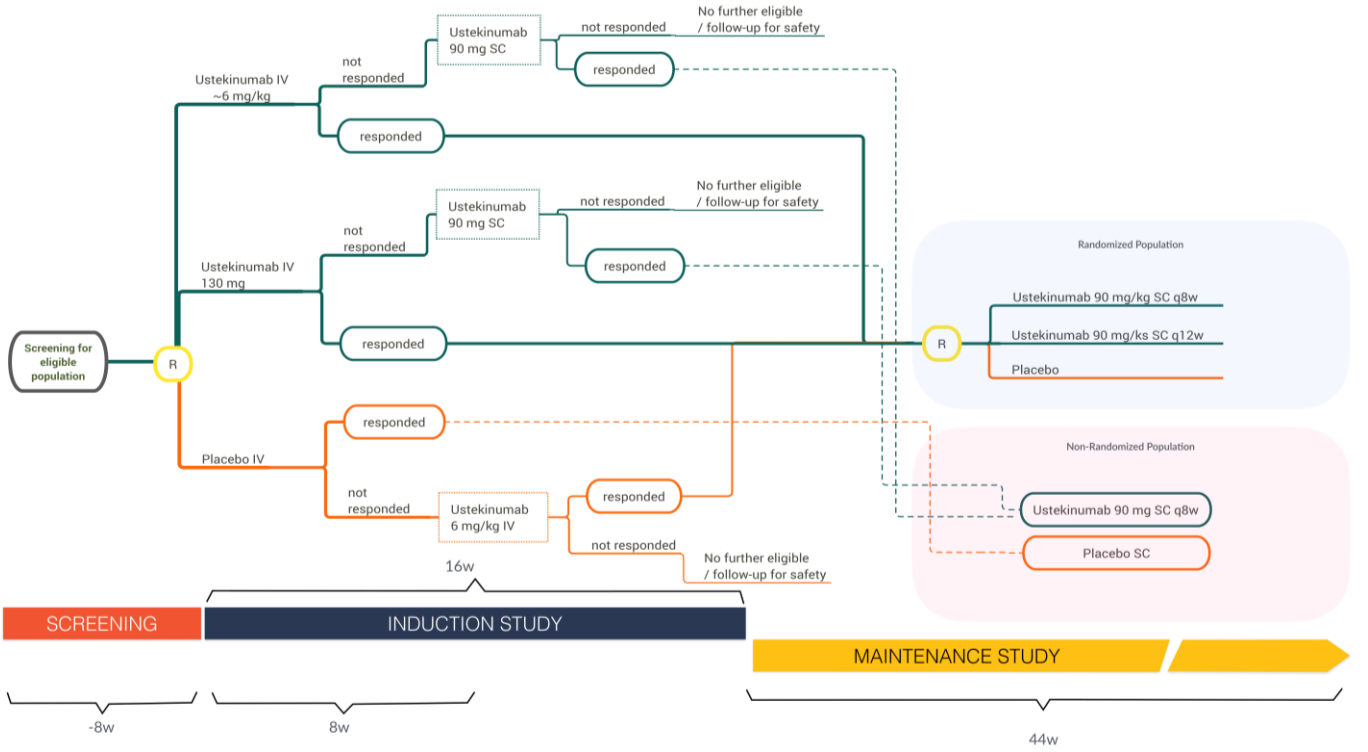
Note: Two additional reports included Clinical Study Report for the UNIFI induction study[1] and UNIFI maintenance study.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance studies.[2]

Description of Studies

One study was included that directly compared ustekinumab versus placebo. The UNIFI study[1,2] is a double-blind, placebo-controlled trial, subdivided into two main phases: induction and maintenance conducted in several countries (including Canada) (detailed in Table 4). A visual summary of both phases is depicted in Figure 2.

Figure 2: Design of the Induction and Maintenance Phases of the UNIFI Study



q8w. = every eight weeks; q12w = every 12 weeks; R = randomization; SC = subcutaneous; w = weeks.

Note: Ustekinumab-induction responders entered the maintenance study at induction study week 8, while ustekinumab delayed responders entered maintenance phase at week 16. Times indicate the duration of each phase more than the timing of initiation of therapies.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance studies.[2]

Induction Study

The first phase, or induction trial, was an eight-week double-blind, randomized, placebo-controlled trial, with another eight-week extension in patients not responding to ustekinumab. The primary objectives of the induction study were to evaluate the efficacy in inducing clinical remission and safety of IV ustekinumab in patients with moderately to severely active UC. Secondary objectives included evaluating the efficacy of IV ustekinumab in inducing endoscopic healing (i.e., improvement in the endoscopic appearance of the mucosa), in inducing clinical response, on disease-specific health-related quality of life, on mucosal healing (a combination of endoscopic healing and histologic healing), and to evaluate the pharmacokinetics, immunogenicity, and pharmacodynamics of ustekinumab-induction therapy, including changes in CRP, fecal

calprotectin, fecal lactoferrin, and other pharmacodynamics biomarkers. The study aimed to evaluate the efficacy of ustekinumab IV by biologic failure status.

Patients were randomized in a 1:1:1 ratio at week 0 to receive a single IV fixed dose of 130 mg of ustekinumab, a weight range–based dose of ustekinumab of approximately 6 mg/kg (i.e., 260 mg [weight ≤ 55 kg], 390 mg [weight > 55 kg and ≤ 85 kg], or 520 mg [weight > 85 kg]), or placebo.

Patients who had a clinical response to either dose of IV ustekinumab at week 8 entered the maintenance randomized phase, as well as patients in the placebo group who did not respond at week 8 but responded at week 16 after receiving an IV dose of 6 mg/kg of ustekinumab (Figure 2). Clinical response was defined as a decrease in the total Mayo score of 30% or more and 3 or more points from baseline, with an accompanying decrease of 1 or more points on the rectal bleeding component of the Mayo scale or a rectal bleeding subscore of 0 or 1.

Those patients who did not respond to an initial dose of IV ustekinumab (either 130 mg or approximately 6 mg/kg) received a second dose of 90 mg of ustekinumab subcutaneously at week 8. Patients who responded to this second dose entered the maintenance study as well as those who had entered the placebo-induction phase and responded at week 8 (without any additional dose of ustekinumab). All of these patients entered into a non-randomized subpopulation of the maintenance study, as described subsequently and in Figure 2. Patients who did not respond were followed up for safety.

All patients were randomized using permuted blocks, with stratification according to previous treatment failure with biologic drugs (yes or no) and geographic region (eastern Europe, Asia, or rest of world). The randomization schedule was concealed with a central randomization scheme under the supervision of the sponsor using an interactive web response system (IWRS) to generate a treatment assignment. The IWRS assigned a treatment code that dictated the treatment assignment and matching study-drug kit for each patient. Blinding of patients, investigators, and clinicians was obtained through dispensing identical packages with labels that did not identify the container contents.

Detailed descriptions of the included patients are presented in Table 5.

Maintenance Study

The second phase, or maintenance study, was a 44-week double-blind, randomized, placebo-controlled trial that included patients from the induction study who responded to ustekinumab (either 130 mg or approximately 6 mg/kg) at week 8, or those in the placebo group from the induction study who did not respond at week 8 but responded to a dose approximating 6 mg/kg at week 16 (Figure 2). Patients from this **randomized population** were assigned, in a 1:1:1 ratio, to receive SC injections of 90 mg of ustekinumab every 12 weeks, 90 mg of ustekinumab every eight weeks, or placebo through week 44. The maintenance study also aimed to evaluate the efficacy of maintenance therapy by biologic failure status.

The primary objectives of the maintenance study were to evaluate clinical remission for SC maintenance regimens of ustekinumab in patients induced into clinical response with ustekinumab, and to evaluate the safety of SC maintenance regimens of ustekinumab. Secondary objectives included evaluating the efficacy of ustekinumab in maintaining clinical response, endoscopic healing (i.e., improvement in the endoscopic appearance of the mucosa), achieving corticosteroid-free clinical remission, maintaining clinical remission,

effect on mucosal healing (i.e., a combination of endoscopic healing and histologic healing), the impact on disease-specific, patient-reported health-related quality of life, the pharmacokinetics and immunogenicity of ustekinumab, as well as changes in levels of CRP, fecal calprotectin, fecal lactoferrin, and other pharmacodynamic biomarkers.

In this maintenance phase, patients from the induction study could also be enrolled into a **non-randomized population** if they were members of the placebo IV induction group who responded at week 8 (these patients were assigned to placebo SC every eight weeks for 44 weeks). Also, if they were from the ustekinumab IV groups (either 130 mg or approximating 6 mg/kg) and did not respond at week 8 but responded to a dose of ustekinumab at week 16 (i.e., patients with a delayed response to ustekinumab), they were assigned to receive ustekinumab SC 90 mg every eight weeks. Non-randomized patients were followed for both efficacy and safety but were not included in the key efficacy analyses.

Eligible patients were randomized using a permuted block randomization schedule and stratified by their status of clinical remission (defined as a Mayo score of ≤ 2 points with no individual subscore > 1) at maintenance baseline (yes or no), oral corticosteroid use at maintenance baseline (yes or no), and induction treatment (placebo IV [induction week 0] moving to 6 mg/kg IV of ustekinumab at induction week 8; ustekinumab 130 mg IV [at induction week 0]; or 6 mg/kg IV of ustekinumab [at induction week 0]) as variables.

The randomization schedule was concealed with a central randomization scheme under the supervision of the sponsor using an IWRS for a treatment assignment. The IWRS assigned a treatment code that dictated the treatment assignment and matching study-drug kit for each patient. The blinding of patients, investigators, and clinicians was obtained by dispensing identical packages with labels that did not identify the container contents. Treatment assignment blinding was maintained (for both the induction and maintenance studies) for investigative sites, site monitors, and patients participating in this protocol until the week 44 analyses were completed.

Detailed descriptions of the included patients are presented in Table 6.

Populations

Inclusion and Exclusion Criteria

For the induction study, adult patients over 18 years of age with moderately to severely active UC, as defined by a Mayo score of 6 to 12, inclusive, at week 0 of the study, including an endoscopy subscore of 2 or higher as assessed during the central review of the video of the endoscopy, were eligible. Patients could have received treatment with one or more TNF antagonists or vedolizumab at a dose approved for the treatment of UC and, if they either did not respond initially, responded initially but then lost response, or were intolerant of the medication, they were considered to have a “biologic failure” and were classified in this subgroup. Patients may also have been biologic-naive or may have been exposed to biologic therapy but did not demonstrate an adequate response or demonstrated intolerance to treatment with a biologic drug. These patients must have demonstrated an inadequate response to, or failed to tolerate, at least one of the following conventional UC therapies: oral or IV corticosteroids or the immunomodulators azathioprine or 6-mercaptopurine. Patients who demonstrated corticosteroid dependence (i.e., an inability to successfully taper corticosteroids without a return of the symptoms of UC) were also eligible for entry into the study.

For the maintenance study, patients were required to be in clinical response to treatment during the induction study; this included patients in clinical response to IV ustekinumab induction, in clinical response to IV placebo, or in delayed clinical response to ustekinumab, as described earlier and in Figure 2.

Baseline Characteristics

Key baseline demographic characteristics from the induction and maintenance studies (randomized and non-randomized populations) are summarized in Table 5, Table 6, and Table 7, respectively. All data were obtained from the efficacy populations.

In the induction study, the variables measured, such as age, sex, weight, and race, were similar in their distribution between the ustekinumab and placebo groups. Overall, the median duration of disease was 5.97 years, with 45.7% of patients with extensive disease and a median Mayo score of 9.0. Also, at baseline in the induction study, 84.4% of patients had moderate UC (Mayo score of 6 to 10), 15.3% had severe disease (Mayo score > 10), and 51.1% of patients had a history of biologic failure. Fecal lactoferrin and fecal calprotectin median values were greater in patients in the ustekinumab groups than in the placebo group (data not shown). However, the percentage of patients with abnormal levels of inflammatory biomarkers (e.g., CRP) at baseline were similar across study groups.

At baseline, 90.2% of patients in the induction study had a concomitant medication with similar distribution between the study groups, except for those receiving aminosalicylates, with a greater proportion of patients receiving aminosalicylates in the group receiving 6 mg/kg of ustekinumab compared with the ustekinumab 130 mg and placebo groups. As well, fewer patients in the placebo group (49.2%) were receiving corticosteroids at baseline versus those in the ustekinumab groups (53.1% combined). Overall, 51.8%, 28.2%, and 68.7% of patients in the induction study were receiving, respectively, corticosteroids, immunomodulatory drugs, and aminosalicylates. A majority of patients (94.0%) had either an inadequate response to or were intolerant to corticosteroids, 6-mercaptopurine, or azathioprine, or demonstrated corticosteroid dependence; the proportions were similar between groups (data not shown). The proportions of patients who had a history of documented biologic failure were similar among treatment groups.

Table 5: Summary of Baseline Characteristics – UNIFI Induction Study

	Ustekinumab IV			Placebo N = 319
	130 mg N = 320	~6 mg/kg N = 322	Combined N = 642	
Sex female, n (%)	130 (40.6)	127 (39.4)	257 (40.0)	122 (38.2)
Age, years, mean (SD)	42.2 (13.94)	41.7 (13.6)	41.9 (13.80)	41.2 (13.50)
Race, n (%)				
White Black Asian Native American Pacific Islander Other Not reported Unknown	239 (74.7) 6 (1.9) 46 (14.4) 0 0 9 (2.8) 18 (5.6) 2 (2.6)	243 (75.5) 0 49 (15.2) 1 (0.3) 0 12 (3.7) 16 (5.0) 1 (0.3)	482 (75.1) 6 (0.9) 95 (14.8) 1 (0.2) 0 21 (3.3) 34 (5.3) 3 (0.5)	248 (77.7) 3 (0.9) 48 (15) 0 0 8 (2.5) 12 (3.8) 0
Weight (kg), mean (SD)	73.67 (16.80)	73.02 (19.25)	73.34 (18.06)	72.91 (16.77)
Height (cm), mean (SD)	171.28 (9.33)	171.49 (9.73)	171.39 (9.53)	172.31 (10.03)
Duration of disease, years, Mean (SD)	8.13 (7.17)	8.17 (7.82)	8.15 (7.50)	8.01 (7.19)
Extent of disease				
N	318	320	638	316
Limited to left side of colon Extensive	183 (57.5) 135 (42.5)	168 (52.5) 152 (47.5)	351 (55.0) 287 (45.0)	167 (52.8) 149 (47.2)
Mayo score of 0 to 12				
N	320	321	641	319
Mean (SD)	8.9 (1.57)	8.9 (1.51)	8.9 (1.54)	8.9 (1.62)
Severity of disease				
N	320	321	641	319
Moderate (Mayo score of 6 to 10), n (%)	271 (84.7)	276 (86.0)	547 (85.3)	263 (82.4)
Severe (Mayo score > 10), n (%)	48 (15.0)	45 (14.0)	93 (14.5)	54 (16.9)
Concomitant medications for UC at baseline, n (%)				
Any UC medication	290 (90.6)	294 (91.3)	584 (91.0)	283 (88.7)
Corticosteroids	173 (54.1)	168 (52.2)	341 (53.1)	157 (49.2)
Immunomodulators				
6-MP/AZA	88 (27.5)	85 (26.4)	173 (26.9)	88 (27.6)
MTX	5 (1.6)	4 (1.2)	9 (1.4)	1 (0.3)
Aminosalicylates	215 (67.2)	238 (73.9)	453 (70.6)	207 (64.9)
Biologic failure status				
N	320	321	641	319
Biologic status failure, n (%)	164 (51.3)	166 (51.6)	330 (51.4)	161 (50.5)

6-MP = 6-mercaptopurine; AZA = azathioprine; MTX = methotrexate; SD = standard deviation; UC = ulcerative colitis.

[a] All percentages and values for each variable are based on the total N from the ITT population unless otherwise specified.

Source: Clinical Study Report of the UNIFI induction study.[1]

The maintenance study was composed of randomized and non-randomized populations. Among the randomized population, variables were well balanced across treatment groups, except for sex (a greater proportion of patients in the placebo group [61.1%] were male compared with the ustekinumab every eight weeks group [53.4%] and ustekinumab every 12 weeks group [55.8%]), race (a greater proportion of patients in the placebo group [19.4%] were Asian compared with the ustekinumab every 12 weeks group [14.0%]), and region, due to the placebo group (17.7%) having a greater proportion of patients from Asia compared with the ustekinumab every 12 weeks group (12.2%), while a greater proportion of patients in the ustekinumab every 12 weeks group (46.5%) were from Eastern Europe compared with the placebo group (38.9%) and the ustekinumab every eight weeks group (38.1%). Among all randomized patients, 56.8% were male and 74.0% were white, with a median age of 40.0 years and a median weight of 70.0 kg. The mean duration of disease was 6.05 years, while the proportion of patients with extensive UC was 47.1%. In the same randomized population, clinical characteristics were, in general, equally distributed among groups, except for the median fecal lactoferrin and fecal calprotectin levels, which were higher in the ustekinumab every eight weeks group.

All patients in the maintenance study were from the induction study, denoting that the previous and concomitant medications were well controlled and stably maintained. However, some differences were noted in the proportion of patients receiving glucocorticoids at induction baseline, which was lower in the ustekinumab every 12 weeks group (48.3%) compared with the every eight weeks group (54.0%) and the placebo group (54.3%). Also, a difference was noted in patients receiving aminosalicylates at induction baseline in the ustekinumab every 12 weeks group (77.9%) versus the ustekinumab every eight weeks (63.6%) group and the placebo group (70.9%).

In the non-randomized population, a greater proportion of white males was noted among the ustekinumab-induction responders. All other variables were similarly distributed between groups. However, CRP concentrations were different between groups of study, with a higher median CRP concentration in the ustekinumab-induction delayed responders (5.30 mg/L) compared with the primary population (3.58 mg/L) (data not shown). Furthermore, the proportion of patients with extraintestinal manifestations was higher in the ustekinumab induction delayed-responder group (32.5%) when compared with the primary population (26.4%). Clinical disease characteristics at maintenance baseline in the ustekinumab group reflected a higher level of disease activity compared with the primary population. The ustekinumab group had fewer patients in clinical remission (13.4% versus 23.5%) and demonstrated less endoscopic healing (22.9% versus 37.5%). These patients also had higher median Mayo scores (5.0 versus 4.0), higher median CRP levels (2.2 mg/L versus 1.58 mg/L), higher median lactoferrin levels (54.82 mcg/g versus 42.48 mcg/g), and higher median fecal calprotectin levels (500.0 mg/kg versus 426.0 mg/kg).

Concurrent medication had a similar distribution between groups in the non-randomized population, except for a greater proportion (34.4%) of patients reporting use of immunomodulatory drugs at induction baseline in the ustekinumab group (delayed responders) compared with the primary population (26.6%). A larger proportion of patients in the ustekinumab (delayed responders) group were biologic failures and refractory to, dependent on, or intolerant of corticosteroid treatment.

Table 6: Summary of Baseline Characteristics – UNIFI Maintenance Study, Randomized Population

	Ustekinumab SC			Placebo N = 175
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 176	Combined N = 348	
Sex female, n (%) [a]	76 (44.2)	82 (46.6)	158 (45.4)	68 (38.9)
Age, years, mean (SD)	40.7 (13.47)	39.5 (13.32)	40.1 (13.38)	42.0 (13.85)
Race, n (%)				
White Black Asian Native American Pacific Islander Other Not reported Unknown	135 (78.5) 0 24 (14.0) 0 0 9 (5.2) 3 (1.7) 1 (0.6)	127 (72.2) 3 (1.7) 29 (16.5) 0 0 5 (2.8) 12 (6.8) 0	262 (75.3) 3 (0.9) 53 (15.2) 0 0 14 (4) 15 (4.3) 1 (0.3)	125 (71.4) 3 (1.7) 34 (19.4) 0 0 3 (1.7) 9 (5.1) 1 (0.6)
Weight (kg), mean (SD)	73.27 (18.90)	72.04 (19.11)	72.64 (18.99)	71.68 (14.61)
Height (cm), mean (SD)	171.32 (9.68)	170.91 (9.96)	171.11 (9.81)	171.02 (10.07)
Duration of disease (years), mean (SD)	8.60 (8.30)	8.08 (6.57)	8.34 (7.47)	7.48 (6.79)
Extent of disease				
N	172	175	347	175
Limited to left side of colon Extensive	92 (53.5) 80 (46.5)	95 (54.3) 80 (45.7)	187 (53.9) 160 (46.1)	89 (50.9) 86 (49.1)
Mayo score 0 to 12				
Mean (SD)	8.9 (1.58)	8.9 (1.55)	8.9 (1.56)	8.7 (1.52)
Severity of disease				
Moderate (Mayo score ≤ 10), n (%)	150 (87.2)	147 (84.5)	297 (85.8)	156 (89.1)
Severe (Mayo score from 11 to 12), n (%)	22 (12.8)	27 (15.5)	49 (14.2)	19 (10.9)
Concomitant medications for UC at induction baseline, n (%)				
Any UC medication	154 (89.5)	155 (88.1)	309 (88.8)	160 (91.4)
Corticosteroids	83 (48.3)	95 (54.0)	178 (51.1)	95 (54.3)
Immunomodulators				
6-MP/AZA	43 (25.0)	45 (25.6)	88 (25.3)	47 (26.9)
MTX	1 (0.6)	1 (0.6)	2 (0.6)	2 (1.1)
Aminosalicylates	134 (77.9)	112 (63.6)	246 (70.7)	124 (70.9)
Biologic failure status, n (%)	70 (40.6)	91 (51.7)	161 (46.2)	88 (50.2)

6-MP = 6-mercaptopurine; AZA = azathioprine; MTX = methotrexate; q.8.w. = every eight weeks; q.12.w. = every 12 weeks; SD = standard deviation.

[a] All percentages and values for each variable are based on the total N from the randomized population, unless otherwise specified.

All values represent the baseline characteristics of patients when entered the induction study and are now broken down based on which group they ended up being assigned to.

Source: Clinical Study Report[2] for the UNIFI maintenance study.

Table 7: Summary of Baseline Characteristics – UNIFI Maintenance Study, Non-Randomized Population

	Ustekinumab 90 mg SC q.8.w.[a] N = 157	Placebo[b] N = 103
Sex female, n (%)	55 (35.0)	38 (36.9)
Age, years, mean (SD)	43.9 (13.60)	43.6 (14.14)
Race, n (%)		
White Black Asian Native American Pacific Islander Other Not reported Unknown	124 (79.0) 1 (0.6) 22 (14.0) 0 0 2 (1.3) 7 (4.5) 1 (0.6)	83 (80.6) 1 (1.0) 12 (11.7) 0 0 2 (1.9) 5 (4.9) 0
Weight (kg), mean (SD)	74.94 (19.25)	74.33 (18.15)
Height (cm), mean (SD)	172.60 (9.25)	172.39 (10.18)
Duration of disease (years), mean (SD)	8.49 (7.58)	9.01 (8.87)
Extent of disease		
N	155	103
Limited to left side of colon Extensive	89 (57.4) 66 (42.6)	62 (60.2) 41 (39.8)
Mayo score 0 to 12, mean (SD)	9.0 (1.53)	8.7 (1.61)
Severity of disease		
Moderate (Mayo score ≤ 10), n (%)	129 (82.2)	89 (87.3)
Severe (Mayo score 11 to 12), n (%)	28 (17.8)	13 (12.7)
Concomitant medications for UC at induction baseline, n (%)		
Any UC medication	148 (94.3)	98 (95.1)
Corticosteroids	82 (52.2)	59 (57.3)
Immunomodulators		
6-MP/AZA	52 (33.1)	33 (32.0)
MTX	2 (1.3)	1 (1.0)
Aminosalicylates	118 (75.2)	75 (72.8)
Biologic failure status,[c] n (%)	82 (52.2)	46 (44.6)

6-MP = 6-mercaptopurine; AZA = azathioprine; MTX = methotrexate; q.8.w. = every eight weeks; SD = standard deviation; UC = ulcerative colitis.

Note: All values represent the baseline characteristics of patients when they entered the induction study and are now broken down based on which group they ended up being assigned to.

[a] Delayed responders, i.e., patients who were not in clinical response to ustekinumab IV at week 8 but were in clinical response at week 16 after subcutaneous administration of ustekinumab at week 8.

[b] Responders to placebo IV induction.

[c] Biologic failures at the baseline of the induction study.

Source: UNIFI maintenance study[2] Clinical Study Report.

Interventions

Induction Study

In the induction phase of the UNIFI study, patients received a single IV dose of ustekinumab or placebo at week 0. The interventions were administered in a one-hour lapse and completed within four hours of preparation of the study drug. Those patients who did not have a clinical response at week 8 received an additional dose at week 8 (Figure 2). An approximate dose of 6 mg/kg was administered using a range of weight-based doses: patients weighing 55 kg or less received ustekinumab 260 mg; those weighing from 55 kg to 85 kg received ustekinumab 390 mg; and those weighing more than 85 kg received ustekinumab 520 mg. The investigators considered the latter group to be the "high" dose group. They also utilized a "low" dose of 130 mg IV based on previous studies conducted in patients with Crohn disease. The placebo was an IV infusion of 10 mM L-histidine, 8.5% sucrose (weight by volume), 0.04% polysorbate 80 (weight by volume), 0.4 mg/mL L-methionine, and 20 mcg/mL disodium salt dihydrate at pH 6.0, supplied as a single-use, sterile solution in 30 mL vials with a nominal volume of 26 mL. According to the investigators, the placebo had the same appearance as the ustekinumab solutions.

Maintenance Study

For the maintenance phase, ustekinumab was supplied as sterile liquid for SC injection in a single-use pre-filled syringe with 90 mg (1.0 mL fill volume of liquid) ustekinumab. Placebo was supplied as a sterile liquid for SC injection at a fill volume of 1.0 mL in a single-use pre-filled syringe. Each placebo syringe contained L-histidine, sucrose, and polysorbate 80 at pH 6.0. Placebo solutions had the same appearance as the ustekinumab preparations.

Ustekinumab was used at a dosage of 90 mg every eight weeks or every 12 weeks to create the two groups of ustekinumab (i.e., the every eight weeks and every 12 weeks groups, respectively). Patients started their assigned dose of the SC study drug at the week 0 visit. Next, all patients received each study drug at all scheduled administration visits to maintain blinding of patients and researchers. Investigators used a schedule of every eight weeks based on previous evidence from the Crohn disease patients. Furthermore, they wanted to investigate a lower dose regimen of the same dose every 12 weeks that would provide enough systemic exposure and meet safety and efficacy requirements. Concomitant medications included oral 5-aminosalicylate compounds, oral corticosteroids, or immunomodulators (i.e., 6-mercaptopurine, azathioprine, or methotrexate) and were allowed from week 0 of the induction study through week 0 of the maintenance study unless the therapy had to be discontinued or reduced in dose because of toxicity or other medical reason. Therapy was not to be restarted if this occurred.

Any medication (e.g., glucocorticoids, 6-mercaptopurine, azathioprine, methotrexate, 5-aminosalicylates) for the treatment of UC that needed to be initiated or increased in any phase of the study was considered a rescue medication.

Outcomes

The primary outcome of the UNIFI study was to evaluate the efficacy of ustekinumab in inducing clinical remission in patients with moderately to severely active UC as well as evaluate safety. For this, several outcomes were assessed in both the induction and maintenance phases of the UNIFI study.

Clinical remission: Investigators used two definitions for clinical remission: the US definition, i.e., an absolute stool number of 3 or less, a rectal bleeding subscore of 0, and a Mayo endoscopy subscore of 0 or 1; and the global definition, i.e., a Mayo score or 2 or higher, with no individual subscore higher than 1. Both were evaluated in all patients at week 8.

Clinical response: This was defined as a decrease from induction baseline in the Mayo score of 30% or more and a decrease of 3 or more points, with either a decrease from baseline in the rectal bleeding subscore of 1 or more or a rectal bleeding subscore of 0 or 1.

Endoscopic healing: An improvement in the endoscopic appearance of the mucosa defined as a Mayo endoscopy subscore of 0 or 1.

Histologic healing: Based on features of the Geboes score, this was defined as neutrophil infiltration in less than 5% of crypts, no crypt destruction, and no erosions, ulcerations, or granulation tissue.

Mucosal healing: A combination of endoscopic healing and histologic healing.

Normal or inactive mucosal disease: A Mayo endoscopy subscore of 0.

Symptomatic remission: A Mayo stool frequency subscore of 0 or 1 and a rectal bleeding subscore of 0.

Normalization of CRP concentration: A CRP concentration of 3 mg/L or less.

Normalization of fecal lactoferrin concentration: A fecal lactoferrin concentration of 7.24 mcg/g or less.

Normalization of fecal calprotectin concentration: A fecal calprotectin concentration of 250 mg/kg or less.

For the evaluation of these outcomes, several instruments are noted, and detailed descriptions may be found in Appendix 4: Description and Appraisal of Outcome Measures.

- Mayo score and partial Mayo score
- Inflammatory Bowel Disease Questionnaire (IBDQ)
- Short Form (36) Health Survey (SF-36)
- EuroQol 5-Dimensions (EQ-5D)

Safety was evaluated based on the frequency of AEs, which were usually willingly reported by the patient or by observation or interview by the clinician or investigator.

AEs were coded in accordance with the Medical Dictionary for Regulatory Activities (MedDRA) Version 20.0 and were monitored through week 44 by treatment group. Investigators evaluated:

- any AE, i.e., any event with a “very likely,” “probable,” or “possible” relationship to the study drug
- serious AEs
- reasonably related AEs
- AEs leading to discontinuation of the study drug
- infections, including infections requiring oral or parenteral antibiotic treatment

- serious infections
- injection-site reactions

Statistical Analysis

Induction Study

For the induction study, sample size and power were based on the chi-square test to detect a significant difference in clinical remission at week 8 between the ustekinumab and placebo groups. The authors assumed a different baseline clinical remission rate for the global versus the US definition for remission. For the global definition, they used a 7% clinical remission rate as the baseline rate in the placebo group and considered an effect difference of 12%, that is, a hypothesized rate of 19% in the ustekinumab group (for both the 6 mg/kg [approximately] and 130 mg doses). For the US definition, they considered a baseline rate of 12% in the placebo group versus 25% in both intervention groups (difference of 13%). With these values, 220 patients per group (660 patients in total) and 135 patients per group (405 patients in total) were required for the US and global definitions, respectively. To provide a sufficient number of patients for the primary population of the maintenance phase, 951 patients (317 per group) would need to be enrolled in the induction study.

The major secondary outcomes per the global analysis were the percentage of patients with endoscopic healing, the percentage of patients achieving clinical response, and the change from baseline in the IBDQ total score to week 8. These outcomes were the same for the US-defined population, except for the change from baseline in the IBDQ total score to week 8, per advice from the FDA. The percentage of patients with mucosal healing (endoscopic plus histologic) was another important secondary outcome.

Tests for multiplicity were also different for the US versus the other countries. For instance, in the US, the Bonferroni method was used, while a step-up Hochberg procedure was performed in the other countries. Power was calculated to reach 90% using a step-up Hochberg approach at the 0.05 (two-sided) level in the global definition group, while the same power in the US definition group was reached at a significance level of 0.025 (two-sided) based on the Bonferroni testing approach using the outcome of clinical remission. Assumed remission rates were based on previous data from the golimumab study (induction phase).[38] In the induction study, multiplicity was assessed using a Cochran-Mantel-Haenszel (CMH) chi-square test for comparing proportions and stratified by biologic failure status (yes or no) and region (Eastern Europe, Asia, or rest of world) for all outcomes except for the third major secondary end point of change from baseline in the IBDQ score at week 8. The treatment groups were compared using an analysis of covariance for continuous outcomes on the van der Waerden normal scores with baseline IBDQ score, biologic failure status, region, and treatment group as covariates.

For the Hochberg step-up adjustment for multiplicity of the global primary outcome, if the P values for the ustekinumab 130 mg group versus the placebo group, and the ustekinumab 6 mg/kg (approximately) group versus the placebo group, were less than 0.05, then it was concluded that the treatments used in both ustekinumab groups were effective compared with the placebo group. Otherwise, the smaller of the two P values was compared with 0.025: if the smaller P value was less than 0.025, then it was concluded that the treatment used in the ustekinumab group associated with the smaller of the two P values was effective compared with the placebo group. The study could be declared positive based on a statistically significant test of the primary outcome for at least one ustekinumab group.

The first major secondary outcome (i.e., endoscopic healing) was tested only if the primary outcome for the ustekinumab group was positive, per the global testing procedure. Subsequent major secondary outcomes for a dose were tested only if the preceding outcome for that dose in the hierarchy was positive at the 0.05 level of significance. If all the primary and major secondary outcomes tested positive for a dose, testing would continue for that dose to the other multiplicity-controlled outcome, mucosal healing at week 8. All other outcomes were not controlled for multiple comparisons.

All efficacy analyses were based on the ITT principle. All patients randomized in the induction study formed the primary efficacy set for analysis. For the safety set of patients, all those who received at least one dose of ustekinumab were considered and analyzed. For patients with missing data, the last observation was carried forward for continuous outcomes and, for dichotomous outcomes, patients with missing data were considered not to have achieved the outcome. Investigators compared the proportion of patients in clinical remission among the ustekinumab and placebo groups using a CMH chi-square test stratified by biologic failure status (yes or no) and region (Eastern Europe, Asia, or rest of world).

Sensitivity analyses were conducted to test the robustness of the primary outcome analyses using a modified ITT analysis (excluding patients who were randomized but not treated) and by using observed cases, last observation carried forward, nonresponder if any missing subscore, and on a "per-protocol" analysis, by using logistic regression to analyze the primary outcome, and based on multiple-imputation methods, on endoscopy subscores, on "worst-case" scenario, and on region (US versus non-US location).

Subgroup analyses were used based on baseline demographics, baseline UC clinical disease characteristics, baseline UC-related concomitant medication use, and UC-related medication history.

Maintenance Study

The efficacy analyses in the maintenance study were based on the primary population. Hence, the primary definitions of the outcomes were based on those outcomes that were also present in the induction study. Likewise, the different definitions between the US and non-US countries and the multiplicity testing approaches were used.

Sample size and power calculations were based on the chi-square test to detect a difference between patients receiving ustekinumab 90 mg SC every eight weeks and those receiving placebo. The sample size calculation was based on previous studies using other anti-TNFs (i.e., golimumab and vedolizumab). Therefore, a baseline clinical remission rate of 20% was assumed for the placebo group (at week 44, based on the US definition), and a rate of 40% was assumed for the ustekinumab every eight weeks group (a risk difference of 20%). Assuming this difference for the global and US definitions, a total of 327 patients were needed (109 in each group) to reach 90% power at a significance level of 0.05. Given that the population for the maintenance study was obtained directly from the induction study, with 317 patients planned for each induction-treatment group, the number of patients in the primary population of the maintenance study was expected to be at least 327.

Analyses using the ITT principle were also planned. The primary outcome analysis set comprised all of the patients randomized at the baseline of the maintenance study, i.e., patients who responded to IV ustekinumab induction at week 8 of the induction study, and patients who were not in clinical response to IV placebo induction at week 8 of the

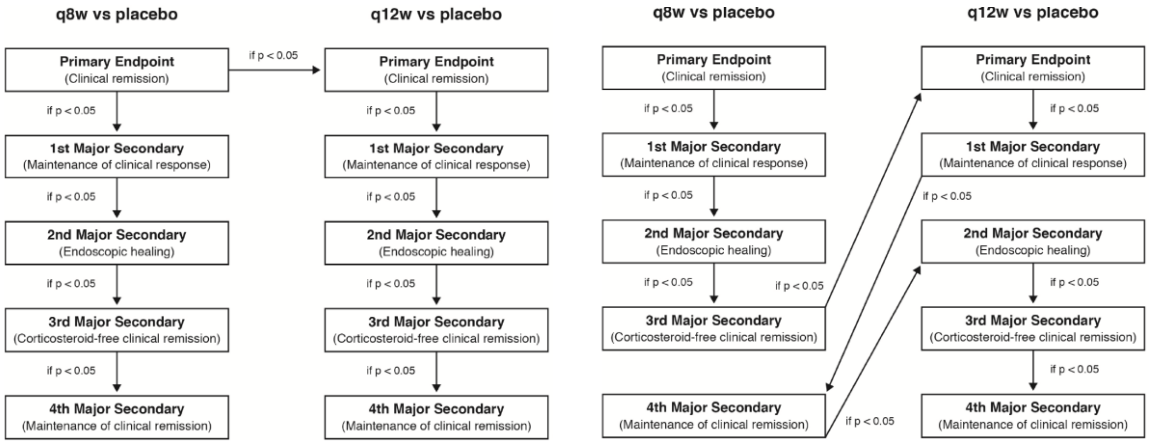
induction study but were in clinical response at induction week 16 after receiving an induction dose of IV ustekinumab at week 8.

Pre-specified efficacy analyses were also conducted in the non-randomized analysis set, which included patients who achieved clinical response to placebo IV induction dosing at week 8 of the induction study, and patients who were delayed responders to ustekinumab induction. The same definitions used in the induction study were used for the missing-data rules. At week 44 of the maintenance study, the proportion of patients in clinical remission was compared between the ustekinumab (every eight weeks or every 12 weeks) and placebo groups using a CMH chi-square test stratified by clinical remission status (yes or no based on the global definition) at maintenance baseline versus the induction-treatment group.

A fixed-sequence testing procedure was used to adjust for multiplicity at the 0.05 level for the global primary outcome. Ustekinumab 90 mg SC every eight weeks was considered statistically significant versus placebo if the P value was less than 0.05, and ustekinumab 90 mg SC every 12 weeks was statistically significant versus placebo if the P values for both the every eight weeks and every 12 weeks maintenance-dose groups were less than 0.05. A positive study was defined as a statistically significant test for the ustekinumab 90 mg SC every eight weeks dosage versus placebo for clinical remission at week 44, regardless of the result of the test for the ustekinumab 90 mg SC every 12 weeks group versus placebo.

For the US-based analyses, a fixed-sequence testing procedure at the 0.05 level across the primary and all four major secondary end points and across the two ustekinumab dosages was conducted, starting with the every eight weeks regimen group versus placebo for clinical remission. The procedure is diagrammed below.

Figure 3: Testing Procedure for the Primary and Major Secondary Outcomes in the UNIFI Maintenance Study



Global Testing Procedure for the Primary and Major Secondary Outcomes

US-Specific Testing Procedure for the Primary and Major Secondary Outcomes

q8w = every eight weeks q12w = every 12 weeks; vs = versus.

Source: UNIFI maintenance study[2] Clinical Study Report.

Analyses of multiplicity-controlled end points, except for the fourth major secondary end point related to maintenance of clinical remission, were conducted using a CMH chi-square test stratified by clinical remission (global definition), status at maintenance baseline (yes or no), and induction treatment. For the fourth major secondary end point (maintenance of clinical remission), a CMH chi-square test stratified by induction treatment was used.

Sensitivity analyses were also conducted to test the robustness of the primary outcome analyses by using observed cases, nonresponder if any missing subscore, and based on multiple-imputation methods, on "worst-case" scenario, on a modified ITT analysis, and on a "per-protocol" analysis, by using logistic regression to analyze the primary outcome, and based on endoscopy subscores.

As in the induction study, subgroup analyses were performed based on demographics and UC clinical disease characteristics, UC-related concomitant medication usage, and UC-related medication history, all at week 0 of the induction study, as well as maintenance-stratification factors and UC clinical disease characteristics at maintenance baseline.

Analysis Populations

Three analysis sets were used in the induction study. The efficacy analysis set, consisting of all patients randomized in the induction study (and analyzed under the ITT principle); the safety analysis set, which included patients who received at least one dose of the study drug, including a partial dose; and the treated analysis set, which consisted of patients who were not in clinical response at week 8 and received an additional dose of ustekinumab SC at week 8.

For the maintenance study, the primary efficacy analysis set consisted of all patients randomized at week 0 of the maintenance study, i.e., in clinical response to IV ustekinumab induction. Also, the non-randomized analysis set was used, which included patients who achieved clinical response to placebo IV induction dosing at week 8 of the induction study and those who were delayed responders to ustekinumab induction. The safety analysis set was defined as it was in the induction study.

Results

Patient Disposition

Induction Study

The number of patients screened for eligibility to enter the induction studies was not reported in the Clinical Study Reports.

In the induction study, 961 patients were included in the randomization schedule, of which 319, 320, and 322 were included in the placebo, ustekinumab 130 mg, and ustekinumab 6 mg/kg (approximately) groups, respectively. Only three patients did not receive the treatment to which they were assigned at week 0. In total, 417 out of 961 patients (43.3%) who were not in clinical response at week 8 received an additional dose of the study drug at week 8. A total of 912 (94.9%) patients completed study participation: 783 (81.5%) entered the maintenance phase and 129 (13.4%) who did not enter the maintenance phase completed the final safety visit (Table 8 and Figure 4). A total of 49 (5.1%) patients terminated study participation (yellow boxes in Figure 4). Twenty patients (2.1%) terminated study participation before week 8. The most common cause was withdrawal of consent, which was done by 14 patients (1.5%): none in the 6 mg/kg group, 5 (1.6%) in the 130 mg

group, and 9 (2.8%) in the placebo group. Among the rest of the patients (29) who terminated study participation: 4 terminated at week 8, and 25 terminated after week 8.

Table 8: Patient Disposition – UNIFI Induction Study (Primary Efficacy Set)

	Ustekinumab IV			Placebo
	130 mg N = 320	~6 mg/kg N = 322	Combined N = 642	N = 319
Screened, N	961			
Randomized, N (%)	320	322	642	319
Discontinued study participation, N (%)	11 (3.43)	15 (4.65)	26 (4.0)	23 (7.2)
Reason for discontinuation of study participation, N (%)				
Adverse events	0	1 (0.2)	1 (0.2)	3 (0.9)
Lost to follow-up	0	1 (0.31)	1 (0.2)	0
Lack of efficacy	–	–	–	–
Withdrawal of consent	9 (2.8)	7 (2.2)	16 (2.5)	17 (5.3)
Sponsor decision	1 (0.3)	0	1 (0.2)	1 (0.3)
Death	0	1 (0.3)	1 (0.2)	0
Other	1 (0.3)	5 (1.6)	6 (0.9)	2 (0.6)
ITT, N	320	322	642	319
PP, N	311	305	616	310
Safety, N	321	320	641	319

ITT = intention to treat; PP = per protocol.

Source: Clinical Study Report for the UNIFI induction[1] study.

Maintenance Study

A total of 783 patients who completed the induction study and were in clinical response to the induction study drug were enrolled in the maintenance study (Figure 4 and Figure 5). Of these, 523 (66.7%) were able to be allocated to the targeted primary population, thus forming the randomized population of the maintenance phase and consisting of those in clinical response to IV ustekinumab induction. Of these 523 patients, 176 were distributed randomly to the ustekinumab 90 mg SC every eight weeks group, 172 to the ustekinumab 90 mg SC every 12 weeks group, and 175 patients to the placebo SC group.

On the other hand, 260 patients were not randomized; these were placebo-induction responders and ustekinumab-induction delayed responders (Figure 5). Of these 260 patients, 103 who were in clinical response to placebo IV induction at induction week 8 (placebo-induction responders) received placebo SC (forming the placebo group of the non-randomized population), while 157 who were ustekinumab-induction delayed responders (i.e., were not in clinical response to ustekinumab at induction week 8 but were in clinical response at induction week 16) received ustekinumab 90 mg SC every eight weeks, forming the ustekinumab group of the non-randomized population.

Prior to week 40, a total of 85 (16.3%) patients from the primary population discontinued the study drug (Figure 5, Table 9). The percentage of patients who discontinued the study drug was greater in the placebo group (24.6%) than in the ustekinumab every eight weeks and every 12 weeks groups (10.2% and 14.0%, respectively). The most common reasons for discontinuation of the study drug were lack of efficacy (6 [3.4%], 7 [4.1%], and 15 [8.6%]

patients in the ustekinumab every eight weeks, ustekinumab every 12 weeks, and placebo groups, respectively) and AE due to worsening of UC (in 0 [0.0%], 4 [2.3%], and 16 [9.1%] patients in the ustekinumab every eight weeks, ustekinumab every 12 weeks, and placebo groups, respectively).

Prior to week 40, 56 (21.5%) patients from the non-randomized population also discontinued the study drug. Among 260 patients from the non-randomized population, 29 (18.5%) from the ustekinumab group and 27 (26.2%) from the placebo group discontinued the study drug. The most common reasons for discontinuation of the study drug among the ustekinumab-induction delayed responders and placebo-induction responders were AEs due to worsening of UC (reported in 8 [5.1%] patients and 8 [7.8%] patients, respectively) and lack of efficacy (reported in 10 [6.4%] patients and 9 [8.7%] patients, respectively) (Table 10).

Table 9: Patient Disposition – UNIFI Maintenance Study, Randomized Population

	Ustekinumab SC			Placebo N = 175
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 176	Combined N = 348	
Screened, N	783			
Randomized, N (%)	172	176	348	175
Discontinued study participation before week 44, N (%)	11 (6.39)	8 (4.54)	19 (5.45)	10 (5.71)
Reason for discontinuing study participation before week 44, N (%)				
Withdraw of consent	10 (5.8)	5 (2.8)	15 (4.3)	8 (4.6)
Lost to follow-up	0	1 (0.6)	1 (0.3)	1 (0.6)
Sponsor decision	0	1 (0.6)	1 (0.3)	1 (0.6)
Death	0	0	0	0
Other	1 (0.6)	1 (0.6)	2 (0.6)	0
Discontinued study drug before week 40, N (%)	24 (14.0)	18 (10.2)	42 (12.1)[a]	43 (24.6)[a]
Reason for discontinuing study drug before week 40, N (%)				
Adverse events	8 (4.7)	4 (2.3)	12 (3.4)	19 (10.9)
Lost to follow-up	0	0	0	1 (0.5)[b]
Lack of efficacy	7 (4.1)	6 (3.4)	13 (3.7)	15 (8.6)
Withdrawal of consent	10 (5.8)	5 (2.8)	15 (4.3)[b]	8 (4.5)[b]
Sponsor decision	0	1 (0.3)	1 (0.2)[b]	1 (0.5)[b]
Death	0	0	0	0
Other	9 (5.23)	8 (4.54)	17 (4.9)	9 (5.2)
ITT, N	172	176	348	175
PP, N	151	157	308	152
Safety, N[c]	172	333	505	277

ITT = intention to treat; PP = per protocol; q.8.w. = every eight weeks; q.12.w. = every 12 weeks.

[a] Total number of patients who discontinued the study drug prior to week 40 are included in the maintenance phase.

[b] Ten patients in the placebo group and 19 in the intervention group terminated study participation prior to week 44 but were already counted as discontinuations.

[c] Includes patients from both randomized and non-randomized sets of the maintenance study.

Source: UNIFI maintenance[2] study Clinical Study Report.

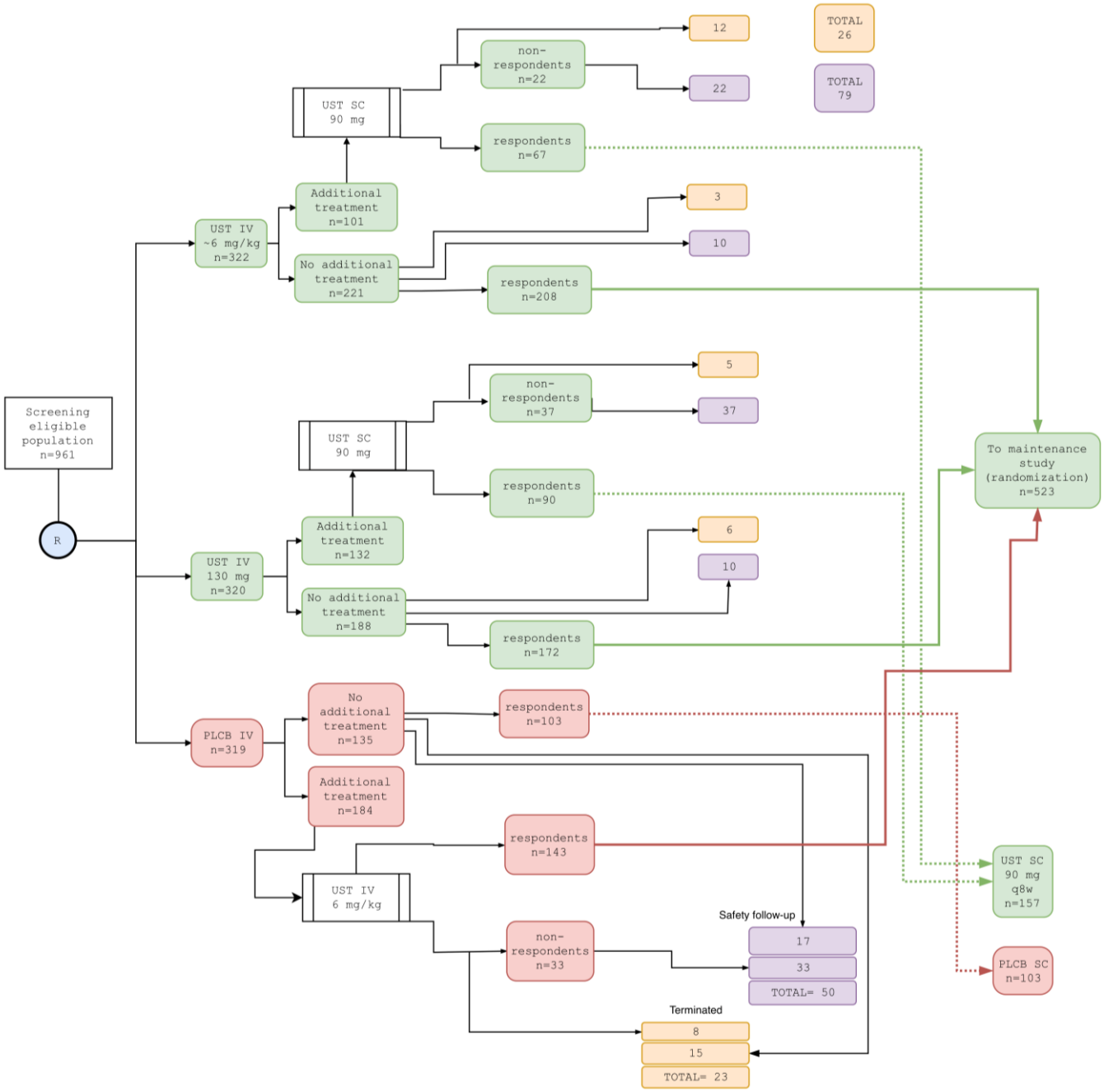
Table 10: Patient Disposition – UNIFI Maintenance Study, Non-Randomized Population

	Ustekinumab 90 mg SC (delayed responders) N = 157	Placebo (induction responders) N = 103
Allocated to group, N (%)	157	103
Discontinued study participation before week 44, N (%)	8 (5.09)	9 (8.73)
Reason for discontinuing study participation before week 44, N (%)		
Withdraw of consent	5 (3.2)	7 (6.8)
Lost to follow-up	1 (0.6)	0
Sponsor decision	0	0
Death	1 (0.6)	0
Other	1 (0.6)	2 (1.9)
Discontinued study drug before week 40, N (%)	29 (18.5)	27 (26.2)
Reason for discontinuing study drug before week 40, N (%)		
Adverse events	10 (6.4)	11 (10.7)
Lost to follow-up	0	0
Lack of efficacy	10 (6.4)	9 (8.7)
Withdrawal of consent	0	0
Death	1 (0.6)	0
Other	6 (3.8)	4 (3.9)
ITT, N	157	103
PP, N	149	94
Safety, N	157	103

ITT = intention to treat; PP = per protocol; SC = subcutaneous.

Source: UNIFI maintenance study[2] Clinical Study Report.

Figure 4: Disposition of Patients in the Induction Phase of the UNIFI Study

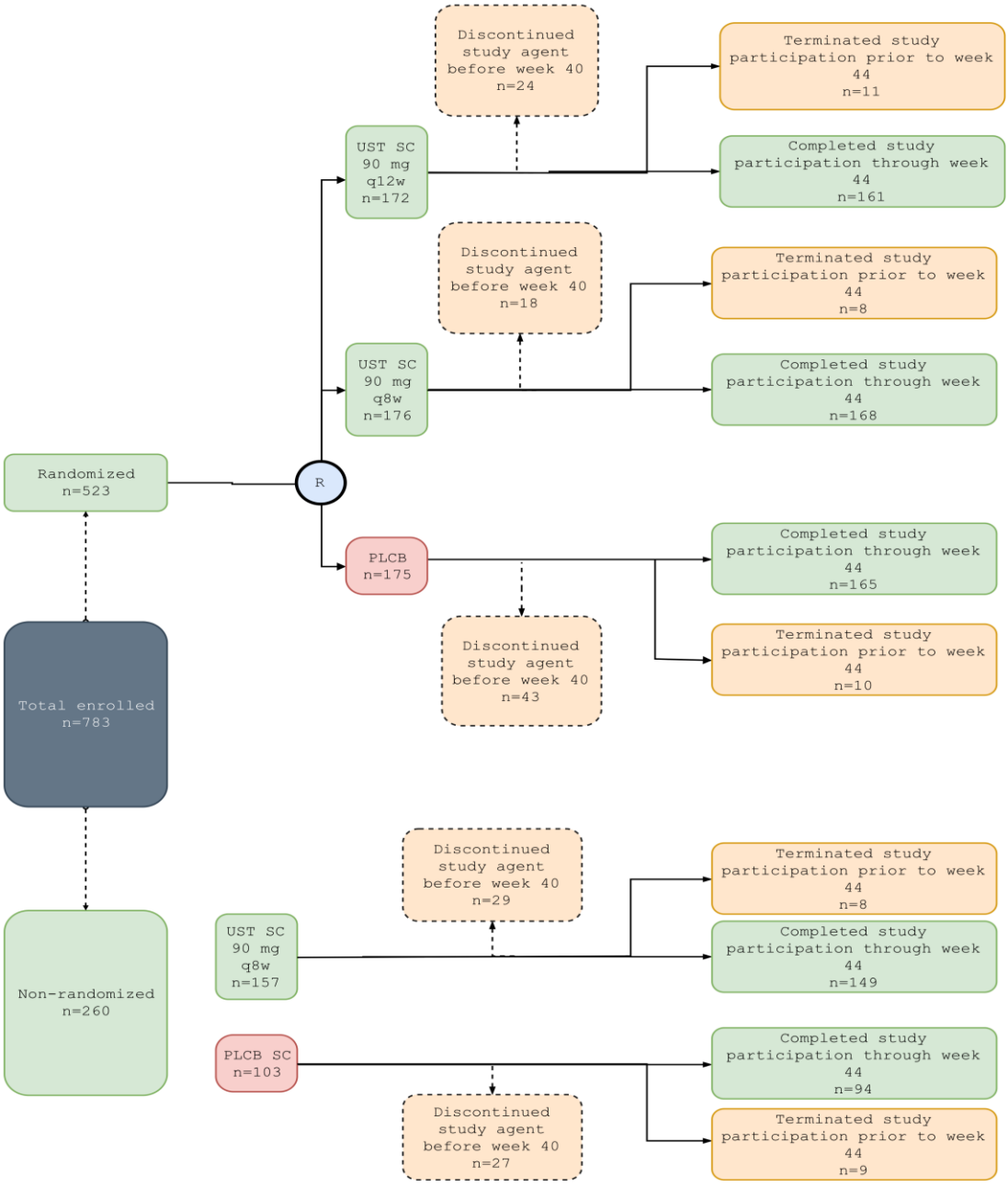


PLCB = placebo; q8w = every eight weeks; SC = subcutaneous; UST = ustekinumab.

Note: The screening period lasted eight weeks and the induction study lasted 16 weeks in total, divided into two eight-week periods. At the end of the first period, a clinical response was assessed, as described in the text. Purple boxes indicate those followed up for safety. Yellow boxes indicate patients (n = 49) who terminated study participation (of which 20 terminated before week 8, four terminated at week 8, and 25 terminated after week 8).

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 5: Disposition of Patients in the Maintenance Phase of the UNIFI Study, Randomized and Non-Randomized Populations



PLCB = placebo; q8w = every eight weeks; q12w = every 12 weeks; SC = subcutaneous; UST = ustekinumab.

Note: Patients from the induction study (Figure 4) who went into the maintenance study were followed up for up to 44 weeks.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Exposure to Study Treatments

There were no divergences between the groups of study in terms of exposure to treatments. In the induction study, 960 of 961 patients received an IV administration of either ustekinumab or placebo at week 0: 641 patients received ustekinumab doses (either 130 mg [n = 321] or 6 mg/kg [n = 320]), while 319 received placebo. One patient was randomized to the 130 mg group but did not receive any study drug, and two patients were randomized to the 6 mg/kg group but received a ustekinumab dose that was closer to 130 mg (these two patients were included in the 130 mg group for the safety analyses).

A total of 417 patients who were not in clinical response at week 8 received an additional single dose of the study drug at week 8 as follows:

- At week 8, 184 patients who received placebo at week 0 received one dose of 6 mg/kg IV of ustekinumab.
- At week 8, 233 patients who received ustekinumab at week 0 received one dose of ustekinumab 90 mg SC as follows:
  - 132 patients who received ustekinumab 130 mg IV at week 0 received one dose of ustekinumab 90 mg SC at week 8.
  - 101 patients who received 6 mg/kg IV of ustekinumab at week 0 received one dose of ustekinumab 90 mg SC at week 8.

A total of 825 randomized patients received at least one dose of ustekinumab during the induction study; all 825 patients received a dose of IV ustekinumab and 233 received a dose of ustekinumab 90 mg SC in addition to a dose of IV ustekinumab. As this was a single-dose drug, there were no concerns in terms of further adherence until the next phase of the study.

For the maintenance study, all 783 enrolled patients received an SC administration of either ustekinumab or placebo at maintenance baseline. A total of 523 patients were randomized in the primary population and 260 patients were not randomized. All enrolled patients received the treatment to which they were assigned at maintenance baseline.

Those patients who were randomized to ustekinumab received an intervention as follows:

- 90 mg every 12 weeks group: 172 patients received a median cumulative dose of 360.0 mg
- 90 mg every eight weeks group: 176 patients received a median cumulative dose of 540.0 mg

In the non-randomized population, the 157 patients in the ustekinumab-induction delayed-responders group (receiving ustekinumab 90 mg SC every eight weeks) received a median cumulative dose of 540.0 mg through week 44. In total, 505 patients received at least one dose of ustekinumab during the maintenance study.

Efficacy

Only those efficacy outcomes and analyses of subgroups identified in the review protocol are reported subsequently.

Induction Study

Clinical remission at week 8: Using the global definition, i.e., a Mayo score of 2 points or less with no individual subscore higher than 1 at week 8, a statistically significant greater

proportion of patients in the 6 mg/kg (approximately) and 130 mg groups achieved clinical remission (15.5% and 15.6%, respectively) compared with patients in the placebo group (5.3%; P < 0.001 for both comparisons) (Table 11). Similar results were reported for the US analysis (Appendix 3). Sensitivity analyses supported the robustness of the primary analyses for both ustekinumab treatment groups versus placebo.

Overall, the subgroup analyses were consistent with the primary analysis for the full study population, with a greater percentage of patients achieving clinical remission at week 8 with ustekinumab than with placebo (Appendix 3).

Clinical response at week 8: Clinical response was defined as a decrease from baseline in the Mayo score of 30% or more and a decrease of 3 points or more, with either a decrease from baseline in the rectal bleeding subscore of 1 point or more or a rectal bleeding subscore of 0 or 1. At week 8, a statistically significant greater proportion of patients in the 6 mg/kg (approximately) and 130 mg groups achieved clinical response (61.8% and 51.3%, respectively) compared with patients in the placebo group (31.3%; P < 0.001 for both comparisons) (Table 11).

Health-related quality of life (HRQoL): At baseline, median IBDQ scores were similar across all treatment groups (Table 11). At week 8, the median improvements from baseline in the IBDQ scores were statistically significantly greater in the 6 mg/kg (approximately) and 130 mg groups (31.0 and 31.5, respectively) compared with the placebo group (10.0; P < 0.001 for both comparisons).

At week 8, greater proportions of patients in the 6 mg/kg (approximately) and 130 mg groups (45.3% and 48.3%, respectively) had a an improvement of 5 points or more in the SF-36 Physical Component Summary (PCS) compared with patients in the placebo group (26.0%; P < 0.001 for both comparisons) (Table 11).

At baseline, the mean EQ-5D index and health state Visual Analogue Scale (VAS) scores were similar across all treatment groups. At week 8, the mean changes from baseline in EQ-5D and the health state VAS were greater for patients in the 130 mg and 6 mg/kg groups compared with those in the placebo group with [redacted] a change from baseline of 13.64 and 13.51 units in favour of the intervention (P < 0.001) for health state VAS. At baseline, the distributions for each of the five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) were generally consistent across treatment groups. At week 8, a greater proportion of patients had improvement in the dimensions of usual activities, pain/discomfort, and anxiety/depression for each ustekinumab group compared with placebo (P ≤ 0.002). A difference in the self-care dimension was also noted in the 6 mg/kg (approximately) group (P = 0.044) compared with the placebo group, but this was not observed in the 130 mg group. Improvement in mobility was not observed.

Need for colectomy at week 8: Only two patients in the placebo group and none in the ustekinumab groups required colectomy.

Mucosal healing: At week 8, statistically significant greater proportions of patients in the 6 mg/kg (approximately) and 130 mg groups achieved endoscopic healing (27.0% and 26.3%, respectively) compared with patients in the placebo group (13.8%; P < 0.001 for both comparisons) (Table 11). Mucosal healing was defined as a combination of endoscopic and histologic healing (i.e., neutrophil infiltration in < 5% of crypts, no crypt

destruction, and no erosions, ulcerations, or granulation tissue). Statistically significantly greater proportions of patients in the 6 mg/kg (approximately) and 130 mg groups achieved mucosal healing (18.4% and 20.3%, respectively) compared with patients in the placebo group (8.9%; P < 0.001 for both comparisons) at week 8.

Work productivity at week 8: Mean decreases from baseline were greater for patients in the 6 mg/kg (approximately) and 130 mg groups in each of the four Work Productivity and Activity Questionnaire (WPAI) categories compared with the placebo group, as follows (Table 11):

- Percentage of time missed from work at week 8: Mean decreases from baseline were 9.1% and 5.9% in the 6 mg/kg (approximately) and 130 mg groups, respectively, compared with 3.7% in the placebo group (P = 0.039 and P = 0.001, respectively).
- Percentage of impairment while working due to health at week 8: Mean decreases from baseline were 20.4% and 15.1% in the 6 mg/kg (approximately) and 130 mg groups respectively, compared with 6.9% in the placebo group (P < 0.001 and P = 0.019, respectively).
- Percentage of overall work impairment due to health at week 8: Mean decreases from baseline were 21.8% and 17.2% in the 6 mg/kg (approximately) and 130 mg groups, respectively, compared with 8.0% in the placebo group (P < 0.001 and P = 0.006, respectively).
- Percentage of activity impairment due to health at week 8: Mean decreases from baseline were 20.8% and 17.7% in the 6 mg/kg (approximately) and 130 mg groups, respectively, compared with 10.9% in the placebo group (P < 0.001 and P = 0.003, respectively) (see also Table 11).

Table 11: Efficacy Outcomes – UNIFI Induction Study, Efficacy Population

	Ustekinumab IV			Placebo
	130 mg N = 320	~6 mg/kg N = 322	Combined N = 642	N = 319
Clinical remission (global definition) at week 8[a]				
Number of patients in clinical remission, n (%)	50 (15.6)	50 (15.5)	100 (15.6)	17 (5.3)
Risk difference against placebo, (95% CI); P value	10.3 (5.7 to 14.9); < 0.001	10.2 (5.6 to 14.8); < 0.001	10.2 (6.6 to 13.9); < 0.001	–
Clinical response at week 8[a]				
Patients in clinical remission, n (%)	164 (51.3)	199 (61.8)	363 (56.5)	100 (31.3)
Risk difference against placebo, (95% CI); P value	19.9 (12.5c 27.3); < 0.001	30.5 (23.2 to 37.8); < 0.001	25.2 (18.9 to 31.5); < 0.001	–
Endoscopic healing at week 8[a]				
Number of patients with mucosal endoscopy healing, n (%)	84 (26.3)	87 (27.0)	171 (26.6)	44 (13.8)
Risk difference against placebo, (95% CI); P value	12.4 (6.5 to 18.4); < 0.001	13.3 (7.3 to 9.3); < 0.001	12.8 (7.9 to 17.8); < 0.001	–
IBDQ score at week 8[a]				
Total IBDQ score				
Baseline (n)	316	321	637	317

	Ustekinumab IV			Placebo
	130 mg N = 320	~6 mg/kg N = 322	Combined N = 642	N = 319
Mean score at baseline (SD)	126.0 (33.14)	127.0 (33.27)	126.5 (33.19)	127.4 (34.45)
Mean score at week 8 (n)	319	322	641	319
Mean at week 8 (SD)	159.2 (37.16)	161.9 (35.64)	160.6 (36.40)	143.5 (39.96)
Change from baseline, (n)	316	321	637	317
Change from baseline, mean (SD); P value	33.4 (32.53); < 0.001	35.0 (31.86); < 0.001	34.2 (32.18); < 0.001	16.1 (31.39)
Number of patients with an improvement of > 20 points from baseline, n (%)	196 (61.3)	200 (62.1)	396 (61.7)	118 (37.0)
Risk difference against placebo, (95% CI); P value	24.3 (16.61 to 31.57); < 0.001	25.1 (17.43 to 32.34); < 0.001	24.7 (18.0 to 31.0); < 0.001	–
SF-36 score at week 8[a]				
Physical Component Summary				
Baseline (n)	318	322	640	319
Mean score at baseline (SD)	43.1 (7.85)	43.1 (7.73)	43.1 (7.79)	43.6 (7.96)
Mean score at week 8 (n)	318	322	640	319
Change from baseline, mean (SD); P value	4.7 (6.49); < 0.001	5.2 (6.16); < 0.001	4.9 (6.33); < 0.001	2.1 (6.39)
Number of patients with an improvement from baseline in the PCS of at least 5 points, n (%)	154 (48.3)	146 (45.3)	300 (46.8)	83 (26.0)
Risk difference against placebo, (95% CI); P value	22.3 (14.87 to 29.38); < 0.001	19.3 (11.92 to 26.38); < 0.001	20.8 (14.43 to 26.73); < 0.001	–
Mental Component Summary				
Baseline (n)	318	322	640	319
Mean score at baseline (SD)	40.1 (10.85)	40.5 (10.59)	40.3 (10.71)	40.5 (11.43)
Mean score at week 8 (n)	318	322	640	319
Change from baseline, mean (SD); P value	5.3 (9.63); < 0.001	5.1 (9.72); < 0.001	5.2 (9.67); < 0.001	2.2 (10.20)
Improvement from baseline in the MCS of at least 5 points, n (%)	140 (43.9)	143 (44.4)	283 (44.1)	100 (31.3)
Risk difference against placebo, (95% CI); P value	12.6 (5.09 to 19.91); < 0.001	12.8 (5.29 to 20.09); < 0.001	12.8 (6.29 to 18.99); < 0.001	–
EQ-5D score at week 8[a]				
EQ-5D index				
Baseline (n)	319	322	641	317
Mean score at baseline (SD)	0.67 (0.204)	0.67 (0.195)	0.67 (0.199)	0.66 (0.208)
Mean score at week 8 (n)	319	322	641	317
Change from baseline, mean (SD); P value	0.09 (0.182); < 0.001	0.11 (0.172); < 0.001	0.10 (0.177); < 0.001	0.04 (0.182)

	Ustekinumab IV			Placebo
	130 mg N = 320	~6 mg/kg N = 322	Combined N = 642	N = 319
Health state VAS				
Baseline (n)	319	322	641	317
Mean score at baseline (SD)	54.14 (20.54)	55.76 (19.33)	54.95 (19.94)	55.11 (20.81)
Week 8 (n)	319	322	641	317
Change from baseline, mean (SD); P value	13.64 (20.39); < 0.001	13.51 (18.44); < 0.001	13.58 (19.42); < 0.001	5.71 (19.58)
Need for colectomy at week 8[a]				
Number of patients needing a colectomy, n (%)	0	0	0	2 (0.6)
Risk difference against placebo, (95% CI); P value	0.6 (−0.13 to 2.21); 0.11	0.6 (−0.13 to 2.21); 0.11	0.6 (−0.13 to 2.21); 0.11	–
Work productivity at week 8[a]				
Percentage of work time missed due to health				
Baseline (n)	195	208	403	182
Mean at baseline (SD)	18.0 (30.22)	17.7 (29.07)	17.8 (29.59)	19.3 (32.32)
Week 8 (n)	192	207	399	165
Mean at week 8 (SD)	9.7 (24.02)	7.5 (19.50)	8.6 (21.79)	14.5 (29.19)
Change from baseline, (n)	173	190	363	147
Change from baseline, mean (SD); P value	−5.9 (31.39); 0.039	−9.1 (23.84); 0.001	−7.6 (27.70); 0.002	−3.7 (30.41)
Percentage of impairment while working due to health				
Baseline (n)	176	192	368	164
Mean at baseline (SD)	43.5 (25.61)	45.3 (24.77)	44.4 (25.16)	39.1 (25.49)
Week 8 (n)	185	203	388	153
Mean at week 8 (SD)	27.4 (24.76)	26.2 (22.29)	26.8 (23.48)	31.0 (25.46)
Change from baseline, (n)	156	178	334	129
Change from baseline, mean (SD); P value	−15.1 (29.17); 0.019	−20.4 (24.11); < 0.001	−17.9 (26.68); < 0.001	-6.9 (21.89)
Percentage of overall work impairment due to health				
Baseline (n)	176	191	367	164
Mean at baseline (SD)	47.5 (27.52)	49.1 (26.32)	48.3 (26.88)	43.7 (27.48)
Week 8 (n)	184	203	387	152
Mean at week 8 (SD)	29.3 (27.01)	29.0 (24.69)	29.1 (25.79)	34.4 (27.96)
Change from baseline, (n)	155	178	333	128
Change from baseline, mean (SD); P value	−17.2 (30.36); 0.006	−21.8 (26.26); < 0.001	−19.7 (28.29); < 0.001	−8.0 (24.83);
Percentage of activity impairment due to health				
Baseline (n)	319	318	637	315
Mean at baseline (SD)	52.8 (27.07)	52.4 (26.90)	52.6 (26.96)	51.8 (26.37)
Week 8 (n)	303	310	613	293

	Ustekinumab IV			Placebo
	130 mg N = 320	~6 mg/kg N = 322	Combined N = 642	N = 319
Mean at week 8 (SD)	35.1 (27.24)	31.3 (26.04)	33.2 (26.69)	40.1 (27.73)
Change from baseline, (n)	303	307	610	289
Change from baseline, mean (SD); P value	−17.7 (29.45); 0.003	−20.8 (26.27); < 0.001	−19.3 (27.92); < 0.001	−10.9 (28.66)

CI = confidence interval; EQ-5D = EuroQol 5-Dimensions; IBDQ = Inflammatory Bowel Disease Questionnaire; MCS = Mental Component Summary; PCS = Physical Component Summary; SD = standard deviation; SF-36 = Short Form (36) Health Survey; VAS = Visual Analogue Scale.

Note: P values represent comparison against placebo.

[a] All percentages and values calculated for each variable are based on the primary efficacy population, unless stated otherwise.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Maintenance Study

Clinical remission at week 44: Using the global definition (a Mayo score of ≤ 2 points with no individual subscore > 1 at week 44), a significantly greater proportion of patients in the ustekinumab every eight weeks group and ustekinumab every 12 weeks group reached clinical remission (43.8% and 38.4%, respectively) compared with the placebo group (24.0%; P < 0.001 and P = 0.002, respectively) (Table 12). Similar results were found using the US-based definition for clinical remission (Appendix 3). Sensitivity analyses supported the primary analysis.

Subgroup analyses were also generally consistent with the primary analysis for the full population. However, it was reported in the UNIFI maintenance study that for the induction-treatment subgroups (ustekinumab 6 mg/kg IV [approximately], 130 mg IV, or placebo IV), there may be a lower maintenance-treatment effect on clinical remission (particularly for the every 12 weeks regimen) for patients who had received the 130 mg IV induction treatment or the placebo IV induction treatment. The sample sizes for these analyses were relatively small and estimates were imprecise.

Corticosteroid-free clinical remission at week 44: A statistically significant greater proportion of patients were in clinical remission and not receiving concomitant corticosteroids at week 44 in the ustekinumab every eight weeks and every 12 weeks groups (42.0% and 37.8%, respectively), compared with 23.4% in the placebo group (P < 0.001 and P = 0.002, respectively) (Table 12). Similar results were reported using the US definition of clinical remission (Appendix 3).

Maintenance of clinical response at week 44: This was defined as a decrease from induction baseline in the Mayo score of 30% or more and a decrease of 3 points or more, with either a decrease from induction baseline in the rectal bleeding subscore of 1 point or more or a rectal bleeding subscore of 0 or 1. A statistically significant difference was observed in the proportion of patients in the ustekinumab every eight weeks and every 12 weeks groups who sustained clinical response through week 44 (71.0% and 68.0%, respectively) versus the placebo group (44.6%; P < 0.001 for both comparisons) (Table 12).

HRQoL: Up until week 44, each of the four dimensions in scores related to quality of life using the IBDQ decreased (deteriorated) for patients in the placebo group; nonetheless, the improvements detected at maintenance baseline were sustained in each of the ustekinumab groups (P ≤ 0.002 for all dimensions), with a change from maintenance baseline (standard deviation) of 172.3 (40.9) and 178.2 (32.7) points in the every 12 weeks and every eight weeks groups, respectively, when compared with the placebo group.

Greater proportions of patients in the ustekinumab every eight weeks and every 12 weeks groups had a greater than 20-point improvement from induction baseline in the IBDQ score at maintenance week 44 (69.9% and 66.3%, respectively) compared with patients in the placebo group (42.9%; P < 0.001 for both comparisons) (Table 12). All differences were statistically significant.

SF-36 physical and mental scores were similar across all treatment groups at baseline. At week 44, the SF-36 physical and mental scores increased in the ustekinumab every eight weeks group, were maintained in the ustekinumab every 12 weeks group, and decreased in the placebo groups. For instance, among patients with an improvement of 5 points or more (from induction baseline) in the SF-36 physical score at maintenance baseline, significantly greater proportions of the ustekinumab every eight weeks and every 12 weeks groups maintained an improvement of 5 points or more through maintenance week 44 (62.4% and 59.5%, respectively) compared with patients in the placebo group (38.3%; P = 0.002 and P = 0.004, respectively) (Table 12); and, among patients with a an improvement of 5 points or more (from induction baseline) in the SF-36 mental score at maintenance baseline, significantly greater proportions of those receiving ustekinumab every eight weeks or every 12 weeks maintained this level of improvement through maintenance week 44 (59.8% and 58.3%, respectively) compared with patients in the placebo group (36.1%; P = 0.001 and P = 0.002, respectively) (Table 12).

The mean EQ-5D index and EQ-5D health state VAS scores were similar across all treatment groups at baseline. Through week 44, the EQ-5D index and EQ-5D health state VAS scores were maintained for patients in the ustekinumab every eight weeks and every 12 weeks groups and decreased (worsened) for patients in the placebo group (Table 12).

Need for colectomy through week 44: One patient in each ustekinumab group (eight weeks and every 12 weeks) required a colectomy, while three in the placebo group underwent this surgery (P = 0.220) (Table 12).

Mucosal endoscopy healing through week 44: The outcome of mucosal healing was defined as a combination of endoscopic healing and histologic healing of note; in the data displays, histologic healing is referred to as 0% to less than 5% neutrophils in epithelium, no crypt destruction, and no erosions or ulcerations or granulations. At week 44, significantly greater proportions of patients in the ustekinumab every eight weeks and every 12 weeks groups achieved mucosal healing (45.9% and 38.8%, respectively) compared with patients in the placebo group (24.1%; P < 0.001 and P = 0.002, respectively) (Table 12).

Work productivity: At maintenance baseline, the mean percentages were similar across all treatment groups within each of the four Work Productivity and Activity Impairment Questionnaire – General Health (WPAI-GH) domains. At week 44, WPAI-GH percentages were maintained from maintenance baseline for the ustekinumab groups in all four WPAI domains, with additional improvement (i.e., decrease) observed in patients in the ustekinumab every eight weeks group in percentage of impairment while working due to health, percentage of overall work impairment due to health, and percentage of activity impairment due to health. For patients in the placebo group, percentages for all four WPAI-GH domains worsened (i.e., increased).

Other parameters were as follows:

- Percentage of work time missed due to health: Mean changes from baseline were −2.0%, 2.1%, and 4.7% in the ustekinumab every 12 weeks, ustekinumab every eight weeks, and placebo groups, respectively (Table 12).
- Percentage of impairment while working due to health: Mean changes from baseline were −1.6% and −6.4% in the ustekinumab every 12 weeks and every eight weeks groups, respectively, compared with 7.4% in the placebo group.
- Percentage of overall work impairment due to health: Mean changes from baseline were −2.2% and −6.1% in the ustekinumab every 12 weeks and every eight weeks groups, respectively, compared with 7.7% in the placebo group.
- Percentage of activity impairment due to health: Mean changes from baseline were 0.8% and −4.2% in the ustekinumab every 12 weeks and ustekinumab every eight weeks groups, respectively, compared with 9.3% in the placebo group.

Table 12: Efficacy Outcomes – UNIFI Maintenance Study, Randomized Population

	Ustekinumab SC			Placebo N = 175
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 176	Combined N = 348	
Clinical remission (global definition) at week 44[a]				
Number of patients in clinical remission, n (%)	66 (38.4)	77 (43.8)	143 (41.1)	42 (24.0)
Risk difference against placebo, (95% CI); P value	14.5 (5.5 to 23.6 23.79); 0.002	19.7 (10.3 to 29.0); < 0.001	17.1 (9.3 to 24.9); < 0.001	–
Clinical response at week 44[a]				
Number of patients with clinical response, n (%)	117 (68.0)	125 (71.0)	242 (69.5)	78 (44.6)
Risk difference against placebo, (95% CI); P value	23.4 (13.00 to 33.08); < 0.001	26.4 (16.14 to 35.87); < 0.001	25.0 (16.4 to 33.6); < 0.001	–
Corticosteroid-free clinical remission at week 44[a]				
Number of patients in corticosteroid-free clinical remission, n (%)	65 (37.8)	74 (42.0)	139 (39.9)	41 (23.4)
Risk difference against placebo, (95% CI); P value	14.5 (5.5 to 23.6); 0.002	18.5 (9.3 to 27.8); < 0.001	16.5 (8.8 to 24.3); < 0.001	–
IBDQ scores at week 44[a]				
Total IBDQ score				
Maintenance baseline (n)	172	174	346	174
Mean score at maintenance baseline (SD)	175.4 (29.75)	174.1 (26.76)	174.7 (28.25)	174.3 (29.15)
Week 44 (n)	172	176	348	174
Mean score at week 44 (SD)	172.3 (40.97)	178.2 (32.71)	175.3 (37.09)	159.3 (40.67)
Change from maintenance baseline, (n)	172	174	346	173
Change from maintenance baseline, mean (SD); P value[b]	−3.0 (32.89); < 0.001	3.9 (31.54); < 0.001	0.5 (32.36); < 0.001	−15.1 (35.43)
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

	Ustekinumab SC			Placebo N = 175
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 176	Combined N = 348	
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
Number of patients with an IBDQ improvement of > 20 points from induction baseline, n (%)	114 (66.3)	123 (69.9)	237 (68.1)	75 (42.9)
Risk difference against placebo, (95% CI); P value	23.4 (12.96 to 33.12); < 0.001	27.0 (16.70 to 36.48); < 0.001	25.2 (16.22 to 33.70); < 0.001	–
SF-36 scores at week 44[a]				
Physical Component Summary				
Maintenance baseline (n)	172	175	347	173
Mean at maintenance baseline (SD)	50.7 (6.86)	50.0 (6.88)	50.3 (6.87)	50.0 (6.65)
Week 44 (n)	172	176	348	175
Mean at week 44 (SD)	50.3 (7.82)	51.3 (7.14)	50.8 (7.49)	48.3 (7.67)
Change from maintenance baseline, (n)	172	175	347	173
Change from maintenance baseline, mean (SD); P value[b]	−0.4 (7.14); 0.009	1.3 (5.68); < 0.001	0.5 (6.49); < 0.001	−1.7 (6.45)

	Ustekinumab SC			Placebo N = 175
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 176	Combined N = 348	
Number of patients with an improvement from maintenance baseline in the PCS of at least 5 points, n (%)	66 (59.5)	58 (62.4)	124 (60.8)	31 (38.3)
Difference against placebo, (95% CI); P value	21.2 (10.70 to 31.04); < 0.001	24.1 (13.68 to 33.78); < 0.001	22.5 (13.46 to 30.99); < 0.001	–
Mental Component Summary				
Maintenance baseline (n)	172	175	347	173
Mean score at baseline (SD)	47.1 (9.99)	48.1 (8.63)	47.6 (9.33)	47.6 (9.41)
Week 44 (n)	172	176	348	175
Mean score at week 44 (SD)	47.4 (10.65)	48.5 (9.69)	48.0 (10.18)	45.3 (11.16)
Change from maintenance baseline, (n)	172	175	347	173
Change from maintenance baseline, mean (SD); P value[b]	0.3 (8.41); 0.006	0.3 (9.51); 0.002	0.3 (8.97); < 0.001	−2.4 (9.89)
Improvement from maintenance baseline in the MCS of at least 5 points, n (%)	56 (58.3)	58 (59.8)	114 (59.1)	30 (36.1)
Risk difference against placebo, (95% CI); P value	22.2 (11.72 to 32.00); < 0.001	23.7 (12.28 to 33.39); < 0.001	23 (13.97 to 31.42); < 0.001	–
EQ-5D scores at week 44[a]				
EQ-5D index score				
Maintenance baseline (n)	172	175	347	173
Mean score at baseline (SD)	0.810 (0.1563)	0.801 (0.1588)	0.8006 (0.1574)	0.820 (0.15)
Week 44 (n)	172	175	347	175
Mean score at week 44 (SD)	0.819 (0.1759)	0.827 (0.1612)	0.823 (0.1684)	0.773 (0.1739)
Change from maintenance baseline, mean (SD); P value[b]	0.008 (0.1656); 0.001	0.025 (0.1674); < 0.001	0.017 (0.166); < 0.001	−0.048 (0.158)
Health state VAS				
Maintenance baseline (n)	172	175	347	173
Mean score at baseline (SD)	75.7 (16.28)	73.2 (16.24)	74.4 (16.28)	75.2 (13.57)
Week 44 (n)	172	176	348	175
Mean score at week 44 (SD)	73.5 (21.90)	75.6 (17.37)	74.6 (19.74)	67.4 (20.07)
Change from maintenance baseline, (n)	172	175	347	173
Change from maintenance baseline, mean (SD); P value[b]	−2.2 (19.87); < 0.001	2.4 (17.28); < 0.001	0.1 (18.72); < 0.001	−7.7 (18.75)
Need for colectomy at week 44[a]				
Number of patients needing a colectomy, n (%)	1 (0.6)	1 (0.6)	2 (0.6)	3 (1.7)
Difference against placebo, (95% CI); P value	1.1 (−1.77 to 4.33); 0.38	1.1 (−1.73 to 4.33); 0.33	1.1 (−0.78 to 4.32); 0.22	–
Mucosal endoscopy healing at week 44[a]				
Number of patients with mucosal endoscopic healing, n (%)	66 (38.8)	79 (45.9)	145 (42.4)	41 (24.1)

	Ustekinumab SC			Placebo N = 175
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 176	Combined N = 348	
Difference against placebo, (95% CI); P value	14.7 (4.83 to 24.1); 0.002	21.8 (11.7 to 31.2); < 0.001	18.3 (9.6 to 26.1); < 0.001	–
Work productivity at week 44[a]				
Percentage work time missed due to health				
Maintenance baseline (n)	112	113	225	111
Mean at baseline (SD)	9.4 (25.65)	2.8 (8.56)	6.1 (19.33)	6.5 (17.13)
Week 44 (n)	108	109	217	108
Mean at week 44 (SD)	9.3 (24.83)	4.4 (16.16)	6.8 (21.03)	11.6 (26.27)
Change from maintenance baseline, (n)	95	98	193	99
Change from maintenance baseline, mean (SD); P value[b]	−2.0 (22.16); 0.133	2.1 (19.07); 0.172	0 (20.70); 0.096	4.7 (21.83)
Percentage of impairment while working due to health				
Maintenance baseline (n)	105	113	218	109
Mean at baseline (SD)	21.9 (20.24)	21.2 (18.74)	21.5 (19.44)	21.0 (22.69)
Week 44 (n)	103	107	210	101
Mean at week 44 (SD)	18.9 (24.89)	14.3 (18.12)	16.6 (21.78)	26.9 (25.87)
Change from maintenance baseline, (n)	88	96	184	95
Change from maintenance baseline, mean (SD); P value[b]	−1.6 (24.02); 0.017	−6.4 (23.80); < 0.001	−4.1 (23.96); < 0.001	7.4 (30.50)
Percentage of overall work impairment due to health				
Maintenance baseline (n)	105	113	218	109
Mean at baseline (SD)	23.7 (22.51)	23.0 (20.42)	23.3 (21.40)	23.6 (24.92)
Week 44 (n)	103	107	210	101
Mean at week 44 (SD)	20.7 (27.78)	16.2 (20.16)	18.4 (24.25)	29.3 (28.46)
Change from maintenance baseline, (n)	88	96	184	95
Change from maintenance baseline, mean (SD); P value[b]	−2.2 (24.65); 0.013	−6.1 (26.60); < 0.001	−4.2 (25.69); < 0.001	7.7 (32.77)
Percentage of activity impairment due to health				
Maintenance baseline (n)	172	175	347	172
Mean at baseline (SD)	24.9 (23.23)	26.3 (22.90)	25.6 (23.05)	26.3 (23.62)
Week 44 (n)	160	159	319	166
Mean at week 44 (SD)	25.9 (28.25)	21.2 (25.14)	23.5 (26.81)	35.3 (29.78)
Change from maintenance baseline, (n)	160	158	318	163
Change from maintenance baseline, mean (SD); P value[b]	0.8 (26.65); 0.002	−4.2 (25.42); < 0.001	−1.7 (26.12); < 0.001	9.3 (31.71)

CI = confidence interval; EQ-5D = EuroQol 5-Dimensions; IBDQ = Inflammatory Bowel Disease Questionnaire; MCS = Mental Component Summary; PCS = Physical Component Summary; q.8.w. = every eight weeks; q.12.w. = every 12 weeks; SC = subcutaneous; SD = standard deviation; SF-36 = Short Form (36) Health Survey; VAS = Visual Analogue Scale.

[a] All percentages and values calculated for each variable are based on the primary efficacy population unless stated otherwise.

[b] Comparison against placebo.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Harms

For the induction study, summaries of AEs and other safety data were based on the safety analysis set, i.e., patients who received at least one dose of the study drug, including a partial dose. Patients were evaluated according to the treatment received. The final safety visit was at 20 weeks after the last administration of the study drug. Among the 960 patients in the safety analysis set, one or more AEs was reported through week 8 for 50.0%, 41.4%, and 48.0% of patients in the 6 mg/kg (approximately), 130 mg, and placebo groups, respectively (Table 13). Through week 8, serious AEs were reported for 3.1%, 3.7%, and 6.6% of patients in the 6 mg/kg (approximately), 130 mg, and placebo groups, respectively. The percentage of patients with one or more AEs within one hour of infusion were 0.9%, 2.2%, and 1.9% in the 6 mg/kg (approximately), 130 mg, and placebo groups, respectively. The proportions of patients with one or more infections were 15.3%, 15.9%, and 15.0% in the 6 mg/kg (approximately), 130 mg, and placebo groups, respectively. Serious infections were reported for 0.3%, 0.6%, and 1.3% of patients in the 6 mg/kg (approximately), 130 mg, and placebo groups, respectively. One patient in the 6 mg/kg (approximately) group died on study day 42 due to an unrelated event, esophageal varices hemorrhage. The most frequently reported AEs were infections; among these, viral upper respiratory tract infection was the most common presentation (5.0%, 4.0%, and 2.5% in the 6 mg/kg [approximately], 130 mg, and placebo groups, respectively). In the gastrointestinal disorders, the most frequently reported AEs were UC (2.2%, 2.8%, and 5.6% of patients, respectively), nausea (2.2%, 2.5%, 1.9% of patients, respectively), and abdominal pain (1.9%, 2.5%, 2.2% of patients, respectively).

For the maintenance study, the safety data were analyzed on all treated patients (randomized and non-randomized populations). In total, 783 patients were treated in this maintenance study; of these, 523 patients were randomized and 260 were not randomized. Of the 783 patients who received at least one administration of the study drug in the maintenance study, one or more AEs was reported through week 44 for 73.7% of the all-ustekinumab group and 75.5% of the placebo group (Table 14). One or more serious AEs were reported for 9.7%, 3.5%, and 9.3% in patients in the placebo, ustekinumab every 12 weeks, and ustekinumab every eight weeks groups, respectively. The most common AE reported was UC. More patients stopped treatment due to AE in the placebo group (11.6%) versus those in the ustekinumab every 12 weeks (5.2%) and every eight weeks (5.1%) groups. Of the notable harms, among all treated patients, serious infections were reported in 11 (2.2%) patients in the all-ustekinumab group and 5 (1.8%) in the placebo group.

Only one death was reported in the ustekinumab every eight weeks group. See Table 13 and Table 14 for detailed harms data.

Table 13: Summary of Harms – UNIFI Induction Study, Safety Analysis Set

	Ustekinumab IV			Placebo N = 319
	130 mg N = 321	~6 mg/kg N = 320	Combined N = 641	
Patients with ≥ 1 adverse event				
n (%)	133 (41.4)	160 (50.0)	293 (45.7)	153 (48)
Most common events[a]				
Headache	22 (6.9)	13 (4.1)	35 (5.5)	14 (4.4)
Upper respiratory tract infection (viral)	13 (4.0)	16 (5.0)	29 (4.5)	8 (2.5)

	Ustekinumab IV			Placebo N = 319
	130 mg N = 321	~6 mg/kg N = 320	Combined N = 641	
Colitis ulcerative	9 (2.8)	7 (2.2)	16 (2.5)	18 (5.6)
Anemia	7 (2.2)	8 (2.5)	15 (2.3)	11 (3.4)
Nausea	8 (2.5)	6 (1.9)	15 (2.3)	6 (1.9)
Nasopharyngitis	1 (0.3)	8 (2.5)	9 (1.4)	1 (0.3)
Pruritis	8 (2.5)	3 (0.9)	11 (1.7)	4 (1.3)
Arthralgia	3 (0.9)	6 (1.9)	9 (1.4)	2 (0.6)
Influenza	2 (0.6)	1 (0.3)	3 (0.5)	0
Patients with ≥ 1 serious adverse event				
n (%)	12 (3.7)	10 (3.1)	22 (3.4)	21 (6.6)
Most common events				
Colitis ulcerative	3 (0.9)	2 (0.6)	5 (0.8)	8 (2.5)
Infections	2 (0.6)	1 (0.3)	3 (0.5)	4 (1.3)
Patients who stopped treatment due to adverse events[b]				
n (%)	NA	NA	NA	NA
Deaths				
n (%)	0	1 (0.3)	1 (0.15)	0
Notable harms				
Serious infections, n (%)	2 (0.6)	1 (0.3)	3 (0.5)	4 (1.3)
Malignancies, n (%)	0	0	0	0
Hypersensitivity or anaphylactic reactions, n (%)	0	0	0	0
MACE, n (%)	0	0	0	0
Thrombosis (any kind), n (%)	0	2 (0.6)	2 (0.3)	0

MACE = major adverse cardiovascular event; NA = not applicable.

[a] Frequency > 5% in any of both phases or studies.

[b] Data are shown as not applicable because patients received only one dose of the IV medication.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Table 14: Summary of Harms – UNIFI Maintenance Study, Randomized and Non-Randomized Populations

	Ustekinumab SC			Placebo N = 277
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 333	Combined N = 505	
Patients with ≥ 1 adverse event				
n (%)	119 (69.2)	253 (76.0)	372 (73.7)	209 (75.5)
Most common events[a]				
Nasopharyngitis	31 (18.0)	45 (13.5)	76 (15.0)	35 (12.6)
Colitis, ulcerative	19 (11.0)	44 (13.2)	63 (12.5)	74 (26.7)
Headache	11 (6.4)	27 (8.1)	38 (7.5)	9 (3.2)
Arthralgia	15 (8.7)	21 (6.3)	36 (7.1)	20 (7.2)
Upper respiratory tract infection	5 (2.9)	23 (6.9)	28 (5.5)	12 (4.3)

	Ustekinumab SC			Placebo N = 277
	90 mg q.12.w. N = 172	90 mg q.8.w. N = 333	Combined N = 505	
Anemia	9 (5.2)	16 (4.8)	25 (5.0)	18 (6.5)
Influenza	6 (3.5)	17 (5.1)	23 (4.6)	12 (4.3)
Patients with ≥ 1 serious adverse event				
n (%)	6 (3.5)	31 (9.3)	37 (7.3)	27 (9.7)
Most common events				
Colitis, ulcerative	1 (0.6)	9 (2.7)	10 (2.0)	10 (3.6)
Infections	0	4 (1.2)	4 (0.8)	4 (1.4)
Patients who stopped treatment due to adverse events				
n (%)	9 (5.2)	17 (5.1)	26 (5.1)	32 (11.6)
Deaths				
n (%)	0	1 (0.3)	1 (0.2)	0
Notable harms				
Serious infections, n (%)	6 (3.5)	5 (1.5)	11 (2.2)	5 (1.8)
Malignancies, n (%)	1 (0.6)	1 (0.3)	2 (0.4)	1 (0.4)
Hypersensitivity / anaphylactic reactions, n (%)	0	0	0	0
MACE, n (%)	0	0	0	1 (0.4)
Thrombosis (any kind), n (%)	1 (0.6)	0	1 (0.2)	0

MACE = major adverse cardiovascular event; q.8.w. = every eight weeks; q.12.w. = every 12 weeks; SC = subcutaneous.

[a] Frequency > 5% in any of both phases/studies.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Critical Appraisal

Internal Validity

Both phases of the study had a proper randomization process. The generation of the randomization sequence was adequate, and the allocation sequence was concealed until participants were enrolled and assigned to the interventions. Furthermore, the differences noted in baseline characteristics in the induction and maintenance phases were small and unlikely to have a meaningful impact on the validity of the results. The blinding of participants, clinicians, and researchers was achieved through identical placebo and ustekinumab presentations, which avoided important and unbalanced deviations from the intended interventions. There is no clear evidence that participants were aware of their assigned intervention during the trial. Patients who discontinued or deviated from the interventions were properly analyzed based on the ITT principle.

Subgroup analyses were adequately performed to examine the consistency of the treatment effect observed for the outcomes of clinical remission at week 8 (global definition) when the sample size of a subgroup did not preclude interpretation of the data. Although the clinical rationale for the analysis of subgroups is sound, some analyses were considered underpowered to detect a significant effect from modifiers. Multiplicity was considered and adequate tests were conducted (i.e., the CMH chi-square test) for all outcomes in both phases of the study.

Overall, follow-up was relatively complete for the primary end point, with more than 94% of patients completing the induction trial. Even when 81.5% (783 of 961) of the patients initially randomized in the induction study entered the maintenance study (randomized and non-randomized population), the rest of the 129 of patients who did not enter the maintenance phase (13.4%) completed the final safety visit.

Overall, the analyses appeared to adhere to the ITT approach. Differences in missing data between groups of study were small and unlikely to affect the final results.

Outcomes were objectively obtained with validated tools (Appendix 4) and the processes to accomplish outcome measurements were well described and assessed in a blinded fashion. There is a low risk of bias due to selection of the reported results. A protocol is well described for both studies and the results analyzed are in accordance with the pre-specified analysis plan.

The UNIFI maintenance phase used re-randomization at week 8 for the ustekinumab patients who responded to induction therapy in the induction phase. The strength of this design is that it allows evaluation of whether the response is maintained in the absence or presence of continued ustekinumab therapy. The use of separate induction and maintenance studies is consistent with European Medicines Agency guidance for the development of drugs for the treatment of UC. However, a limitation of this approach includes the fact that all patients enrolled in the maintenance phase were a select population: they were responders to induction therapy and, as well, were able to tolerate treatment with ustekinumab. However, this design is reasonable because these are also the patients who would be continued on treatment in clinical practice; nonetheless, from a research perspective, this design may obscure the true effectiveness and occurrence of AEs. This also has implications for the central "placebo" connection in the indirect comparison analyses (see indirect evidence section that follows).

External Validity

The populations included in the induction and maintenance phases of the UNIFI trial are similar overall to what is encountered in clinical practice and relevant to the population of interest for this review. Also, the doses of ustekinumab administered are in accordance with what would be expected in real-life practice. Adherence could be overstated, however, as it is a high proportion of patients who adhered to treatment is usual in well-controlled randomized trials, and generalizability might be an issue when the medication is applied in clinical settings. One concern about the design of the maintenance phase is that responders to ustekinumab at week 8 and placebo nonresponders who responded at week 16 to dose of ustekinumab at week 8 were both included in the study. This creates a patient population with two subgroups that had different treatment schedules at the start of the maintenance phase. Since the randomization of the maintenance arms was stratified by the assigned treatment during the induction phase, this issue would not introduce bias to comparisons during the maintenance phase but could increase the estimated precision of the treatment effect, thus reducing power to detect differences.

One concern from the clinical expert consulted for this review was the large number of patients (157 out of 233 [67%]) who did not respond in the induction study — at week 8 — but later responded at week 16 following a second dose of ustekinumab SC 90 mg. It was suggested that this number may be larger than one would likely see and thus begets uncertainty about whether clinicians would wait another eight weeks to assess whether patients achieve a late response before administering a second dose of ustekinumab at

week 8 to try to induce remission. The clinical expert noted that this decision would be based on several factors, including other indicators of clinical response (e.g., reduced symptoms), drug levels, and levels of UC-relevant markers (e.g., fecal calprotectin).

Indirect Evidence

The supplemental literature search conducted for the ITC obtained 62 references which, after title and abstract screening, were sifted to obtain one final study: a systematic review and NMA prepared for and submitted by the sponsor.

Objectives and Methods for the Summary of Indirect Evidence

Several interventions (biologics) have indications similar to that of ustekinumab for the treatment of UC. Up to the writing of this manuscript, no studies of head-to-head comparisons of ustekinumab versus other active therapies in UC have been conducted.

The objective of the ITC review was to synthesize evidence supporting the effects of ustekinumab against its relevant comparators in the treatment of moderate-to-severe active UC. These included:

- Conducting a systematic literature review (SLR) of clinical trials assessing the efficacy and safety of treatments used in moderate-to-severe active UC.
- Assessing the comparative efficacy of ustekinumab against its relevant comparators through an NMA.

Description of Indirect Treatment Comparison(s)

The authors of the review performed an NMA to assess the efficacy indirectly compared with other interventions: infliximab, adalimumab, vedolizumab, ustekinumab, golimumab, tofacitinib, and placebo. They evaluated three outcomes: Clinical remission, clinical response, and mucosal healing, all based on subgroups of patients considered biologic or non-biologic failures, and also in the subgroups of induction and maintenance phases of drug administration. Table 15 summarizes the study selection criteria and methods for the ITCs.

Table 15: Study Selection Criteria and Methods for ITCs

	Indirect treatment comparison (network meta-analysis)
Population	Patients with moderate-to-severe active ulcerative colitis that failed to respond to conventional therapy as well as patients whose condition did not respond to prior biologic(s)
Intervention	Ustekinumab • Induction: Solution for intravenous single-use infusion, either at a weight-based dose (approximately 6 mg/kg) of 260 mg, 390 mg, or 520 mg, depending on body weight, or at a fixed dose of 130 mg. • Maintenance: Subcutaneous injection of 90 mg every 8 or 12 weeks (90 mg/1.0 mL vial).
Comparators	Infliximab, adalimumab, golimumab, vedolizumab, tofacitinib, and placebo
Outcomes	Efficacy: • clinical response • clinical remission • mucosal healing Safety:

	Indirect treatment comparison (network meta-analysis)
	• not assessed
Study design	Randomized controlled trials (excluding phase I studies)
Exclusion criteria	Naive patients; mild and active ulcerative colitis; studies not assessing efficacy, safety, or quality-of-life outcomes; single-arm trials; non-randomized trials; and studies published in languages other than English.
Databases searched	MEDLINE, Embase, Cochrane Library, hand searches of clinical trials and data from recent studies not yet published using the NICE guide.
Selection process	Titles and, where available, abstracts of studies retrieved by the search were reviewed by two independent reviewers. Full-text articles were then reviewed to assess their eligibility according to pre-specified inclusion and exclusion criteria. Publications identified as potentially relevant during the first phase of the screening were then reviewed in full and assessed for inclusion according to the list of pre-specified inclusion and exclusion criteria.
Data extraction process	Following identification of full-text articles, relevant data were extracted according to pre-specified template. The following fields were extracted: • publication details (title, authors, date of publication, journal, volume, and reference page) • study characteristics, including objective, interventions (including dose and mode of administration), phases description (induction or maintenance), blinding, sample size, length of follow-up, treatment duration, allowed concomitant therapies, primary and secondary end points, country or location, statistical methods of data analysis, relevant biases • patient characteristics (disease phenotype, demographics, medical history, treatment history, average age of disease onset, average disease duration, severity, number of prior treatments received, and description of those treatments) • details of the results of interest (efficacy results, safety results, quality-of-life results, sub-populations with available results, study limitations specified by the authors when reported).
Quality assessment	Authors conducted a quality assessment at the study level with the NICE STA template and the guidance by the Centre for Reviews and Dissemination at the University of York for the NICE submission. They also followed the PRISMA template for reporting systematic reviews of multiple comparisons (network meta-analysis).

ITC = indirect treatment comparison; NICE National Institute of Clinical Excellence; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analysis; STA = single technology appraisal.

Source: Stelara network meta-analysis study report.[39]

Methods of the Sponsor-Submitted NMA

Objectives

The review focused on patients with moderate-to-severe active UC who failed conventional therapy and patients who failed prior biologic(s). The list of comparators has included all the biologics currently approved for moderate-to-severe UC, including biosimilars.

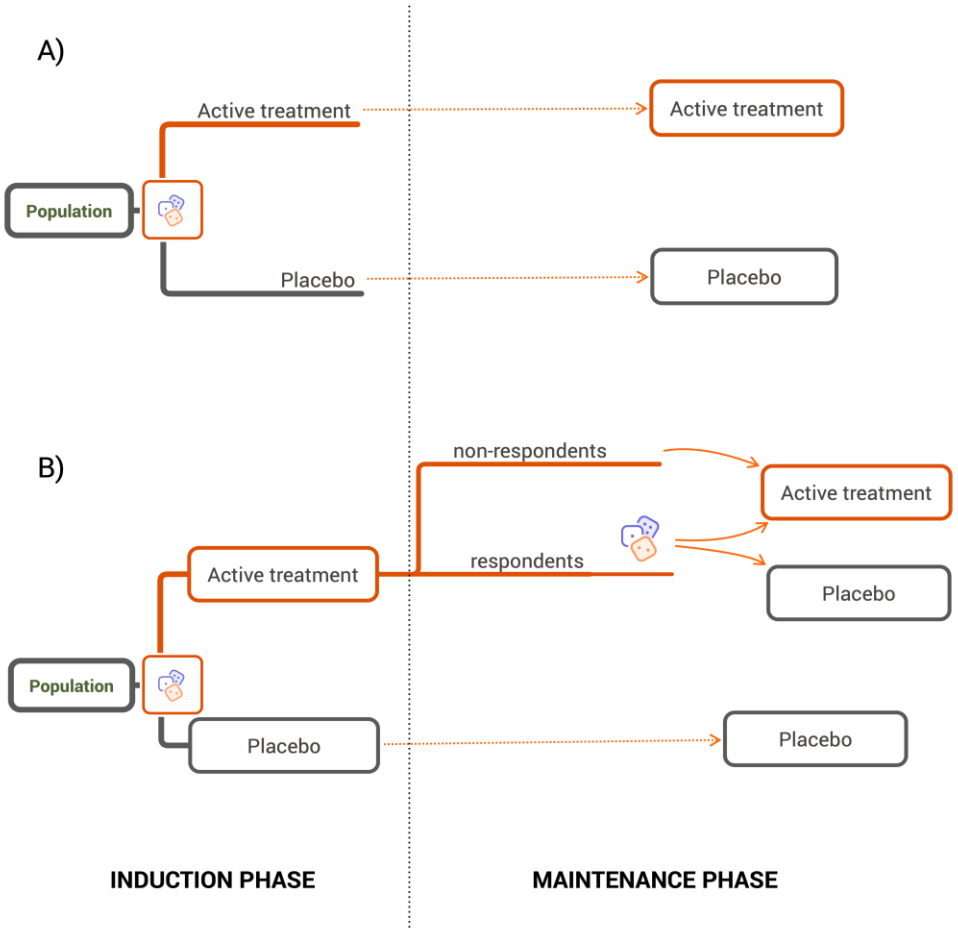
Study Selection Methods

The protocol for the SLR was drafted a priori and subsequently registered with PROSPERO (CRD42019131015). The search strategy was run in four databases including: MEDLINE, MEDLINE In-Process, Embase, and Cochrane library on August 14, 2018, and January 22 and March 28, 2019. 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 as well as conference abstracts. No time restriction was applied to the first search conducted on August 14, 2018. The second and third update searches restricted publication dates to reduce the number of overlapping articles with the first search. Search terms were developed using a combination of MeSH and Emtree terms and free-text terms. In addition to searches of electronic databases, hand searches were conducted to capture data not reported in the main publication of clinical trials and data from recent studies not yet published.

ITC Analysis Methods

Due to the specific characteristics of UC treatment and the clinical evolution of patients, randomized trials evaluating anti-TNF therapies for UC have to consider two phases: an induction phase to reach remission and a maintenance phase for continuing remission. Therefore, the design of these studies has evolved over time into two main categories: The first is a **treat-through** standard design in which patients in the induction phase are randomized to intervention or placebo and then continue with the same assigned intervention during the entirety of the maintenance phase. Second, a **response-based** or **re-randomized** study design, where patients who respond to the intervention during the induction phase are re-randomized to treatment or placebo. In the latter type of design, there can be several variations, like withdrawal of patients who do not respond, or patients treated through with the drug on an open-label phase (Figure 6).

Figure 6: Two Types of Study Designs for Trials Evaluating Anti-TNF Therapies in Patients With UC



TNF = tumour necrosis factor; UC = ulcerative colitis.

Note: A) Treat-through design, where patients in the induction phase continue to the same treatment arm regardless whether they respond to the intervention. B) Re-randomized design, where patients can change study group. Depending on their response in the induction phase, these studies can show variations in the maintenance phase, as nonresponders can either be reallocated to an open-label design (treated with the intervention) or withdrawn from the trial.

Source: Stelara network meta-analysis study report.[39]

Studies included in this NMA were known to fall into either category (treat-through or re-randomized design), and the authors expected that the key assumptions necessary to perform an NMA — and the pooled maintenance data in UC — were not valid. According to the authors, performing a regular NMA with these two different designs would be prone to selection bias and ignore patients who may not respond to treatment at the end of the induction phase but could respond after this. Heterogeneity exists between the maintenance phases of the studies due to differences in designs and reporting of maintenance outcomes. The placebo arms reported from re-randomized response-based trials are not "true" placebo arms, since participants are exposed to treatment in the

induction phase of the study. Hence, the data will be heterogeneous due to differences in carryover effects of active induction therapy.

Based on the sources of heterogeneity and bias in maintenance outcomes, authors considered that a standard NMA to compare maintenance phases would not adhere to best practices and an approach that compares the full one-year regimens of treatments was necessary. The authors considered this as the most rigorous method to mitigate uncertainty, given the necessity to perform an NMA. The approach involved first comparing treat-through data between all trials, and recalculating data from re-randomized response-based trials to mimic a treat-through design to be able to pool the data. That is, the authors decided to convert all studies to a treat-through design, as this was considered the least prone to bias for evaluating estimates of effects. In terms of study design and patients' characteristics, induction trials — or the induction phase of the trials — are similar enough to be pooled in a regular NMA, while maintenance trials require a conversion to emulate a treat-through design. To achieve this goal, the authors imputed data for studies that had a re-randomized design. This imputation was done on a case-by-case basis for each study, since some studies excluded different groups of patients in the maintenance phase. Sample sizes for re-randomized studies were also recalculated.

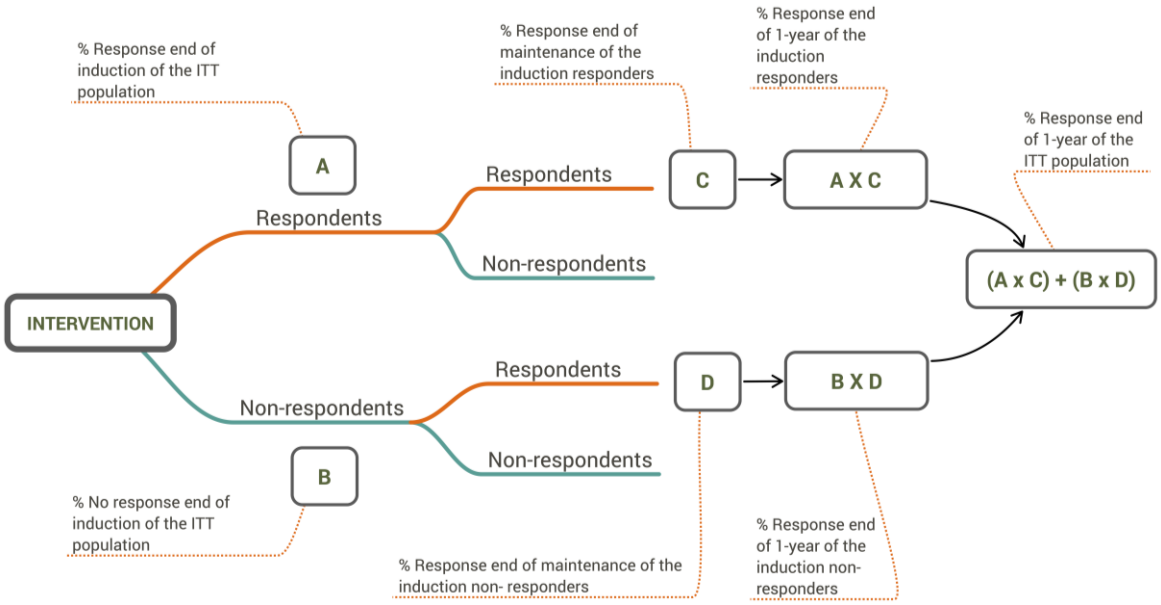
The authors had two options to achieve the conversion and equalization of studies:

- mimicking a treat-through design, that is, all studies that did not have a treat-through design would be converted to resemble one, or
- mimicking a response-based (re-randomized) design in which the authors would convert all studies without a re-randomized design to resemble one.

They decided to use the first approach (mimicking a treat-through design) because it was considered to be more closely related to clinical practice and reflected an ITT approach, with less risk of selection bias, and it is feasible with this approach to include more studies. Besides this analysis, the authors also decided to run a sensitivity analysis based on mimicking a response-based design. Calculations for mimicking a treat-through approach were based on responders and nonresponders in both the induction and maintenance phase, as presented in Figure 7.

As part of the recalculation of sample sizes for mimicking a treat-through design, patients who were assigned to the study drug during induction and assigned to placebo during maintenance were excluded from the analysis. These patients would have been (hypothetically) randomized to placebo under this specific design.

Figure 7: Calculations to Mimic a Treat-Through Design



ITT = intention to treat.

Note: For the sensitivity analysis (i.e., to mimic a response-based, re-randomized design), authors used only the induction respondents (A × C).

Source: Stelara network meta-analysis Clinical Study Report.[39]

Once recalculations were made and all studies were equivalent and able to be pooled, authors used a Bayesian hierarchical model for the NMA. They subgroup populations were based on biologic failure status. NMAs were conducted for the analysis of both the induction data and one-year data (induction and maintenance). A standard NMA was conducted of the induction phases, whereby studies were connected via a common placebo arm.

This allowed for treatment comparisons to be made between the efficacy of full one-year regimens. Other detailed methods are portrayed in Table 16.

Table 16: ITC Analysis Methods

	Sponsor-Submitted NMA[a]
ITC methods	Bayesian hierarchical model for the NMA.
Priors	Non-informative prior distributions were used for unknown parameters to get results driven by data. The following priors were used for the base-case analysis: • Normal distribution with a mean of 0 and a variance of 10,000 for treatment effects. • Uniform distribution for the between-trial standard deviation, with a range of 0 to 2 for binary outcomes.
Assessment of model fit	The relative goodness of fit of the model was assessed using the DIC. The fixed-effects and random-effects models were developed and the one associated with the lowest DIC was selected (with a difference of at least three points in DIC).
Assessment of Consistency	Authors did not expect closed loops to measure inconsistency (statistically), yet it is not mentioned in the paper. There is no mention on the specific model used for the only closed loop found in the NMA. The authors crafted an approach to decrease the differences between the induction and maintenance phases of the studies included, recalculating data to mimic a treat-through design for all studies.

	Sponsor-Submitted NMA[a]
Assessment of convergence	Convergence was assessed through the inspection of diagnostic plots including: trace, history, autocorrelation, and density plots. The authors used the Markov Chain Monte Carlo simulation method for the NMA; three chains were simulated, and their convergence was assessed by examining the history and the Gelman-Rubin plots.
Outcomes	(Note any outcomes that were used to inform the pharmacoeconomic analysis for the submission.)
Follow-up time points	Induction studies (reports) assessed outcomes at 6 to 8 weeks of follow-up. Maintenance studies (reports) were recalculated to assess outcomes at one year (full regimen) when using the approach to mimic a treat-through (response-based) design.
Construction of nodes	Nodes were created fitting classifications in relation to drug classes and dosages, as shown in Figure 7.
Sensitivity analyses	The following sensitivity analyses were performed. For the induction phase: • by trials focusing on Japanese and Chinese populations For the one-year analysis: • by trials focusing on Japanese and Chinese populations • by mimicking a response-based trial design.
Subgroup analysis	Separate analyses were performed for trials conducted in patients who had not responded to biologic therapy (biologic failures) and patients who had not responded to conventional therapy (non-biologic failures).
Methods for pairwise meta-analysis	Standard pairwise meta-analyses were used to obtain direct comparison odds ratios as measurement of effects. Heterogeneity was evaluated for each pairwise comparison with the Cochran's Q test and the I^2 statistic. The Cochran's Q test was considered with a significance level of 10%, or I^2 higher than 50%. The forest plot was generated to illustrate a suspected heterogeneity. The inverse variance–weighted method was used to analyze the end points using a random-effects model.

DIC = deviance information criterion; ITC = indirect treatment comparison; NMA = network meta-analysis.

[a] Methods applied for both induction and maintenance phases.

Source: Stelara network meta-analysis Clinical Study Report.[39]

Results of the Sponsor-Submitted NMA

Summary of Included Studies

Only trials assessing at least one intervention of interest (biologics or conventional therapy for UC) were included in the NMA. A total of 49 publications were identified (including 32 full articles, 15 abstracts, and two posters) for six comparators: infliximab, adalimumab, vedolizumab, golimumab, tofacitinib, and placebo. For ustekinumab, six abstracts were identified through electronic search, though the clinical results of ustekinumab were primarily extracted from the Clinical Study Reports provided by Janssen. The number of trials identified for each treatment were six for infliximab, four for tofacitinib, five for adalimumab, three for golimumab, four for vedolizumab, and one for ustekinumab. The authors excluded 21 trials; 19 were used and contributed toward the narrative analysis of the results. In Figure 8 to Figure 13, a main overview and overviews of the trials included by subgroups are presented. In Table 17, a detail description of each trial with the interventions included is presented.

The authors assessed the risk of bias of 17 individual studies. Two studies (11.7%) had an unclear randomization process, 6 (35%) had unclear allocation concealment, 8 (47%) had unclear blinding, 1 (5.8%) had imbalances in dropouts between groups, another (5.8%) had likely selective reporting, while 8 (47%) did not perform an adequate ITT analysis.

Table 17: Characteristics of the Studies Included for the Comparisons in the ITC

Study or report	Location	Trial phase	Treatment phase	Blinding	Treatments, doses, and sample size
ULTRA 1 (ADA)	Multinational	III	Induction	Double blind	ADA 80 mg/40 mg (n = 130) ADA 160 mg/80 mg (n = 130) Placebo (n = 130)
		IV	Maintenance	Open label	ADA 40 mg (n = 390)
ULTRA 2 (ADA)	Multinational	III	Induction	Double blind	ADA 160 mg/80 mg (n = 258) Placebo (n = 260)
			Maintenance		ADA 40 mg (n = 258) Placebo (n = 260)
NCT00853099	Japan	II and III	Induction	Double blind	ADA 80 mg/40 mg (n = 87) ADA 160 mg/80 mg (n = 90) IFX (n = 21)
Jiang 2015 (IFX)	China	NA	Induction	Double blind	IFX 3.5 mg (n = 41) IFX 5 mg (n = 41)
			Maintenance		Placebo (n = 41)
Probert 2003 (IFX)	UK and Germany	NA	Induction	Double blind	IFX 5 mg (n = 23) Placebo (n = 20)
Japis CTI060297 (IFX)	Japan	III	Induction	Double blind	IFX 5 mg (n = 104)
			Maintenance		Placebo (n = 104)
ACT 1	Multinational	III	Induction	Double blind	IFX 5 mg/kg (n = 121) IFX 10 mg/kg (n = 122)
			Maintenance		Placebo (n = 121)
ACT 2	Multinational	III	Induction	Double blind	IFX 5 mg/kg (n = 121) IFX 10 mg/kg (n = 120)
			Maintenance		Placebo (n = 123)
PURSUIT-M (GOL)	Multinational	III	Maintenance	Double blind	GOL 50 mg (n = 154)
					GOL 100 mg (n = 154) Placebo (n = 156)
PURSUIT-J (GOL)	Japan	III	Induction	Open label	GOL 200 mg (n = 144)
			Maintenance	Double blind	GOL 100 mg (n = 32) GOL 100 mg (n = 60) Placebo (n = 31)
PURSUIT-SC (GOL)	Multinational	II	Induction	Double blind	GOL 100 mg/50 mg (n = 42) GOL 200 mg/100 mg (n = 42) GOL 400 mg/200 mg (n = 43) Placebo (n = 42)
		III	Induction		GOL 200 mg/100 mg (n = 258) GOL 400 mg/200 mg (n = 258) Placebo (n = 258)
NCT00787202 (TOF)	Multinational	II	Induction	Double blind	TOF 0.5 mg (n = 31)
					TOF 3 mg (n = 33)
					TOF 10 mg (n = 33)
					TOF 15 mg (n = 49)
					Placebo (n = 48)

Study or report	Location	Trial phase	Treatment phase	Blinding	Treatments, doses, and sample size
OCTAVE Induction 1 (OCTAVE-I1)	Multinational	III	Induction	Double blind	TOF 10 mg (n = 476) Placebo (n = 122)
OCTAVE Induction 2 (OCTAVE-I2)	Multinational	III	Induction	Double blind	TOF 10 mg (n = 429) Placebo (n = 112)
OCTAVE Sustain	Multinational	III	Maintenance	Double blind	TOF 5 mg (n = 198) TOF 10 mg (n = 197) Placebo (n = 198)
GEMINI 1	Multinational	III	Induction	Double blind and open label	Placebo (n = 149) (cohort 1)
					VDZ 300 mg (n = 225) (cohort 1)
					VDZ 300 mg (n = 521) (cohort 2)
			Maintenance	Double blind	VDZ 300 mg q.4.w. (n = 122) VDZ 300 mg q.8.w. (n = 125) Placebo (n = 126)
NCT02039505	Japan	III	Induction	Double blind open label	VDZ: 300 mg q.8.w. (n = 246) Placebo (n = 83)
			Maintenance	Double blind	VDZ: 300 mg (n = 246) Placebo (n = 83)
VARSITY	Multinational	III	Maintenance	Double blind	VDZ: 300 mg (n = 111) ADA: 160 mg/80 mg/40 mg (n = 120)
UNIFI	Multinational	III	Induction	Double blind	UST 130 mg (n = 320)
					UST 6 mg/kg (n = 322)
					Placebo (n = 319)
			Maintenance		UST 90 mg q.12.w. (n = 172)
					UST 90 mg q.8.w. (n = 176)
					Placebo (n = 175)

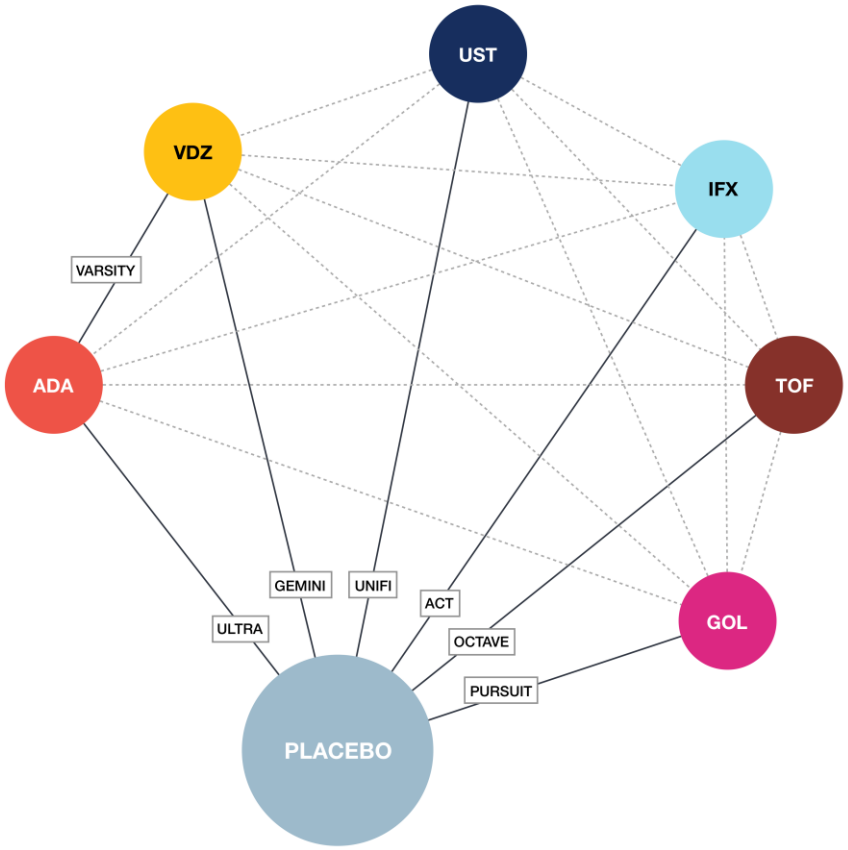
ADA = adalimumab; GOL = golimumab; IFX = infliximab; ITC = indirect treatment comparison; NA = not applicable; q.4.w. = every four weeks; q.8.w. = every eight weeks; q.12.w. = every 12 weeks; TOF = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Indirect comparison interventions: UST, IFX, ADA, VDZ, GOL, TOF, and placebo.

For the NMA of the maintenance phase, three of the studies (VARSITY, ULTRA, and ACT) used a treat-through design and one-year outcome data to mimic a treat-through design was directly available, while four (GEMINI, UNIFI, OCTAVE, AND PURSUIT) utilized a re-randomized, response-based strategy and required a recalculation of data to mimic a treat-through design.

Important characteristics of the effect modifiers and variables describing heterogeneity of the body of evidence is presented in Table 18.

Figure 8: Overview of the Clinical Studies and Comparisons in the Network



ADA = adalimumab; GOL = golimumab; IFX = infliximab; PLC = placebo; TOF = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Note: Continuous lines represent the direct comparisons with the name of the main study. Interrupted lines represent the indirect comparisons.

The VARSITY, ULTRA, and ACT studies used a treat-through design, while GEMINI, UNIFI, OCTAVE, and PURSUIT utilized a re-randomized, response-based strategy.

Source: Stelara network meta-analysis Clinical Study Report.[39]

Table 18: Assessment of Homogeneity for the ITCs

	Description and handling of potential effect modifiers
Disease severity	The authors note the main differences across studies to be disease duration at inclusion, C-reactive protein level at baseline, and Mayo score at baseline. The rest of the baseline populations' characteristics across studies were considered to be comparable in their clinical and laboratory values. All studies included patients with moderate-to-severe UC.
Treatment history	Prior anti-TNF therapy exposure was found to be a potential source of heterogeneity. For this, the authors performed subgroup analyses based on this variable. The following characteristics were assessed and deemed comparable across trials: duration of disease, age and weight at baseline, proportion of males and females, C-reactive protein level, and Mayo score at baseline.
Clinical trial eligibility criteria	All studies had comparable patient inclusion criteria. However, different study designs for the maintenance phase were found, as described in the text (ITC analysis methods).

	Description and handling of potential effect modifiers
Dosing of comparators	Doses of comparators are in keeping with the indication for moderate-to-severe UC.
Definitions of end points	Clinical remission was defined differently by two studies: one definition was used only in the induction phase of the NMA, while the other used remission instead of clinical remission, defined as a total Mayo score of 0 to 2, with no subscore exceeding 1 point and a rectal bleeding subscore of 0. UNIFI defined mucosal healing as "having both endoscopic healing and histologic healing," and endoscopic healing as a "Mayo endoscopy subscore of 0 or 1." To align with the definitions across the other trials for mucosal healing, the data corresponding to endoscopic healing were included for the analysis of this end point and referred to as "mucosal healing" hereafter.
Timing of end point evaluation or trial duration	The efficacy end points across studies included clinical response, clinical remission, and mucosal healing at week 6, 8, and 10 for the induction phase, and from week 30 to week 60 for the maintenance phase. When multiple time points were reported, similar times of assessment were selected for each intervention. End points reported at 6 weeks were used as inputs for golimumab and vedolizumab, and end point at 8 weeks were used as inputs for ustekinumab, tofacitinib, adalimumab, and infliximab. The authors considered the results obtained between 6 weeks and 8 weeks to be comparable.
Clinical trial setting	All settings were similar, with the inclusion of patients in both ambulatory and hospital settings.
Study design	All studies were randomized controlled trials against placebo (except for one head-to-head comparison study). As described in the ITC methods section, trial designs conducted in UC have evolved over time from standard **treat-through** designs for anti-TNF therapies (including infliximab and adalimumab) to designs based on response to treatment (**response-based re-randomized**) for the newer therapies. The authors adapted and recalculated the data to make these two design types equivalent so they could be pooled in the NMA; this adaptation was necessary only for the analysis of outcomes at one year.

ITC = indirect treatment comparison; NMA = network meta-analysis; TNF = tumour necrosis factor; UC = ulcerative colitis.

Source: Stelara NMA study report.[39]

Results

In Figure 9 to Figure 14, a description of the different studies included for each network is presented. Based on the subdivision of patients who failed biologic therapy versus those who did not fail biologic therapy, the results are presented in Figure 9 to Figure 14, as well as in Table 19 to Table 24.

Induction Phase

For the induction phase, the fixed-effects model was chosen on the three outcomes (clinical response, clinical remission, and mucosal healing) for both the non-biologic and biologic failure groups. This was based on the smallest deviance information criterion (DIC) detected in the Bayesian framework for model selection.

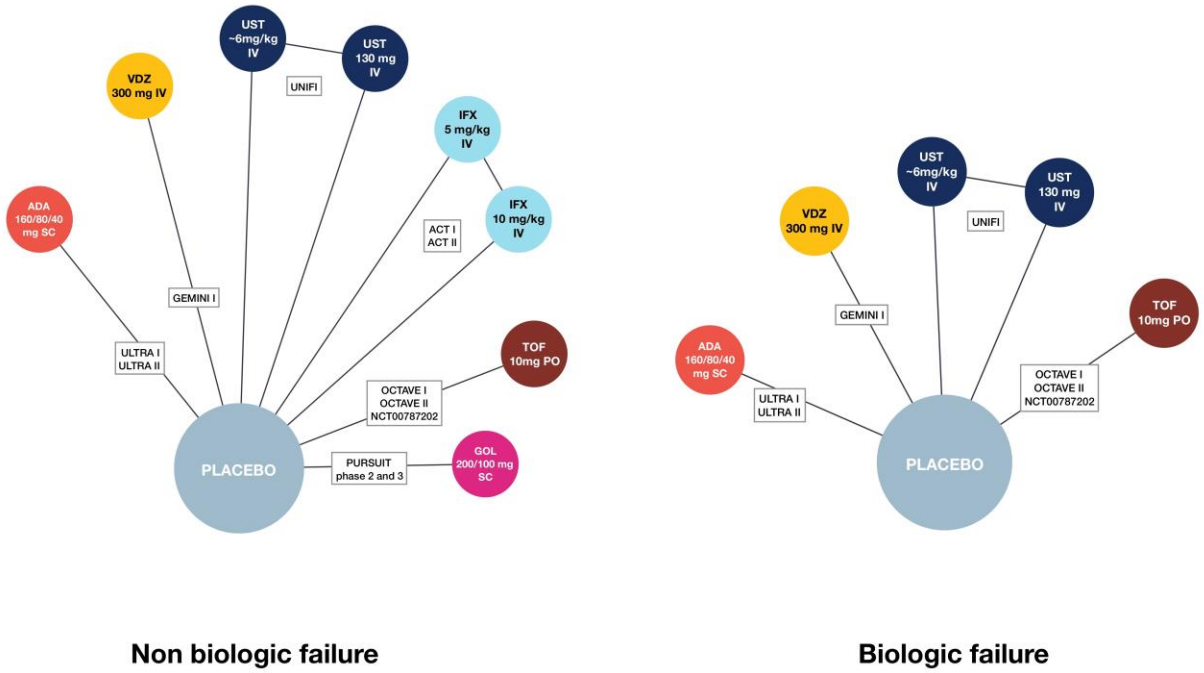
Clinical Response

In the non-biologic failure group, 10 studies assessing the outcome of clinical response were included (Figure 8). The sample size of the studies ranged from 56 (NTC00787202) to 550 (OCTAVE I and II combined). The response rates ranged from 43.9% in PURSUIT-SC phase II (GOL 200 mg or 100 mg, n = 41) to 69.4% in ACT I (infliximab 5 mg, n = 121). UNIFI reported response rates of 57.7% and 66.7% in the 130 mg and 6 mg/kg treatment arms, respectively. The clinical response rates in the placebo arms across all trials ranged from 26.3% in GEMINI I (n = 76) to 45.5% in NTC00787202 (n = 33); UNIFI reported a 35.4% response rate in the placebo arm. Based on the direct comparisons estimates, all interventions were better than placebo, although heterogeneity was present, with I^2 values ranging from 0% to 69.9%. From the NMA estimates (Table 19), ustekinumab at a dose of

6 mg/kg was not superior to the other comparisons, with the exception of placebo (median odds ratio [OR] = 3.66; credible interval [CrI], 2.31 to 5.88) and adalimumab 160 mg or 80 mg (median OR = 1.94; 95% CrI, 1.10 to 3.45). At a 130 mg dose, ustekinumab was only better than placebo (OR = 2.49; CrI, 1.58 to 3.96).

In the biologic failure group, ustekinumab at 130 mg/kg and 6 mg/kg doses was better than placebo (OR = 2.20; CrI, 1.39 to 3.53 and OR = 3.58; CrI, 2.27 to 5.74, respectively) and the 6 mg/kg dose of ustekinumab had an increased response when compared with adalimumab (OR = 2.48; CrI, 1.17 to 5.31). It was not, however, better or worse than the other comparisons for this outcome (Table 19).

Figure 9: Clinical Response (Induction) Studies in the Network



ADA = adalimumab; GOL = golimumab; IFX = infliximab; PO = orally; SC = subcutaneous; TOF = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Table 19: Clinical Response – Induction Study

	Non-biologic failure patients		Biologic failure patients	
	OR (95% CrI) for ustekinumab 130 mg versus comparator	OR (95% CrI) for ustekinumab 6 mg/kg versus comparator	OR (95% CrI) for ustekinumab 130 mg versus comparator	OR (95% CrI) for ustekinumab 6 mg/kg versus comparator
Infliximab 5 mg/kg	0.61 (0.33 to 1.10)	0.89 (0.49 to 1.63)		
Infliximab 10 mg/kg	0.65 (0.36 to 1.19)	0.96 (0.53 to 1.76)		
Adalimumab 160 mg/80 mg	1.32 (0.75 to 2.33)	1.94 (1.10 to 3.45)	1.52 (0.71 to 3.25)	2.48 (1.17 to 5.31)
Vedolizumab 300 mg	0.78 (0.36 to 1.68)	1.14 (0.52 to 2.47)	0.87 (0.35 to 2.11)	1.43 (0.58 to 3.43)
Golimumab 200 mg/100 mg	1.09 (0.61 to 1.93)	1.60 (0.90 to 2.84)		
Tofacitinib 10 mg	0.92 (0.50 to 1.71)	1.36 (0.74 to 2.53)	0.64 (0.34 to 1.21)	1.05 (0.55 to 1.98)
Placebo	2.49 (1.58 to 3.96)	3.66 (2.31 to 5.88)	2.20 (1.39 to 3.53)	3.58 (2.27 to 5.74)

CrI = credible interval; OR = odds ratio.

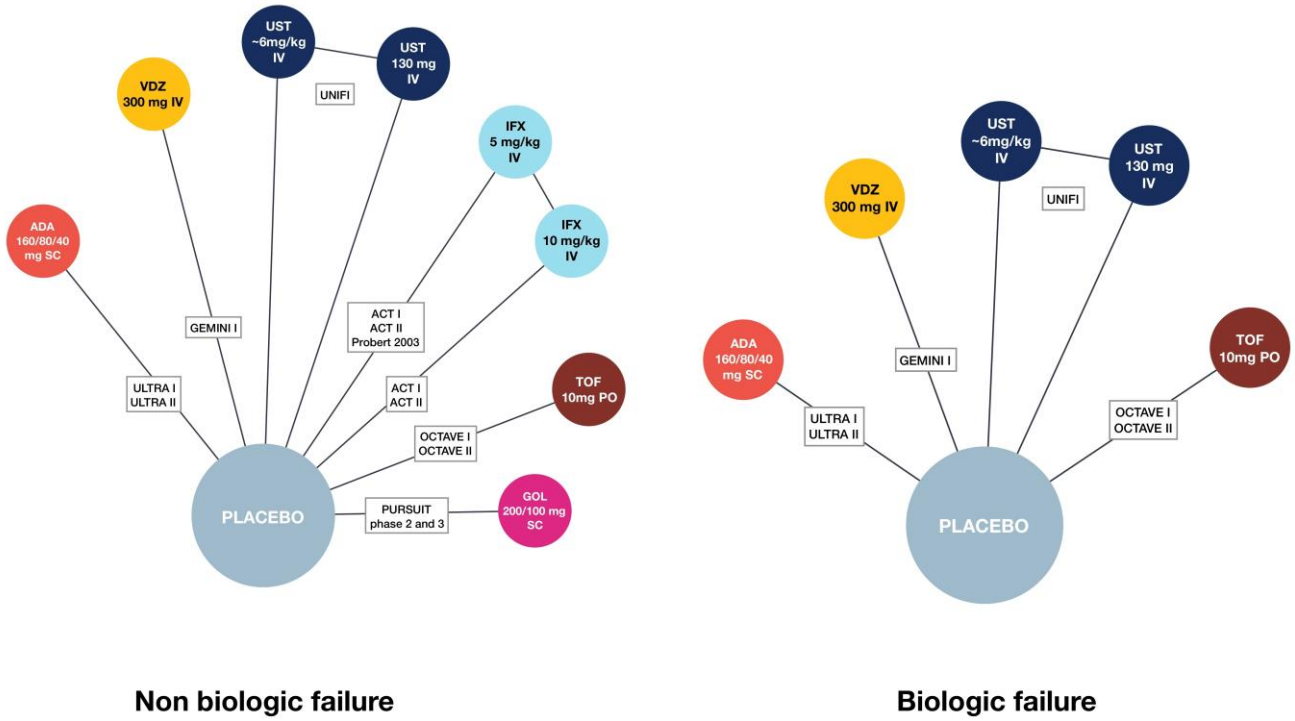
Source: Stelara network meta-analysis Clinical Study Report.[39]

Clinical Remission

In the non-biologic failure group, a total of 11 studies were included (Figure 10). The sample size of the studies ranged from 43 (Probert 2003) to 254 (PURSUIT-SC phase III). The remission rates ranged from 17.1% in PURSUIT-SC phase II (GOL 200 mg/100 mg, n = 41) to 39% in Probert 2003 (infliximab 5 mg/kg, n = 23). UNIFI reported remission rates of 19.9% and 18.6% in the 130 mg and in the 6 mg/kg treatment arm, respectively. The clinical remission rates in the placebo arms across all trials ranged from 5.7% in ACT II (n = 123) to 30% in Probert 2003 (n = 20); UNFI reported 9.5% remission rate in placebo arm. On the direct comparisons, all interventions performed better than placebo, with two comparisons presenting significant heterogeneity (infliximab 10 mg/kg versus placebo and infliximab 5 mg/kg versus placebo, I^2 of 58.9% and 62.7%, respectively). On the NMA estimates, ustekinumab did not show differences against any of the other interventions (Table 20), either at doses of 130 mg or 6 mg/kg, except against placebo (OR = 2.38 [CrI, 1.24 to 4.78] and OR = 2.19 [CrI, 1.14 to 4.39], respectively).

In the biologic failure group, the results were similar. With four studies included (Figure 10) the sample size ranged from 145 (GEMINI I) to 589 (OCTAVE I and II combined). The remission rates ranged from 9.2% in ULTRA II (adalimumab 160 mg or 80 mg; n = 98) to 12.7% in UNIFI (ustekinumab 6 mg/kg; n = 166). The clinical remission rates in the placebo arms across all trials ranged from 0.8% in OCTAVE I and II combined (n = 124) to 6.9% in ULTRA II (n = 101); UNIFI reported a 1.2% remission rate in the placebo arm. From the NMA results, ustekinumab had a better effect than adalimumab at both doses of 130 mg and 6 mg/kg, although with wide CrIs (OR = 9.01; CrI, 1.58 to 80.08 and OR = 9.97; CrI, 1.77 to 88.37) and a better effect against placebo (Table 21).

Figure 10: Clinical Remission Studies Included in the Network – Induction Study



ADA = adalimumab; GOL = golimumab; IFX = infliximab; PO = orally; SC = subcutaneous; TOF = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Table 20: Clinical Remission – Induction Study

	Non-biologic failure patients		Biologic failure patients	
	OR (95% CrI) ustekinumab 130 mg versus comparator	OR (95% CrI) ustekinumab 6 mg/kg versus comparator	OR (95% CrI) ustekinumab 130 mg versus comparator	OR (95% CrI) ustekinumab 6 mg/kg versus comparator
Infliximab 5 mg/kg	0.54 (0.24 to 1.22)	0.49 (0.22 to 1.14)	–	–
Infliximab 10 mg/kg	0.70 (0.31 to 1.62)	0.64 (0.28 to 1.48)	–	–
Adalimumab 160 mg/80 mg	1.08 (0.47 to 2.49)	0.99 (0.43 to 2.30)	9.01 (1.58 to 80.08)	9.97 (1.77 to 88.37)
Vedolizumab 300 mg	0.52 (0.14 to 1.70)	0.48 (0.13 to 1.58)	3.27 (0.29 to 36.81)	3.60 (0.32 to 40.71)
Golimumab 200 mg/100 mg	0.80 (0.34 to 1.93)	0.74 (0.31 to 1.78)	–	–
Tofacitinib 10 mg	0.98 (0.38 to 2.42)	0.90 (0.35 to 2.24)	0.54 (0.02 to 7.18)	0.59 (0.02 to 7.92)
Placebo	2.38 (1.24 to 4.78)	2.19 (1.14 to 4.39)	12.12 (3.24 to 86.24)	13.41 (3.62 to 94.58)

CrI = credible interval; OR = odds ratio.

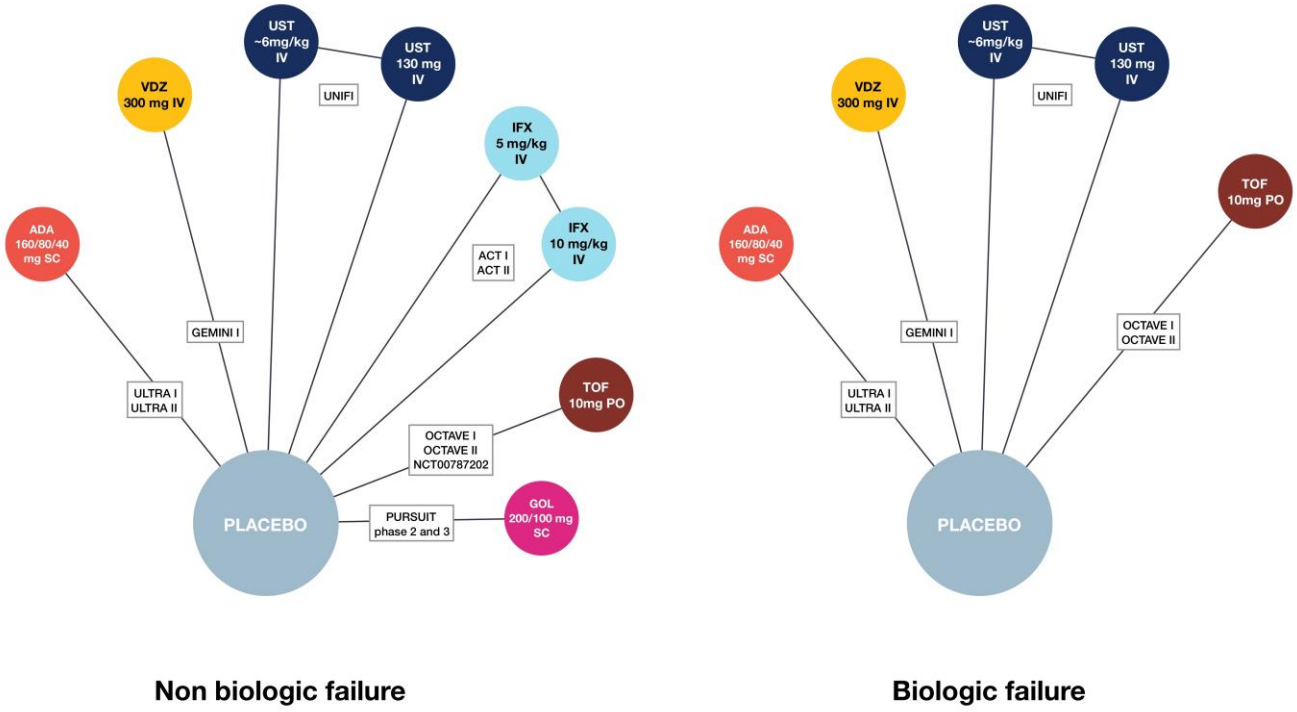
Source: Stelara network meta-analysis Clinical Study Report.[39]

Mucosal Healing

In the non-biologic failure group, authors included nine studies (Figure 11). The sample size of these studies ranged from 82 (PURSUIT-SC phase II) to 550 (OCTAVE I and II - combined). The mucosal healing rates ranged from 33.3% in UNIFI (ustekinumab 6 mg/kg, n = 156) to 62% in ACT I (infliximab 5 mg/kg, n = 121). The mucosal healing rates in the placebo arms across all trials ranged from 20.9% in UNIFI (n = 158) to 41.5% in ULTRA I (n = 130). From the pairwise direct comparisons, all interventions had better effects than placebo, with no significant heterogeneity. In the NMA, ustekinumab was better than placebo at both doses of 130 mg and 6 mg/kg (OR = 2.01 [CrI, 1.22 to 3.40] and OR = 1.90 [CrI, 1.15 to 3.20], respectively) but did not present differences when compared with any of the other interventions (Table 21).

In the biologic failure group, four studies are presented in Figure 11. The sample size of the studies ranged from 145 (GEMINI I) to 589 (OCTAVE I and II combined). The mucosal healing rates ranged from 18.3% in UNIFI (ustekinumab 130 mg, n = 164) to 30.5% in GEMINI I (vedolizumab 300 mg, n = 82). The mucosal healing rates in the placebo arms across all trials ranged from 6.5% in OCTAVE I and II combined (n = 124) to 26.7% in ULTRA II (n = 101); UNFI reported a mucosal healing rate of 6.8% in the placebo arm. When assessing the indirect (NMA) estimates of effect, ustekinumab was better than placebo (OR = 3.12 [CrI, 1.53 to 6.78] and OR = 3.73 [CrI, 1.86 to 8.04] for the 130 mg and 6 mg/kg doses, respectively) and adalimumab (OR = 2.85 [CrI, 1.10 to 7.68], and OR = 3.42 [1.33 to 9.12] for the 130 mg and 6 mg/kg doses, respectively) but no different than tofacitinib and vedolizumab (Table 21).

Figure 11: Mucosal Healing (Induction) Studies in the Network



ADA = adalimumab; GOL = golimumab; IFX = infliximab; PO = orally; SC = subcutaneous; TOF = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Table 21: Mucosal Healing – Induction Study

	Non-biologic failure patients		Biologic failure patients	
	OR (95% CrI) ustekinumab 130 mg versus comparator[a]	OR (95% CrI) ustekinumab 6 mg/kg versus comparator[a]	OR (95% CrI) ustekinumab 130 mg versus comparator[a]	OR (95% CrI) ustekinumab 6 mg/kg versus comparator[a]
Infliximab 5 mg/kg	0.61 (0.32 to 1.15)	0.57 (0.30 to 1.10)	–	–
Infliximab 10 mg/kg	0.63 (0.33 to 1.20)	0.59 (0.32 to 1.13)	–	–
Adalimumab 160 mg/80 mg	1.33 (0.72 to 2.49)	1.26 (0.68 to 2.35)	2.85 (1.10 to 7.68)	3.42 (1.33 to 9.12)
Vedolizumab 300 mg	0.68 (0.30 to 1.52)	0.64 (0.29 to 1.45)	1.83 (0.63 to 5.40)	2.19 (0.76 to 6.41)
Golimumab 200 mg/100 mg	1.12 (0.60 to 2.09)	1.06 (0.57 to 1.98)	–	–
Tofacitinib 10 mg	0.90 (0.44 to 1.81)	0.85 (0.41 to 1.72)	0.73 (0.24 to 2.07)	0.87 (0.29 to 2.46)
Placebo	2.01 (1.22 to 3.40)	1.90 (1.15 to 3.20)	3.12 (1.53 to 6.78)	3.73 (1.86 to 8.04)

CrI = credible interval; OR = odds ratio.

Results reported from a network meta-analysis with a fixed-effects model under a Bayesian framework.

Source: Stelara network meta-analysis Clinical Study Report.[39]

Maintenance Phase

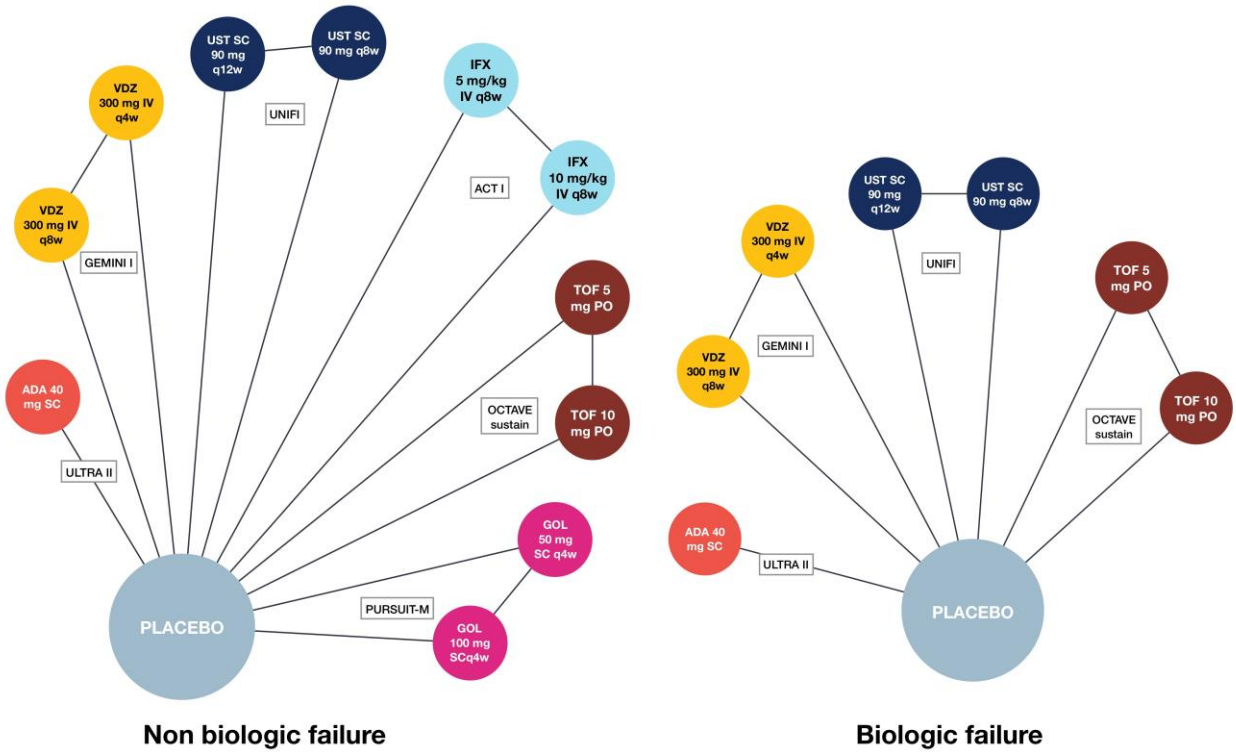
The results of the one-year NMAs are presented for the base-case approach mimicking a treat-through design. Only fixed-effects model results were presented due to data being obtained from one study informing each pair of treatments in the networks and therefore no data to inform the random-effects parameter.

Clinical Response

For the non-biologic failure population, doses were pooled for treatment arms (as no dose-response relationship was observed) to increase statistical power. Six studies were included in the analysis. The network of studies for the clinical response is displayed in Figure 12. Ustekinumab presented higher odds of clinical response against adalimumab, golimumab, tofacitinib, and placebo, but not against vedolizumab 300 mg (OR = 1.93; CrI, 0.75 to 4.82) (Table 22).

Among the biologic failure patients, doses for treatment arms were not pooled, as a dose-response relationship was evident. An analysis of mucosal healing in the biologic failure group was not feasible, as the imputation data needed for placebo were not available in this population. In total, four studies were used for evaluating this outcome (Figure 12). In the NMA, ustekinumab demonstrated higher odds of clinical response than placebo (OR = 4.83 [CrI, 2.56 to 9.25] and OR = 4.82 [CrI, 2.28 to 10.30] for the every eight week and every 12 week doses, respectively) but showed no effect when compared with the other interventions (Table 22).

Figure 12: Clinical Response (One-Year Base Case) Studies in the Network



ADA = adalimumab; GOL = golimumab; IFX = infliximab; PO = orally; q4w = every four weeks; q8w = every eight weeks; q12w = every 12 weeks; SC = subcutaneous; TOF = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Table 22: Clinical Response (One Year Mimicking of a Treat-Through Approach)[a]

	Non-biologic failure patients	Biologic failure patients	
	OR (95% CrI) ustekinumab 6 mg/kg – ustekinumab 90 mg pooled	OR (95% CrI) ustekinumab 6 mg/kg – ustekinumab 90 mg q.8.w.	OR (95% CrI) Ustekinumab 6 mg/kg – Ustekinumab 90 mg q.12.w.
Infliximab pooled – Infliximab pooled[b]	2.62 (1.22 to 5.60)		
Adalimumab 160 mg/80 mg/40 mg – adalimumab 40 mg[b]	4.76 (2.25 to 10.16)	2.03 (0.70 to 5.72)	2.02 (0.65 to 6.14)
Vedolizumab 300 mg – vedolizumab 300 mg pooled	1.93 (0.75 to 4.82)	1.76 (0.51 to 6.00)	1.75 (0.48 to 6.35)
Golimumab 200 mg/100 mg – golimumab pooled	3.76 (1.90 to 7.57)		
Tofacitinib 10 mg – tofacitinib pooled	2.27 (1.06 to 4.86)	1.66 (0.69 to 3.94)	1.65 (0.63 to 4.28)
Placebo – placebo	8.70 (5.03 to 15.40)	4.83 (2.56 to 9.25)	4.82 (2.28 to 10.30)

CrI = credible interval; OR = odds ratio; q.8.w. = every eight weeks; q.12.w. = every 12 weeks.

[a] All reported results were generated based on the mimicking of a treat-through design; hence, sample size was recalculated to correspond with number of participants at the baseline of the induction phase.

[b] For infliximab and adalimumab, no data recalculation was performed, as it comes directly from treat-through design studies.

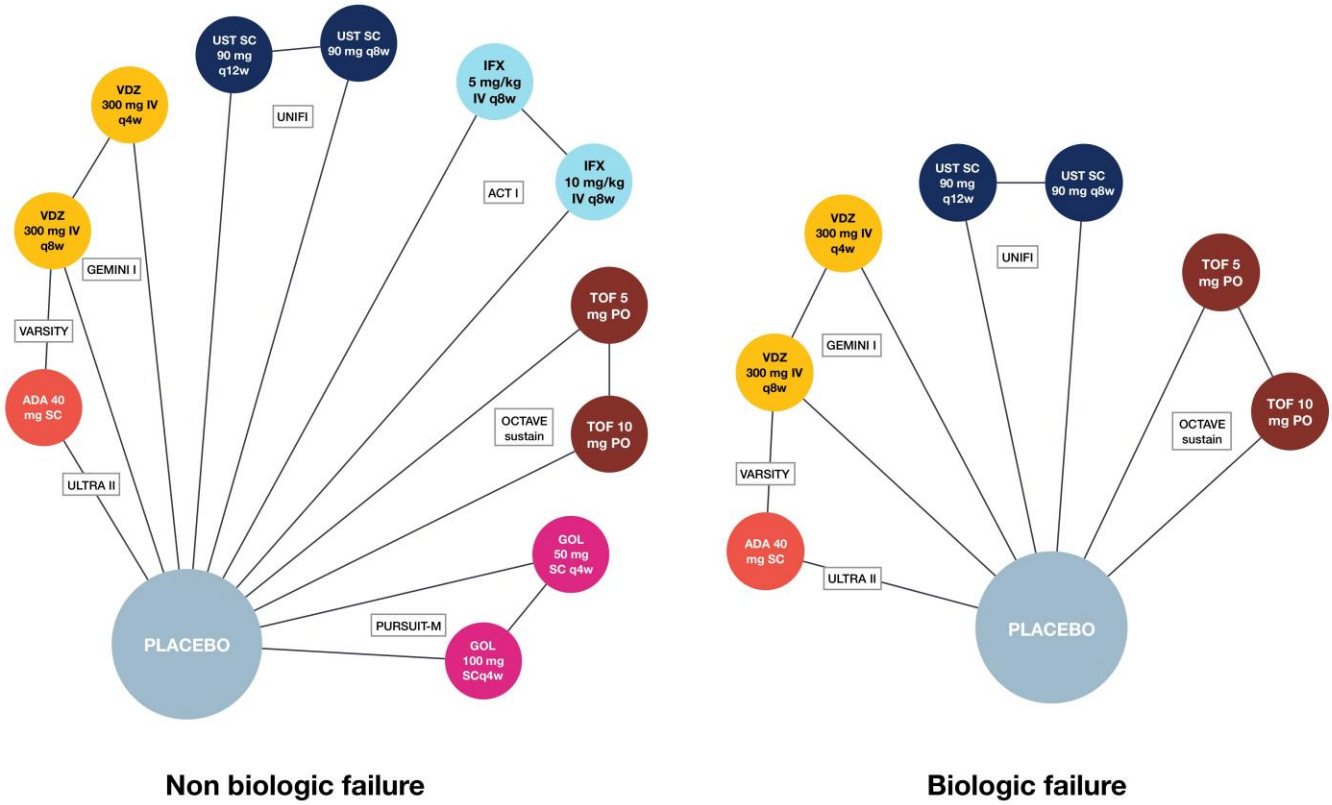
Source: Stelara network meta-analysis Clinical Study Report.[39]

Clinical Remission

In the non-biologic failure group, seven studies were included in the analysis. The network of studies for the clinical remission is displayed in Figure 13. Based on the results in the individual trials, no dose-response relationship was observed for clinical remission; therefore, the authors considered it appropriate to pool the doses for the same treatment. Ustekinumab presented higher odds of clinical remission than placebo (OR = 5.11; CrI, 2.83 to 9.52), golimumab (OR = 2.40; CrI, 1.13 to 5.22), and adalimumab (OR = 2.43; CrI, 1.10 to 5.42), but no statistical difference when compared with vedolizumab, infliximab, or tofacitinib (Table 23).

In the biologic failure group, five studies were included (Figure 13). The authors only present the results for un-pooled doses, as a dose-response relationship was detected. In this group, ustekinumab had better odds for clinical remission than placebo but not against the other interventions (adalimumab, vedolizumab, and tofacitinib) (Table 23).

Figure 13: Clinical Remission (One-Year Base Case) Studies in the Network



ADA = adalimumab; GOL = golimumab; IFX = infliximab; PO = orally; q4w = every four weeks; q8w = every eight weeks; q12w = every 12 weeks; SC = subcutaneous; TOF = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Table 23: Clinical Remission (One Year Mimicking a Treat-Through Approach)

	Non-biologic failure patients	Biologic failure patients	
	OR (95% CrI) ustekinumab 6 mg/kg – ustekinumab 90 mg pooled	OR (95% CrI) ustekinumab 6 mg/kg – ustekinumab 90 mg q.8.w.	OR (95% CrI) ustekinumab 6 mg/kg – Ustekinumab 90 mg q.12.w.
Infliximab pooled – infliximab pooled[a]	1.89 (0.83 to 4.29)		
Adalimumab 160 mg/80 mg/40 mg – adalimumab 40 mg[a]	2.43 (1.10 to 5.42)	1.71 (0.42 to 6.55)	1.32 (0.29 to 5.48)
Vedolizumab 300 mg – vedolizumab 300 mg pooled	1.47 (0.65 to 3.33)	1.26 (0.31 to 4.91)	0.97 (0.22 to 4.11)
Golimumab 200 mg/100 mg – golimumab	2.40 (1.13 to 5.22)		
Tofacitinib 10 mg – tofacitinib pooled	1.51 (0.64 to 3.51)	1.57 (0.44 to 5.36)	1.21 (0.31 to 4.52)
Placebo – placebo	5.11 (2.83 to 9.52)	6.89 (2.98 to 16.90)	5.34 (1.97 to 14.62)

CrI = credible interval; OR = odds ratio; q.8.w. = every eight weeks; q.12.w. = every 12 weeks.

Note: All reported results were generated based on mimicking a treat-through design; hence, sample size was recalculated to correspond with number of participants at baseline of the induction phase.

[a] For infliximab and adalimumab, no data recalculation was performed, as this data comes directly from the treat-through design studies.

Source: Stelara network meta-analysis Clinical Study Report.[39]

Mucosal Healing

For the non-biologic failure response, six studies were included (Figure 14). In these comparisons, authors pooled the doses of the same treatments, as there was no dose-response relationship observed. Ustekinumab was superior to placebo (OR = 5.57; CrI, 3.19 to 9.92), adalimumab (OR = 2.91; CrI, 1.33 to 6.39), and golimumab (OR = 2.79; CrI, 1.39 to 5.69), but it did not present a difference when compared with infliximab, tofacitinib, or vedolizumab (Table 24).

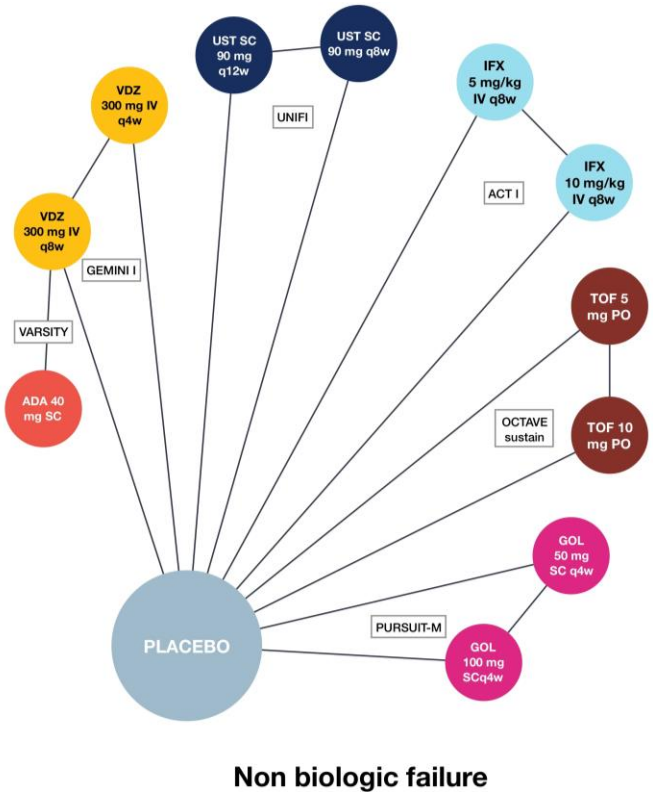
There were no data available for assessing the group of biologic failures for the outcome of mucosal healing, as the imputation data needed for placebo were not available in this population.

Sensitivity Analyses

Authors performed a priori determined sensitivity analyses to test the robustness of the results obtained in the base-case analysis. For this, in the induction phase, they included trials focusing on Japanese and Chinese populations, perceiving these as possible effect modifiers. No effect difference was noted on this sensitivity analysis. Also, a sensitivity analysis that included or excluded open-label trials was conducted, but no open-label trials were identified (that, with the exception of GEMINI I, also presented double-blind data); therefore, authors did not use this sensitivity analysis, as it was not applicable. A sensitivity analysis reporting the results of the fixed-effects model was presented if the random-effect model was selected in the base-case analysis. The authors always selected the fixed-effect model and this sensitivity analysis was not applicable, according to their methods.

For the one-year NMA results, authors also focused on trials of Japanese and Chinese populations where no differences were detected. Also, authors performed an analysis of all NMA estimates based on mimicking a response-based approach. None of these showed a significant difference in the sensitivity analysis.

Figure 14: Mucosal Healing (One-Year Base Case) Studies in the Network



ADA = adalimumab; GOL = golimumab; IFX = infliximab; PO = orally; q4w = every four weeks; q8w = every eight weeks; q12w = every 12 weeks; SC = subcutaneous; TOF = tofacitinib; UST = ustekinumab; VDZ = vedolizumab.

Table 24: Mucosal Healing (One Year Mimicking a Treat-Through Approach)

	Non-biologic failure patients	Biologic failure patients
	OR (95% CrI) ustekinumab 6 mg/kg – ustekinumab 90 mg pooled	OR (95% CrI) Ustekinumab 6 mg/kg – ustekinumab 90 mg pooled
Infliximab pooled – infliximab pooled[a]	1.43 (0.66 to 3.09)	NA
Adalimumab 160 mg/80 mg/40 mg – adalimumab 40 mg[a]	2.91 (1.33 to 6.39)	NA
Vedolizumab 300 mg – vedolizumab 300 mg pooled	1.60 (0.69 to 3.77)	NA
Golimumab 200 mg/100 mg – golimumab pooled	2.79 (1.39 to 5.69)	NA
Tofacitinib 10 mg – tofacitinib pooled	1.94 (0.88 to 4.25)	NA
Placebo – placebo	5.57 (3.19 to 9.92)	NA

CrI = credible interval; NA = not applicable; OR = odds ratio.

Note: All reported results were generated based on mimicking a treat-through design; hence, sample size was recalculated to correspond with number of participants at baseline of the induction phase.

[a] For infliximab and adalimumab, no data recalculation was performed, as this data comes directly from the treat-through design studies.

Source: Stelara network meta-analysis Clinical Study Report.[39]

Critical Appraisal of the ITC

This systematic review and NMA of ITCs was performed under the PRISMA report checklist, with an appropriate search strategy and based on an available protocol. The reviewers performed an appropriate data extraction and analysis. Several limitations were noted.

Significant heterogeneity (inconsistency) was considered due to differences in the design of the studies in the maintenance phase. Even though the authors performed an imputation method for mimicking a treat-through design (needed for obtaining pooled estimates of effect and assessing heterogeneity from pairwise comparisons), no overall narrative or statistical assessment of heterogeneity or the inconsistency of the network is presented. The variations in placebo-effect estimates across studies support these concerns about heterogeneity and indicate possible violations of the assumptions of transitivity for the NMA. The likely explanation for these variations is due to the primary outcomes being based on a subjective measure (Mayo score). Different routes of drug administration and dose and regimen plans could provide different placebo-effect estimates.

Risk of bias from the individual studies was detected in the systematic review provided by the sponsor; for instance, bias due to an unclear randomization process was present in 35% of the studies, 47% had unclear blinding, 5.8% had unbalanced dropout rates, and 8% had no ITT analysis. For the NMA and its estimates, the included studies are reported to have moderate risk of bias (i.e., low or unclear) in the allocation concealment domain (OCTAVE, PURSUIT), in the blinding of patients or outcome assessors (OCTAVE and PURSUIT), and in using an adequate ITT analysis or missing data (OCTAVE), although these risks of bias are not described in detail for each outcome.

Although the NMAs conducted in non-biologic failure patients at one year were robust, involved fewer imputations, and presented one loop for the network of clinical remission and mucosal healing, there is imprecision in the effect estimates on other comparisons, with a considerable number of wide CrIs around several NMA effect estimates, suggesting low numbers of events or issues with the stability of the model results.

The authors present results based on the fixed-effects model, and one analysis that used a random-effects model is presented as an appendix for comparison. Given the heterogeneity suspected, using a random-effects model would have been desirable, as would presenting both results with fixed- and random-effects models for proper comparison.

The authors' recalculation of data in the mimicking of a treat-through design is a novel approach with one reference for validation.[40] They based their results on an imputation process that might be prone to errors and a risk of bias, and the severity and direction of any potential bias is unknown. The precision of their comparisons was likely overestimated because they used a single imputation approach and did a recalculation of the sample size to correspond to the number of patients recruited during the induction phase of the study. By doing so, the authors were assuming they have more observed data than actually exists in each study, which can produce an overestimated precision. The authors provided a multiple-imputation sensitivity analysis that showed an overall consistent result and no change in conclusions. However, the shortcomings of this analysis are that the distributions used to impute data did not account for the heterogeneity observed across studies, and recalculations did not account for patients who dropped out of the studies. These shortcomings can both overestimate precision.

Only three outcomes were included in this NMA. Furthermore, no AEs were evaluated. Although the authors clearly state their reasons for not addressing AEs, a narrative statement of all of the important outcomes (including AEs) would be helpful to provide a comprehensive overview of the desirable and undesirable effects for decision-making.

Summary

This systematic review and NMA of ITCs provides a synthesis to assess the efficacy of ustekinumab when compared indirectly with other interventions, namely, infliximab, adalimumab, vedolizumab, golimumab, tofacitinib, and placebo. It evaluates three outcomes: clinical remission, clinical response, and mucosal healing, all in patients considered biologic or non-biologic failures, and also in the induction and maintenance phases of drug administration. To address the differences in the maintenance phase of the studies, the authors performed a recalculation to equalize studies in terms of design, that is, the data from re-randomized response-based trials were recalculated to mimic data that would have been obtained through a treat-through design, and relative effects and probabilities were calculated at the one-year target efficacy.

Based on the NMA of the induction phase, ustekinumab had higher odds of clinical response, clinical remission, and mucosal healing against placebo and adalimumab (both biologic and non-biologic failure patients were more likely to achieve clinical response, but only biologic failure patients were more likely to achieve both clinical remission and mucosal healing). For the rest of the comparisons, ustekinumab either did not increase or decreased the odds of any of these outcomes when compared with infliximab, vedolizumab, golimumab, and tofacitinib.

In the maintenance phase, ustekinumab had higher odds of clinical response in non-biologic failure patients when compared with adalimumab, golimumab, tofacitinib, and placebo, but not against vedolizumab, while in the biologic failure patients, it was only better than placebo. For clinical remission, ustekinumab provided higher odds against golimumab, adalimumab, and placebo in the non-biologic failure group (but not against vedolizumab, infliximab, or tofacitinib); while in the biologic failure group, ustekinumab was only better than placebo. Lastly, ustekinumab had higher odds of mucosal healing in non-biologic

failure patients than adalimumab, golimumab, and placebo, but it was no better than infliximab, tofacitinib, and vedolizumab.

The limitations of the NMA include uncertainty about the effect estimates, particularly for the one-year outcomes, mostly due to concerns regarding heterogeneity and intransitivity, the potential for bias due to violations in the assumptions of the imputation process, and overestimated precision for reported comparisons. Furthermore, individual studies had a moderate risk of bias, with concerns arising from the randomization process, unclear blinding, and unbalanced dropout rates, and no ITT analysis.

Discussion

Summary of Available Evidence

One study, the UNIFI trial, and an ITC were reviewed. The UNIFI trial represents the only randomized trial available that assessed the use of ustekinumab in patients with moderate-to-severe UC and its results were included into the systematic review of ITCs (NMA).

UNIFI is an RCT with an eight-week induction and a 44-week maintenance period that included a total of 961 patients randomized to ustekinumab (either 130 mg [n = 320 patients] or approximately 6 mg/kg [n = 322]) or placebo (n = 319). Patients who had a response to induction therapy eight weeks after the administration of IV ustekinumab were randomly assigned again to receive SC maintenance injections of 90 mg of ustekinumab (either every 12 weeks [172 patients] or every eight weeks [176]) or placebo (175). The primary end point in the induction trial (week 8) and the maintenance trial (week 44) was clinical remission. In both phases, ustekinumab improved the primary outcomes. For instance, in the induction phase, the groups that received IV ustekinumab at week 8 had a higher proportion of patients with clinical remission, both in the 130 mg (15.6%) group and the 6 mg/kg (15.5%) group when compared with placebo (5.3%) ($P < 0.001$ for both comparisons), while the proportion of patients who had clinical remission at week 44 was significantly higher in the 90 mg every 12 weeks group (38.4%) and every eight weeks (43.8%) group than in the placebo group (24.0%) ($P = 0.002$ and $P < 0.001$, respectively), with no clear differences in the proportion of AEs.

The systematic review of ITCs is an evidence synthesis that was submitted by the sponsor to address the effect of ustekinumab versus other comparators with similar indications. The authors of the review performed an NMA to assess the efficacy indirectly compared with other interventions: infliximab, adalimumab, vedolizumab, golimumab, tofacitinib, and placebo, and evaluated three outcomes: clinical remission, clinical response, and mucosal healing, all in patients considered biologic or non-biologic failures, and also in the induction and maintenance phases of drug administration. The authors performed recalculations to equalize data from re-randomized response-based trials to mimic data that would have been obtained through a treat-through design. Based on the NMA of the induction phase, ustekinumab had higher odds of clinical response, clinical remission, and mucosal healing against placebo and adalimumab (in biologic and non-biologic failure patients for clinical response, but only in biologic failure patients for clinical remission and mucosal healing). For the rest of the comparisons, ustekinumab either did not increase or decreased the odds of any of these outcomes when compared with infliximab, vedolizumab, golimumab, and tofacitinib. In the maintenance phase, ustekinumab had higher odds of clinical response in non-biologic failure patients when compared with adalimumab, golimumab, tofacitinib, and placebo, but not against vedolizumab while, in the biologic failure patients, it was only better

than placebo. For clinical remission, ustekinumab provided higher odds against golimumab, adalimumab, and placebo in the non-biologic failure group (but not against vedolizumab, infliximab, or tofacitinib) while, in the biologic failure group, it was only better than placebo. Lastly, ustekinumab had higher odds of mucosal healing in non-biologic failure patients than adalimumab, golimumab, and placebo, but it was no better than infliximab, tofacitinib, and vedolizumab.

Interpretation of Results

Efficacy

Overall, ustekinumab was more effective than placebo for inducing and maintaining remission in patients with moderate-to-severe UC. When compared with placebo, ustekinumab IV demonstrated superiority in both doses (6 mg/kg and 130 mg) for induction of remission, and this effect was also observed in the maintenance phase, at a dosage of 90 mg SC either every 8 or every 12 weeks. The results were consistently in favour of ustekinumab in both phases, regardless of the definition of clinical remission used (US or global), and sensitivity analyses showed the results to be robust and relatively stable. Subgroup analyses also generally aligned with the full population analysis. However, no statistics related to tests for interaction between subgroups were reported and, given the relatively small sample sizes for the subgroup analyses, any finding is difficult to interpret. It was reported in the UNIFI maintenance study that for the induction-treatment subgroups (ustekinumab 6 mg/kg IV [approximately], 130 mg IV, or placebo IV), there may be a lower maintenance-treatment effect on clinical remission (particularly for the every 12 weeks regimen) for patients who had received the 130 mg IV induction treatment or the placebo IV induction treatment. The sample sizes for these analyses were relatively small and estimates were imprecise.

This superior efficacy with ustekinumab over placebo was consistent in other clinically relevant outcomes, such as clinical response, glucocorticoid-free remission, and endoscopic healing. Although the data suggested greater improvements in HRQoL, mucosal healing, and productivity with ustekinumab versus placebo, these outcomes were not included in the hierarchical analysis plan for the maintenance phase and, therefore, not adjusted for inflated type I error. There were too few events related to colectomies (three patients treated with placebo and two patients in the combined ustekinumab group) to draw conclusions on.

Based on ITCs, there is no clear superiority of ustekinumab against other common comparators with the same indication, with inconsistency in the body of evidence and risk of bias that decrease our confidence in this result. When comparing ustekinumab with other similar comparators with the same indication, it is difficult to address the relative effects of ustekinumab and its superiority, as most ORs and predictive intervals are close to the unity and imprecision was frequent.

An important consideration is that the evidence for the efficacy of ustekinumab arises from a single RCT. This is important, considering a proportion of patients treated with ustekinumab may require a second dose to address those patients considered late responders. The timing to assess the late response is a matter of clinical discussion and one that requires additional studies to investigate.

Harms

Overall, no unexpected AEs were detected in the UNIFI trial. Important concerns such as cancer and serious infections were similar between the ustekinumab and placebo groups, although there were relatively few events to form strong conclusions on. Cancers developed in seven patients who received ustekinumab (including three cases of non-melanoma skin cancer) and in one patient who received placebo. Potential opportunistic infections developed in four patients who received ustekinumab. There were no cases of anaphylaxis or serious hypersensitivity reactions in patients who received ustekinumab.

The systematic review and NMA did not address AEs due to how exposure might be related to efficacy and to the fact the placebo groups in the different studies (including UNIFI) are not "true" placebo groups. The authors provide data from previous studies for their rationale, although this was a concern for our review team, as not including AEs will provide an incomplete picture for decision-making, based on a systematic review of ITCs.

Conclusions

Based on one trial, ustekinumab is more effective than placebo for inducing and maintaining clinical remission and clinical response, maintaining a corticosteroid-free remission, and inducing and maintaining endoscopic healing in patients who have moderate-to-severe UC, despite current or previous treatment with conventional or biologic therapy.

Based on one review of ITCs, although with better odds for all outcomes when compared with placebo, ustekinumab had no clear superiority over other common comparators with the same indication, although there is still uncertainty due to inconsistency in the body of evidence and risk of bias that decreases our confidence in this result.

Although AEs were not different between ustekinumab and placebo, the number of events were low and more long-term studies are needed to assess possible harms.

Appendix 1: Literature Search Strategy

Clinical Literature Search

OVERVIEW	
Interface:	Ovid
Databases:	MEDLINE All (1946-present) Embase (1974-present) Cochrane Central Register of Controlled Trials (CCTR) **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	September 10, 2019
Alerts:	Weekly search updates until project completion
Study Types:	No study limits used
Limits:	Publication date limit: None used Conference abstracts: excluded
SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
MeSH	Medical Subject Heading
.fs	Floating subheading
exp	Explode a subject heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
#	Truncation symbol for one character
?	Truncation symbol for one or no characters only
adj#	Requires terms to be adjacent to each other within # number of words (in any order)
.ti	Title
.ab	Abstract
.hw	Heading word; usually includes subject headings and controlled vocabulary
.kf	Author keyword heading word (MEDLINE)
.kw	Author keyword (Embase);
.pt	Publication type
.mp	Mapped term
.rn	Registry number
.yr	Publication year
.jw	Journal word title
freq =#	Requires terms to occur # number of times in the specified fields
medall	Ovid database code: MEDLINE All, 1946 to present, updated daily
oemezd	Ovid database code; Embase, 1974 to present, updated daily
cctr	Ovid database code; Cochrane Central Register of Controlled Trials

MULTI-DATABASE STRATEGY	
1	ustekinumab/
2	(Stelara* or ustekinumab* or CNTO 1275 or CNTO1275 or FU77B4U5Z0 or UNIIU77B4U5Z0).ti,ab,rn,nm,kf,ot.
3	Colitis, ulcerative/
4	(colitis or colorectitis or proctocolitis).ti,ab,kf.
5	1 or 2
6	3 or 4
7	5 and 6
8	7 use medall
9	*ustekinumab/
10	(Stelara* or ustekinumab* or CNTO 1275 or CNTO1275).ti,ab,kw,dq.
11	exp ulcerative colitis/
12	(colitis or colorectitis or proctocolitis).ti,ab,kw.
13	9 or 10
14	11 or 12
15	13 and 14
16	15 use oemezd
17	16 or 8
18	conference abstract.pt.
19	conference review.pt.
20	18 or 19
21	17 NOT 20
22	remove duplicates from 21

CLINICAL TRIAL REGISTRIES		
ClinicalTrials.gov	Produced by the US National Library of Medicine. Targeted search used to capture registered clinical trials. Terms used: Stelara AND ulcerative colitis	
WHO ICTRP	International Clinical Trials Registry Platform, produced by the World Health Organization. Targeted search used to capture registered clinical trials. Terms used: Stelara AND ulcerative colitis	

OTHER DATABASES		
PubMed	Searched to capture records not found in MEDLINE. Same MeSH, keywords, limits, and study types used as per MEDLINE search, with appropriate syntax used.	

Dates for Search:	No date limits used
Keywords:	Stelara, ustekinumab, ulcerative colitis
Limits:	None used

Relevant websites from the following sections of the CADTH grey literature checklist *Grey Matters: A Practical Tool For Searching Health-Related Grey Literature* (https://www.cadth.ca/grey-matters) were searched:

- Health Technology Assessment Agencies
- Health Economics
- Clinical Practice Guidelines
- Drug and Device Regulatory Approvals
- Advisories and Warnings
- Drug Class Reviews
- Clinical Trial Registries
- Databases (free)
- Health Statistics
- Internet Search

Appendix 2: Excluded Studies

Table 25: Excluded Studies

Reference	Reason for Exclusion
Paul C, Griffiths CEM, van de Kerkhof PCM, et al. Ixekizumab provides superior efficacy compared with ustekinumab over 52 weeks of treatment: Results from IXORA-S, a phase 3 study. J Am Acad Dermatol. 2019;80(1):70-79.e73.	Different population assessed (patients with psoriasis).

Appendix 3: Detailed Outcome Data

Table 26: Efficacy Outcomes – UNIFI Study, Efficacy Population (US Definitions)

Induction phase	Ustekinumab IV			Placebo
	130 mg N = 320	~6 mg/kg N = 322	Combined N = 642	N = 319
Clinical remission (US definition) at week 8[a]				
Number of patients in clinical remission, n (%)	53 (16.6)	61 (18.9)	114 (17.8)	20 (6.3)
Risk difference against placebo, (95% CI);[b] P value[c]	10.3 (4.8 to 15.8); < 0.001	12.7 (7.0 to 18.4); < 0.001	11.5 (7.0 to 16.0); < 0.001	–
Maintenance phase	**Ustekinumab SC**			**Placebo**
	90 mg q.12.w. N = 172	**90 mg q.8.w. N = 176**	**Combined N = 348**	**N = 175**
Clinical remission (US definition) at week 44[a]				
Number of patients in clinical remission, n (%)	68 (39.5)	75 (42.6)	143 (41.1)	43 (24.6)
Percentage difference against placebo, (95% CI);[b] P value[c]	15.1 (6.0 to 24.2); 0.002	17.9 (8.6 to 27.2); < 0.001	16.5 (8.7 to 24.3); < 0.001	–
Corticosteroid-free clinical remission (US definition) at week 44[a]				
Number of patients in clinical remission, n (%)	67 (39.0)	74 (40.9)	139 (39.9)	42 (24.0)
Percentage difference against placebo, (95% CI);[b] P value[c]	15.1 (6.1 to 24.2); 0.002	16.8 (7.6 to 26.0); < 0.001	15.9 (8.2 to 23.7); < 0.001	–

CI = confidence interval; q.8.w. = every eight weeks; q.12.w. = every 12 weeks; SC = subcutaneous.

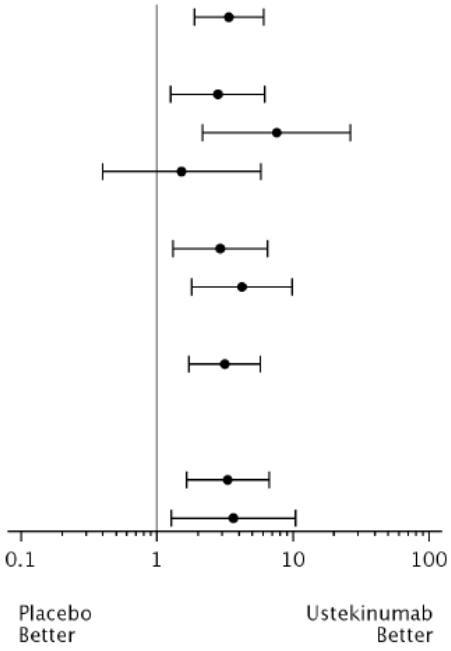
[a] An absolute stool number ≤ 3, a Mayo rectal bleeding subscore of 0, and a Mayo endoscopy subscore of 0 or 1.

[b] The CIs were based on the Wald statistic with Mantel-Haenszel weight.

[c] The P values were based on the Cochran-Mantel-Haenszel test.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 15: Clinical Remission (Global Definition) at Week 8 in the Ustekinumab 130 mg Group Versus Placebo Group for Extent of Disease and Disease Severity Subgroups (Induction Primary Efficacy Analysis Set)

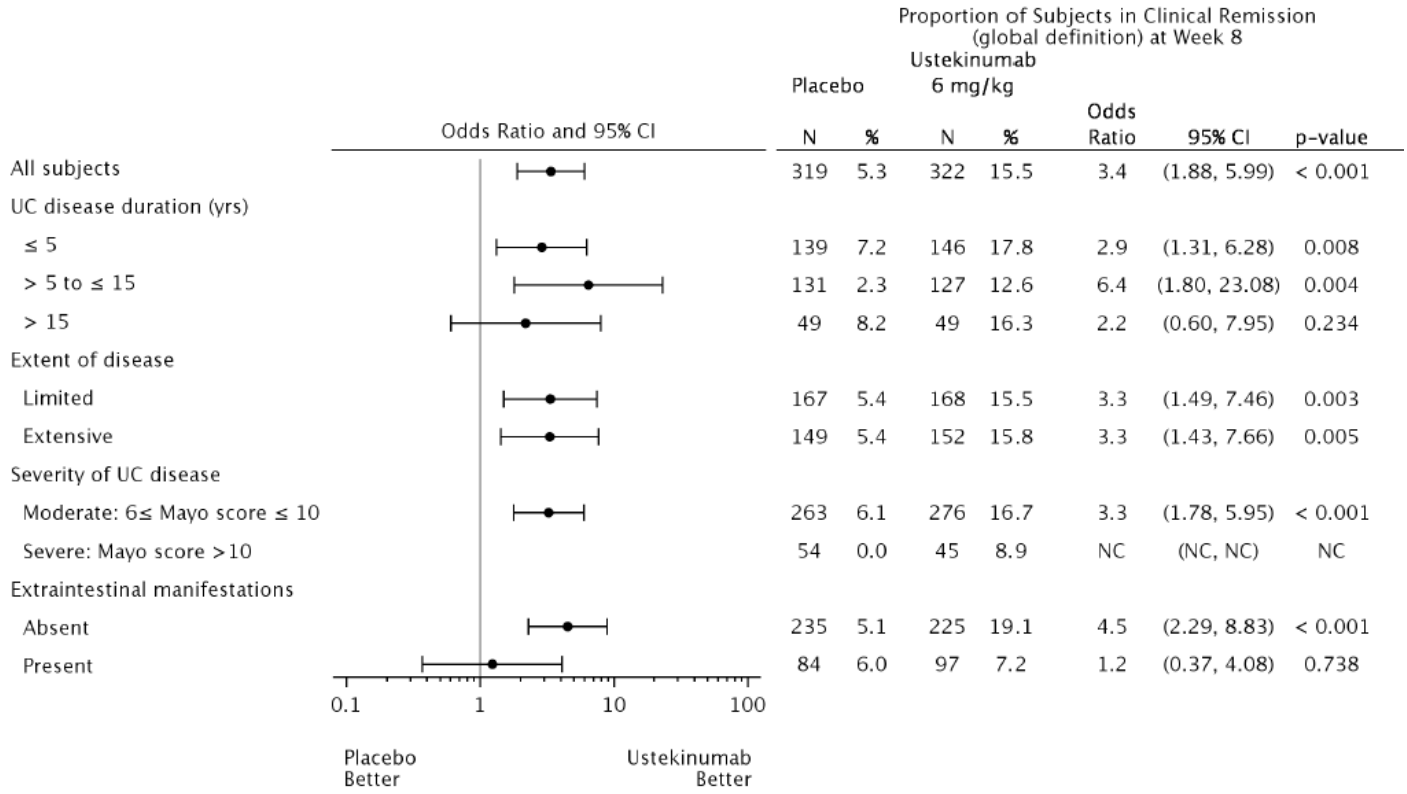


	Odds Ratio and 95% CI	Proportion of Subjects in Clinical Remission (global definition) at Week 8						
		Placebo		Ustekinumab 130 mg				
		N	%	N	%	Odds Ratio	95% CI	p-value
All subjects		319	5.3	320	15.6	3.4	(1.89, 6.04)	< 0.001
UC disease duration (yrs)								
≤ 5		139	7.2	137	17.5	2.8	(1.26, 6.17)	0.011
> 5 to ≤ 15		131	2.3	131	14.5	7.5	(2.15, 26.48)	0.002
> 15		49	8.2	52	13.5	1.5	(0.40, 5.76)	0.541
Extent of disease								
Limited		167	5.4	183	13.7	2.9	(1.31, 6.50)	0.009
Extensive		149	5.4	135	18.5	4.2	(1.79, 9.84)	< 0.001
Severity of UC disease								
Moderate: 6≤ Mayo score ≤ 10		263	6.1	271	16.6	3.1	(1.71, 5.75)	< 0.001
Severe: Mayo score >10		54	0.0	48	8.3	NC	(NC, NC)	NC
Extraintestinal manifestations								
Absent		235	5.1	230	14.3	3.3	(1.64, 6.67)	< 0.001
Present		84	6.0	90	18.9	3.6	(1.27, 10.44)	0.016

CI = confidence interval; NC = no change; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

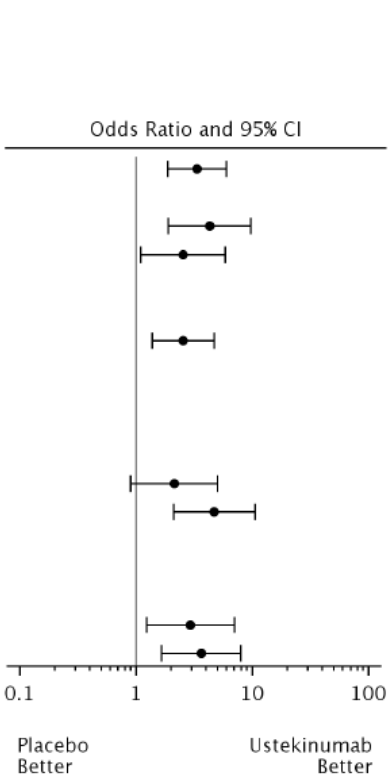
Figure 16: Clinical Remission (Global Definition) at Week 8 in the Group Receiving Approximately 6 mg/kg of Ustekinumab Versus the Placebo Group for Extent of Disease and Disease Severity Subgroups (Induction Primary Efficacy Analysis Set)



CI = confidence interval; NC = no change; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 17: Clinical Remission (Global Definition) at Week 8 in the Group Receiving Approximately 6 mg/kg of Ustekinumab Versus the Placebo Group for History of Conventional Therapy for UC Subgroups (Induction Primary Efficacy Analysis Set)

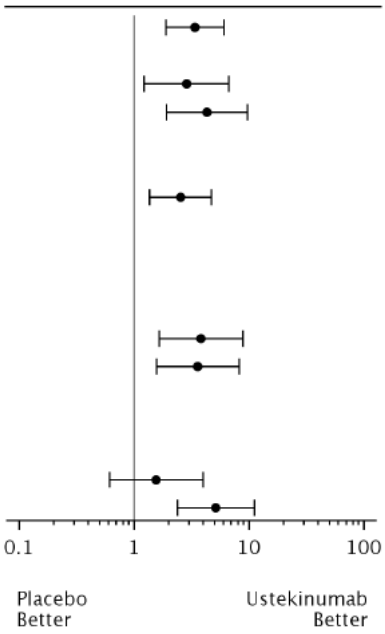


	Proportion of Subjects in Clinical Remission (global definition) at Week 8						
	Placebo		Ustekinumab 6 mg/kg				
	N	%	N	%	Odds Ratio	95% CI	p-value
All subjects	319	5.3	322	15.5	3.4	(1.88, 5.99)	< 0.001
Refractory or intolerant to 6-MP/AZA							
Yes	173	4.6	175	17.1	4.3	(1.89, 9.73)	< 0.001
No	146	6.2	147	13.6	2.5	(1.10, 5.83)	0.029
Refractory, dependent or intolerant to oral or IV corticosteroids							
Yes	267	6.4	259	14.3	2.5	(1.38, 4.68)	0.003
No	52	0.0	63	20.6	NC	(NC, NC)	NC
Refractory, dependent or intolerant to oral or IV corticosteroids, but not refractory or intolerant to 6-MP/AZA							
Yes	126	7.1	129	13.2	2.1	(0.90, 5.03)	0.087
No	193	4.1	193	17.1	4.7	(2.10, 10.57)	< 0.001
Refractory, dependent or intolerant to oral or IV corticosteroids, AND refractory or intolerant to 6-MP/AZA							
Yes	141	5.7	130	15.4	2.9	(1.23, 7.07)	0.016
No	178	5.1	192	15.6	3.6	(1.66, 7.95)	0.001

6-MP = 6-mercaptopurine; AZA = azathioprine; CI = confidence interval; NC = no change; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 18: Clinical Remission (Global Definition) at Week 8 in the Ustekinumab 130 mg Group Versus Placebo Group for History of Conventional Therapy for UC Subgroups (Induction Primary Efficacy Analysis Set)

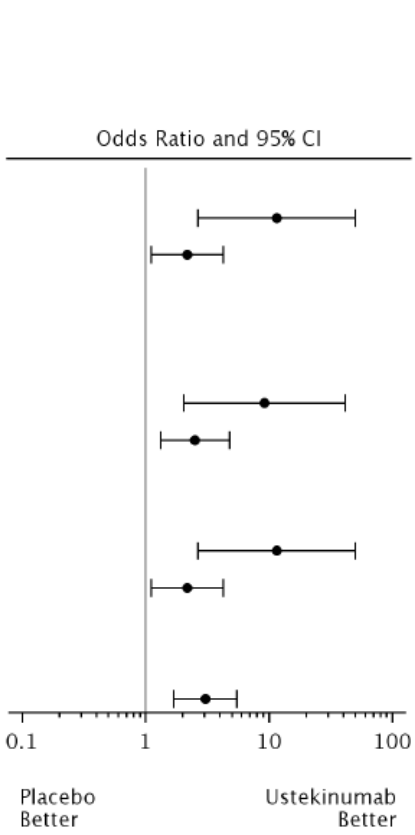


	Odds Ratio and 95% CI	Proportion of Subjects in Clinical Remission (global definition) at Week 8						
		Placebo		Ustekinumab 130 mg				
		N	%	N	%	Odds Ratio	95% CI	p-value
All subjects		319	5.3	320	15.6	3.4	(1.89, 6.04)	< 0.001
Refractory or intolerant to 6-MP/AZA								
Yes		173	4.6	182	12.6	2.9	(1.22, 6.68)	0.016
No		146	6.2	138	19.6	4.3	(1.91, 9.72)	< 0.001
Refractory, dependent or intolerant to oral or IV corticosteroids								
Yes		267	6.4	250	14.4	2.5	(1.37, 4.69)	0.003
No		52	0.0	69	20.3	NC	(NC, NC)	NC
Refractory, dependent or intolerant to oral or IV corticosteroids, but not refractory or intolerant to 6-MP/AZA								
Yes		126	7.1	118	19.5	3.8	(1.65, 8.84)	0.002
No		193	4.1	201	13.4	3.6	(1.56, 8.17)	0.003
Refractory, dependent or intolerant to oral or IV corticosteroids, AND refractory or intolerant to 6-MP/AZA								
Yes		141	5.7	132	9.8	1.6	(0.61, 3.96)	0.358
No		178	5.1	187	19.8	5.1	(2.37, 11.10)	< 0.001

6-MP = 6-mercaptopurine; AZA = azathioprine; CI = confidence interval; NC = no change; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 19: Clinical Remission (Global Definition) at Week 8 in the Group Receiving Approximately 6 mg/kg of Ustekinumab Versus the Placebo Group for History of Biologics for UC Subgroups (Induction Primary Efficacy Analysis Set)

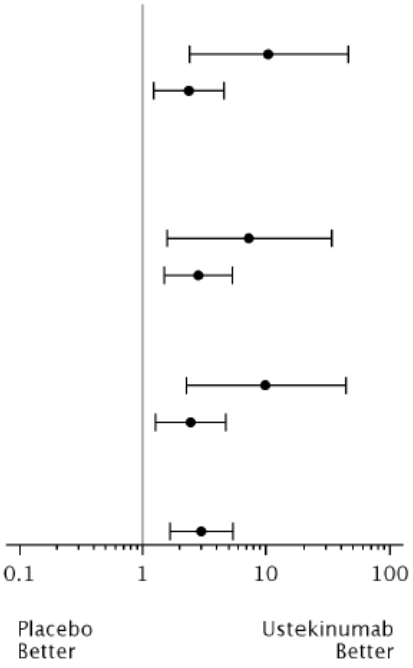


	Proportion of Subjects in Clinical Remission (global definition) at Week 8						
	Placebo		Ustekinumab 6 mg/kg				
	N	%	N	%	Odds Ratio	95% CI	p-value
Biologic failure status							
Yes	161	1.2	166	12.7	11.5	(2.66, 50.05)	0.001
No	158	9.5	156	18.6	2.2	(1.11, 4.25)	0.024
Subjects with biologic failures to							
Only anti-TNF (NOT to vedolizumab)							
Yes	112	1.8	106	14.2	9.2	(2.04, 41.24)	0.004
No	207	7.2	216	16.2	2.5	(1.32, 4.81)	0.005
At least one anti-TNF (regardless of vedolizumab)							
Yes	159	1.3	164	12.8	11.6	(2.66, 50.16)	0.001
No	160	9.4	158	18.4	2.2	(1.11, 4.26)	0.023
Any anti-TNF and vedolizumab							
Yes	47	0.0	58	10.3	NC	(NC, NC)	NC
No	272	6.3	264	16.7	3.0	(1.69, 5.50)	< 0.001

CI = confidence interval; NC = no change; TNF = tumour necrosis factor; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 20: Clinical Remission (Global Definition) at Week 8 in the Ustekinumab 130 mg Group Versus Placebo Group for History of Biologics for UC Subgroups (Induction Primary Efficacy Analysis Set)

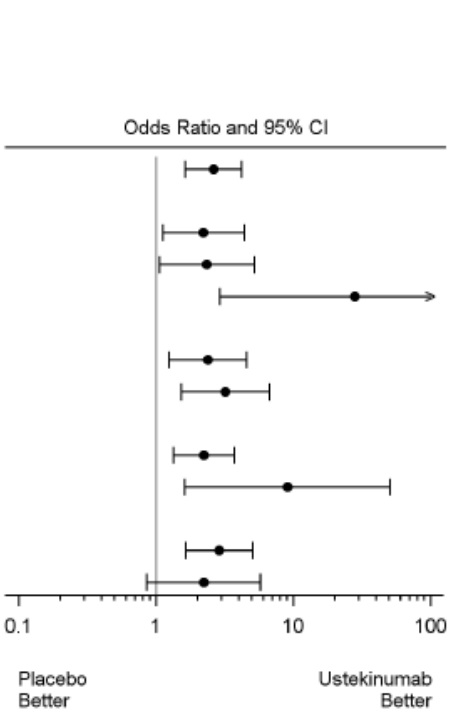


	Proportion of Subjects in Clinical Remission (global definition) at Week 8						
	Placebo		Ustekinumab 130 mg				
	N	%	N	%	Odds Ratio	95% CI	p-value
Biologic failure status							
Yes	161	1.2	164	11.6	10.5	(2.40, 45.88)	0.002
No	158	9.5	156	19.9	2.4	(1.22, 4.60)	0.011
Subjects with biologic failures to							
Only anti-TNF (NOT to vedolizumab)							
Yes	112	1.8	107	11.2	7.3	(1.58, 33.56)	0.011
No	207	7.2	213	17.8	2.8	(1.50, 5.36)	0.001
At least one anti-TNF (regardless of vedolizumab)							
Yes	159	1.3	162	11.1	10.0	(2.27, 43.77)	0.002
No	160	9.4	158	20.3	2.5	(1.27, 4.76)	0.007
Any anti-TNF and vedolizumab							
Yes	47	0.0	55	10.9	NC	(NC, NC)	NC
No	272	6.3	265	16.6	3.0	(1.66, 5.41)	< 0.001

CI = confidence interval; NC = no change; TNF = tumour necrosis factor; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 21: Clinical Remission (Global Definition) at Week 44 in the Ustekinumab 90 mg SC Every Eight Weeks Group Versus Placebo Group for Extent of Disease and Disease Severity Subgroups (Maintenance Primary Efficacy Analysis Set; From Induction Baseline)

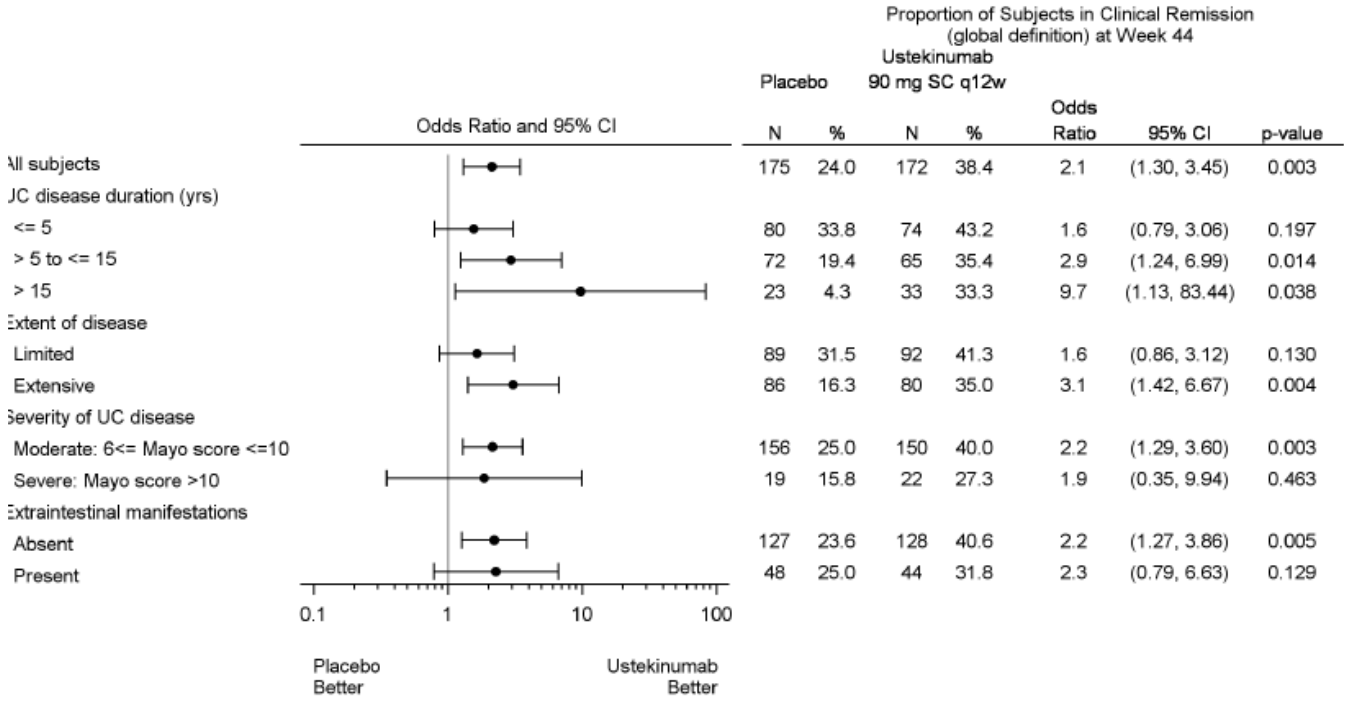


	Proportion of Subjects in Clinical Remission (global definition) at Week 44						
	Placebo		Ustekinumab 90 mg SC q8w				
	N	%	N	%	Odds Ratio	95% CI	p-value
All subjects	175	24.0	176	43.8	2.6	(1.63, 4.21)	< 0.001
UC disease duration (yrs)							
<= 5	80	33.8	70	50.0	2.2	(1.12, 4.39)	0.023
> 5 to <= 15	72	19.4	81	37.0	2.3	(1.05, 5.23)	0.037
> 15	23	4.3	25	48.0	28.1	(2.92, 269.73)	0.004
Extent of disease							
Limited	89	31.5	95	46.3	2.4	(1.25, 4.58)	0.009
Extensive	86	16.3	80	41.3	3.2	(1.52, 6.72)	0.002
Severity of UC disease							
Moderate: 6<= Mayo score <=10	156	25.0	147	41.5	2.2	(1.34, 3.72)	0.002
Severe: Mayo score >10	19	15.8	27	55.6	9.1	(1.61, 50.75)	0.012
Extraintestinal manifestations							
Absent	127	23.6	130	44.6	2.9	(1.65, 5.06)	< 0.001
Present	48	25.0	46	41.3	2.2	(0.86, 5.76)	0.099

CI = confidence interval; q8w = every eight weeks; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

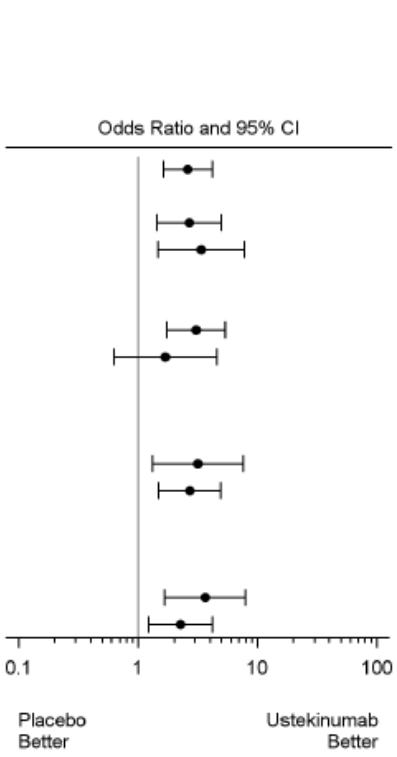
Figure 22: Clinical Remission (Global Definition) at Week 44 in the Ustekinumab 90 mg SC Every 12 Weeks Group Versus the Placebo Group for Extent of Disease and Disease Severity Subgroups (Maintenance Primary Efficacy Analysis Set; From Induction Baseline)



CI = confidence interval; q12w = every 12 weeks; SC = subcutaneous; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 23: Clinical Remission (Global Definition) at Week 8 in the Ustekinumab 90 mg SC Every Eight Weeks Group Versus Placebo Group for History of Conventional Therapy for UC Subgroups (Maintenance Primary Efficacy Analysis Set)

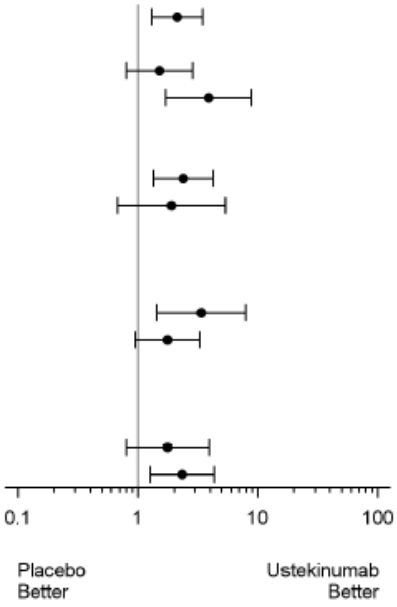


	Proportion of Subjects in Clinical Remission (global definition) at Week 44						
	Placebo		Ustekinumab 90 mg SC q8w				
	N	%	N	%	Odds Ratio	95% CI	p-value
All subjects	175	24.0	176	43.8	2.6	(1.63, 4.21)	< 0.001
Refractory or intolerant to 6-MP/AZA							
Yes	100	31.0	94	52.1	2.7	(1.44, 4.97)	0.002
No	75	14.7	82	34.1	3.4	(1.46, 7.74)	0.004
Refractory, dependent or intolerant to oral or IV corticosteroids							
Yes	133	19.5	135	42.2	3.1	(1.75, 5.37)	< 0.001
No	42	38.1	41	48.8	1.7	(0.62, 4.57)	0.301
Refractory, dependent or intolerant to oral or IV corticosteroids, but not refractory or intolerant to 6-MP/AZA							
Yes	62	16.1	69	36.2	3.2	(1.31, 7.59)	0.010
No	113	28.3	107	48.6	2.7	(1.49, 4.92)	0.001
Refractory, dependent or intolerant to oral or IV corticosteroids, AND refractory or intolerant to 6-MP/AZA							
Yes	71	22.5	66	48.5	3.6	(1.68, 7.91)	0.001
No	104	25.0	110	40.9	2.3	(1.22, 4.19)	0.009

6-MP = 6-mercaptopurine; AZA = azathioprine; CI = confidence interval; q8w = every eight weeks; SC = subcutaneous; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 24: Clinical Remission (Global Definition) at Week 8 in the Ustekinumab 90 mg SC Every 12 Weeks Group Versus the Placebo Group for History of Conventional Therapy for UC Subgroups (Maintenance Primary Efficacy Analysis Set)

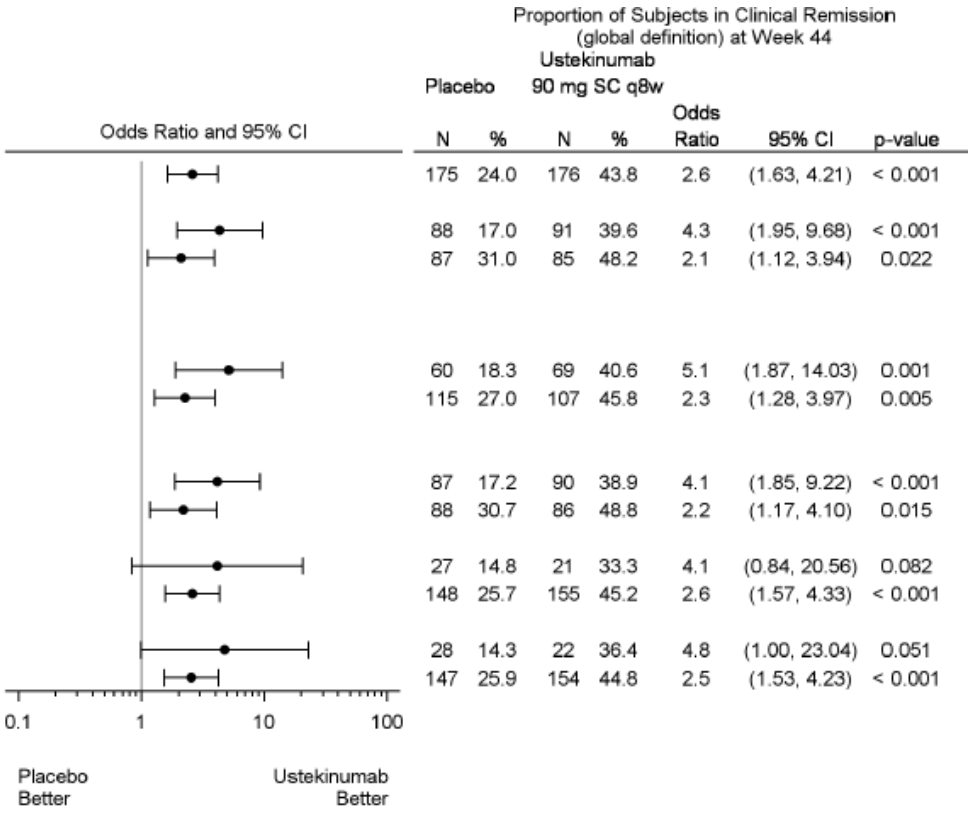


	Proportion of Subjects in Clinical Remission (global definition) at Week 44						
	Placebo		Ustekinumab 90 mg SC q12w				
	N	%	N	%	Odds Ratio	95% CI	p-value
All subjects	175	24.0	172	38.4	2.1	(1.30, 3.45)	0.003
Refractory or intolerant to 6-MP/AZA							
Yes	100	31.0	90	38.9	1.5	(0.80, 2.86)	0.205
No	75	14.7	82	37.8	3.9	(1.71, 8.83)	0.001
Refractory, dependent or intolerant to oral or IV corticosteroids							
Yes	133	19.5	137	35.8	2.4	(1.34, 4.22)	0.003
No	42	38.1	35	48.6	1.9	(0.68, 5.39)	0.222
Refractory, dependent or intolerant to oral or IV corticosteroids, but not refractory or intolerant to 6-MP/AZA							
Yes	62	16.1	74	37.8	3.4	(1.43, 7.96)	0.006
No	113	28.3	98	38.8	1.8	(0.95, 3.28)	0.070
Refractory, dependent or intolerant to oral or IV corticosteroids, AND refractory or intolerant to 6-MP/AZA							
Yes	71	22.5	63	33.3	1.8	(0.80, 3.94)	0.161
No	104	25.0	109	41.3	2.3	(1.26, 4.35)	0.007

6-MP = 6-mercaptopurine; AZA = azathioprine; CI = confidence interval; q12w = every 12 weeks; SC = subcutaneous; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

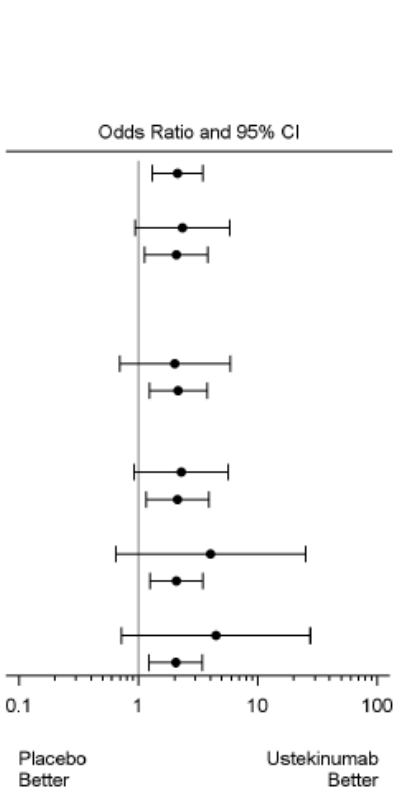
Figure 25: Clinical Remission (Global Definition) at Week 8 in the Ustekinumab 90 mg SC Every Eight Weeks Group Versus Placebo Group for History of Biologics for UC Subgroups (Maintenance Primary Efficacy Analysis Set)



CI = confidence interval; q8w = every eight weeks; TNF = tumour necrosis factor; SC = subcutaneous; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 26: Clinical Remission (Global Definition) at Week 8 in the Ustekinumab 90 mg SC Every 12 Weeks Group Versus the Placebo Group for History of Biologics for UC Subgroups (Maintenance Primary Efficacy Analysis Set)

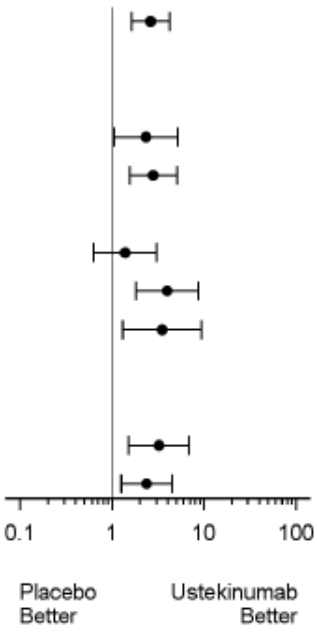


	Proportion of Subjects in Clinical Remission (global definition) at Week 44						
	Placebo		Ustekinumab 90 mg SC q12w				
	N	%	N	%	Odds Ratio	95% CI	p-value
All subjects	175	24.0	172	38.4	2.1	(1.30, 3.45)	0.003
Biologic failure status							
Yes	88	17.0	70	22.9	2.3	(0.94, 5.78)	0.066
No	87	31.0	102	49.0	2.1	(1.12, 3.82)	0.020
Subjects with biologic failures to							
Only anti-TNF (NOT to vedolizumab)							
Yes	60	18.3	48	22.9	2.0	(0.69, 5.85)	0.198
No	115	27.0	124	44.4	2.2	(1.23, 3.77)	0.007
At least one anti-TNF (regardless of vedolizumab)							
Yes	87	17.2	70	22.9	2.3	(0.92, 5.64)	0.075
No	88	30.7	102	49.0	2.1	(1.15, 3.90)	0.016
Any anti-TNF and vedolizumab							
Yes	27	14.8	22	22.7	4.0	(0.64, 25.15)	0.136
No	148	25.7	150	40.7	2.1	(1.24, 3.46)	0.005
Vedolizumab							
Yes	28	14.3	22	22.7	4.5	(0.72, 27.64)	0.109
No	147	25.9	150	40.7	2.0	(1.22, 3.41)	0.006

CI = confidence interval; q12w = every 12 weeks; TNF = tumour necrosis factor; SC = subcutaneous; UC = ulcerative colitis.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 27: Clinical Remission (Global Definition) at Week 8 in the Ustekinumab 90 mg SC Every Eight Weeks Group Versus Placebo Group for Maintenance Baseline Stratification Variables (Maintenance Primary Efficacy Analysis Set)

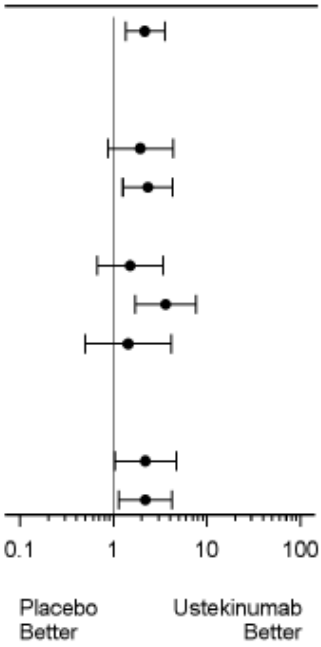


	Proportion of Subjects in Clinical Remission (global definition) at Week 44						
	Placebo		Ustekinumab 90 mg SC q8w				
	N	%	N	%	Odds Ratio	95% CI	p-value
All subjects	175	24.0	176	43.8	2.6	(1.63, 4.21)	< 0.001
Clinical remission status at maintenance baseline as determined by the IWRS							
Yes	50	42.0	51	62.7	2.3	(1.05, 5.18)	0.038
No	125	16.8	125	36.0	2.8	(1.54, 5.07)	< 0.001
Induction treatment							
Ustekinumab 130 mg IV	58	32.8	58	39.7	1.4	(0.63, 3.01)	0.427
Ustekinumab 6 mg/kg IV	69	20.3	70	48.6	4.0	(1.82, 8.68)	< 0.001
Placebo IV to Ustekinumab 6 mg/kg IV	48	18.8	48	41.7	3.5	(1.30, 9.34)	0.013
Oral corticosteroid use at maintenance baseline as recorded in the IWRS							
Yes	84	21.4	84	41.7	3.2	(1.51, 6.76)	0.002
No	91	26.4	92	45.7	2.4	(1.26, 4.46)	0.007

CI = confidence interval; IWRS = interactive web response system; q8w = every eight weeks; SC = subcutaneous.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Figure 28: Clinical Remission (Global Definition) at Week 8 in the Ustekinumab 90 mg SC Every 12 Weeks Group Versus the Placebo Group for Maintenance Baseline Stratification Variables (Maintenance Primary Efficacy Analysis Set)



	Proportion of Subjects in Clinical Remission (US definition) at Week 44						
	Placebo		Ustekinumab 90 mg SC q12w				
	N	%	N	%	Odds Ratio	95% CI	p-value
All subjects	175	24.6	172	39.5	2.2	(1.34, 3.54)	0.002
Clinical remission status at maintenance baseline as determined by the IWRS							
Yes	50	44.0	48	60.4	1.9	(0.87, 4.34)	0.107
No	125	16.8	124	31.5	2.3	(1.26, 4.31)	0.007
Induction treatment							
Ustekinumab 130 mg IV	58	29.3	56	37.5	1.5	(0.67, 3.39)	0.328
Ustekinumab 6 mg/kg IV	69	24.6	69	52.2	3.6	(1.69, 7.59)	< 0.001
Placebo IV to Ustekinumab 6 mg/kg IV	48	18.8	47	23.4	1.4	(0.50, 4.13)	0.507
Oral corticosteroid use at maintenance baseline as recorded in the IWRS							
Yes	84	22.6	81	35.8	2.2	(1.03, 4.67)	0.041
No	91	26.4	91	42.9	2.2	(1.15, 4.19)	0.017

CI = confidence interval; IWRS = interactive web response system; q12w = every 12 weeks; SC = subcutaneous.

Source: Clinical Study Reports for the UNIFI induction[1] and maintenance[2] studies.

Appendix 4: Description and Appraisal of Outcome Measures

Aim

To describe the following outcome measures and review their measurement properties (validity, reliability, responsiveness to change, and minimal clinically important difference) (MCID):

- Mayo scoring system
- IBDQ
- SF-36
- EQ-5D
- WPAI-GH

The WPAI-GH was measured in the UNIFI induction and UNIFI maintenance studies with the objective of informing the pharmacoeconomic model. The Ulcerative Colitis Endoscopic Index of Severity was also measured in the UNIFI studies as an exploratory outcome and, therefore, is not reviewed in this section.

Table 27: Outcome Measures Included in Each Study

Outcome measure	UNIFI induction	UNIFI maintenance
Mayo score	Primary	Primary
IBDQ	Secondary	Secondary
SF-36	Secondary	Secondary
EQ-5D	Secondary	Secondary
WPAI-GH	Other	Other

EQ-5D-5L = EuroQol 5-Dimensions 5-Levels; IBDQ = Inflammatory Bowel Disease Questionnaire; SF-36 = Short Form (36) Health Survey; WPAI-GH = Work Productivity and Activity Impairment Questionnaire – General Health.

Findings

The validity, reliability, responsiveness, and MCID of each outcome measure were summarized and evaluated. Interpretation of the reliability and validity metrics were based on the following criteria:

- Inter-rater reliability, kappa statistics (level of agreement):[41]
  - less than 0 = poor agreement
  - 0.00 to 0.21 = slight agreement
  - 0.21 to 0.40 = fair agreement
  - 0.41 to 0.60 = moderate agreement
  - 0.61 to 0.8 = substantia
  - 0.81 to 1.00 = almost perfect agreement
- Internal consistency (Cronbach's alpha) and test-retest reliability (≥ 0.7 is considered acceptable)[42]
- Validity, i.e., between-scale comparison (correlation coefficient, r):[43]
  - 0.3 or less = weak

- 0.3 up to 0.5 = moderate
- more than 0.5 = strong

Table 28: Summary of Outcome Measures and Their Measurement Properties

Outcome measure	Type	Conclusions about measurement properties	MID
Mayo score	Disease-specific physician-measured score with parts: rectal bleeding, stool frequency, PGA, and endoscopy findings.	**Validity:** There was limited evidence on the validity for the total Mayo score. Construct validity of the Mayo endoscopic subscore was found to be strongly correlated with the total Mayo score (Spearman's rho ≥ 0.97), as well as two histologic indices (Pearson's r ≥ 0.55).[44] **Reliability and responsiveness:** The endoscopic subscore was found to have moderate-to-substantial agreement in the inter-rater reliability estimates, as well as responsiveness of the subscore to change over time with treatment.[44-47]	**Clinical response:** Reduction in total Mayo score of ≥ 3 points. **Clinical remission:** Total Mayo score of ≤ 2 points, with or without an individual subscore of > 1.[47]
IBDQ	Disease-specific Likert-based questionnaire consisting of 32 items classified into four dimensions: bowel symptoms, systemic symptoms, emotional function, and social function. The IBDQ can be administered by an interviewer or self-administered.	**Validity:** There was limited evidence on the validity of the IBDQ in the UC population. **Reliability and responsiveness:** The IBDQ was shown to be highly reliable through evaluation of internal consistency (Cronbach's alpha 0.7) and test-retest assessments (ICC 0.9 to 0.99 or r ≥ 0.8). The IBDQ was also shown to be responsive to change in IBD patients.[48,49]	Absolute score change of ≥ 30 points, or a score of ≥ 15 points above the placebo score among IBD patients.[50]
SF-36	Generic self-reported questionnaire consisting of eight domains: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health.	**Validity:** Construct validity was demonstrated through strong moderate-to-strong correlations (r > 0.4) between the eight subscales of the SF-36 and corresponding domains of five patient-reported clinical constructs. The scale showed evidence of discriminative validity.[51] **Reliability and responsiveness:** The SF-36 was found to be reliable through internal consistency for all eight subscales (Cronbach's alpha > 0.7) and test-retest assessments for six of the eight subscales (ICC > 0.7).[51] The scale and its subscores were found to be responsive to treatment-related changes.[51]	≥ 3 to ≥ 5 points in PCS, MCS, and individual subscore.[50]
EQ-5D	Generic preference-based HRQoL instrument consisting of a VAS and a composite index score of five dimensions: mobility, self-care, usual activities,	**Validity:** Stark et al. assessed the validity, reliability, and responsiveness of EQ-5D in a German population of IBD patients (including UC). Construct validity was supported by strong correlation of the scores with the CAI (Spearman rank correlation, between 0.65 and 0.67). The CAI score and VAS as well as all but one domain of the scale (self-care domain) showed discriminative	Not found in UC patients. Among IBD patients: VAS of 10.9 and index score of 0.05 for improved health; VAS of −14.4 and index score of −0.067 for deteriorated health.[53]

Outcome measure	Type	Conclusions about measurement properties	MID
	pain/discomfort, and anxiety/depression.	validity. Konig et al. also demonstrated strong correlation between the EQ-5D VAS and index scores and with the IBDQ total score (0.70 and 0.62, respectively), and a moderate-to-strong correlation with the SF-36 subscores (0.37 to 0.72).[52] **Reliability and responsiveness:** Test-retest reliability was generally high for the index score (0.67 ≤ ICC ≤ 0.73), VAS (ICC = 0.93), and all five items of the scale (0.67 ≤ kappa ≤ 1.00). Konig et al. reported similar results (ICC of 0.89 for the index score, and 0.77 for the VAS score).[52] Both the index score and VAS were shown to be responsive to detecting change in health status.[53]	
WPAI-GH	Self-rated disease-specific questionnaire consisting of six items divided into four domains: absenteeism, presenteeism, percentage of overall work impairment, and regular activities impairment.	**Validity:** Convergent validity was demonstrated for all WPAI domains between the SIBDQ bowel symptoms (Spearman rank-order coefficient of −0.47 to −0.68) and SF-12v2 bodily pain (−0.52 to −0.55) subscores, as well as between the WPAI and measures of disease activity (median 0.45).[54] Known-group validity data demonstrated that patients with worse health outcomes scored worse on the WPAI than patients with better health outcomes, based on partial Mayo, SCCAI, UC-DAI, and FACIT-Fatigue disease severity measures.[54] **Reliability and responsiveness:** Test-retest assessment demonstrated that differences in each domain were < 5% over a 12-month period; however, no ICC was reported for these data.[54] One study demonstrated that patients with active UC disease who achieved remission at week 8 reported a 25% to 30% decrease in presenteeism, OWI, and activity impairment, and a 9% decrease in absenteeism. Responsiveness of the WPAI domains to effective treatment was demonstrated with an approximate 20% decrease in presenteeism, OWI, and activity impairment, and an 8% decrease in absenteeism.[54]	Not found in UC patients; however, a 7-point change has been estimated in Crohn disease.[55]

CAI = Clinical Activity Index; EQ-5D = EuroQol 5-Dimensions; EQ-5D-5L = EuroQol 5-Dimensions 5-Levels; HRQoL = health-related quality of life; FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy – Fatigue scale; IBD = inflammatory bowel disease; IBDQ = Inflammatory Bowel Disease Questionnaire; ICC = intraclass correlation; MID = minimal important difference; SF-12v2 = Short Form (12) Health Survey, version 2; SF-36 = Short Form (36) Health Survey; OWI = overall work impairment; PGA = Physician's Global Assessment; SCCAI = Simple Clinical Colitis Activity Index; SIBDQ = Short Inflammatory Bowel Disease Questionnaire; VAS = Visual Analogue Scale; WPAI-GH = Work Productivity and Activity Impairment Questionnaire – General Health; UC = ulcerative colitis; UC-DAI = UC Disease Activity Index.

Mayo Score

The Mayo scoring system is a combined endoscopic and clinical scale used to assess the severity of UC. It was first developed by Dr. Schroeder in 1987 and is now one of the most commonly used disease activity indices in UC.[47,56] In its complete form, the Mayo score is composed of four components: rectal bleeding, stool frequency, Physician's Global Assessment (PGA), and endoscopy findings. Each part is rated from 0 to 3, yielding a total

score of 0 to 12. A score of 3 to 5 points indicates mildly active disease, while a score of 6 to 10 points indicates moderately active disease, and a score of 11 to 12 points indicates severe disease. Two abridged versions have been developed and validated: the partial Mayo score that excludes the endoscopy subscore, and the non-invasive six-point score comprising only the bleeding and stool frequency subscores.[44] Mucosal healing has been defined as a Mayo endoscopic subscore of 0 or 1 in major trials of biological therapies in UC. The grading of each component is defined in Table 29.

Table 29: Components and Grading of the Mayo Score in Ulcerative Colitis

Component	Grading
Stool frequency	0 = Normal 1 = 1 to 2 stools per day more than normal 2 = 3 to 4 stools per day more than normal 3 = More than 4 stools per day more than normal
Rectal bleeding	0 = None 1 = Visible blood with stool less than half the time 2 = Visible blood with stool half of the time or more 3 = Passing blood alone
Mucosal appearance at endoscopy[a]	0 = Normal or inactive disease 1 = Mild disease (erythema, decreased vascular pattern, mild friability) 2 = Moderate disease (marked erythema, absent vascular pattern, friability, erosions) 3 = Severe disease (spontaneous bleeding, ulceration)
Physician rating of disease activity	0 = Normal 1 = Mild 2 = Moderate 3 = Severe

[a] The mucosal appearance at endoscopy score is not included in the partial Mayo score.

Validity

A recent Cochrane systematic review by Mohammed et al. assessed the validity, reliability, and responsiveness of endoscopic-scoring incidences for evaluation of disease activity in UC, which included six of 20 studies evaluating the Mayo score.[44] None of the included studies assessed content validity.[44] Construct validity of the Mayo endoscopic subscore was reported in two studies with UC patients, and a strong correlation was found between the endoscopic subscore and two histologic indices (the Riley score and Rubin histologic score, $r \geq 0.55$ for both). However, the endoscopic subscore was shown to fail in discriminating between patients who achieved remission and response compared with those who did not.[44] Dhanda et al. also demonstrated a strong correlation between the partial and total Mayo scores (rho ≥ 0.97 at weeks 4 and 8).[57]

Reliability and Responsiveness

The endoscopic subscore was evaluated for reliability and responsiveness in a placebo-controlled trial designed to assess change in UC disease activity with mesalamine treatment.[45] The authors reported excellent inter- and intra-observer reliability (intraclass correlation [ICC] 0.79 and 0.89, respectively) as well as responsiveness of the subscore to change over time with treatment.[45] Mohammed et al. reported a moderate-to-substantial agreement in the inter-rater reliability estimates (range 0.45 to 0.75) and a substantial agreement in the intra-rater reliability estimates (0.75) for the endoscopic subscore.[44] Another study by Walsh et al. evaluated the comparative inter-rater variation for three UC

disease-activity indices, including the Mayo score.[46] The inter-rater agreement for the total Mayo score was high (kappa = 0.72); however, the agreement was lower for the relatively subjective PGA and endoscopic subscores (kappa = 0.56 and 0.38, respectively). The Mayo score has been demonstrated to correlate with patient assessment of change in UC activity,[47] as well as to correlate with improvement in quality-of-life measures.[58]

Minimal Clinically Important Difference

Lewis et al. reported that a reduction of at least 3.5 points in the total Mayo score reflected an optimum cut point for clinical improvement or response (based on sensitivity, specificity, and area under the curve [AUC]) in UC, using patient's rating of the improvement as an anchor.[47] The optimum cut point for clinical remission varies; Lewis et al. reported a cut point of 4.5 (based on sensitivity, specificity, and AUC), although other cut points ranging from a Mayo score of 0.6 to 2 or less were reported in clinical trials.[47] The FDA defines clinical remission in relation to the Mayo score as a total score of 2 or less with no individual subscore great than 1, a rectal bleeding subscore of 0, a stool frequency subscore of 0 (a ≤ 1 point decrease in the stool frequency subscore from baseline and achieving a score of 1 is considered), and a Mayo endoscopy subscore of 0 or 1. Clinical response is defined as a reduction in total Mayo score of 30% or more and a decrease of 3 points or more from baseline, with a rectal bleeding subscore of less than 1.25.

Limitations

Although the Mayo score is a widely recognized UC activity index and is accepted by Canadian and American regulatory bodies, it may not be optimal. Cooney et al. argued that two components of the Mayo score — the PGA and the endoscopy subscore — are subjective and introduce variability and lack of precision into the index. The PGA also includes a sigmoidoscopy score, which introduces double counts of some elements.[59] Additionally, a single general item in the PGA is not sensitive enough to adequately capture benefits in all or some of the important signs and symptoms. As a result, the FDA does not recommend the PGA subscore or the full Mayo score as end point measures to support a marketing decision; however, it does recommend the endoscopy, stool frequency, and rectal bleeding subscores as end point measures for clinical trials until the availability of well-defined and reliable end points.[60]

Inflammatory Bowel Disease Questionnaire

Developed by Guyatt et al., the IBDQ is an interviewer- or self-administered questionnaire to assess HRQoL in patients with IBD.[61,62] It is a 32-item Likert-based questionnaire divided into four dimensions: bowel symptoms (10 items), systemic symptoms (five items), emotional function (12 items), and social function (five items). Patients are asked to recall symptoms and quality of life from the last two weeks with response graded on a seven-point Likert scale (1 being the worst situation, 7 being the best) with the total IBDQ score ranging from 32 to 224 (i.e., higher scores representing better quality of life). A total IBDQ score of at least 170 points or higher is considered clinical remission. This questionnaire has been validated in a variety of settings, countries, and languages, and is available in a 9-, 10-, and 36-item form.[63]

Validity

Two systematic reviews published in the last three years reported the measurement properties and methodological quality of a number of IBD-specific HRQoL instruments, including the IBDQ.[48,49] Overall, the IBDQ was proven to be a valid, reliable, and responsive

scale; however, the methodological quality was poor to fair for some of these measurement properties. The IBDQ demonstrated content validity, as it was developed through patient interviews and covered the most frequent and important items. Results from factor analysis showed the items/domains of the scale explained at least 50% of the variance. The scale showed strong correlation with the Crohn's Disease Activity Index (r = −0.67), proving convergent validity. In addition, criterion validity was proven, as there was similar correlation with changes in IBDQ and other measures. The scale showed lower discriminant validity, particularly in patients who required surgery.[48,49]

Reliability and Responsiveness

The reliability parameters showed high internal consistency (Cronbach's alpha 0.7), test-retest reliability (ICC, 0.9 to 0.99 or Pearson's r ≥ 0.8), and low measurement error (i.e., the standard deviations of the score changes were of similar magnitude and the smallest detectable change was less than the MCID). Responsiveness was satisfactory, as the scale was sensitive to change corresponding to clinical improvement or deterioration. Floor and ceiling effects were not found, as less than 15% of the respondents achieved the highest or lowest possible score.[48,49]

Minimal Clinically Important Difference

Irvine et al. reported that a change of 30 or more points in actual score or an improvement of 15 or more points above the placebo score is associated with clinical benefits in IBD patients, including those with UC.[50] Several other studies have reported an increase of 15 to 32 points from baseline as clinically meaningful improvement.[64]

Short Form (36) Health Survey

The SF-36 is a generic self-reported health assessment questionnaire that has been used in clinical trials to study the impact of chronic disease on HRQoL. The original version (SF-36v1) was released in 1992; however, a revised version (SF-36v2), released in 1996, is used more commonly. The SF-36 consists of eight domains: physical functioning, role limitations due to physical health problems, bodily pain, general health, vitality, social functioning, role limitations due to emotional health problems, and mental health. The SF-36 also provides two component summaries: the PCS and the Mental Component Summary (MCS), which are scores created by aggregating the eight domains. The SF-36 PCS and MCS and individual domains are each measured on a scale of 0 to 100, with an increase in score indicating improvement in health status.[65]

Validity

A recently published systematic review assessed the construct validity, reliability, and responsiveness of the SF-36v2 among UC patients.[51] Construct validity was demonstrated by more than two dozen studies in which the correlations between the eight subscales of SF-36 and corresponding domains of five patient-reported clinical constructs (the IBDQ, IBD Quality of Life Questionnaire, Brief Pain Inventory, Short Health Scale, and Rating Form of IBD Patient Concerns) were found to be in the same hypothesized direction and of moderate-to-high strength (r > 0.4) overall. The scale showed evidence of discriminative validity, as there were clinically meaningful differences in most SF-36 subscores between subgroups of patients classified by disease activity, symptom status, and comorbidity status.

Reliability and Responsiveness

Yarlas et al. found one study that evaluated the reliability of the SF-36, and found evidence supporting internal consistency for all eight subscales (Cronbach's alpha > 0.7) and high test-retest reliability for six of the eight subscales (ICC > 0.7). The role physical and role emotional subscales had a lower ICC of 0.64 and 0.63, respectively; the authors indicated high floor and ceiling effect as a possible reason for this.[51] The scale and its subscores were found to be responsive to treatment-related changes, as evidenced by clinically meaningful changes in most SF-36 subscores over time following effective treatment in non-comparative trials or among treated patients relative to controls in RCTs.[51]

Minimal Clinically Important Difference

For both the PCS and MCS as well as the individual subscale scores in the SF-36, an absolute score increase of 3 to 5 points was shown to capture MCIDs in various conditions, including colitis.[50]

EuroQol 5-Dimensions 3-Levels

The EQ-5D-3L is a generic, preference-based, HRQoL measure consisting of descriptive questions and a VAS.[66] The EQ-5D-3L has been applied to a wide range of health conditions and treatments, including IBD.[66,67] The descriptive questions comprise five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension is divided into three levels (1, 2, 3) representing "no problems," "some problems," and "extreme problems," respectively. Respondents (aged ≥ 12 years) are asked to choose one level that reflects their own health state for each of the five dimensions. The five questions are scored and together contribute to an EQ-5D index (utility) score between 0 and 1, where 0 represents death, and 1 represents perfect health. Different utility functions are available that reflect the preferences of specific populations (e.g., US, UK). The second part of the tool records the patient's self-rated health on a 20 cm scale with end points 0 and 100, with respective anchors of "the worst health you can imagine" and "the best health you can imagine," respectively.

Validity

Stark et al. assessed the validity, reliability, and responsiveness of the EQ-5D in a German population of IBD patients (including those with UC).[53] Respondents completed the EQ-5D twice, four weeks apart. At the four-week follow-up, patients were asked in a transition question to report whether their health status was better, worse, or the same. Construct validity was evaluated in two methods: assessing the correlation between the EQ-5D index and VAS scores with disease activity, and comparing responses between patients with active disease versus those in remission.[53] Construct validity of the EQ-5D index score and VAS was supported by the strong correlation of these scores with the Clinical Activity Index (Spearman rank correlation, r, between 0.65 and 0.67). The EQ-5D index score and VAS as well as all but one domain of the scale (self-care) showed discriminative validity by correctly differentiating patients in remission and active disease. A smaller study, Konig et al. (29 patients with UC; two-week recall period), also demonstrated strong correlation between the EQ-5D VAS and index scores with the IBDQ total score (0.70 and 0.62, respectively), and moderate-to-strong correlation with the SF-36 subscores (0.37 to 0.72).[52]

Reliability and Responsiveness

Stark et al. assessed test-retest reliability by comparing baseline and follow-up measurements of the EQ-5D in the subset of patients who indicated no change in HRQoL in the transition question. Test-retest reliability was generally high for the index score (0.67 ≤ ICC ≤ 0.73), VAS (ICC of 0.93), and all five items of the scale (0.67 ≤ kappa ≤ 1.00). Konig et al. reported similar results (ICC of 0.89 for the index score and 0.77 for the VAS score).[52] Responsiveness (sensitivity to change) of the EQ-5D VAS scores and the index scores was tested in patients indicating a change in their health status in the transition question with paired t-tests, effect size, and standardized response mean.[53] Both the index score and VAS were shown to be responsive to detecting change in health status; however, the VAS was found to be more responsive for detecting deterioration in health than for improvement in health and was more responsive than the index score.[53]

Minimal Clinically Important Difference

Stark et al. estimated a disease-specific MCID using a regression model; the MCIDs for improved health were reported to be 10.9 for the VAS, and 0.050 (European Union) and 0.076 (UK) for the index score.[53] This is within the range of other reported MCIDs for the index score of 0.033 to 0.074.[68]

Work Productivity and Activity Impairment Questionnaire – General Health

The WPAI-GH is one of the most frequently used patient-reported, work-related outcome measure.[54,69] The WPAI-GH measures the impact of health problems on absenteeism (missing work), presenteeism (impaired productivity at work), overall work performance (combined absenteeism and presenteeism), and non-work activities (activity impairment).[54] It is a self-administered six-item questionnaire with a recall period of seven days.[69] Scores from all domains are expressed as percentages (0% to 100%) of impairment, with lower values indicating less impairment due to the health problem.[54] The WPAI has been shown to be reliable, valid, and responsive when used with patients across several disease areas, including other gastrointestinal conditions such as irritable bowel syndrome, gastroesophageal reflux disease, and Crohn disease.[54]

Validity

A recent systematic review by Yarlas et al. evaluated eight articles and five posters evaluating the psychometric validation of the WPAI in UC.[54] One study was found that assessed convergent validity between the WPAI domains and other HRQoL measures, including the Short Inflammatory Bowel Disease Questionnaire (SIBDQ) and the SF-12v2.[54] The strongest evidence for convergent validity was reported between all WPAI domains and the SIBDQ bowel symptoms (Spearman rank-order coefficient −0.47 to −0.68) and SF-12v2 bodily pain (−0.52 to −0.55) subscores. With the exception of absenteeism, the WPAI domains also converged with the SIBDQ social function, and SF-12v2 role physical and role emotional subscores.[54] Convergent validity was also assessed between the WPAI and measures of disease activity, specifically, the Simple Clinical Colitis Activity Index (SCCAI), the UC Disease Activity Index (UC-DAI), and the partial Mayo score in three individual studies.[54] Inter-scale correlations between the WPAI domains and disease-activity measures ranged from 0.32 to 0.85 (median 0.45). Across the three studies, convergence with disease activity was supported for presenteeism, overall work impairment (OWI) and activity impairment (0.43 to 0.60), although the median correlation for absenteeism was not far behind (0.39).[54] Furthermore, a known-group validity assessment

demonstrated that patients with worse health outcomes scored worse on the WPAI than patients with better health outcomes, based on partial Mayo, SCCAI, UC-DAI, and the Functional Assessment of Chronic Illness Therapy – Fatigue scale (FACIT-Fatigue) disease severity measures.[54]

Reliability and Responsiveness

Test-retest reliability of the WPAI domains was assessed in one study by Yarlas et al. in 2015 (N = 98) that compared scores at the start and end of an open-label maintenance-treatment period in patients whose remission status was unchanged (as determined by the UC-DAI).[54] The results demonstrated that the differences in each domain were less than 5% over a 12-month period, with none of these differences exceeding the proposed MCID of 7% for Crohn disease; however, no ICC was reported for this data.[54] The ability of WPAI domains to detect changes was evaluated by assessing the magnitude of change in the WPAI domains for patients demonstrating changes in disease states (i.e., change from active disease to remission, or vice-versa) in one study by Yarlas et al.[54] The study demonstrated that patients with active UC disease who achieved remission at week 8 reported a 25% to 30% decrease in presenteeism, OWI, and activity impairment, and a 9% decrease in absenteeism. The inverse was also found in patients with disease relapse.[54] Responsiveness of the WPAI domains to effective treatment was also demonstrated with data from three RCTs investigating either multi-matrix mesalamine treatment or adalimumab in UC patients; results indicated that patients reported an approximate 20% decrease in presenteeism, OWI and activity impairment, and an 8% decrease in absenteeism.[54]

Minimal Clinically Important Difference

There is currently no MCID defined for the WPAI in UC patients. However, the MCID estimated for Crohn disease is a decrease of seven points.[55]

References

1. Clinical Study Report: CNTO1275UCO3001-INDUCTION STUDY. 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. [CONFIDENTIAL internal manufacturer's report]. 2018.
2. Clinical Study Report: CNTO1275UCO3001-MAINTENANCE STUDY. 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[CONFIDENTIAL internal manufacturer's report]. 2018.
3. Feuerstein JD, Cheifetz AS. Ulcerative colitis: epidemiology, diagnosis, and management. *Mayo Clin Proc.* 2014;89(11):1553-1563.
4. Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG Clinical Guideline: Ulcerative Colitis in Adults. *Am J Gastroenterol.* 2019;114(3):384-413.
5. Danese S, Fiocchi C. Ulcerative colitis. *N Engl J Med.* 2011;365(18):1713-1725.
6. Kaplan GG, Bernstein CN, Coward S, et al. The Impact of Inflammatory Bowel Disease in Canada 2018: Epidemiology. *J Can Assoc Gastroenterol.* 2019;2(Suppl 1):S6-S16.
7. Bernstein CN, Nugent Z, Targownik LE, Singh H, Lix LM. Predictors and risks for death in a population-based study of persons with IBD in Manitoba. *Gut.* 2015;64(9):1403-1411.
8. Kawalec P, Malinowski KP. Indirect health costs in ulcerative colitis and Crohn's disease: a systematic review and meta-analysis. *Expert Rev Pharmacoecon Outcomes Res.* 2015;15(2):253-266.
9. Longobardi T, Jacobs P, Wu L, Bernstein CN. Work losses related to inflammatory bowel disease in Canada: results from a National Population Health Survey. *Am J Gastroenterol.* 2003;98(4):844-849.
10. Stange EF, Travis SP, Vermeire S, et al. European evidence-based Consensus on the diagnosis and management of ulcerative colitis: Definitions and diagnosis. *J Crohns Colitis.* 2008;2(1):1-23.
11. Solberg IC, Lygren I, Jahnsen J, et al. Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (IBSEN Study). *Scand J Gastroenterol.* 2009;44(4):431-440.
12. Bressler B, Marshall JK, Bernstein CN, et al. Clinical practice guidelines for the medical management of nonhospitalized ulcerative colitis: the Toronto consensus. *Gastroenterology.* 2015;148(5):1035-1058 e1033.
13. Ko CW, Singh S, Feuerstein JD, et al. AGA Clinical Practice Guidelines on the Management of Mild-to-Moderate Ulcerative Colitis. *Gastroenterology.* 2019;156(3):748-764.
14. Stelara (ustekinumab): 45 mg/0.5 mL and 90 mg/1.0 mL solution for subcutaneous injection; 130 mg/26 ML (5 mg/mL) solution for intravenous infusion [DRAFT product monograph]. Toronto (ON): Janssen Inc.; 2019.
15. CADTH Common Drug Review: Ustekinumab (Stelara) Indication: Crohn's disease. Ottawa (ON): CADTH; 2017: https://www.cadth.ca/sites/default/files/cdr/clinical/SR0501_Stelara_CL_Report.pdf. Accessed Nov 25 2019.
16. CADTH Canadian Drug Expert Committee (CDEC) final recommendation: Stelara (ustekinumab). Indication: Chronic moderate to severe plaque psoriasis 2009: https://cadth.ca/sites/default/files/cdr/complete/cdr_complete_Stelara_June-17-2009.pdf. Accessed 2019 Nov 25.
17. CADTH Canadian Drug Expert Committee (CDEC) final recommendation: drug name: Stelara (ustekinuma). Indication: Psoriatic arthritis. Ottawa (ON): CADTH; 2014: https://www.cadth.ca/sites/default/files/cdr/complete/cdr_complete_SR0359_Stelara_Oct-22-14.pdf. Accessed 2019 Oct 20.
18. CADTH Canadian Drug Expert Committee (CDEC) final recommendation: Ustekinumab (Stelara) [Internet]. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health; 2017 Apr. CDEC FINAL RECOMMENDATION. Available from: https://www.ncbi.nlm.nih.gov/books/NBK476180/. 2017.
19. Desch M. Prevention: Icosapent ethyl reduces cardiovascular risk. *Arzneimitteltherapie.* 2019;37(4):133-135.
20. Chan LN. Cardiovascular Risk Reduction with Icosapent Ethyl. *N Engl J Med.* 2019;380(17):1677.
21. Anonymous. Comparison table: Some lipid-lowering drugs. *Med Lett Drugs Ther.* 2019;61(1565):e24-e30.
22. Anonymous. Lipid-lowering drugs. *Med Lett Drugs Ther.* 2019;61(1565):17-24.
23. Xeljanz (tofacitinib): Tofacitinib tablets / 5 mg tofacitinib and 10 mg; Tofacitinib extended release tablets / 11 mg [product monograph]. Kirkland (Quebec): Pzifer, Canada; 2019.
24. Xeljanz, Xeljanz XR (tofacitinib): Drug Safety Communication - Due to an increased risk of blood clots and death with higher dose. Vol 2019. U.S. Food & Drug Administration: U.S. Food & Drug Administration (FDA); 2019: https://www.fda.gov/safety/medwatch-safety-alerts-human-medical-products/xeljanz-xeljanz-xr-tofacitinib-drug-safety-communication-due-increased-risk-blood-clots-and-death. Accessed July 26, 2019.
25. Lee MC, Jiang X, Brenna JT, Abbaspourrad A. Oleogel-structured composite for the stabilization of omega3 fatty acids in fish oil. *Food Funct.* 2018;9(11):5598-5606.
26. Coward S, Clement F, Benchimol EI, et al. The rising prevalence of inflammatory bowel disease in Canada: Analyzing the past to predict the futur. *Journal of the Canadian Association of Gastroenterology.* 2018;1(Suppl 2):47-48.

27. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol.* 2016;75:40-46.

28. Grey matters: a practical tool for searching health-related grey literature. Ottawa (ON): CADTH; 2019: https://www.cadth.ca/grey-matters. Accessed 2019 Nov 25.

29. Sands BE, Sandborn WJ, Panaccione R, et al. Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis. *N Engl J Med.* 2019;381(13):1201-1214.

30. Sands BE, Sandborn WJ, Panaccione R, et al. Safety and efficacy of ustekinumab induction therapy in patients with moderate to severe ulcerative colitis: Results from the phase 3 UNIFI study. Abstract presented at: American College of Gastroenterology Annual Scientific Meeting & Postgraduate Course. *United European Gastroenterol.* 2018;6(Supplement 1).

31. Bonovas S, Nikolopoulos GK, Piovani D, et al. Comparative assessment of budesonide-MMX and mesalamine in active, mild-to-moderate ulcerative colitis: A systematic review and network meta-analysis. *Br J Clin Pharmacol.* 2019;03:03.

32. Adedokun OJ, Xu Z, Marano CW, et al. Pharmacokinetics and Exposure-Response Relationships of Intravenously Administered Ustekinumab During Induction Treatment in Patients with Ulcerative Colitis: Results from the Unifi Induction Study. *Gastroenterology.* 2019;156(6):S-1110.

33. Sands BE, Peyrin-Biroulet L, Marano CW, et al. Efficacy in Biologic-Failure and Nonbiologic-Failure Populations in a Phase 3 Study of Ustekinumab in Moderatesevere Ulcerative Colitis: Unifi. *Gastroenterology.* 2019;156(6):S-181-S-182.

34. Bhatt DL, Steg PG, Miller M, et al. Cardiovascular Risk Reduction with Icosapent Ethyl for Hypertriglyceridemia. *N Engl J Med.* 2019;380(1):11-22.

35. Sands BE, Han C, Zhang H, et al. Ustekinumab Therapy Induced Clinically Meaningful Improvement and Remission As Measured by the Inflammatory Bowel Disease Questionnaire in Patients with Moderate to Severe Ulcerative Colitis: Results from the Phase 3 Unifi Induction and Maintenance Studies. *Gastroenterology.* 2019;156(6):S-442-S-443.

36. Sandborn WJ, Sands BE, Panaccione R, et al. OP37 Efficacy and safety of ustekinumab as maintenance therapy in ulcerative colitis: Week 44 results from UNIFI. *Journal of Crohn's and Colitis.* 2019;13(Supplement_1):S025-S026.

37. Van Assche GA, Targan SR, Baker T, et al. Sustained Remission in Patients with Moderate to Severe Ulcerative Colitis: Results from the Phase 3 Unifi Maintenance Study. *Gastroenterology.* 2019;156(6):S-1106-S-1107.

38. Sandborn WJ, Feagan BG, Marano C, et al. Subcutaneous golimumab induces clinical response and remission in patients with moderate-to-severe ulcerative colitis. *Gastroenterology.* 2014;146(1):85-95; quiz e14-85.

39. Systematic review and network meta-analysis to assess the relative efficacy and safety of Stelara in ulcerative colitis In: CDR submission: Stelara (ustekinumab), 90 mg/1.0 mL solution for subcutaneous injection, 130 mg/26 mL solution for intravenous infusion [CONFIDENTIAL manufacturer's submission]. Toronto (ON): Janssen Inc.; 2019.

40. Thorlund K, Druyts E, Toor K, Jansen JP, Mills EJ. Incorporating alternative design clinical trials in network meta-analyses. *Clin Epidemiol.* 2015;7:29-35.

41. Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.

42. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res.* 2013;22(8):1889-1905.

43. Cohen J. A power primer. *Psychol Bull.* 1992;112(1):155-159.

44. Mohammed Vashist N, Samaan M, Mosli MH, et al. Endoscopic scoring indices for evaluation of disease activity in ulcerative colitis. *Cochrane Database Syst Rev.* 2018;1:CD011450.

45. Levesque B, Pola S, King D, al. e. Responsiveness of central endoscopic assessment of disease activity using the Modified Mayo Clinic Score in ulcerative colitis. *Gastroenterol Hepatol (N Y).* 2013;11:S767.

46. Walsh AJ, Ghosh A, Brain AO, et al. Comparing disease activity indices in ulcerative colitis. *J Crohns Colitis.* 2014;8(4):318-325.

47. Lewis J, Chuai S, Nessel L, Lichtenstein G, Aberra F, Ellenberg J. Use of the non-invasive components of the Mayo Score to assess clinical response in ulcerative colitis. *Inflammatory bowel diseases.* 2008;14(12):1660–1666.

48. Chen XL, Zhong LH, Wen Y, et al. Inflammatory bowel disease-specific health-related quality of life instruments: a systematic review of measurement properties. *Health Qual Life Outcomes.* 2017;15(1):177.

49. Alrubaiy L, Rikaby I, Dodds P, Hutchings HA, Williams JG. Systematic review of health-related quality of life measures for inflammatory bowel disease. *J Crohns Colitis.* 2015;9(3):284-292.

50. Irvine J. Quality of life of patients with ulcerative colitis: Past, present, and future. *Inflamm Bowel Dis.* 2008;14(4):554-565.

51. Yarlas A, Bayliss M, Cappelleri JC, et al. Psychometric validation of the SF-36((R)) Health Survey in ulcerative colitis: results from a systematic literature review. *Qual Life Res.* 2018;27(2):273-290.

52. Konig HH, Ulshofer A, Gregor M, et al. Validation of the EuroQol questionnaire in patients with inflammatory bowel disease. *Eur J Gastroenterol Hepatol.* 2002;14(11):1205-1215.

53. Stark RG, Reitmeir P, Leidl R, Konig HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. *Inflammatory bowel diseases.* 2010;16(1):42-51.

54. Yarlas A, Maher SM, Bayliss MS, Lovley A, Cappelleri JC, DiBonaventura MD. Psychometric validation of the work productivity and activity impairment questionnaire in ulcerative colitis: results from a systematic literature review. *Journal of Patientreported Outcomes.* 2018;2(1):62.

55. Reilly M, Brown M, Brabant Y. Minimally important difference for WPAI: CD scores: Defining relevant impact on work productivity in active Crohn's disease (Abst 962) [Conference Abstract]. American College of Gastroenterology Annual Scientific Meeting, Philadelphia2007.

56. Schroeder KW, Tremaine WJ, Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *N Engl J Med.* 1987;317(26):1625-1629.

57. Dhanda AD, Creed TJ, Greenwood R, Sands BE, Probert CS. Can endoscopy be avoided in the assessment of ulcerative colitis in clinical trials? *Inflammatory bowel diseases.* 2012;18(11):2056-2062.

58. Samaan MA, Mosli MH, Sandborn WJ, et al. A systematic review of the measurement of endoscopic healing in ulcerative colitis clinical trials: recommendations and implications for future research. *Inflammatory bowel diseases.* 2014;20(8):1465-1471.

59. Cooney RM, Warren BF, Altman DG, Abreu MT, Travis SP. Outcome measurement in clinical trials for ulcerative colitis: towards standardisation. *Trials.* 2007;8:17.

60. Center for Drug Evaluation and Research. Ulcerative colitis: clinical trial endpoints guidance for industry: draft guidance. Rockville (MD): U.S. Food and Drug Administration; 2016 Aug: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM515143.pdf. Accessed 2019 Oct 10.

61. Guyatt G, Mitchell A, Irvine EJ, et al. A new measure of health status for clinical trials in inflammatory bowel disease. *Gastroenterology.* 1989;96(3):804-810.

62. Irvine EJ. Development and subsequent refinement of the inflammatory bowel disease questionnaire: a quality-of-life instrument for adult patients with inflammatory bowel disease. *J Pediatr Gastroenterol Nutr.* 1999;28(4):S23-27.

63. Pallis AG, Mouzas IA, Vlachonikolis IG. The inflammatory bowel disease questionnaire: a review of its national validation studies. *Inflamm Bowel Dis.* 2004;10(3):261-269.

64. LeBlanc K, Mosli MH, Parker CE, MacDonald JK. The impact of biological interventions for ulcerative colitis on health-related quality of life. *Cochrane Database Syst Rev.* 2015(9):CD008655.

65. Bernklev T, Jahnsen J, Lygren I, Henriksen M, Vatn M, Moum B. Health-related quality of life in patients with inflammatory bowel disease measured with the short form-36: psychometric assessments and a comparison with general population norms. *Inflammatory bowel diseases.* 2005;11(10):909-918.

66. EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. *Health policy (Amsterdam, Netherlands).* 1990;16(3):199-208.

67. Brooks R. EuroQol: the current state of play. *Health policy (Amsterdam, Netherlands).* 1996;37(1):53-72.

68. Sinnott P, Joyce V, Barnett P. Guidebook: preference measurement in economic analysis Menlo Park (CA): Health Economics Research Center; 2007: http://www.herc.research.va.gov/files/BOOK_419.pdf. Accessed 2019 Oct 10.

69. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics.* 1993;4(5):353-365.