Common Drug Review
Clinical Review Report

April 2017

Drug	Ustekinumab (Stelara)
Indication	For the treatment of adult patients with moderately to severely active Crohn's disease who have had an inadequate response, loss of response to, or were intolerant to either immunomodulators or one or more tumour necrosis factor-alpha antagonists, or have had an inadequate response, intolerance or demonstrated dependence on corticosteroids.
Reimbursement request	As per indication.
Dosage form(s)	130 mg solution for intravenous infusion (for induction period) 90 mg solution for subcutaneous injection (for maintenance period)
NOC Date	December 12, 2016
Manufacturer	Janssen Inc.

This review report was prepared by the Canadian Agency for Drugs and Technologies in Health (CADTH). In addition to CADTH staff, the review team included a clinical expert in gastroenterology who provided input on the conduct of the review and the interpretation of findings

TABLE OF CONTENTS

Tables

Figures

ABBREVIATIONS

6-MP	6-mercaptopurine
AE	adverse event
AZA	azathioprine
CAG	Canadian Association of Gastroenterology
CDAI	Crohn's Disease Activity Index
CDEC	CADTH Canadian Drug Expert Committee
CDR	CADTH Common Drug Review
CEDAC	CADTH Canadian Expert Drug Advisory Committee
CRP	C-reactive protein
EMA	European Medicines Agency
ES	effect size
IBDQ	Inflammatory Bowel Disease Questionnaire
IL	interleukin
IV	intravenous
IVRS	interactive voice response system
IWRS	interactive web response system
LOCF	last observation carried forward
MCID	minimal clinically important difference
MCS	mental component score
MTX	methotrexate
NMA	network meta-analysis
PCS	physical component score
SAE	serious adverse event
SC	subcutaneous
SF-36	Short Form (36) Health Survey
TB	tuberculosis
TNF	tumour necrosis factor
VAS	visual analogue scale
WDAE	withdrawal due to adverse event
WLQ	Work Limitations Questionnaire

EXECUTIVE SUMMARY

Introduction

Crohn's disease is a chronic form of inflammatory bowel disease that can affect any part of the gastrointestinal tract but most commonly affects the ileum, colon, and rectum. Common gastrointestinal symptoms experienced by patients with Crohn's disease include abdominal pain, rectal bleeding, fatigue, vomiting, diarrhea, perianal disease, weight loss, and bloating.[1-3] According to Crohn's and Colitis Canada, there are approximately 129,000 Canadians living with Crohn's disease (one in 150 people), and it is estimated that 5,700 new cases of Crohn's disease are diagnosed each year.[1]

Ustekinumab (Stelara) is a fully human immunoglobulin G1 kappa monoclonal antibody that binds to the shared p40 subunit of interleukin (IL)-12 and IL-23.[4] Ustekinumab is already 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, alone or in combination with methotrexate.[4]

The current indication under review is for the treatment of adult patients with moderately to severely active Crohn's disease, who have had an inadequate response with, loss of response to, or intolerance to either immunomodulators or one or more tumour necrosis factor (TNF) alpha antagonists, or have had an inadequate response with, intolerance to, or demonstrated dependence on corticosteroids. The recommended dosage for ustekinumab in the treatment of Crohn's disease is an initial single intravenous (IV) induction dose based on body weight (approximating 6 mg/kg), followed by a 90 mg subcutaneous (SC) maintenance dose eight weeks later, then one dose every eight weeks thereafter as maintenance treatment. 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 inadequately respond to the 90 mg SC every 12 weeks regimen may be switched to the every eight weeks regimen. Immunomodulators and/or corticosteroids may be continued during treatment with ustekinumab. The product monograph recommends that, in patients who have responded to treatment with ustekinumab, corticosteroids may be reduced or discontinued in accordance with standard of care.[4]

The objective of this report is to perform a systematic review of the beneficial and harmful effects of ustekinumab in accordance with the Health Canada–approved indication for the treatment of Crohn's disease. Only Health Canada–approved dosage regimens for ustekinumab for Crohn's disease were included in this review. Ustekinumab has been previously reviewed through the CADTH Common Drug Review (CDR) for the treatment of plaque psoriasis and psoriatic arthritis.[5,6]

Results and Interpretation

Included Studies

The CDR review included four, multi-centre, multinational, double-blind, randomized placebo-controlled trials: two phase III induction-treatment studies, UNITI-1 (N = 769) and UNITI-2 (N = 640);[7,8] one phase III maintenance-treatment study, IM-UNITI (N = 397);[9] and one phase II induction and maintenance study, CERTIFI (N = 526).[10,11] The phase III studies were designed as superiority studies, whereas the phase II study was a dose-ranging study. The results from the UNITI studies and IM-UNITI are the focus of the CDR review; results from the induction phase of CERTIFI were considered supportive. All of the studies enrolled adult patients with moderately to severely active Crohn's disease.

UNITI-1 and UNITI-2 were identically designed studies to evaluate the efficacy and safety of IV dosage regimens of ustekinumab (tiered weight-based dose approximating 6 mg/kg or 130 mg [not Health Canada–approved]) versus placebo for inducing clinical response (reduction from baseline in the Crohn's Disease Activity Index [CDAI] score of ≥ 100 points) at six weeks (primary outcome) in patients with moderately to severely active Crohn's disease. UNITI-1 included patients who had had an inadequate response with or were intolerant to one or more TNF antagonist therapies, whereas UNITI-2 included patients who had had an inadequate response with or were intolerant to conventional therapy only (i.e., corticosteroids or immunomodulators such as 6-mercaptopurine, azathioprine, and methotrexate). Patients in UNITI-2 could have previously received TNF antagonists but could not have failed treatment. The same patient population as UNITI-1 was enrolled in CERTIFI for IV induction.

The IM-UNITI study was designed to evaluate the efficacy (clinical remission) and safety of two SC maintenance regimens of ustekinumab (90 mg every eight weeks or every 12 weeks) in patients with moderately to severely active Crohn's disease who had had a clinical response with ustekinumab in the induction studies, UNITI-1, and UNITI-2. The primary outcome was the proportion of patients in clinical remission (defined as a CDAI score < 150 points) at week 44.

Although the treatment groups in each study were generally similar with respect to baseline demographic and disease characteristics, there were notable potential differences, such as in the concomitant use of oral corticosteroids at baseline in UNITI-2 and IM-UNITI. Nonetheless, the clinical expert consulted by CDR considered the populations to be representative of those seen in practice in Canada. Other key limitations of the included studies were as follows:

- Uncertainty as to the extent of exposure to corticosteroids during IM-UNITI
- Greater than 20% treatment discontinuation in IM-UNITI
- Not a true intention-to-treat analysis for all outcomes
- Potential misclassification of patients as having been adequately treated with TNF antagonist and having subsequently failed therapy
- Lack of direct comparison between ustekinumab and TNF antagonists and between ustekinumab and vedolizumab
- Inclusion of patients who had a clinical response with a non–Health Canada–approved dose of ustekinumab (IV 130 mg)

Efficacy

Clinical Remission

A statistically significantly higher proportion of patients treated with ustekinumab 6 mg/kg (20.9% and 40.2%) than with placebo (7.3% and 19.6%) were in remission at week 8 in UNITI-1 and UNITI-2, respectively (Table 1). Likewise, statistically significantly higher proportions of patients treated with ustekinumab every 12 weeks (48.8% and 42.6%) and ustekinumab every eight weeks (53.1% and 46.9%) were in clinical remission and corticosteroid-free remission, respectively, at week 44 of IM-UNITI than with placebo (35.9% and 29.8%) (Table 2). [redacted] The clinical expert consulted by CDR noted that, although the between-group differences are not large, they likely represent clinically meaningful results, especially in UNITI-1, in which patients had experienced a failure of TNF antagonist treatment.

Clinical Response

The primary outcome for the UNITI induction studies was achieved: the proportion of patients with a clinical response at week 6 was statistically significantly higher in the ustekinumab groups (33.7% and 55.5%) than in the placebo groups (21.5% and 28.7%) in UNITI-1 and UNITI-2, respectively (Table 1).

Almost 60% of patients randomized to ustekinumab maintenance treatments in IM-UNITI were responders at week 44, whereas 44% of those assigned to placebo achieved clinical response. The comparison versus placebo was statistically significant for both ustekinumab regimens (Table 2).

The results from CERTIFI were supportive of these.

Health-Related Quality of Life, Functional and Disability Outcomes

[redacted] (Table 1). [redacted] (Table 2).

[redacted] Table 2, [redacted]

[redacted]

Need for Surgery for Crohn's Disease

[redacted] (Table 1). [redacted]

[redacted] (Table 2).

In the absence of head-to-head trials comparing the efficacy of ustekinumab with TNF antagonists or vedolizumab, CDR examined the comparative effectiveness results of two network meta-analyses

(NMAs, Appendix 7).[12,13] The manufacturer submitted an NMA of ustekinumab versus infliximab, adalimumab, and vedolizumab.[12] There were no statistically significant differences between ustekinumab and vedolizumab or adalimumab for clinical response and remission. Infliximab may be superior to ustekinumab for inducing clinical response and remission among patients who have failed conventional therapies. The indirect comparisons between drugs for induction have several limitations, in particular large differences in placebo response rates and a high degree of heterogeneity associated with the adalimumab and especially the infliximab studies compared with the ustekinumab studies. Given the limitations of the available indirect comparisons and the heterogeneity across studies, the effectiveness of ustekinumab compared with infliximab, adalimumab, and vedolizumab is uncertain for the treatment-sequence analysis (induction plus maintenance phases).

Harms

The proportions of patients who experienced at least one adverse event or serious adverse event were similar between the ustekinumab and placebo groups across all of the included studies. Nasopharyngitis and upper respiratory tract infections appeared to be more frequent with ustekinumab treatment than with placebo. As may be expected, patients treated with ustekinumab tended to report more administration-related reactions than those on placebo; however, there were no reports of anaphylaxis in any of the studies. [redacted]

The CDR review summarized an NMA conducted by Mocko et al.[14,15] that reported no statistically significant differences in the incidence of adverse events, serious adverse events, discontinuations due to adverse events, or some of the more prominent adverse events (e.g., infections, injection-site reactions, nausea, headache, arthralgia, etc.) among adalimumab, ustekinumab, or vedolizumab during induction therapy and among adalimumab, infliximab, and vedolizumab during maintenance therapy in patients with Crohn's disease (Appendix 7). However, several major limitations associated with the conduct of this NMA introduce a very high degree of uncertainty regarding the results. Hence, caution is required when interpreting the authors' observations that there are no differences in safety between these drugs during the induction and maintenance phases of therapy for patients with Crohn's disease.

Patient groups expressed an understanding of the potential risks associated with biologic treatments and noted that those living with Crohn's disease are often willing to accept these risks rather than undergo surgery, which they consider to be a last resort.

Potential Place in Therapy

The information in this section is based on that provided in draft form by the clinical expert consulted by CDR reviewers for the purpose of this review.

Based on current standards of practice with existing therapies, the clinical expert consulted by CDR indicated that there are several areas of unmet need where ustekinumab may play a role:

1. It may provide primary induction therapy for Crohn's disease for patients who experience primary nonresponse to either conventional therapy with immunomodulators or TNF antagonists.
2. It may also be useful in secondary nonresponse during maintenance therapy. An important proportion of patients with Crohn's disease lose response to TNF antagonist therapy during maintenance, either owing to formation of anti-drug antibodies or to inflammatory mechanisms

that are independent of TNF. Evidence summarized in this review suggests that ustekinumab may provide clinically meaningful benefit in this patient group.

3. It may also provide salvage therapy for patients who respond to therapy with immunomodulators or TNF antagonists but who develop adverse effects. Immunomodulators such as azathioprine and methotrexate are generally safe medications; however, there are well-known side effects, including pancreatitis, neutropenia, hepatitis, and neoplasia (e.g., skin cancers). TNF antagonists can be associated with severe allergic reactions, psoriatic skin diseases, neurological complications, congestive heart failure, lupus, and severe infections. In these situations, ustekinumab therapy may be safer and allow for continued treatment of the disease.

Conclusions

Three phase III, randomized, placebo-controlled, double-blind trials investigated the effects of ustekinumab on treatment induction (UNITI-1 and UNITI-2) or maintenance (IM-UNITI) in patients with moderate-to-severe Crohn's disease. A single IV dose of ustekinumab (approximating 6 mg/kg) appears to be significantly superior to placebo for inducing clinical response after six weeks of therapy. Likewise, both the ustekinumab 90 mg SC every 12 weeks and every eight weeks maintenance-treatment regimens were statistically significantly superior to placebo in achieving clinical remission and corticosteroid-free remission in patients who had a clinical response at week 8 of induction therapy. Moreover, these results for induction and maintenance therapy with ustekinumab were reported in subpopulations of patients with Crohn's disease who had experienced failure of failed conventional therapies only or of TNF antagonist therapies. These findings were considered likely to be clinically meaningful by the clinician expert consulted by CDR. [redacted]

The proportion of patients who experienced at least one adverse event or serious adverse event was similar between the ustekinumab and placebo groups across all of the included studies. Nasopharyngitis and upper respiratory tract infection were reported more frequently in ustekinumab-treated patients than in placebo-treated patients, but these did not lead to discontinuation of treatment. Administration-related reactions were relatively rare.

There were no studies in which ustekinumab has been compared directly with the approved TNF antagonists or vedolizumab for induction or maintenance treatment of Crohn's disease. Three indirect comparisons reviewed by CDR, including one submitted by the manufacturer, were challenging to interpret because of numerous limitations related to the source data and the NMA methods used to compare treatments. These limitations precluded any definitive conclusions regarding the efficacy and safety of ustekinumab compared with TNF antagonists and vedolizumab.

TABLE 1: SUMMARY OF KEY RESULTS FROM THE INDUCTION STUDIES

Parameter	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
Efficacy outcomes				
Clinical remission at week 8, n (%)	18 (7.3)	52 (20.9)	41 (19.6)	84 (40.2)
P value		< 0.001		< 0.001
Clinical response at week 6, n (%)	53 (21.5)	84 (33.7)	60 (28.7)	116 (55.5)
P value		0.003		< 0.001
Change in IBDQ total score at week 8				
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]
Change in SF-36 PCS score at week 8	[redacted]	[redacted]	[redacted]	[redacted]
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]
Change in SF-36 MCS score at week 8	[redacted]	[redacted]	[redacted]	[redacted]
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]
Number of patients with CD-related surgery through week 8				
n (%)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]
Harms outcomes, n (%)	**N = 245**	**N = 249**	**N = 208**	**N = 207**
AEs	159 (64.9)	164 (65.9)	113 (54.3)	115 (55.6)
SAEs	15 (6.1)	18 (7.2)	12 (5.8)	6 (2.9)
WDAEs	[redacted]	[redacted]	[redacted]	[redacted]
Deaths	0	0	0	0
AEs of interest				
Infusion reactions	5 (2.0)	9 (3.6)	6 (2.9)	3 (1.4)
Anaphylaxis	0	0	0	0
Infections[b]	58 (23.7)	64 (25.7)	48 (23.1)	45 (21.7)
Serious infections[b]	3 (1.2)	7 (2.8)	3 (1.4)	1 (0.5)
Malignancy	0	0	0	0
Major cardiovascular events	0	0	0	0
Neurological	[redacted]	[redacted]	[redacted]	[redacted]

AE = adverse event; CD = Crohn's disease; IBDQ = Inflammatory Bowel Disease Questionnaire; MCS = mental component score; PCS = physical component score; PLA = placebo; SAE = serious adverse event; SD = standard deviation; SF-36 = Short Form (36) Health Survey; UST = ustekinumab; WDAE = withdrawal due to adverse event.
Source: Clinical Study Reports for UNITI-1 and UNITI-2.[7,8]

TABLE 2: SUMMARY OF KEY RESULTS FROM THE MAINTENANCE STUDY

Parameter	IM-UNITI		
	PLA (N = 131)	UST q.12.w. (N = 129)	UST q.8.w. (N = 128)
Efficacy outcomes			
Clinical remission at week 44, n (%)	47 (35.9)	63 (48.8)	68 (53.1)
P value		0.04	0.005
CS-free clinical remission at week 44, n (%)	39 (29.8)	55 (42.6)	60 (46.9)
P value		0.035	0.004
Clinical response at week 44, n (%)	58 (44.3)	75 (58.1)	76 (59.4)
P value		0.033	0.018
Change in IBDQ total score at week 44	[redacted]	[redacted]	[redacted]
Baseline score, median (IQR)	167.0 ([redacted])	172.0 ([redacted])	176.5 ([redacted])
Change from baseline, median (IQR)[a]	–14.5 [redacted])	–2.5 ([redacted])	–2.0 ([redacted])
P value		[redacted]	[redacted]
Change in SF-36 PCS score at week 44	[redacted]	[redacted]	[redacted]
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)[a]	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Change in SF-36 MCS score at week 44	[redacted]	[redacted]	[redacted]
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)[a]	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Number of patients with CD-related surgery through week 44			
n (%)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Harms outcomes, n (%)	**N = 133**	**N = 132**	**N = 131**
AEs	111 (83.5)	106 (80.3)	107 (81.7)
SAEs	20 (15.0)	16 (12.1)	13 (9.9)
WDAEs	8 (6.0)	10 (7.6)	4 (3.1)
Deaths	0	0	0
AEs of interest			
Injection-site reactions	1 (0.8)	3 (2.3)	9 (6.9)
Anaphylaxis	0	0	0
Infections	66 (49.6)	61 (46.2)	63 (48.1)
Serious infections	3 (2.3)	7 (5.3)	3 (2.3)
Malignancy	1	0	1
Major cardiovascular events	0	0	0
Neurological	[redacted]	[redacted]	[redacted]

AE = adverse event; CD = Crohn's disease; CS = corticosteroid; IBDQ = Inflammatory Bowel Disease Questionnaire; MCS = mental component score; PCS = physical component score; PLA = placebo; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; SAE = serious adverse event; SD = standard deviation; SF-36 = Short Form (36) Health Survey; UST = ustekinumab; WDAE = withdrawal due to adverse event.

[a] Negative values indicate worsening.

Source: Clinical Study Reports for UNITI-1 and UNITI-2.[9]

1. INTRODUCTION

1.1 Disease Prevalence and Incidence

Crohn's disease is a chronic form of inflammatory bowel disease that can affect any part of the gastrointestinal tract, but most commonly affects the ileum (i.e., small intestine), colon (i.e., beginning of the large intestine), and rectum. Common gastrointestinal symptoms experienced by patients with Crohn's disease include abdominal pain, rectal bleeding, fatigue, vomiting, diarrhea, perianal disease, weight loss, and bloating.[1-3] Crohn's disease–associated inflammation can also manifest outside the gastrointestinal tract, affecting the joints, eyes, and skin of the patient. Complications associated with Crohn's disease can include malnutrition, weight loss, anemia, bowel obstructions, fistulas, anal fissures, intra-abdominal and other abscesses, and ulcers.[3] In addition, patients with colonic Crohn's disease have been shown to have an increased risk of colon cancer.[3] According to Crohn's and Colitis Canada, there are approximately 129,000 Canadians living with Crohn's disease (one in 150 people), and it is estimated that 5,700 new cases of Crohn's disease are diagnosed each year.[2]

According to patients — based on patient-group input for this review (0) — Crohn's disease has a profound effect on physical, emotional, and social well-being. It affects interactions with others and patients' work life.

TABLE 3: CLASSIFICATION OF DISEASE SEVERITY IN CROHN'S DISEASE

Status	CDAI Score	Description from ACG Guidelines
Remission	< 150	Asymptomatic or without any symptomatic inflammatory sequelae
Mild-to-moderate	150 to 220	Ambulatory and able to tolerate oral alimentation without manifestations of dehydration, systemic toxicity, abdominal tenderness, painful mass, intestinal obstruction, or > 10% weight loss
Moderate-to-severe	220 to 450	Failed to respond to treatment for mild-to-moderate disease, or those with more prominent symptoms of fever, significant weight loss, abdominal pain or tenderness, intermittent nausea or vomiting, or significant anemia
Severe	> 450	Persistent symptoms despite the introduction of conventional corticosteroids or biologic drugs as outpatients, or individuals presenting with high fevers, persistent vomiting, evidence of intestinal obstruction, significant peritoneal signs such as involuntary guarding or rebound tenderness, cachexia, or evidence of an abscess

ACG = American College of Gastroenterology; CDAI = Crohn's Disease Activity Index.
Source: American College of Gastroenterology.[16]

1.2 Standards of Therapy

Currently, there is no cure for Crohn's disease, and the therapeutic goals include inducing and maintaining clinical and endoscopic remission, reducing the need for long-term corticosteroid use, and preventing colon cancer. Several drug classes are used in the treatment of Crohn's disease, including aminosalicylates, immunomodulators (e.g., azathioprine [AZA], cyclosporine, methotrexate [MTX], and 6-mercaptopurine [6-MP]), corticosteroids (e.g., prednisone), tumour necrosis factor (TNF) alpha antagonists (e.g., infliximab and adalimumab), and integrin inhibitors (e.g., vedolizumab).[3] With the exception of the TNF alpha antagonists and vedolizumab, all are commonly referred to as conventional therapies. Medical management is based on a stepwise approach, with treatments used sequentially and escalated to either newer therapies or higher doses as patients fail to respond to each step of treatment.[16] Most drugs have important adverse effects that may have short-term or long-term

consequences.[3] Surgery, including total colectomy and ileostomy, may be considered for patients with serious complications or medically refractory disease.[16]

1.3 Drug

Ustekinumab (Stelara) is a fully human immunoglobulin G1 kappa monoclonal antibody that binds to the shared p40 subunit of interleukin (IL)-12 and IL-23.[4] Ustekinumab is previously 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, alone or in combination with methotrexate.[4]

The current indication under review is for the treatment of adult patients with moderately to severely active Crohn's disease, who have had an inadequate response with, loss of response to, or intolerance to either immunomodulators or one or more TNF antagonists, or have had an inadequate response with, intolerance to, or demonstrated dependence on corticosteroids. The recommended dosage for ustekinumab in the treatment of Crohn's disease is as a single intravenous (IV) induction dose based on body weight (approximating 6 mg/kg) followed by a 90 mg subcutaneous (SC) maintenance dose eight weeks later, then every eight weeks thereafter as maintenance treatment. For some patients (e.g., "those with low inflammatory burden," per the product monograph), an alternative maintenance regimen of 90 mg SC every 12 weeks may be administered at the discretion of the treating physician. Patients who have an inadequate response with 90 mg SC every 12 weeks may be switched to the every eight weeks regimen. Immunomodulators and/or corticosteroids may be continued during treatment with ustekinumab. The product monograph recommends that, in patients who have responded to treatment with ustekinumab, corticosteroids may be reduced or discontinued in accordance with standard of care.[4] The product monograph notes that ustekinumab should be used only by physicians who have sufficient knowledge of the indication for which it is being considered (e.g., Crohn's disease) and who have fully familiarized themselves with the efficacy and safety profile of the drug.[4]

Indication under review
For the treatment of adult patients with moderately to severely active Crohn's disease, who have had an inadequate response, loss of response to, or were intolerant to either immunomodulators or one or more TNF antagonists, or have had an inadequate response, intolerance or demonstrated dependence on corticosteroids.
Reimbursement criteria requested by sponsor
As per indication

The objective of this report is to perform a systematic review of the beneficial and harmful effects of ustekinumab in accordance with the Health Canada–approved indication for the treatment of Crohn's disease.

1.4 Key Comparators

Ustekinumab is the first IL inhibitor approved for the treatment of Crohn's disease in Canada. At the time of this review, there are two TNF antagonists (infliximab and adalimumab) and one integrin inhibitor (vedolizumab) approved for the treatment of Crohn's disease in Canada. Infliximab and vedolizumab are administered via IV infusion only, whereas adalimumab is administered SC only. The Crohn's disease indications for ustekinumab and vedolizumab are limited to adult patients, which is

more restrictive than the indications for infliximab and adalimumab (Table 4). Infliximab currently has the broadest indication for use in the treatment of Crohn's disease, being approved for use in the treatment of adults, children, and patients with fistulizing Crohn's disease. Adalimumab is approved for use in both adults and children with Crohn's disease.

The Health Canada–approved dosage regimens are similar for vedolizumab and infliximab, with administration at weeks 0, 2, and 6 during the induction phase and every eight weeks during maintenance treatment.[17-19] Administration of adalimumab is more frequently during maintenance treatment (i.e., once every two weeks).[20] The dose of infliximab is calculated based on the patient's weight (i.e., 5 mg/kg), whereas the dose of vedolizumab and adalimumab is not adjusted based on the weight of the patient. The product monographs for adalimumab and infliximab indicate that the dosage of these products can be escalated in the event of nonresponse, incomplete response, and/or a disease flare.[18-20] In contrast, the dosage and administration section of the product monograph for vedolizumab does not specify that the dosage can be escalated.[17] The product monograph for ustekinumab indicates that the maintenance administration interval may be shortened to every eight weeks among patients who received the every 12 weeks regimen after induction but did not adequately respond. However, the product monograph for ustekinumab does not specify that the dose or interval may be modified (escalated) beyond 90 mg every eight weeks in patients not meeting treatment goals with this regimen.[4]

1.5 Previous Reviews by CADTH Common Drug Review

Ustekinumab has previously been reviewed twice through the CADTH Common Drug Review (CDR) process 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.[6] The CADTH Canadian Drug Expert Committee (CDEC) recommended that ustekinumab not be reimbursed at the submitted price for the treatment of psoriatic arthritis.[5] Adalimumab for the treatment of moderately to severely active Crohn's disease was reviewed through the CDR process in 2007, and CEDAC recommended that it be reimbursed with clinical criteria and conditions.[21,22] Infliximab (Remicade) has not been reviewed through the CDR process for the treatment of Crohn's disease. However, a subsequent entry biologic of infliximab (Inflectra) was reviewed by CDR, and a subsequent CDEC recommendation that it be reimbursed for the treatment of Crohn's disease, with a clinical criterion and conditions, was issued in 2016.[23]

TABLE 4: KEY CHARACTERISTICS OF USTEKINUMAB, VEDOLIZUMAB, INFLIXIMAB, AND ADALIMUMAB

	Ustekinumab	Vedolizumab	Infliximab	Adalimumab
Mechanism	IL-12 and IL-23 inhibitor	Integrin inhibitor	TNF alpha antagonist	
Indications[a]	**Adult CD** Adult patients with moderately to severely active Crohn's disease who have had an inadequate response with, loss of response to, or intolerance to either conventional therapy (CS or immunomodulators) or one or more TNF antagonist, or who were CS dependent	**Adult CD** Adult patients with moderately to severely active CD who have had an inadequate response with, loss of response to, or intolerance to immunomodulators or a TNF antagonist; or who have had an inadequate response to, intolerance to, or demonstrated dependence on CS	**Adult CD** Reduction of signs and symptoms, induction and maintenance of clinical remission and mucosal healing and reduction of CS use in adults with moderately to severely active CD who have had an inadequate response with a CS and/or aminosalicylate **Pediatric CD** Reduction of signs and symptoms and induction and maintenance of clinical remission in pediatrics with moderately to severely active CD who have had an inadequate response with conventional therapy **Fistulizing CD** Adults with fistulizing CD who have not responded despite conventional treatment	**Adult CD** • Reduction of signs and symptoms and induction and maintenance of clinical remission in adults with moderately to severely active CD who have had an inadequate response with conventional therapy • Reduction of signs and symptoms and induction of clinical remission in adults with moderately to severely active CD who have loss of response to or intolerance to infliximab **Pediatric CD** Reduction of signs and symptoms and induction and maintenance of clinical remission in severely active CD and/or who have had an inadequate response with or intolerance to conventional therapy and/or a TNF antagonist
Administration	IV (induction) and SC (maintenance)	IV		SC
Recommended Dose	• Induction: tiered weight-based dose approximating 6 mg/kg IV at week 0 • Maintenance: 90 mg SC at week 8 and q.8.w. thereafter • Alternative maintenance: 90 mg SC at week 12 and q.12.w. thereafter; may switch to q.8.w. for inadequate response	**Adults (moderate-to-severe CD)** • Induction: 300 mg at weeks 0, 2, and 6 • Maintenance: 300 mg q.8.w. starting at week 6	**Adults (moderate-to-severe CD)** • Induction: 5 mg/kg at weeks 0, 2, and 6 • Maintenance: 5 mg/kg q.8.w.; 10 mg/kg for incomplete responders **Adults (fistulizing CD)** • Induction: 5 mg/kg at weeks 0, 2, and 6 • Maintenance: 5 mg/kg q.8.w. or 10 mg/kg q.8.w. for those with relapse following an initial response	**Adult CD** • Induction: 160 mg at week 0, 80 mg at week 2 • Maintenance: 40 mg q.2.w. beginning at week 4; dose escalation for patients with a disease flare or nonresponse **Pediatrics CD** • Induction: 160 mg at week 0, 80 mg at week 2

	Ustekinumab	Vedolizumab	Infliximab	Adalimumab
			Pediatrics (moderate-to-severe CD) • Induction: 5 mg/kg at weeks 0, 2, and 6 • Maintenance: 5 mg/kg q.8.w.	• Maintenance: 20 mg q.2.w. beginning at week 4, 40 mg q.2.w. for patients with a disease flare or nonresponse
Serious Side Effects / Safety Issues	• Infections and reactivation of latent infections • Administration site reactions • Malignancy	• Serious infections • Infusion and serious allergic reactions	• Serious infections • Malignancy • Infusion and serious allergic reactions	• Serious infections • Malignancies, particularly lymphoma • Administration-site reactions

CD = Crohn's disease; CS = corticosteroids; IL = interleukin; IV = intravenous; SC = subcutaneous; TNF = tumour necrosis factor; q.2.w. = every two weeks; q.8.w. = every eight weeks; q.12.w. = every 12 weeks.
[a] Health Canada indication.
Source: Ustekinumab product monograph,[4] vedolizumab product monograph,[24] infliximab product monograph,[18,19] and adalimumab product monograph.[20]

2. OBJECTIVES AND METHODS

2.1 Objectives

To perform a systematic review of the beneficial and harmful effects of ustekinumab for the induction and maintenance treatment of adult patients with moderately to severely active Crohn's disease who have had an inadequate response with, loss of response to, or intolerance to either conventional therapy (corticosteroids or immunomodulators) or one or more TNF alpha antagonist, or who were corticosteroid dependent.

2.2 Methods

All manufacturer-provided trials considered to be pivotal were included in the systematic review. Phase III studies were selected for inclusion based on the selection criteria presented in Table 5.

TABLE 5: INCLUSION CRITERIA FOR THE SYSTEMATIC REVIEW

Patient Population	Adult patients with moderately to severely active Crohn's disease who have had an inadequate response with, loss of response to, or intolerance to either conventional therapy (corticosteroids or immunomodulators) or one or more TNF alpha antagonist, or who were corticosteroid dependent.
	
	Subgroups:
	• Disease severity at baseline
	• No prior experience with a TNF alpha antagonist or conventional therapy
	• Previous therapy with a TNF alpha antagonist or conventional therapy:
	  ○ Failure or intolerance to a TNF alpha antagonist or conventional therapy
	  ○ No failure or intolerance to a TNF alpha antagonist or conventional therapy.

Intervention	Ustekinumab single IV induction infusion based on body weight at week 0; followed by 90 mg SC 8 weeks later, then every 8 or 12 weeks thereafter. Induction recommended IV dose (approximately 6 mg/kg): • 260 mg if body weight is ≤ 55 kg • 390 mg if body weight is > 55 kg to ≤ 85 kg • 520 mg if body weight is > 85 kg.
Comparators	• Adalimumab • Infliximab • Vedolizumab

Outcomes	**Efficacy outcomes:**
	• Clinical remission (e.g., using the CDAI score)[a]
	• Clinical response (e.g., using the CDAI score)
	• Health-related quality of life, functional and disability outcomes (e.g., IBDQ, SF-36) [a]
	• Mucosal healing determined by histology or endoscopy
	• Need for surgery for Crohn's disease [a]
	
	Harms outcomes:
	• Mortality
	• SAEs
	• WDAEs
	• AEs, including but not limited to
	  ○ Infusion or injection-site reactions
	  ○ Hypersensitivity reactions

	○ Infections ○ Malignancies ○ Major cardiovascular events ○ Neurological AEs (new onset or exacerbation of MS, PN, or GBS)
Study Design	Published and unpublished phase III RCT

AE = adverse events; CDAI = Crohn's Disease Activity Index; GBS = Guillain–Barré syndrome; IBDQ = Inflammatory Bowel Disease Questionnaire; IV = intravenous; MS = multiple sclerosis; PN = peripheral neuropathy; RCT = randomized controlled trial; SAE = serious adverse events; SC = subcutaneous injection; SF-36 = Short Form (36) Health Survey; TNF = tumour necrosis factor; WDAE = withdrawal due to adverse events.
[a] These outcomes were identified as being of particular importance to patients in the input received by CADTH from patient groups.

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

Published literature was identified by searching the following bibliographic databases: MEDLINE (1946–) with in-process records and daily updates via Ovid; Embase (1974–) via Ovid; and PubMed. The search strategy consisted of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were Stelara (ustekinumab) and Crohn's disease.

No filters were applied to limit retrieval by study type. Retrieval was not limited by publication year 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 October 7, 2016. Regular alerts were established to update the search until the meeting of the CADTH Canadian Drug Expert Committee (CDEC) on February 15, 2017. Regular search updates were performed on databases that do not provide alert services.

Grey literature (literature that is not commercially published) was identified by searching relevant websites from the following sections of the *Grey Matters* checklist (https://www.cadth.ca/grey-matters): health technology assessment agencies, health economics, clinical practice guidelines, drug and device regulatory approvals, advisories and warnings, drug class reviews, databases (free) and Internet search. Google and other Internet search engines were used to search for additional Web-based materials. These searches were supplemented by reviewing the bibliographies of key papers and through contacts with appropriate experts. In addition, the manufacturer of the drug was contacted for information regarding unpublished studies.

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. Included studies are presented in Table 6, Table 7, and Table 8; excluded studies (with reasons) are presented in 0.

3. RESULTS

3.1 Findings from the Literature

A total of four studies were identified from the literature for inclusion in the systematic review (Figure 1). The included studies are summarized in Table 6 and described in Section 3.2. A list of excluded studies is presented in 0.

FIGURE 1: FLOW DIAGRAM FOR INCLUSION AND EXCLUSION OF STUDIES

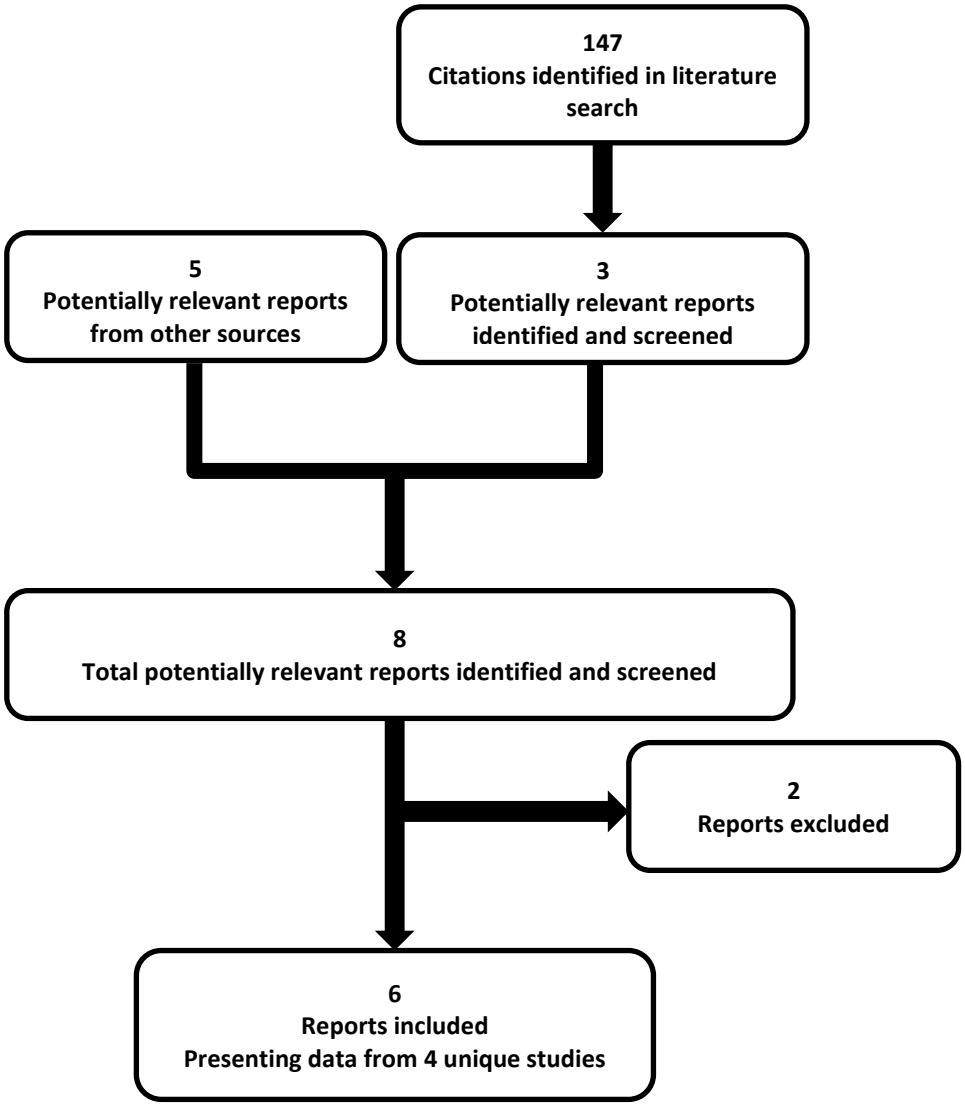


TABLE 6: DETAILS OF INCLUDED INDUCTION STUDIES

		UNITI-1	UNITI-2
DESIGNS & POPULATIONS	**Study Design**	Phase III superiority DB RCT	
	Locations	North America (including Canada), Europe, Asia, the Asia-Pacific region, South Africa, and Brazil (178 sites)	North America (including Canada), Europe, Asia, the Asia-Pacific region, South Africa, and Brazil (175 sites)
	Randomized (N)	N = 769 randomized N = 741 after study restart[a]	N = 640 N = 628 after study restart[a]
	Inclusion Criteria	• Age ≥ 18 • Moderate-to-severe active[b] CD or fistulizing CD of ≥ 3 months' duration with colitis, ileitis, or ileocolitis (radiographic, histologic, and/or endoscopic confirmation) • Inadequate response or intolerance[d] to ≥ 1 TNF antagonist(s)[e] • Meet criteria for concomitant medication stability, screening laboratory test results, and TB history and testing results	• Age ≥ 18 • Moderate-to-severe active[c] CD or fistulizing CD of ≥ 3 months' duration with colitis, ileitis, or ileocolitis (radiographic, histologic, and/or endoscopic confirmation) • Receiving CS and/or MTX, AZA, or 6-MP • Inadequate response or intolerance[d] to conventional therapy (CS and/or MTX, AZA, or 6-MP) or is CS dependent • Not had inadequate response or intolerance to ≥ 1 TNF antagonist(s)[d,e] • Meet criteria for concomitant medication stability, screening laboratory test results, and TB history and testing results
	Exclusion Criteria	• CD complications requiring surgery or precluding use of CDAI to assess response • Intra-abdominal abscess within 8 weeks of randomization • Bowel resection or diversion within 6 months, or other intra-abdominal surgery within 3 months of randomization • Draining stoma or ostomy • Positive test for enteric pathogens within 4 months of randomization • Previous treatment with: IL-12 or IL-23 inhibitor (e.g., UST or BRI) • Received parenteral CS within 3 weeks, immunomodulators other than AZA, 6-MP, or MTX within 6 weeks, biologics within 8 weeks, or TPN within 3 weeks Active or latent TB, opportunistic infection, HIV, or hepatitis B or C infection	
DRUGS	**Intervention**	A single dose at week 0 of: UST IV infusion 130 mg or UST IV infusion (approximately 6 mg/kg): • 260 mg if body weight is ≤ 55 kg • 390 mg if body weight is > 55 to ≤ 85 kg • 520 mg if body weight is > 85 kg	
	Comparator(s)	Placebo	
DURATION	**Phase**		
	Run-in	5 weeks (used for screening)	
	Double-blind	20 weeks • Week 6 primary outcome assessment • Week 8 assessment for entry to the maintenance-treatment study (IM-UNITI) • Week 9 to 20 follow-up for those not entering the maintenance-treatment study	
	Follow-up	44-week maintenance study (IM-UNITI)	

		UNITI-1	UNITI-2
OUTCOMES	**Primary End Point**	Clinical response at 6 weeks	
	Other End Points	• Clinical remission at week 8 • IBDQ and SF-36 at week 6 • WLQ, productivity VAS, time lost from work • CD-related surgery • AE, SAE, WDAE	
NOTES	**Publications**	Feagan et al. 2016[25]	

6-MP = 6-mercaptopurine; AE = adverse event; AZA = azathioprine; BRI = briakinumab; CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CRP = C-reactive protein; CS = corticosteroid; DB = double-blind; IBDQ = Inflammatory Bowel Disease Questionnaire; IL = interleukin; IV = intravenous; MTX = methotrexate; RCT = randomized controlled trial; SAE = serious adverse event; SF-36 = Short Form (36) Health Survey; TB = tuberculosis; TNF = tumour necrosis factor; TPN = total parenteral nutrition; UST = ustekinumab; VAS = visual analogue scale; WDAE = withdrawal due to adverse event; WLQ = Work Limitations Questionnaire.

[a] Twenty-eight patients in UNITI-1 and 12 patients in UNITI-2 were excluded following a study protocol amendment (see Section 3.2.1).

[b] Active disease was defined as a CDAI score ≥ 220 but ≤ 450 points.

[c] Active disease was defined as a CDAI score ≥ 220 but ≤ 450 points and at least one of the following: 1) an abnormal CRP (> 3.0 mg/L); 2) fecal calprotectin > 250 mg/kg at screening; 3) endoscopy (within 3 months before baseline) with evidence of active Crohn's disease during the current disease flare.

[d] Inadequate response and intolerance to TNF antagonists as defined in the study are presented in Table 9 of this report. Definitions for inadequate response and intolerance to CS and/or immunomodulators were not provided.

[e] Protocol-specified TNF antagonists: infliximab, adalimumab, or certolizumab pegol at approved doses.

Source: Clinical Study Reports for UNITI-1 and UNITI-2.[7,8]

TABLE 7: DETAILS OF INCLUDED MAINTENANCE STUDY

		IM-UNITI
DESIGNS & POPULATIONS	**Study Design**	Phase III superiority DB RCT
	Locations	North America (including Canada), Europe, Asia, the Asia-Pacific region, South Africa, and Brazil (137 sites)
	Randomized (N)	N = 1,281 enrolled from induction studies N = 397 randomized
	Inclusion Criteria	• Age ≥ 18 years • Moderate-to-severe active CD in clinical response to IV UST at week 8 of induction studies (UNITI-1 and UNITI-2)
	Exclusion Criteria	• Specific changes to patients' concomitant medications due to CD (i.e., lack of efficacy) since week 0 of induction studies[a] • Initiated protocol-prohibited medication since week 0 of induction studies • CD-related surgery since week 0 of induction studies • Signs or symptoms, or diagnosis of any medical condition which would have precluded enrolment in induction studies
DRUGS	**Intervention**	UST 90 mg SC q.12.w. (with last dose at week 36) UST 90 mg SC q.8.w. (with last dose at week 40)
	Comparator(s)	Placebo
DURATION	**Phase**	
	Run-in	Not applicable
	Double-blind	44 weeks
	Follow-up	Extension study database lock at week 272

		IM-UNITI
Outcomes	**Primary End Point**	Clinical remission (CDAI score < 150 points) at week 44
	Other End Points	Major secondary outcomes: • Clinical response (reduction from week 0 and induction study in CDAI score ≥ 100 points) at week 44 • Clinical remission at week 44 among patients in clinical remission to UST at week 0 • Corticosteroid-free remission at week 44 • Clinical remission at week 44 in the subset of patients who were refractory or intolerant to TNF antagonist therapy (i.e., patients from UNITI-1) Other: • IBDQ • SF-36 • WLQ, productivity VAS, time lost from work • Mucosal healing (endoscopic) • CD-related surgery • AEs, SAEs, WDAEs
Notes	**Publications**	Feagan et al. 2016[25]

AE = adverse event; CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; DB = double-blind; IBDQ = Inflammatory Bowel Disease Questionnaire; IV = intravenous; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; RCT = randomized controlled trial; SAE = serious adverse event; SC = subcutaneous; SF-36 = Short Form Health Survey; TNF = tumour necrosis factor; UST = ustekinumab; VAS = visual analogue scale; WDAE = withdrawal due to adverse event; WLQ = Work Limitations Questionnaire.

a Changes to concomitant medications: increase in daily dose of oral corticosteroids of > 5 mg of prednisone (or equivalent increase in prednisone-equivalent dose of other corticosteroids), initiation of oral budesonide or increase in daily dose, initiation of parenteral or oral corticosteroids for CD except for dose equivalent substitutions among oral corticosteroids, and initiation or increased daily dose of MTX, 6-MP, or AZA, except for dose equivalent substitutions.

Source: Clinical Study Report for IM-UNITI.[9]

Table 8: Details of Included Phase II Study

		CERTIFI
Designs & Populations	**Study Design**	Phase IIb dose-ranging DB RCT
	Locations	North America (including Canada), Europe, and Australia (153 sites)
	Randomized (N)	N = 526
	Inclusion Criteria	• Age ≥ 18 • Moderate-to-severe active[a] CD or fistulizing CD of ≥ 3 months' duration with colitis, ileitis, or ileocolitis (radiographic and/or endoscopic confirmation) • Inadequate response or intolerance to ≥ 1 TNF antagonist(s)[b] • Meet criteria for: concomitant medication stability, screening laboratory test results, and TB history and testing results
	Exclusion Criteria	• CD complications requiring surgery or preclude use of CDAI to assess response • Intra-abdominal abscess within 8 weeks of randomization • Bowel resection or diversion within 6 months of randomization, or other intra-abdominal surgery within 3 months of randomization • A draining stoma or ostomy • Positive for enteric pathogens within 4 months of randomization • Previous treatment with: IL-12 or IL-23 inhibitor • Received parenteral CS within 3 weeks; immunomodulators other than AZA, 6-MP, or

		CERTIFI
		MTX within 8 weeks; biologics within 8 weeks; randomization or TPN within 2 weeks of screening • Active or latent TB, or opportunistic infection
DRUGS	**Intervention**	**Induction phase weeks 0 to 6 (single dose at week 0):** • UST 1 mg/kg IV • UST 3 mg/kg IV • UST 6 mg/kg IV **Maintenance phase weeks 8 to 36:** Originally randomized to UST • UST 90 mg SC (doses at week 8 and 16) Originally randomized to placebo and not in clinical response at week 6 • UST 270 mg SC (week 8) and then UST 90 mg SC (week 16)
	Comparator(s)	Placebo (all phases)
DURATION	**Phase**	
	Run-in	Duration not reported
	Double-blind	36 weeks • Induction phase weeks 0 to 6 • Maintenance phase weeks 8 to 36
	Follow-up	3 days after any final visit for AE information
OUTCOMES	**Primary End Point**	Clinical response at 6 weeks
	Other End Points	**Major secondary outcomes:** • Clinical remission at week 6 • Clinical remission at week 22 • Clinical response at week 22 **Other secondary outcomes:** • Mucosal healing (endoscopic) at week 6 • IBDQ at week 6 • Clinical remission at weeks 28 and 36 • Clinical remission at every visit through weeks 22, 28, and 36 • Clinical response at weeks 28 and 36 • Clinical response at every visit through weeks 22, 28, and 36 • Productivity VAS, time lost from work • CD-related surgery • AEs, SAEs, WDAEs
NOTES	**Publications**	Sandborn et al. 2012[11]

6-MP = 6-mercaptopurine; AE = adverse event; AZA = azathioprine; CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CS = corticosteroid; DB = double-blind; IBDQ = Inflammatory Bowel Disease Questionnaire; IL = interleukin; IV = intravenous; MTX = methotrexate; RCT = randomized controlled trial; SAE = serious adverse event; SC = subcutaneous; TB = tuberculosis; TNF = tumour necrosis factor; TPN = total parenteral nutrition; UST = ustekinumab; VAS = visual analogue scale; WDAE = withdrawal due to adverse event.

Note: One additional report was included (CDR submission[26]). One additional report, Feagan et al.,[25] was identified after the initial literature search was conducted during regular literature search updates.

[a] Active disease was defined as a CDAI score ≥ 220 but ≤ 450 points.

[b] Inadequate response and intolerance as defined in the study are presented in Table 9 of this report.

[c] Protocol-specified TNF antagonists: infliximab, adalimumab, or certolizumab pegol at approved doses.

Source: Clinical Study Report for CERTIFI.[10]

3.2 Included Studies

3.2.1 Description of Studies

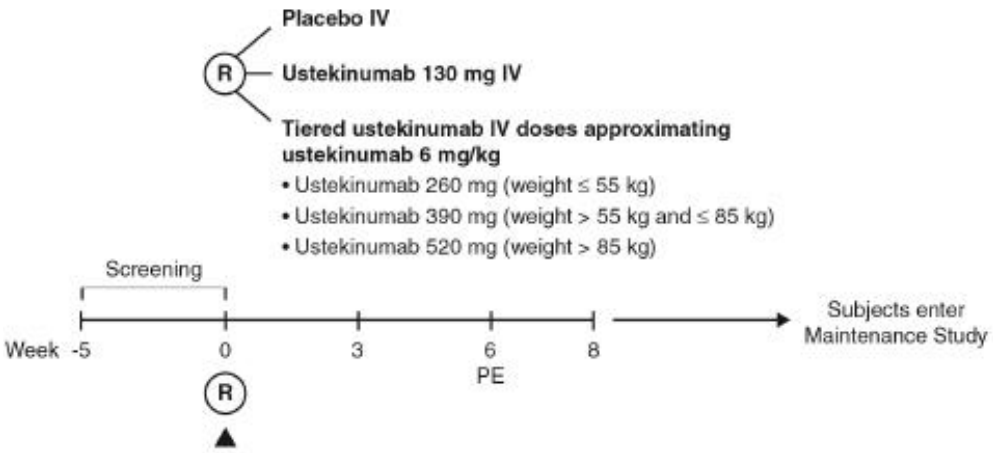
The CDR review included four, multi-centre, multinational, double-blind, randomized, placebo-controlled trials: two phase III induction-treatment studies, UNITI-1 (N = 769) and UNITI-2 (N = 640); one phase III maintenance-treatment study, IM-UNITI (N = 397); and one phase II induction and maintenance study, CERTIFI (N = 526). The phase III studies were designed as superiority studies, whereas the phase II study was a dose-ranging study. All four studies were submitted to CDR by the manufacturer as pivotal studies.

a) Induction Studies

UNITI-1 and UNITI-2 were designed — with identical protocols — to evaluate the efficacy and safety of IV induction regimens of ustekinumab in inducing clinical response in patients with moderately to severely active Crohn's disease. However, the studies differed with respect to patients' previous Crohn's disease treatment experience: UNITI-1 included patients who had had an inadequate response or were intolerant to one or more TNF antagonist therapies, whereas UNITI-2 included patients who had had an inadequate response or were intolerant to conventional therapy only (i.e., corticosteroids or immunomodulators such as 6-MP, AZA, and MTX). Patients in UNITI-2 could have previously received TNF antagonists but could not have failed treatment. Detailed descriptions of the populations are located in Section 3.2.2 of this review.

Patients were randomized in a 1:1:1 ratio to receive a single IV administration of either placebo or one of two induction doses of ustekinumab at week 0, as shown in Figure 2. Patients were allocated to a treatment group using a permuted-block randomization with study region (Asia, Eastern Europe, or rest of world) and Crohn's Disease Activity Index (CDAI) score (≤ 300 or > 300 points) as the stratification variables. In UNITI-1, randomization was further stratified by initial response to TNF antagonist therapy (yes or no). For patients who had previously received multiple TNF antagonist therapies, their initial response status (yes or no) was determined by whether they had initially responded to the first TNF antagonist therapy received. Allocation to treatment group was done using a central randomization centre by means of an interactive voice response system (IVRS) and/or interactive web response system (IWRS).

FIGURE 2: STUDY DESIGN SCHEMATIC FOR INDUCTION STUDIES UNITI-1 AND UNITI-2



IV = intravenous; PE = primary end point; R = randomization.
Note: The triangle indicates study drug administration.
Source: Clinical Study Reports for UNITI-1 and UNITI-2..[7,8]

Patients randomized to ustekinumab induction therapy who achieved clinical response were eligible to enter the maintenance-treatment study, IM-UNITI, at week 8. Patients who did not enter the maintenance study continued to be followed as part of the induction studies and had a safety follow-up visit at week 20.

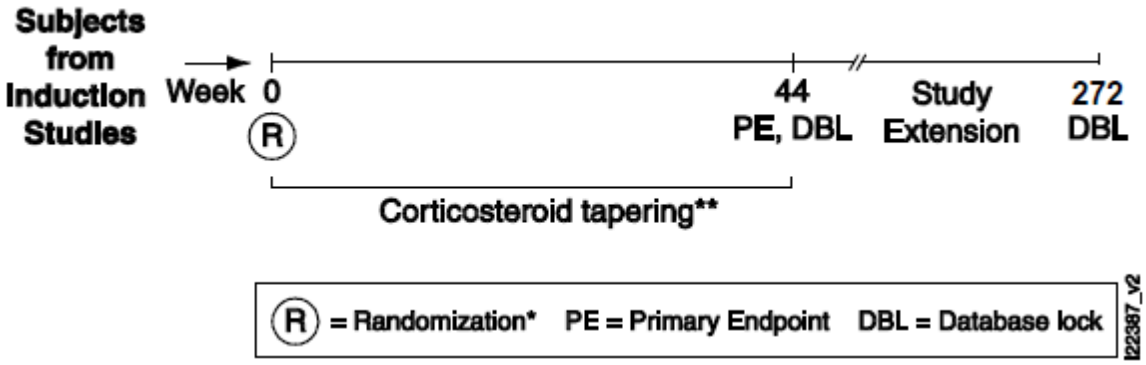
After 28 (UNITI-1) and 12 (UNITI-2) patients had been randomized to treatments in the induction studies, the manufacturer temporarily suspended drug administration to patients because of a stability issue with the IV formulation of ustekinumab (130 mg ustekinumab in 26 mL [5 mg/mL; 27 mL fill of liquid]) used in the studies. The manufacturer substituted the 90 mg ustekinumab formulation (which is already approved for SC injection for other indications) for the protocol-specified IV induction administrations. The protocols for the induction (and maintenance) studies were amended to incorporate the use of the 90 mg formulation. Data from the 28 (UNITI-1) and 12 (UNITI-2) patients who were randomized before the study was temporarily stopped were not used in the planned analyses because knowledge of the stability issue could potentially bias the assessments. The manufacturer restarted the randomization with new blocks for each stratum when the study was restarted.

The induction studies ended on the date either the last patient who entered the maintenance study completed the week 8 visit or the last patient who did not enter the maintenance study completed his or her final safety visit at week 20, whichever occurred later. Once the study ended and the database was locked, selected manufacturer personnel were unblinded to induction-treatment assignment, although exactly which personnel this referred to was not specified. The clinical study reports indicate that, in order to protect the integrity of the maintenance study, treatment assignment blinding in the induction studies was maintained for sites, site monitors, and patients until the week 44 analyses for the maintenance study were completed.

b) Maintenance Study

The IM-UNITI study was designed to evaluate the efficacy (clinical remission) and safety of two SC maintenance regimens of ustekinumab (90 mg every eight weeks or every 12 weeks) in patients with moderately to severely active Crohn's disease who had a clinical response with ustekinumab in the induction studies, UNITI-1 and UNITI-2. Figure 3 shows the study design.

FIGURE 3: STUDY DESIGN SCHEMATIC OF MAINTENANCE STUDY IM-UNITI



* Patients in clinical response to ustekinumab induction dosage regimen.
** Required for patients in clinical response.
Source: Clinical Study Report for IM-UNITI.[9]

Patients who had a clinical response to ustekinumab induction at week 8 in either UNITI-1 or UNITI-2 were randomly assigned in a 1:1:1 ratio to one of three treatment groups (placebo, ustekinumab 90 mg SC every 12 weeks, or ustekinumab 90 mg SC every eight weeks) based on computer-generated randomization schedule defined a priori. Permuted-block randomization with stratification factors of clinical remission at week 0 (yes or no) and ustekinumab induction dose (130 mg or tiered dose approximating 6 mg/kg ustekinumab) were used. Patients who did not achieve clinical response with ustekinumab induction therapy, as well as all patients who were randomized to placebo, irrespective of whether they achieved clinical response in the UNITI studies, were also eligible to enter IM-UNITI at week 8; however, these patients were not randomized to the primary efficacy population. An IVRS/IWRS dictated the treatment assignment for each patient.

As mentioned, the induction studies were temporarily suspended because of a stability issue with the IV formulation of ustekinumab. Consequently, the maintenance study was also temporarily suspended. A total of 40 patients had been randomized in the induction studies before the studies were temporarily suspended. Because knowledge of the stability issue could bias the assessments, data from nine randomized patients enrolled in IM-UNITI were excluded from the efficacy analyses.

The duration of the double-blind maintenance treatment period of IM-UNITI was 44 weeks, meaning that patients randomized to treatment in the induction studies and treated to the end of IM-UNITI had a total of 52 weeks of treatment. Eligible patients at week 44 could continue into the long-term extension phase and be followed to week 220 (see Appendix 6 for preliminary details of this study).

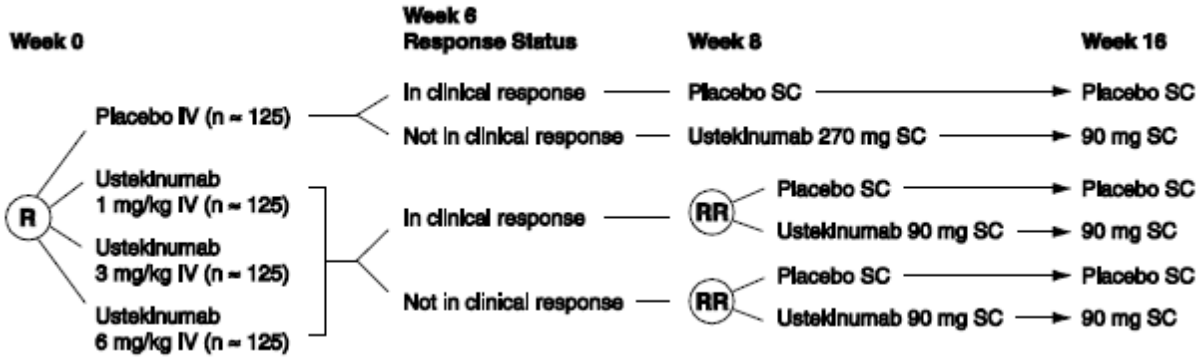
c) Phase II Study

CERTIFI was a dose-ranging study that evaluated the efficacy and safety of ustekinumab induction and maintenance regimens versus placebo in patients with moderately to severely active Crohn's disease who had received treatment with one or more TNF antagonists and who had not responded initially to therapy, who had responded and then lost response to therapy, or who were intolerant to therapy at a dose approved for Crohn's disease.

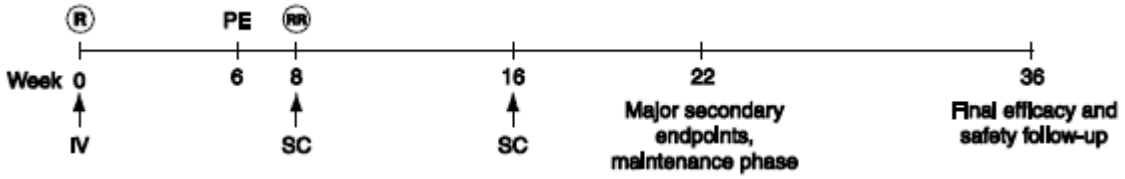
The study design is depicted in Figure 4. Patients were randomized 1:1:1:1 to placebo or one of three IV ustekinumab induction doses: 1 mg/kg, 3 mg/kg, or 6 mg/kg. Randomizations were performed using a central randomization centre and IVRS. Patients were randomized to a treatment regimen using the adaptive randomization procedure of Pocock and Simon,[27,28] with study site and initial response to TNF antagonist therapy (yes or no) as the stratification variables. For patients who had multiple TNF antagonist therapies, their initial response status to TNF antagonist therapy (yes or no) was determined by whether they initially responded to the first TNF antagonist therapy they had received.

FIGURE 4: STUDY DESIGN SCHEMATIC FOR CERTIFI

(a) Study treatments



(b) Key time points



IV = intravenous; SC = subcutaneous; PE = primary end point; R = randomization; RR = randomization only for patients receiving ustekinumab induction therapy.
Note: the up arrow indicates study drug administration points.
Source: Clinical Study Report for CERTIFI.[10]

CDR considered CERTIFI a supportive study because of several serious limitations, including the adaptive and dose-finding design, and the mix of patients who received various induction regimens in the maintenance phase (Section 3.5 Critical Appraisal). Therefore, only clinical response and remission outcomes from the induction phase of CERTIFI (using the Health Canada–approved induction regimen, ustekinumab 6 mg/kg IV) are presented and interpreted in this review.

3.2.2 Populations

a) Inclusion and Exclusion Criteria

Induction Studies

Inclusion and exclusion criteria for the induction studies are summarized in Table 6. As previously mentioned, the main difference between the patient populations in the two induction studies was their experience with previous treatments for Crohn's disease. UNITI-1 included patients who had received TNF antagonist therapies (specifically infliximab, adalimumab, or certolizumab pegol at a dose approved for the treatment of Crohn's disease) and who did not respond initially (primary nonresponse), who responded initially but then lost response (secondary nonresponse), or who were intolerant to the medication. Conversely, UNITI-2 enrolled patients who had failed conventional therapy, and had not previously demonstrated inadequate response or intolerance to one or more TNF antagonist therapies (i.e., infliximab, adalimumab, or certolizumab pegol). Table 9 provides definitions used to identify eligible patients for the UNITI studies based on TNF antagonist failure and/or intolerance.

Key exclusion criteria for both studies were complications of Crohn's disease that required surgery or precluded use of CDAI to assess response; history of bowel resection or diversion or any other intra-

abdominal surgery within specified time periods; previous treatment with an IL-12 or IL-23 inhibitor, received IV corticosteroids, immunomodulators other than AZA, 6-MP, or MTX, biologics, or total parenteral nutrition within specified time periods; or active or latent tuberculosis (TB); opportunistic infection; HIV; or hepatitis B or C infection.

TABLE 9: DEFINITIONS OF INADEQUATE INITIAL RESPONSE, LOSS OF RESPONSE, OR INTOLERANCE TO TUMOUR NECROSIS FACTOR ANTAGONIST THERAPIES IN THE CLINICAL TRIALS

Inadequate Initial Response (Primary Nonresponse)[a]	Loss of Response (Secondary Nonresponse)[a]	Intolerance[a]
Received induction doses of: Infliximab (2 or 3 doses of ≥ 5 mg/kg) or Adalimumab (at a dose of 160 mg followed by a dose ≥ 80 mg **or** at a dose of 80 mg followed by a dose ≥ 40 mg) or Certolizumab pegol (2 or 3 doses of ≥ 400 mg) **Presence of ≥ 1 persistent CD signs or symptoms[b,c]:** • Lack of improvement or worsening in stool frequency • Lack of improvement or worsening in daily abdominal pain • Occurrence, lack of improvement, or worsening of fever thought to be related to CD • Recurring drainage from a previously nondraining fistula or development of a new draining fistula • Lack of improvement or worsening in rectal bleeding • Initiation or increase in antidiarrheal medication. **Provide documentation that:** • Provides the dates and doses of the failed TNF antagonist induction therapy • Consists of medical records, referring physician letter, or other "reason for referral" documents (e.g., insurance authorization form) indicating that the patient had persistent	**Initially responded to induction therapy** **Received ≥ 2 maintenance doses of:** Infliximab (at a dose of ≥ 5 mg/kg) or Adalimumab (at a dose of ≥ 40 mg) or Certolizumab pegol (at a dose of ≥ 400 mg) **Presence of ≥ 1 persistent CD signs or symptoms[b,c]:** • Worsening in stool frequency • Worsening in daily abdominal pain • Occurrence or worsening in fever thought to be related to CD • Recurring drainage from a previously nondraining fistula or development of a new draining fistula • Worsening in rectal bleeding • Initiation or increase in antidiarrheal medication. **Provide documentation that:** • Provides the dates and doses of the failed TNF antagonist induction therapy • Consists of medical records, referring physician letter, or other "reason for referral" documents (e.g., insurance authorization form) indicating that the patient had persistent CD activity following TNF antagonist therapy.	**AE that meets 1 of the following 3 criteria:** • Significant acute infusion or administration reaction • Significant delayed infusion or administration reaction (e.g., delayed hypersensitivity or serum-sickness like reaction) • Significant injection-site reaction. **AEs also must have followed ≥ 1 dose of TNF antagonist and, in the treating physician's opinion, precluded continued use of the therapy.** **Provide documentation that:** • Provides the date of discontinuation of TNF antagonist therapy • Consists of medical records, referring physician letter, or other "reason for referral" documents (e.g., insurance authorization form) indicating that the patient had intolerance to TNF antagonist therapy.

Inadequate Initial Response (Primary Nonresponse)[a]	Loss of Response (Secondary Nonresponse)[a]	Intolerance[a]
CD activity following TNF antagonist therapy.		

AE = adverse event; CD = Crohn's disease; TNF = tumour necrosis factor.
[a] Eligible patients had to satisfy all criteria (bolded).
[b] Assessed by a treating physician.
[c] These signs and symptoms of CD must have occurred ≥ 2 weeks after receiving the last induction dose (or maintenance dose in the case of lost response) of TNF antagonist.
Source: UNITI-1, UNITI-2, and CERTIFI study protocols.[28-31]

Maintenance Study

The randomized population consisted of adult patients with a history of moderately to severely active Crohn's disease who had a clinical response to IV ustekinumab induction therapy (at week 8 of the induction studies, UNITI-1, and UNITI-2).

Patients were excluded from IM-UNITI if there had been specific changes to their concomitant medications due to Crohn's disease (i.e., lack of efficacy) since week 0 of the induction studies (see footnote a, Table 7 for definition), or if they initiated protocol-prohibited medication or underwent Crohn's disease–related surgery since week 0 of induction studies, or had signs or symptoms, or diagnosis of any medical condition that would have precluded enrolment in the induction studies.

Phase II Study

Patients were eligible for inclusion in CERTIFI if they were adults aged 18 years or older with active Crohn's disease (defined as a baseline CDAI score of ≥ 220 and ≤ 450) or fistulizing Crohn's disease of at least three months' duration, with colitis, ileitis, or ileocolitis, confirmed by radiography and/or endoscopy (Table 8). As in UNITI-1, patients had to have received infliximab, adalimumab, or certolizumab pegol at a dose approved for the treatment of Crohn's disease and they did not respond initially, they responded initially but then lost response with continued therapy, or they were intolerant to the medication (Table 9).

Patients were excluded for similar reasons and criteria as in the induction studies.

b) Baseline Characteristics

Of note, the manufacturer defined baseline as the time of randomization in the UNITI and CERTIFI studies. For IM-UNITI, baseline for most efficacy analyses was defined as the time of randomization (IM-UNITI week 0/week 8 of the UNITI studies); however, for the presentation of baseline characteristics and concomitant medication use (and certain efficacy analyses; Section 3.2.5 Statistical Analysis), baseline was defined as the baseline in the UNITI studies.

Induction Studies

Key baseline demographic and Crohn's disease characteristics from the induction phase studies are summarized in Table 10. [redacted] There were more women than men enrolled in the two studies (range: 52% to 59% women); higher proportions of women were randomized to ustekinumab treatment groups than to placebo. The mean age was lower in UNITI-1 (37.3 years in both groups) than in UNITI-2 (placebo: 38.4 years; ustekinumab: 40.2 years), [redacted]

The mean duration of disease was approximately two to four years longer in UNITI-1 than in UNITI-2. Mean baseline CDAI scores were lower in UNITI-2 (302.2 points in both groups) compared with UNITI-1 (placebo: 319.0 points; ustekinumab: 327.6 points). The same pattern was observed for baseline median C-reactive protein (CRP) levels, which ranged from 8.5 mg/L to 9.9 mg/L in UNITI-1 and 7.8 mg/L to 8.5 mg/L in UNITI-2. A majority of patients in both studies had disease activity in both the ileum and the colon, although the proportions were lower in UNITI-2 (56% to 61%) than in UNITI-1 68% to 69%). More patients in UNITI-1 than in UNITI-2 had a history of Crohn's disease complications, such as intra-abdominal abscess in 14% to 15% (versus 11% to 12% in UNITI-2), current or prior sinus tracts or perforation in 7% to 9% (versus 4% to 6% in UNITI-2), current or prior fistulizing disease in 45% to 51% (versus 35% to 37% in UNITI-2), and current or past stricturing in 44% to 46% (versus 28% to 35% in UNITI-2). Half of the patients in both studies (UNITI-1: 49% to 51%; UNITI-2: 56% to 57%) had at least one extra-intestinal manifestation of Crohn's disease, with the most prevalent being arthritis or arthralgia (data not shown).

TABLE 10: SUMMARY OF BASELINE CHARACTERISTICS FROM THE INDUCTION STUDIES

Parameter	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
Sex, n (%)				
Female	129 (52)	148 (59)	111 (53)	119 (57)
Age (years), mean (SD)	37.3 (11.8)	37.3 (12.5)	40.2 (13.1)	38.4 (13.1)
Body weight (kg), mean (SD)	71.5 (17.7)	69.5 (19.5)	74.0 (19.9)	71.9 (18.8)
BMI (kg/m^2), mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Geographic region, n (%)				
Asia	[redacted]	[redacted]	[redacted]	[redacted]
Eastern Europe	[redacted]	[redacted]	[redacted]	[redacted]
Rest of world	[redacted]	[redacted]	[redacted]	[redacted]
Current smoker, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Duration of CD (years)				
mean (SD)	12.1 (8.4)	12.7 (9.2)	10.4 (9.8)	8.7 (8.4)
CDAI, mean (SD)	319.0 (59.7)	327.6 (62.0)	302.2 (61.7)	302.2 (58.9)
CRP (mg/L), median (IQR)	8.5 (3.4 to 21.9)	9.9 (3.7 to 26.6)	8.5 (3.2 to 21.7)	7.8 (3.8 to 21.1)

Parameter	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
Disease localization, n (%)				
Ileum only	28 (11)	37 (15)	44 (21)	49 (23)
Colon only	48 (20)	40 (16)	37 (18)	43 (21)
Ileocolonic	166 (68)	171 (69)	129 (61)	117 (56)
Proximal GI tract	45 (18)	54 (22)	32 (15)	29 (14)
Perianal	107 (44)	107 (43)	57 (27)	61 (29)
CD complications, n (%)				
Intra-abdominal abscess (past)	34 (14)	38 (15)	25 (12)	22 (11)
Sinus tracts or perforation[a]	16 (7)	22 (9)	12 (6)	9 (4)
Fistula[a]	127 (51)	112 (45)	77 (37)	74 (35)
Current	53 (22)	47 (19)	33 (16)	31 (15)
Bowel stricturing[a]	108 (44)	115 (46)	74 (35)	58 (28)
Current	20 (8)	23 (9)	25 (12)	16 (8)
Extra-intestinal manifestations, n (%)	121 (49)	128 (51)	120 (57)	116 (56)

BMI = body mass index; CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CRP = C-reactive protein; GI = gastrointestinal; IQR = interquartile range; PLA = placebo; SD = standard deviation; UST = ustekinumab;
a Current or past.
Source: Clinical Study Reports for UNITI-1 and UNITI-2.[7,8]

At least 70% of patients were receiving one or more concomitant medications for Crohn's disease at baseline of UNITI-1 and UNITI-2 (Table 11). The proportion of patients receiving corticosteroids (including budesonide) was similar in both studies (approximately 44%), except in the placebo group in UNITI-2 (36%). Approximately one-third of patients in both trials were receiving immunomodulators (AZA, 6-MP, or MTX) at baseline. [redacted] In general, the proportions of patients receiving each class of Crohn's disease medication at baseline were balanced across the treatment groups in both studies.

Prior exposure to Crohn's disease treatments for the induction studies is also summarized in Table 11. In accordance with the study designs, prior TNF antagonist exposure was reported for essentially all patients in UNITI-1 and approximately one-third of the patients in UNITI-2. Of the patients in UNITI-1 with prior exposure to a TNF antagonist, inadequate initial response with, loss of response to, or intolerance to one TNF antagonist was reported for approximately one-half of patients in both treatment groups; slightly smaller proportions of patients (placebo: 44%; ustekinumab: 41%) reported inadequate initial response with, loss of response to, or intolerance to two TNF antagonists. In keeping with the inclusion criteria, none of the patients randomized to treatment groups in UNITI-2 had an inadequate initial response with, loss of response to, or intolerance to TNF antagonists. [redacted]

TABLE 11: HISTORY OF MEDICATION USE FOR CROHN'S DISEASE IN THE INDUCTION STUDIES

Parameter, n (%)	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
Concomitant medications for CD at baseline				
≥ 1 concomitant medication(s)	185 (75)	174 (70)	158 (75)	170 (81)
Corticosteroids (including budesonide)	111 (45)	108 (43)	75 (36)	92 (44)
Any immunomodulators	81 (33)	78 (31)	73 (35)	72 (34)
6-MP or AZA	58 (24)	56 (23)	[redacted]	[redacted]
MTX	24 (10)	22 (9)	[redacted]	[redacted]
Aminosalicylates	54 (22)	50 (20)	[redacted]	[redacted]
Antibiotics	21 (9)	24 (10)	[redacted]	[redacted]
CD medication history				
Adequately treated and failed corticosteroids or immunomodulators[a, b]	[redacted]	[redacted]	[redacted]	[redacted]
Ever been treated with a full and adequate course of oral corticosteroids[b]	[redacted]	[redacted]	[redacted]	[redacted]
Failed to respond to corticosteroids	[redacted]	[redacted]	[redacted]	[redacted]
Intolerant to or developed a medical contraindication to corticosteroids	[redacted]	[redacted]	[redacted]	[redacted]
Have been corticosteroid dependent	[redacted]	[redacted]	[redacted]	[redacted]
Ever been treated with a full and adequate course of immunomodulators (AZA, 6-MP, or MTX)	[redacted]	[redacted]	[redacted]	[redacted]
Failed to respond to immunomodulators	[redacted]	[redacted]	[redacted]	[redacted]
Intolerant to or developed a medical contraindication to immunomodulators	[redacted]	[redacted]	[redacted]	[redacted]
Received TNF antagonist therapy	[redacted]	[redacted]	[redacted]	[redacted]
Not failed or intolerant to TNF antagonist therapy per study entry criteria[c]	[redacted]	[redacted]	[redacted]	[redacted]
Inadequate initial response to:	74 (30)	72 (29)	NA	NA
1 TNF antagonist	59 (24)	72 (29)		
2 TNF antagonists	15 (6)	64 (26)		
3 TNF antagonists	0	0		
Response followed by loss to:	170 (69)	171 (69)	NA	NA
1 TNF antagonist	112 (45)	124 (50)		
2 TNF antagonists	52 (21)	39 (16)		
3 TNF antagonists	6 (2)	8 (3)		
Intolerance to:	87 (25)	105 (42)	NA	NA
1 TNF antagonist	69 (28)	87(35)		

Parameter, n (%)	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
2 TNF antagonists	17 (7)	15 (6)		
3 TNF antagonists	1 (< 1)	3 (1)		
Inadequate initial response, loss of response, or intolerance to:	246 (100)	246 (99)	NA	NA
1 TNF antagonist	112 (45)	120 (48)		
2 TNF antagonists	108 (44)	102 (41)		
3 TNF antagonists	26 (11)	24 (10)		

6-MP = 6-mercaptopurine; AZA = azathioprine; CD = Crohn's disease; MTX = methotrexate; NA = not applicable; PLA = placebo; TNF = tumour necrosis factor-alpha; UST = ustekinumab.
[a] Includes patients who have failed to respond or became intolerant to corticosteroids or immunomodulators, or became dependent on corticosteroids.
[b] Includes budesonide.
[c] Denominator is the number of patients who received TNF antagonist therapy.
Source: CSR for UNITI-1 and UNITI-2.[7,8]

Maintenance Study

Key baseline demographic and Crohn's disease characteristics from the maintenance phase study are summarized in Table 12. Baseline data were reported in the clinical study report for IM-UNITI as the baseline values from the induction study (UNITI-1 or UNITI-2) in which patients were enrolled.

IM-UNITI included a higher proportion of women (approximately 57%) than men, with a mean age ranging from 38 to 40 years. [redacted]

The mean duration of Crohn's disease was approximately 10 years in IM-UNITI. Patients randomized to ustekinumab every eight weeks appeared to have somewhat less disease activity based on a lower mean CDAI score (313.1) and median CRP (8.8 mg/L) at baseline versus those randomized to ustekinumab every 12 weeks (CDAI score: 320.4; CRP: 9.1) or placebo (CDAI score: 319.1; CRP: 9.6). Patients in IM-UNITI had predominantly ileocolonic disease. [redacted]

At week 0 of IM-UNITI, 80 (approximately 60%) of patients in each of the three treatment groups had a clinical remission.

TABLE 12: SUMMARY OF BASELINE CHARACTERISTICS FROM THE MAINTENANCE STUDY

Parameter[a]	IM-UNITI		
	PLA (N = 133)	UST q.12.w. (N = 132)	UST q.8.w. (N = 132)
Sex, n (%)			
Female	74 (56)	74 (56)	76 (58)
Age (years), mean (SD)	39.5 (12.7)	38.6 (13.7)	37.9 (13.2)
Body weight (kg), mean (SD)	72.3 (17.3)	70.0 (19.6)	70.6 (16.9)
BMI (kg/m^2), mean (SD)	[redacted]	[redacted]	[redacted]
Geographic region, n (%)			
Asia	[redacted]	[redacted]	[redacted]
Eastern Europe	[redacted]	[redacted]	[redacted]
Rest of world	[redacted]	[redacted]	[redacted]
Current smoker status, n (%)	[redacted]	[redacted]	[redacted]
Duration of CD (years)			
Mean (SD)	10.6 (9.5)	9.5 (8.7)	10.3 (8.7)
CDAI, mean (SD)	319.1 (60.8)	320.4 (66.7)	313.1 (58.0)
CRP (mg/L), median (IQR)	9.6 ([redacted])	9.1 ([redacted])	8.8 ([redacted])
Disease localization, n (%)			
Ileum only	19 (14)	19 (14)	26 (20)
Colon only	28 (21)	29 (22)	23 (17)
Ileocolonic	86 (65)	84 (64)	83 (63)
Proximal GI tract	28 (21)	19 (14)	18 (14)
Perianal	43 (32)	46 (35)	39 (30)
CD complications, n (%)			
Intra-abdominal abscess (past)	[redacted]	[redacted]	[redacted]
Sinus tracts or perforation[b]	[redacted]	[redacted]	[redacted]
Fistula[b]	[redacted]	[redacted]	[redacted]
Current	[redacted]	[redacted]	[redacted]
Bowel stricturing[b]	[redacted]	[redacted]	[redacted]
Current	[redacted]	[redacted]	[redacted]
Extra-intestinal manifestations, n (%)	[redacted]	[redacted]	[redacted]

BMI = body mass index; CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CRP = C-reactive protein; GI = gastrointestinal; IQR = interquartile range; PLA = placebo; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; SD = standard deviation; UST = ustekinumab.

[a] At baseline of the induction studies.

[b] Current or past.

Source: Clinical Study Report for IM-UNITI.[9]

The proportions of patients receiving concomitant Crohn's disease medications (at induction study baseline and enrolment in IM-UNITI), including corticosteroids, immunomodulators, and aminosalicylates, were generally similar across treatment groups (Table 13). Approximately 80% of patients were receiving one or more concomitant Crohn's disease medications at baseline. Higher proportions of patients in the ustekinumab every eight weeks treatment group were receiving corticosteroids (including budesonide) and MTX ([redacted]%) at baseline than in the ustekinumab every 12 weeks (44% and [redacted]%, respectively) and placebo groups (44% and [redacted]%, respectively). [redacted]

[redacted]

Of the 397 randomized patients in IM-UNITI, approximately 45% had had a failure of (inadequate response or intolerance to) TNF antagonist therapy; the proportions were similar across treatment groups.

TABLE 13: SUMMARY OF CONCOMITANT CROHN'S DISEASE MEDICATIONS AT BASELINE FROM THE MAINTENANCE STUDY

Parameter, n (%)	IM-UNITI		
	PLA (N = 133)	UST q.12.w. (N = 132)	UST q.8.w. (N = 132)
Concomitant medications for CD at baseline[a]			
≥ 1 concomitant medication(s)	101 (76)	106 (80)	108 (82)
Corticosteroids (including budesonide)	59 (44)	58 (44)	64 (49)
Any immunomodulators	47 (35)	52 (39)	44 (33)
6-MP or AZA	[redacted]	[redacted]	[redacted]
MTX	[redacted]	[redacted]	[redacted]
Aminosalicylates	46 (35)	47 (36)	49 (37)
Antibiotics	[redacted]	[redacted]	[redacted]
CD medication history			
Adequately treated and failed corticosteroids or immunomodulators[b,c]	[redacted]	[redacted]	[redacted]
Ever been treated with a full and adequate course of oral corticosteroids[c]	[redacted]	[redacted]	[redacted]
Failed to respond to corticosteroids	[redacted]	[redacted]	[redacted]
Intolerant to or developed a medical contraindication to corticosteroids	[redacted]	[redacted]	[redacted]
Have been steroid-dependent	[redacted]	[redacted]	[redacted]
Ever been treated with a full and adequate course of immunomodulators (AZA, 6-MP, or MTX)	[redacted]	[redacted]	[redacted]
Failed to respond to immunomodulators	[redacted]	[redacted]	[redacted]
Intolerant to or developed a medical contraindication to immunomodulators	[redacted]	[redacted]	[redacted]
TNF antagonist refractory (UNITI-1 population)	61 (46)	59 (45)	58 (44)
Not TNF antagonist refractory (UNITI-2 population)	72 (54)	73 (55)	74 (56)
Not failed or intolerant to TNF antagonist therapy per study entry criteria[d]	[redacted]	[redacted]	[redacted]

Parameter, n (%)	IM-UNITI		
	PLA (N = 133)	UST q.12.w. (N = 132)	UST q.8.w. (N = 132)
Not received TNF antagonist therapy before entry into UNITI-2	52 (39)	53 (40)	52 (39)

6-MP = 6-mercaptopurine; AZA = azathioprine; CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CRP = C-reactive protein; MTX = methotrexate; PLA = placebo; SD = standard deviation; TNF = tumour necrosis factor; UST = ustekinumab.
[a] At baseline of the induction studies.
[b] Includes patients who have failed to respond or became intolerant to corticosteroids or immunomodulators, or became dependent on corticosteroids.
[c] Includes budesonide.
[d] Denominator is the number of patients who received TNF antagonist therapy.
Source: Clinical Study Report for IM-UNITI.[9]

Phase II Study

Baseline characteristics of patients randomized to treatment in CERTIFI are presented in Table 14. As in the other studies, a majority of patients enrolled in CERTIFI were women; the proportion of women was higher in the ustekinumab 6 mg/kg group (63%) than in the placebo group (52%). The mean age was similar between groups, at 39 years, and most patients were recruited from North America, with study sites in Canada. The proportion of current smokers was higher in the ustekinumab group (32%) than in the placebo group (23%).

Patients in the ustekinumab group appeared to have worse markers of Crohn's disease severity, based on baseline mean CDAI scores and median CRP. Consistent with the other studies, most patients had ileocolonic disease (placebo: 42%; ustekinumab: 47%) and extra-intestinal manifestations (placebo: 52%; ustekinumab: 63%). The proportion of patients who reported a history of Crohn's disease–related surgery (total or subtotal colectomy and small or large bowel resection) was similar between treatment groups.

TABLE 14: SUMMARY OF BASELINE CHARACTERISTICS FROM THE PHASE II STUDY

Parameter	CERTIFI	
	PLA (N = 132)	UST 6 mg/kg (N = 131)
Sex, n (%)		
Female	68 (52)	83 (63)
Age (years), mean (SD)	39.5 (13.1)	39.4 (13.2)
Body weight (kg); mean (SD)	74.4 (20.5)	74.1 (21.4)
Geographic region, n (%)		
Asia-Pacific	[redacted]	[redacted]
Europe	[redacted]	[redacted]
North America	[redacted]	[redacted]
Smoking status, n (%)		
Current smoker	30 (23)	42 (32)
Former smoker	31 (23)	28 (21)
Duration of CD (years)		
mean (SD)	12.4 (9.1)	12.7 (8.9)
CDAI, mean (SD)	312.4 (64.2)	338.0 (67.3)

Parameter	CERTIFI	
	PLA (N = 132)	UST 6 mg/kg (N = 131)
CRP (mg/L), median (IQR)	9.3 (3.2 to 28.4)	12.6 (3.3 to 34.3)
Disease localization, n (%)		
Ileum only	34 (26)	34 (26)
Colon only	41 (31)	33 (25)
Ileocolonic	55 (42)	62 (47)
Proximal GI tract	14 (11)	12 (9)
CD history, n (%)		
Intra-abdominal abscess	12 (9)	12 (9)
Stoma	8 (6)	3 (2)
Fistula	16 (12)	19 (15)
Small bowel stricture	21 (16)	29 (22)
Total or subtotal colectomy	24 (18)	19 (15)
CD-related small or large bowel resection	55 (42)	59 (45)
Extra-intestinal manifestations, n (%)	68 (52)	82 (63)

CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CRP = C-reactive protein; GI = gastrointestinal; IQR = interquartile range; PLA = placebo; SD = standard deviation; UST = ustekinumab.
Source: Clinical Study Report for CERTIFI.[10]

Table 15 summarizes the baseline concomitant and prior medications used for Crohn's disease by the patients enrolled in CERTIFI. The proportion of patients randomized to ustekinumab 6 mg/kg who were receiving concomitant oral corticosteroids (including budesonide) at baseline was 45% versus 55% for those randomized to placebo. Conversely, the proportion of ustekinumab-treated patients receiving immunomodulators was greater than that for placebo patients (27% versus 23%, respectively), mainly driven by more patients receiving 6-MP or AZA.

Although more patients in the ustekinumab group reported having been previously treated with oral corticosteroids and immunomodulators than in the placebo group, the proportion of patients who reported having a failure with these treatments (i.e., inadequate response, intolerance, or contraindication) was greater in the placebo group than in the ustekinumab group. One-half or more patients had an inadequate initial response with, loss of response to, or intolerance to treatment with one TNF antagonist, while more than one-third reported treatment failure with two TNF antagonists; there were numerical differences between the groups regarding prior TNF antagonist response.

TABLE 15: SUMMARY OF CONCOMITANT CROHN'S DISEASE MEDICATIONS AT BASELINE FROM PHASE II STUDY

Parameter, n (%)	CERTIFI	
	PLA (N = 132)	UST 6 mg/kg (N = 131)
Concomitant medications for CD at baseline		
≥ 1 concomitant medication(s)	101 (77)	92 (70)
Corticosteroids (including budesonide)	73 (55)	59 (45)
Any immunomodulators	30 (23)	35 (27)
6-MP or AZA	[redacted]	[redacted]
MTX	[redacted]	[redacted]

Parameter, n (%)	CERTIFI	
	PLA (N = 132)	**UST 6 mg/kg (N = 131)**
Aminosalicylates	24 (18)	25 (19)
Antibiotics	7 (5)	12 (9)
CD medication history		
Ever been treated with a full and adequate course of oral corticosteroids[a]	[redacted]	[redacted]
Failed to respond, became intolerant, or developed a medical contraindication to corticosteroids[b]	31 (61)	34 (52)
Ever been treated with a full and adequate course of immunomodulators (AZA, 6-MP, or MTX)	113 (86)	119 (91)
Failed to respond, became intolerant, or developed a medical contraindication to immunomodulators[c]	101 (89)	97 (82)
Inadequate initial response to	[redacted]	[redacted]
1 TNF antagonist	[redacted]	[redacted]
2 TNF antagonists	[redacted]	[redacted]
3 TNF antagonists	[redacted]	[redacted]
Response followed by loss to:	[redacted]	[redacted]
1 TNF antagonist	[redacted]	[redacted]
2 TNF antagonists	[redacted]	[redacted]
3 TNF antagonists	[redacted]	[redacted]
Intolerance to:	[redacted]	[redacted]
1 TNF antagonist	[redacted]	[redacted]
2 TNF antagonists	[redacted]	[redacted]
3 TNF antagonists	[redacted]	[redacted]
Inadequate initial response, loss of response, or intolerance to:	[redacted]	[redacted]
1 TNF antagonist	[redacted]	[redacted]
2 TNF antagonists	[redacted]	[redacted]
3 TNF antagonists	[redacted]	[redacted]

6-MP = 6-mercaptopurine; AZA = azathioprine; CD = Crohn's disease; PLA = placebo; TNF = tumour necrosis factor; UST = ustekinumab.
[a] Includes budesonide.
[b] Denominator is the number of patients who have ever been treated with a full and adequate course of corticosteroids.
[c] Denominator is the number of patients who have ever been treated with a full and adequate course of immunomodulators.
Source: Clinical Study Report for CERTIFI.[10]

3.2.3 Interventions

a) Study Treatments

Induction Studies

All patients received one IV administration of placebo or ustekinumab at week 0. Two dose regimens of IV ustekinumab were used: 130 mg and a tiered dose approach approximating 6 mg/kg that allowed administration of complete vials (130 mg per vial) to patients to simplify dose calculation:

- Ustekinumab 260 mg (weight ≤ 55 kg)
- Ustekinumab 390 mg (weight > 55 kg and ≤ 85 kg)
- Ustekinumab 520 mg (weight > 85 kg).

Only the results for the 6 mg/kg tiered regimen group are approved by Health Canada and reported in the CDR review.

The manufacturer stated that the placebo preparation was identical in appearance to ustekinumab.

IV administration was chosen for induction over SC administration based on findings from a manufacturer-sponsored crossover, proof-of-concept study (study C0379T07) in which a single IV and a single SC administration of ustekinumab were evaluated.[26] This study suggested that serum concentrations of inflammation-related markers decreased more rapidly after IV administration than after SC administration. Therefore, IV administration was used in the phase III induction trials of ustekinumab in Crohn's disease. The IV induction doses were based on the results of CERTIFI.

Maintenance Study

Patients who had a clinical response to ustekinumab induction treatment were randomized in a 1:1:1 ratio at week 0 of IM-UNITI/week 8 of the UNITI studies to receive one of the following SC regimens:

- Placebo
- Ustekinumab 90 mg SC every 12 weeks (with last dose at week 36)
- Ustekinumab 90 mg SC every eight weeks (with last dose at week 40).

Randomized patients who lost response at any scheduled visit between week 8 and week 32 (visits occurred every four weeks) were eligible to receive ustekinumab as follows:

- Randomized to placebo: Patients' dosage adjusted to receive ustekinumab 90 mg SC every eight weeks
- Randomized to ustekinumab 90 mg every 12 weeks: Patients' dosage adjusted to receive ustekinumab 90 mg SC every eight weeks
- Randomized to ustekinumab 90 mg every eight weeks: Patients continued on ustekinumab 90 mg SC every eight weeks.

Patients who had their dosage adjusted were assessed 16 weeks after the visit at which the criteria for loss of response were met. Patients who did not show improvement in Crohn's disease activity at that time (as assessed by the investigator) discontinued from study drug administration. Patients assessed by the investigator to have clinical improvement continued to receive the same adjusted dose in a blinded manner; however, patients who had their dosage adjusted between weeks 8 and 32 were coded as nonresponders for the primary analyses.

A total of 27.5% (109 patients) of the randomized population had a dosage adjustment as follows:

- Among patients randomized to placebo, 38.3% (51 patients) had a dosage adjustment to an ustekinumab 90 mg SC every eight weeks regimen
- Among patients randomized to ustekinumab 90 mg SC every 12 weeks, 22.0% (29 patients) had a dosage adjustment to an ustekinumab 90 mg SC every eight weeks regimen
- Among patients randomized to ustekinumab 90 mg SC every eight weeks, 22.0% (29 patients) continued on the same regimen.

[redacted]

Phase II Study
Patients in CERTIFI were randomized (1:1:1:1) to receive one of the following IV induction regimens at week 0:

- Placebo
- Ustekinumab 1 mg/kg
- Ustekinumab 3 mg/kg
- Ustekinumab 6 mg/kg.

In the maintenance phase, patients who were in clinical response to IV ustekinumab induction (irrespective of dose) at week 6 were re-randomized (1:1 ratio) at week 8 to receive an SC maintenance regimen of placebo (N = 73) or ustekinumab 90 mg (N = 72) administered at week 8 and week 16. Hence, the maintenance phase consisted of patients who received induction treatment with dosages that are not Health Canada–approved; therefore, only data from the induction phase of CERTIFI are reported in this review.

Placebo administrations were reported as having the same appearance as the respective ustekinumab administrations.

Treatment adherence was controlled by the study staff administering the drug as an SC injection and/or an IV infusion.

b) Concomitant Medications
Induction Studies
Patients were permitted to receive the following concomitant medications to treat Crohn's disease:

- Oral aminosalicylates (i.e., 5-ASA)
- Oral corticosteroids at a prednisone-equivalent dose of ≥ 40 mg/day or ≤ 9 mg/day of budesonide
- Immunomodulators (i.e., AZA, 6-MP, or MTX)
- Antibiotics, as a primary treatment for Crohn's disease.

Patients had to be receiving stable doses for at least three weeks (four weeks in the case of immunomodulators) and to be taking them for a total of 12 weeks or more before baseline. Patients were required to maintain a stable dose of these concomitant drugs throughout the induction studies.

Patients were not allowed to start any of the aforementioned medications or total parenteral nutrition as a treatment for Crohn's disease during the studies.

Maintenance Study
Patients were permitted to receive the same concomitant medications as in the induction studies, with the same conditions regarding maintaining stable dosages (except for oral corticosteroids) and the same prohibition against starting therapy at any point during IM-UNITI. Patients were permitted to transiently use (i.e., for four weeks or less) increased doses of oral corticosteroids for reasons other than loss of response to treatment for Crohn's disease (e.g., for surgery, asthma, adrenocortical insufficiency, etc.).

Patients receiving corticosteroids at the start of IM-UNITI who had a clinical response to ustekinumab had their corticosteroid tapered. Tapering was mandatory and followed this schedule:

- Oral corticosteroids (other than budesonide)
  - Dose > 15 mg/day prednisone or equivalent: taper daily dose by 5 mg/week until receiving 10 mg/day, then continue tapering by 2.5 mg/week until discontinued.
  - Dose 11 to 15 mg/day prednisone or equivalent: taper daily dose to 10 mg/day for 1 week, then continue tapering by 2.5 mg/week until discontinued.
  - Dose ≤ 10 mg/day prednisone or equivalent: taper daily dose by 2.5 mg/week until discontinued.
- Oral budesonide
  - Daily dose tapered by 3 mg every 3 weeks until discontinued.

Tapering could be suspended if a patient experienced a worsening of disease activity; tapering could then be resumed within four weeks.

3.2.4 Outcomes

The outcomes analyzed in the studies and relevant to this review according to the protocol (Section 2) are summarized in Table 6, Table 7, and Table 8 for the induction, maintenance, and phase II studies, respectively. The following are brief descriptions of the outcomes from the studies. Detailed descriptions may be found in Appendix 5: Validity of Outcomes.

a) Crohn's Disease Activity Index

The CDAI is an instrument used to evaluate and quantify the severity of symptoms for patients with Crohn's disease. The CDAI consists of the following eight factors, each of which is summed after adjustment with a weighting factor:

- Number of liquid or soft stools each day for seven days
- Abdominal pain each day for seven days (0 [none] to 3 [severe])
- General well-being each day for seven days (0 [well] to 4 [terrible])
- Presence of complications
- Requirement to take diphenoxylate/atropine or opiates for diarrhea
- Presence of an abdominal mass (0 [none], 2 [questionable], 5 [definite])
- Hematocrit of < 0.47 in men and < 0.42 in women
- Percentage deviation from standard weight.

The total CDAI score ranges from 0 to 600, with higher scores indicating greater Crohn's disease activity.

Clinical Remission

Clinical remission was defined as a CDAI score < 150 points. However, patients who had any of the following events before the clinical remission assessment end point were not considered in clinical remission, regardless of the CDAI score:

- specified changes in concomitant Crohn's disease medications
- a Crohn's disease–related surgery (with the exception of drainage of a cutaneous or perianal abscess or seton placement)
- discontinuation of treatment due to lack of efficacy or due to an adverse event of worsening Crohn's disease
- in IM-UNITI, loss of clinical response, defined as a CDAI score ≥ 220 points and a ≥ 100-point increase from CDAI score from the week 0 to week 8 in the induction study.

In addition, patients who did not return for evaluation or who had insufficient data to assess their clinical response status (i.e., four or fewer components of the CDAI are available) were also not considered to have achieved clinical response.

Clinical remission at week 44 was the primary outcome of IM-UNITI, and a key secondary outcome at week 8 of the UNITI and CERTIFI studies.

Those who could discontinue corticosteroids and were in clinical remission (i.e., had a CDAI score ≤ 150) were considered to have corticosteroid-free clinical remission. This was a key secondary outcome at week 44 in IM-UNITI.

Clinical Response

Clinical response was defined as a reduction from baseline in the CDAI score of ≥ 100 points. In the patients with a baseline CDAI score of 220 to 248, clinical response was considered to be achieved if the CDAI score was reduced to less than 150. In addition, treatment failure rules were applied to determine each patient's final response status. Patients who had any of the following events before the clinical response assessment end point were not considered to have a clinical response, regardless of the CDAI score:

- a Crohn's disease–related surgery (with the exception of drainage of a cutaneous or perianal abscess or seton placement) that was thought to be a result of lack of efficacy of study treatment
- specified changes in concomitant Crohn's disease medications.

In addition, patients who did not return for evaluation or who had insufficient data to assess their clinical response status (i.e., four or fewer components of the CDAI are available) were not considered to have a clinical response.

Clinical response at week 6 was the primary outcome in the UNITI and CERTIFI studies and a key secondary outcome at week 44 in IM-UNITI.

b) Inflammatory Bowel Disease Questionnaire

The Inflammatory Bowel Disease Questionnaire (IBDQ) is a 32-item questionnaire that aims to capture how the patient felt during the two weeks before the measurement time point. Questions are related to the symptoms of Crohn's disease, how the patient felt in general and his/her mood over the previous two weeks, and social or employment problems that may have resulted from Crohn's disease.[32,33] 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 between 32 and 224. An increase in IBDQ score indicates an improvement in health-related quality of life, while a decrease indicates deterioration. Scores of patients in remission typically range from 170 to 190. The minimal clinically important difference (MCID) for the IBDQ is considered 16 points.[34] IBDQ was a secondary outcome assessment tool in all of the included studies.

c) Short Form (36) Health Survey

The Short Form (36) Health Survey (SF-36) is a generic instrument that was used to assess health-related quality of life in the UNITI and IM-UNITI studies. It assesses eight domains of health-related quality of life: physical function, role physical, general health, mental health, role emotional, social functioning, and vitality. The eight domains are aggregated to create two component summaries: the physical component summary (PCS) and the mental component summary (MCS), with scores ranging from 0 to 100, with higher scores indicating better health-related quality of life. The MCID for the PCS and the

MCS has been estimated at 4.1 and 3.9, respectively, in the Crohn's disease patient population.[35] However, there is uncertainty as to the validity of these estimates. In general use, the MCIDs are two points for the PCS and three points for the MCS.[36]

d) Work Limitations Questionnaire

The Work Limitations Questionnaire (WLQ) is a 25-item self-reported questionnaire that asks respondents to rate their level of difficulty or ability to perform specific job demands to assess health-related loss of work productivity. [37-39] The 25 items are aggregated into four scales: time management, physical demands, mental-interpersonal, and output. Scale scores range from 0 (limited none of the time) to 100 (limited all of the time) and represent the reported amount of time in the prior two weeks that respondents were limited in functioning on the job. The MCID has been estimated in non-Crohn's disease conditions to range from 3.2 to 4.[40,41] The WLQ was a secondary outcome measure in the UNITI and IM-UNITI studies.

e) Productivity Visual Analogue Scale

The impact of disease on patients' daily productivity was measured using a visual analogue scale (VAS, 0 = no impact at all to 10 = impacts productivity very much). No other information about this measure was identified. This was a secondary outcome in all of the included studies.

f) Time Lost From Work

Time lost from work was collected by asking patients, "How many days did you miss from work due to your Crohn's disease in the last four weeks?" No other information about this measure was identified. This was a secondary outcome in all of the included studies.

g) Mucosal Healing

Mucosal healing was defined as the complete absence of mucosal ulcerations. Mucosal healing was assessed using endoscopy (ileocolonoscopy) in a subset of patients from specific study sites enrolled in the UNITI and IM-UNITI studies (endoscopic substudy). Patients who participated in the endoscopic substudy underwent endoscopic assessments at baseline and at week 8 of the UNITI studies, and at the end of the IM-UNITI (week 44 of maintenance). In addition, biopsies were collected to support exploratory histologic evaluation. Video endoscopies were assessed by a central facility that was blinded to treatment group.[26] Changes in the Simplified Endoscopic Disease Severity Score for Crohn's Disease (SES-CD), in addition to endoscopic detection of presence or absence of mucosal ulceration, were used to evaluate mucosal healing. The SES-CD consists of four endoscopic variables: presence and size of ulcers, proportion of surface covered by ulcers, proportion of surface affected by disease, and presence and severity of stenosis.[42] Each endoscopic variable is scored by colon segment (scores range 0 to 3), and the total SES-CD score ranges from 0 to 56, with higher SES-CD scores indicating more severe disease activity.[42] The MCID was not identified.

h) Need for Surgery

Crohn's disease–related surgery included, but was not limited to, total or subtotal colectomy or other partial bowel resection, and perianal surgery. Minor procedures such as placement of a seton or cutaneous drainage of an abscess were excluded. Crohn's disease–related surgery was evaluated in each included study.

i) Safety

Treatment safety was assessed in all included studies by collecting adverse events (AEs), serious AEs (SAEs), withdrawals due to AEs (WDAEs), vital signs, AEs related to drug-administration reactions,

hematology and chemistry parameters, physical examinations, and 12-lead electrocardiograms. None of the studies was designed to specifically evaluate safety outcomes.

The frequency of AEs throughout was summarized by treatment group, the Medical Dictionary for Regulatory Activities (MedDRA) system-organ class, and preferred term. All reported AEs with onset during the treatment phase (i.e., treatment-emergent AEs, and AEs that had worsened since baseline) were included in the analysis.

3.2.5 Statistical Analysis

a) Primary Outcome

Sample Size and Power Calculations

Sample size considerations for all four studies are summarized in Table 16.

The sample size and power calculations for all four studies were based on the chi-square test (two-sided) for detecting a significant difference between patients receiving the higher dosage of ustekinumab (tiered weight-based dosage approximating ustekinumab 6 mg/kg IV in the UNITI studies, 90 mg SC every eight weeks in IM-UNITI, and ustekinumab 6 mg/kg IV in CERTIFI) and those receiving placebo. The assumptions for sample size and power calculations for UNITI-1 were based on data from CERTIFI, and those for UNITI-2 and CERTIFI were based the phase IIa manufacturer-sponsored crossover, proof-of-concept study, study C0379T07;[26] additional assumptions were made based on "recent [inflammatory bowel disease] literature" for CERTIFI.[28] Assumptions for IM-UNITI sample size were based on clinical remission data from randomized controlled trials of maintenance treatment for Crohn's disease with infliximab and adalimumab.[43,44]

TABLE 16: SUMMARY OF SAMPLE SIZE CONSIDERATION FROM THE INCLUDED STUDIES

Study	Hierarchy	Outcome	Assumed Response (%)	Patients Required	Power
UNITI-1 (Induction)	Primary	Clinical response at week 6	PLA (25) UST 6 mg/kg (40)	225 per group	93%
	Secondary	Clinical remission at week 8	PLA (10) UST 6 mg/kg (20)		85%
UNITI-2 (Induction)	Primary	Clinical response at week 6	PLA (33) UST 6 mg/kg (50)	200 per group	93%
	Secondary	Clinical remission at week 8	PLA (12) UST 6 mg/kg (25)		92%
IM-UNITI (Maintenance)	Primary	Clinical remission at week 44	PLA (15) UST 90 mg q.8.w. (35)	100 per group	91%
CERTIFI (Induction)	Primary	Clinical response at week 6	PLA (30) UST (50)	124 per group	90%

PLA = placebo; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; UST = ustekinumab.
Source: Clinical Study Reports for UNITI-1, UNITI-2, IM-UNITI, and CERTIFI[7-10] and study protocols.[28-31]

Primary Outcome Analysis

The proportion of patients in clinical response at week 6 (UNITI studies and CERTIFI) or clinical remission at week 44 (IM-UNITI) was compared between the ustekinumab treatment group and the placebo group using a two-sided Cochran–Mantel–Haenszel chi-square test, stratified by study region (Asia, Eastern Europe, or rest of world) and CDAI score (≤ 300 or > 300) in UNITI-1 and UNITI-2, and by initial response to TNF antagonist therapy (yes or no) in UNITI-1; the latter was the only stratification factor for the CERTIFI analysis. Stratification factors in IM-UNITI were clinical remission status at week 0 (yes or no), ustekinumab induction dose (130 mg or tiered dosage approximating ustekinumab 6 mg/kg), and induction study (UNITI-1 or UNITI-2). Comparisons were made with a significance level of 0.05.

Patients randomized to treatment before UNITI studies were restarted were not included in the primary analysis population of the UNITI studies or IM-UNITI. In addition, data for one patient at a single study site in UNITI-2 ([redacted]) were excluded because the patient was randomized despite major protocol deviations (e.g., not meeting the inclusion criteria for active Crohn's disease).

In the UNITI studies, the CDAI score was calculated for a visit only if four or more of the eight components were available at that visit. Any missing components were imputed by carrying forward the last non-missing component, with the exception of a missing hematocrit value, when at least four of the eight components were available. If the CDAI score could not be calculated (i.e., four or fewer components available) at a visit, the CDAI score was considered missing. Patients with a missing CDAI score at week 6 were not considered to have achieved clinical response at week 6. (See Section 3.5 for more discussion on missing data.)

Sensitivity Analyses

The manufacturer conducted several sensitivity analyses to examine the robustness of the primary outcome analysis in each study. Sensitivity analyses of the primary outcome were conducted using the following data methods: observed case, last observation carried forward (LOCF), multiple imputation, and worst case missing data methods. Treatment failure rules superseded missing data rules, meaning that if a patient had both a treatment failure (i.e., Crohn's disease–related surgery, specified changes in concomitant Crohn's disease medications, discontinued treatment due to lack of efficacy or due to an adverse event, or, in IM-UNITI, a loss of response) before end point assessment and had a missing CDAI score at the end point assessment (i.e., four or fewer components of the CDAI available), the patient was considered a nonresponder in the sensitivity analysis regardless of whether CDAI data were present.

Subgroup Analyses

Subgroup analyses of the efficacy of ustekinumab versus placebo for the primary outcome were carried out for demographic baseline disease characteristics (baseline of the UNITI studies for IM-UNITI), Crohn's disease medication history, concomitant Crohn's disease medication use at baseline (baseline of the UNITI studies for IM-UNITI), study site location, and initial response to TNF antagonist therapy (except UNITI-2). It was reported that subgroup analyses were planned when the number of patients in the subgroups permitted; however, the threshold number of patients was not specified. The odds ratios of each ustekinumab dosage group versus placebo and corresponding 95% confidence intervals were provided for each of the subgroups.

b) Secondary Outcomes

Categorical secondary outcomes (e.g., the proportion of patients with a Crohn's disease–related surgery) were also compared between ustekinumab and placebo using a two-sided Cochran–Mantel–

Haenszel chi-square test (or chi-square test where appropriate; circumstances in which this occurred were not specified), with the same data-handling protocol specified for the primary outcome analysis.

Continuous secondary outcomes were compared using analysis of variance or covariance models. An analysis of variance or covariance on the van der Waerden normal scores or a nonparametric test such as the Kruskal–Wallis test was used if the normality assumption was uncertain. Covariates for the analyses of change from baseline in productivity VAS and WLQ were: study region (Asia, Eastern Europe, or rest of world), baseline CDAI score (≤ 300 or > 300), and initial response to TNF antagonist therapy (yes or no) (except UNITI-2). Covariates for the change from baseline in IBDQ score and SF-36 were not stated.

The IBDQ score and the PCS and MCS scores of the SF-36 were analyzed as change from baseline and as the proportion of patients achieving at least a 16-point improvement (IBDQ) or at least a five-point improvement from baseline (SF-36 components). If any one of the dimensions within the IBDQ could not be calculated (e.g., due to missing items), then the total IBDQ score was not calculated. For the SF-36, if < 50% of the items that make up a subscale were available, the subscale was not calculated. And, if any of the individual subscales that constitute the PCS or the MCS were missing, then the PCS or MCS summary scores were not calculated.

[redacted]

For IBDQ, SF-36, productivity VAS, WLQ, and surgery, treatment failure rules used for the primary analysis applied only to the daily productivity outcome, and no imputation was performed for missing values; missing values were coded as missing.

c) Multiple Comparisons

The primary efficacy comparisons in all the included studies used a fixed-sequence testing procedure to control the overall type I error rate at a significance level of 0.05 (two-sided).

For the induction analyses (UNITI and CERTIFI studies), the comparison between the ustekinumab high-dose group (dose approximating 6 mg/kg ustekinumab) and placebo was made first; if the ustekinumab high-dose group was statistically significantly different from the placebo group, then the ustekinumab low-dose group (130 mg ustekinumab in UNITI; 3 mg/kg followed by 1 mg/kg ustekinumab in CERTIFI) was compared with the placebo group at the two-sided 0.05 level of significance. However, given that this review considers only the Health Canada–approved ustekinumab 6 mg/kg IV induction dose, this procedure has no impact on the interpretation of the results from the induction phase studies because the lower doses are out of scope.

For the IM-UNITI primary analysis, however, this procedure is relevant, as both ustekinumab dosages are approved. In IM-UNITI, the 90 mg every eight weeks ustekinumab group was compared with the placebo group first. The 90 mg every 12 weeks ustekinumab group was compared with the placebo group only if the comparison of the 90 mg every eight weeks group with placebo was statistically significant at the two-sided 0.05 level of significance.

In all included studies, major secondary outcomes (Table 17) were tested in a hierarchical fashion. The first major secondary outcome was tested only if the primary outcome was positive, and the subsequent

outcomes were tested only if the preceding outcome in the hierarchy was positive. Within each major secondary outcome, the fixed-sequence testing procedure was applied: the low dose for a major secondary outcome could not be tested unless the low dose tested positive for the preceding major secondary outcome (or the primary outcome if the outcome being tested was the first major secondary outcome).

It was not reported whether other secondary outcomes (e.g., change in IBDQ score) were included in the hierarchical analysis procedure, and it is assumed that they were not.

TABLE 17: MAJOR SECONDARY OUTCOMES

Study			
UNITI-1	**UNITI-2**	**IM-UNITI**	**CERTIFI**
• Clinical remission at week 8 • Clinical response at week 8[a] • 70-point response at week 6[a] • 70-point response at week 3[a]		• Clinical response at week 44 • Clinical remission at week 44 (among those in clinical remission to UST at week 0) • Corticosteroid-free remission at week 44 • Clinical remission at week 44 (among those refractory or intolerant to TNF antagonist therapy)	• Clinical remission at week 6 • Clinical response at week 4[a] • Clinical remission at week 22 (among those in clinical response to UST at week 6)[a] • Clinical response at week 22 (among those in clinical response to UST at week 6)[a]

[a] Not included in the CDR review.
Source: Clinical Study Reports for UNITI-1 and UNITI-2, IM-UNITI, and CERTIFI.[7-10]

d) Endoscopy Substudy

The primary objectives of the endoscopy substudy were to evaluate:

- The efficacy of ustekinumab compared with placebo to induce endoscopic mucosal healing
- The efficacy of ustekinumab maintenance treatment compared with placebo on the achievement of endoscopic healing of the mucosa among patients who had a clinical response to ustekinumab induction.

The pre-specified primary analysis population for endoscopy outcomes in the induction phase was the combined populations from UNITI-1 and UNITI-2, and the ustekinumab induction-treatment groups (130 mg and ~6 mg/kg) were pooled. This pooled ustekinumab induction-treatment group was compared with the placebo induction-treatment group. For analyses of maintenance endoscopy outcomes, the ustekinumab maintenance-treatment groups (90 mg every 12 weeks and 90 mg every eight weeks) were combined. This pooled ustekinumab maintenance-treatment group was compared with the maintenance placebo group.

Patients had to have the following in order to be eligible for the endoscopic analyses:

- A SES-CD score of three or more at baseline (for evaluation of SES-CD–based end points), or
- Evidence of ulceration in any segment of the colon at baseline (for evaluation of mucosal healing).

The primary outcome in the endoscopy substudy was the change from baseline in the SES-CD score at week 8 of induction. The major secondary outcomes were (in hierarchical order according to the hierarchical testing procedure for the substudy):

- The change from induction baseline in the SES-CD score at week 44 of maintenance
- The proportion of patients with mucosal healing at week 44 of maintenance
- The proportion of patients with mucosal healing at week 8 of induction.

The change from baseline in the SES-CD score at week 8 of induction was compared between the pooled ustekinumab group and the placebo group using an analysis of covariance on the van der Waerden normal scores, with baseline SES-CD score and study as covariates, at a significance level of 0.05. Missing data rules and treatment failure rules were applied. (Note that the manufacturer-provided documents did not specify whether these rules were the same as those applied in the UNITI and IM-UNITI studies.)

The manufacturer reported that the substudy also included pre-specified subgroup analyses of key endoscopic end points by individual induction study, by induction ustekinumab dose, and by maintenance ustekinumab dosage, as well as sensitivity analyses to evaluate the impact of different missing data rules. However, data for the subgroup and sensitivity analyses were not provided in the materials submitted for this review.

e) Analysis Populations

The manufacturer stated that efficacy analyses included patients randomized at week 0 of each study, and were based on an intention-to-treat principle. Therefore, the efficacy data for each patient were analyzed according to the assigned treatment, regardless of the actual treatment received. See Section 3.5 Critical Appraisal for comments on this.

Safety analyses were based on patients who received at least one dose of study treatment. Patients were analyzed according to the actual treatment received.

3.3 Patient Disposition

3.3.1 Induction Studies

The number of patients screened for eligibility to enter the induction studies was not reported in the clinical study reports.

Greater than 90% of patients randomly assigned to ustekinumab or placebo completed both induction studies (Table 18). A higher proportion of patients treated with ustekinumab than given placebo entered the maintenance study (IM-UNITI). More patients in the ustekinumab group in UNITI-1 (6%; placebo: 4%) and more patients in the placebo group in UNITI-2 (6%; ustekinumab: 1%) discontinued the study.

TABLE 18: PATIENT DISPOSITION FROM THE INDUCTION STUDIES

Disposition, n (%)	UNITI-1		UNITI-2	
	PLA	UST	PLA	UST
Randomized[a]	247	249	210	209
Safety population	245	249	208	207
ITT population	247	249	210	209
PP population	NR	NR	NR	NR
Completed study	237 (96)	235 (94)	198 (94)	207 (99)
Entered maintenance study	214 (87)	227 (91)	186 (89)	203 (97)
Completed 20 week follow-up visit	23 (9)	8 (3)	12 (6)	4 (2)
Discontinued	10 (4)	14 (6)	12 (6)	2 (1)
Withdrawal of consent	4 (2)	8 (3)	6 (3)	1 (<1)
Lost to follow-up	1 (< 1)	0	2 (1)	1 (< 1)
Other	2 (< 1)	1 (< 1)	1 (< 1)	0

ITT = intention-to-treat; NR = not reported; PLA = placebo; PP = per-protocol; UST = ustekinumab.
[a] Randomization restart. Excludes 28 and 12 patients from UNITI-1 and UNITI-2, respectively, randomized prior to study restart following the second protocol amendment.
Source: Clinical Study Reports for UNITI-1 and UNITI-2.[7,8]

3.3.2 Maintenance Study

A total of 1,281 patients who completed the ustekinumab induction studies UNITI-1 and UNITI-2 were enrolled in the IM-UNITI maintenance-treatment study (Table 19). A total of 397 of these patients (31.0%) had a clinical response to ustekinumab induction and were randomized to the primary population. [redacted]

Over 44 weeks of treatment, greater than 20% of patients in each treatment group discontinued treatment, and from 7% (ustekinumab every 12 weeks) to 11% (ustekinumab every eight weeks) discontinued the study. The most common reason for discontinuing treatment was lack of efficacy, followed by AEs.

TABLE 19: PATIENT DISPOSITION FROM THE MAINTENANCE STUDY

Disposition, n (%)	IM-UNITI		
	PLA[a]	UST q.12.w.	UST q.8.w.
Completed induction studies	1,281		
Randomized	133	132	132
Safety population	133	132	131
ITT population[b]	131	129	128
PP population	NR	NR	NR
UNITI study induction dosage			
Weight-based 6 mg/kg IV	[redacted]	[redacted]	[redacted]
130 mg IV	[redacted]	[redacted]	[redacted]
Completed study	120 (90)	123 (93)	118 (89)
Entered extension study	96 (72)	103 (78)	99 (75)
Discontinued study	13 (10)	9 (7)	14 (11)

Disposition, n (%)	IM-UNITI		
	PLA[a]	UST q.12.w.	UST q.8.w.
Withdrawal of consent	[redacted]	[redacted]	[redacted]
Lost to follow-up	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]
Discontinued treatment	31 (23)	29 (22)	30 (23)
Adverse event	[redacted]	[redacted]	[redacted]
Lack of efficacy	[redacted]	[redacted]	[redacted]
Protocol violation	[redacted]	[redacted]	[redacted]
Withdrawal of consent	[redacted]	[redacted]	[redacted]
Lost to follow-up	[redacted]	[redacted]	[redacted]

ITT = intention-to-treat; IV = intravenous; NR = not reported; PLA = placebo; PP = per-protocol; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; UST = ustekinumab.
[a] Patients who were in clinical response to ustekinumab IV induction dosing and were randomized to placebo SC on entry into the maintenance study.
[b] Excluding the nine patients who were randomized prior to study restart. Therefore, there are 388 patients in the primary analysis population.
Source: Clinical Study Report for IM-UNITI.[9]

3.3.3 Phase II Study

The number of patients screened for eligibility to enter CERTIFI was not reported in the clinical study report.

A total of 526 patients were randomized to treatment in CERTIFI, with 132 and 131 patients randomized to placebo and ustekinumab 6 mg/kg, respectively (Table 20). A larger proportion of patients receiving placebo (14%) than receiving ustekinumab (6%) discontinued treatment in the eight-week induction phase. Lack of efficacy was the primary reason, as well as the related worsening of Crohn's disease, captured as an AE.

TABLE 20: PATIENT DISPOSITION FROM THE PHASE II STUDY INDUCTION PHASE

Disposition, n (%)	CERTIFI	
	PLA	UST 6 mg/kg
Randomized	132	131
Safety population	132	131
ITT population	133	131
PP population	NR	NR
Completed Study	113 (86)	123 (94)
Discontinued study	19 (14)	8 (6)
Withdrawal of consent	[redacted]	[redacted]
Lost to follow-up	[redacted]	[redacted]
Other	[redacted]	[redacted]
Discontinued treatment	19 (14)	8 (6)
Adverse event	5 (4)	1 (< 1)
Lack of efficacy	9 (7)	3 (2)

Disposition, n (%)	CERTIFI	
	PLA	UST 6 mg/kg
Lost to follow-up	1 (< 1)	0
Other	4 (3)	4 (3)

ITT = intention-to-treat; NR = not reported; PLA = placebo; PP = per-protocol; UST = ustekinumab.
Source: Clinical Study Report for CERTIFI.[10]

3.4 Exposure to Study Treatments

Patients enrolled in the UNITI studies and in the induction phase of CERTIFI received a single IV infusion of the treatment they were randomized to, administered by study staff.

Patients enrolled in IM-UNITI in the primary analysis population who were randomized to ustekinumab received treatment as follows (up to the point of meeting criteria for dosage adjustment due to loss of response):

- 90 mg every 12 weeks: 132 patients received a median cumulative dose of 360.0 mg
- 90 mg every eight weeks: 131 patients received a median cumulative dose of 540.0 mg.

Patients in IM-UNITI had a mean duration of follow-up of 32, 37, and 35 weeks in the placebo, ustekinumab every 12 weeks, and ustekinumab every eight weeks treatment groups. (Note: standard deviations for these means were not reported in the clinical study reports.)

The median average daily prednisone-equivalent oral corticosteroid dosage (excluding budesonide) at IM-UNITI baseline was the same for the ustekinumab groups (20.0 mg/day [IQR: 5.0 to 30.0 mg/day for every 12 weeks group; 10.0 to 25.0 mg/day for every eight weeks group]) and lower in the placebo group (15.0 mg/day [IQR: 10.0 to 25.0 mg/day]). The extent of exposure to oral corticosteroids during the 44 week study was not reported. [redacted]

3.5 Critical Appraisal

3.5.1 Internal Validity

There were a number of potential differences between studies and between the placebo and ustekinumab groups within studies with respect to certain baseline characteristics (Section 3.2.2.2). A notable potential imbalance was in the concomitant use of oral corticosteroids at baseline in UNITI-2, in which 44% of patients treated with ustekinumab and 36% of patients on placebo were receiving these drugs (Table 11). [redacted] The extent of corticosteroid exposure during UNITI-2 was not reported; however, the study protocol required that patients maintain their dose at a stable level throughout. This imbalance in the proportion of patients receiving oral corticosteroids [redacted], and given the documented efficacy of corticosteroids in inducing response and remission in Crohn's disease,[45] could have led to an overestimation of the response and remission outcomes for ustekinumab versus placebo, although there is no direct evidence that this occurred.

The IM-UNITI study protocol included a pre-specified regimen for tapering patients' oral corticosteroid if they had a clinical response and were receiving corticosteroids at week 0. Corticosteroid tapering is recommended by guidance from the European Medicines Agency (EMA) on the design of trials

evaluating treatments for Crohn's disease.[46] The clinical expert consulted by CDR noted that the tapering regimen used in the IM-UNITI trial was a reasonable reflection of clinical practice in Canada. The median average daily prednisone-equivalent oral corticosteroid dose (excluding budesonide) at IM-UNITI baseline was higher in the ustekinumab groups than the placebo group. However, the extent of exposure to oral corticosteroids during the course of the study was not reported, [redacted] There is no direct evidence that oral corticosteroid exposure in IM-UNITI affected outcomes.

The manufacturer indicated that an adaptive randomization approach (Pocock and Simon minimization method) was used in CERTIFI because many sites were expected to enroll very few participants, making it difficult to achieve balance in treatment assignment within each site if a traditional permuted-block randomization were used.[27,28] While the Pocock and Simon minimization method is a recognized approach to adaptive randomization, using adaptive randomization in general creates challenges in interpretation of the study outcome relative to a fixed randomization procedure, and trials of this design are considered to be "less well understood" by the FDA for this reason.[47] Specifically, concern has been expressed that changes to the randomization probabilities could create imbalances in known and unknown patient characteristics at the end of the study, thereby increasing the risk of bias. The FDA recommends caution with the use of adaptive randomization in confirmatory trials.[47]

The IM-UNITI study used re-randomization at week 8 for patients receiving ustekinumab who responded to induction therapy in the UNITI studies. 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 EMA guidance for the development of drugs for the treatment of Crohn's disease.[46] However, a limitation of this approach is that all patients enrolled in IM-UNITI are a selected population: they were responders to induction therapy in the UNITI studies and could 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, it may obscure the true effectiveness and occurrence of AEs.

Study treatments were administered in a double-blind manner in all of the included studies. Given that the maintenance study IM-UNITI included two dosage regimens of ustekinumab (i.e., 90 mg every 12 weeks and 90 mg every eight weeks), patients received an SC administration of study treatment (either placebo or ustekinumab) every four weeks from week 0 to week 40 with the exception of week 4. It is difficult to ascertain from the provided data whether the adverse event profile of ustekinumab could have compromised blinding in IM-UNITI (or the induction studies), given that there were few evident differences between groups regarding frequency of AEs, including the proportion of patients who experienced administration-related reactions.

Greater than 20% of patients prematurely discontinued treatment in IM-UNITI. For all analyses related to clinical remission and clinical response, patients who discontinued for any reason were considered to have failed treatment. This is a common approach to handling missing data in these types of trials, but may bias results in the case of differential withdrawal rates. In IM-UNITI, the overall proportion of withdrawals and the reasons for discontinuation were generally similar between the placebo and ustekinumab groups. Lack of efficacy was the most commonly cited reason for discontinuation in both the placebo and ustekinumab groups (11% in each group), suggesting that classifying patients who discontinued as having failed treatment may be an accurate reflection for half of those who failed to complete the study. High rates of withdrawal are not unusual in longer-term studies for Crohn's disease

and are consistent with the high rates of withdrawal (or early escape) reported in previous pivotal studies for TNF antagonists in the maintenance treatment of Crohn's disease. The primary analysis was supported by a number of sensitivity analyses to investigate alternative approaches for imputing missing data (e.g., observed case, LOCF, multiple imputation, and the worst case missing data method). In general, these analyses yielded results similar to the primary analysis.

The protocols for each study indicated that the primary and major secondary outcome analyses were based on the intention-to-treat principle. However, one patient randomized to treatment in UNITI-2 was excluded from the analysis because of major protocol deviations. This is reasonable, given that it was discovered that the patient did not meet the inclusion criteria for active Crohn's disease, and loss of one patient is unlikely to have affected the validity of the results in that study. However, for certain outcome analyses relevant to this review (i.e., IBDQ, SF-36, WLQ, and surgery) the intention-to-treat principle was not used.

As mentioned previously, the protocols for the UNITI studies and IM-UNITI stated that the CDAI score was calculated for a visit only if four or more of the eight components were available at that visit. When at least four of the eight components were available, any missing components were imputed by carrying forward the last non-missing component (with the exception of a missing hematocrit value). If the CDAI score could not be calculated (i.e., four or fewer components available) at a visit, the CDAI score was considered missing. Patients with a missing CDAI score at study end point were not considered to have achieved clinical response or remission, depending on the study. In UNITI-1, 13 (5%) patients each in the placebo and ustekinumab groups had missing CDAI scores at week 6. In UNITI-2, eight (4%) placebo-treated patients and three (1.4%) ustekinumab-treated patients had missing CDAI scores at week 6. In IM-UNITI, eight (6%), one (2%), and 10 (8%) patients in each of the placebo, every 12 weeks, and every eight weeks groups had missing CDAI scores at week 44; however, almost all of these patients were already coded as treatment failures (i.e., because of Crohn's disease–related surgery or medication adjustment before week 44). Differences between treatment groups may introduce bias; however, only in UNITI-2 and the ustekinumab every 12 weeks group in IM-UNITI were differences in missing data notably different. It is uncertain whether these differences had a measurable effect on the validity of the results in those two studies.

In the UNITI studies and IM-UNITI, analyses of health-related quality of life and functional status were limited for several reasons. [redacted] However, the direction of potential bias is unclear.

There was limited adjustment for multiple comparisons in the UNITI studies and IM-UNITI. The comparisons between treatment groups for health-related quality of life and functional outcomes, as well as surgery and subgroup analyses, were included in the pre-specified hierarchical analysis plan (Table 17). Given the number of outcomes and statistical comparisons made in each study, this is a potential limitation when interpreting study results.

The endoscopic substudy was limited by several factors. First, patients were enrolled only from select study sites within the UNITI study programs. Only 334 patients (142 from UNITI-1 and 192 from UNITI-2) of the total 1,409 enrolled in the UNITI studies were included in the endoscopy substudy. Only 252 of

334 patients met the inclusion criterion of a baseline SES-CD score of three or more points. Second, it was not described how patients were selected, and what impact this had on maintaining randomization. The manufacturer reported potential differences between the pooled ustekinumab group and placebo group with respect to certain baseline characteristics (median CDAI score, median CRP level, and extra-intestinal manifestations of Crohn's disease). It is possible that certain patient or disease factors were no longer adequately balanced between groups, which may have acted as important confounders or effect modifiers of the endoscopy analysis. Third, the primary outcome of the endoscopy substudy was the change from baseline in the SES-CD score at week 8 of induction. Studies have been conducted to determine the instrument's reliability and validity.[42,48] However, a recent Cochrane systematic review reported that, although the SES-CD is increasingly used in randomized controlled trials of interventions for Crohn's disease, its clinical relevance has not been fully elucidated.[49] In particular, the overall validity of the SES-CD has not been fully established, and SES-CD cut-off points used for endoscopic remission and response require additional study. Furthermore, the MCID for the change in SES-CD needs to be determined. Last, the SES-CD was centrally assessed by a single evaluator for all video endoscopies. Although intra- and inter-rater reliability have been reported as high for the SES-CD (interclass correlation coefficients of > 0.9), it has been noted that there were potentially important biases associated with these analyses.[49] Therefore, there is some uncertainty as to how reliable the assessments were.

3.5.2 External Validity

Inclusion and exclusion criteria for both the induction and maintenance studies were generally reflective of patients who would be considered candidates for treatment with ustekinumab in Canada. In its clinical practice guidelines on the use of TNF antagonists in the treatment of Crohn's disease, the Canadian Association of Gastroenterology (CAG) states that moderate-to-severe Crohn's disease should be defined as a CDAI score between 220 and 400.[50] This is consistent with the inclusion criteria of the UNITI studies and CERTIFI.

The study protocols for the included studies specifically defined inclusion criteria for inadequate response, loss of response, or intolerance to a previous TNF antagonist, immunomodulator, or corticosteroid (Table 9). The protocols stated that patients could be considered primary nonresponders to treatment with adalimumab after receiving one 80 mg dose followed by one 40 mg dose. This is half the dose of adalimumab recommended in the Canadian product monograph for inducing remission (i.e., 160 mg at week 0 and 80 mg at week 2).[20] Similarly, patients could be considered primary nonresponders following treatment with infliximab if they had received two or three doses of 5 mg/kg; the Canadian product monograph recommends three doses of 5 mg/kg for induction with infliximab (i.e., at weeks 0, 2, and 6).[18,19] Dosage recommendations in the CAG guidelines on the use of TNF antagonists in the treatment of Crohn's disease are consistent with those noted in the product monographs.[50] The clinical expert consulted by CDR noted that more recent understanding of drug levels has changed clinical practice. For patients with no or poor clinical response after two induction doses (week 0 and 2) of infliximab, most clinicians would provide a rescue dose of 10 mg/kg at weeks 4 and 6 before deciding whether the patient is a primary nonresponder. A similar approach would be taken with adalimumab in practice. In addition, patients could be eligible for enrolment in the UNITI and CERTIFI studies if they failed treatment with certolizumab pegol. Certolizumab has been approved for use in the treatment of Crohn's disease by the FDA, but not by Health Canada or the EMA. The CAG guidelines on the use of TNF antagonists state that certolizumab has been shown to be clinically effective in the treatment of Crohn's disease.[50] However, the FDA and the CAG guidelines recommend that certolizumab (400 mg SC) be administered at weeks 0, 2, and 4 for induction;[50,51] the criteria set by the manufacturer allow primary nonresponse to certolizumab to be declared after only two doses.

Differences between the UNITI studies with respect to baseline disease characteristics (e.g., higher baseline mean CDAI score, CRP, and proportion of patients with Crohn's disease complications in UNITI-1) likely reflect the main difference in the patient populations: namely, that patients with more advanced disease, based on their having failed TNF antagonist therapy, were enrolled in UNITI-1.

As well, a larger proportion of patients in UNITI-2 were receiving concomitant aminosalicylates at baseline (42% to 45%) than in UNITI-1 (20% to 22%). [redacted] However, aminosalicylates are at present not recommended for the treatment of Crohn's disease. [45,52] The clinical expert consulted by CDR did not consider this imbalance likely to be an important source of bias.

The lack of an active control group is a limitation of the studies. There are three other biologics indicated for the treatment of Crohn's disease in Canada that could be considered relevant comparators: infliximab, adalimumab, and vedolizumab. Infliximab and adalimumab are long-established TNF antagonists in the treatment of Crohn's disease, and studies could have been conducted to compare ustekinumab with these. The Crohn's disease development programs for ustekinumab and vedolizumab overlapped, so designing a study comparing these was likely not feasible. In the absence of a study directly comparing ustekinumab with other biologics, the manufacturer submitted an indirect comparison, and CDR identified two other indirect comparisons in addition (0).

Another key limitation is that the maintenance-treatment phases (IM-UNITI and the maintenance phase of CERTIFI) randomized or re-randomized patients who had a response or remission with various strengths of ustekinumab. The UNITI induction studies included two dose regimens for IV ustekinumab: 130 mg and a weight-based tiered dose approximating 6 mg/kg. CERTIFI included three IV ustekinumab doses: 1 mg/kg, 3 mg/kg, and 6 mg/kg. The product monograph for ustekinumab indicates that the approved induction regimen is the weight-based tiered dose approximating 6 mg/kg. However, IM-UNITI and the maintenance phase of CERTIFI randomized patients from the induction studies or phase regardless of the induction dose of ustekinumab used. Table 34 shows the distribution of patients in the IM-UNITI treatment groups by ustekinumab induction dose. [redacted] The proportions of patients who achieved clinical response at week six with ustekinumab 130 mg were 34% and 52%, and with 6 mg/kg were 34% and 56% in UNITI-1 and UNITI-2, respectively. Similarly, the proportions of patients who achieved clinical remission at week eight with ustekinumab 130 mg were 16% and 31%, and with 6 mg/kg were 21% and 40% in UNITI-1 and UNITI-2, respectively. Hence, randomizing patients from the ustekinumab 130 mg IV induction-treatment groups in the UNITI studies to IM-UNITI limits the generalizability of the IM-UNITI results to clinical practice. In practice, the 130 mg dose is not expected to be used because it is not the approved regimen. Of note, the maintenance phase of CERTIFI included patients who had a response and remission with doses not approved by Health Canada. For this reason, as well as others already mentioned, only data from the induction phase of CERTIFI were included in this review.

The CDAI has been validated within the Crohn's disease population (Appendix 5). The clinical expert consulted by CDR noted that CDAI scores are not calculated in clinical practice, although all of the various individual components of the scale are evaluated when assessing the status of a patient with Crohn's disease. The definition of clinical remission (i.e., CDAI score < 150) used in the studies is consistent with guidance from regulatory authorities[46,53] and with guidance from CAG.[50]

The combined duration of the UNITI induction and IM-UNITI maintenance studies was 52 weeks, which is consistent with guidance from regulatory authorities.[46,53] The duration of the pivotal studies may not have provided sufficient exposure to ustekinumab to allow adequate assessment of some of the more rare AEs (e.g., malignancy, progressive multifocal leukoencephalopathy, and serious infections).[54]

The generalizability of the results of the endoscopic substudy is unclear, given the small and select sample, pooling of ustekinumab induction and maintenance dosages, pooling of patients from two populations (i.e., inadequate response to conventional therapy plus inadequate response to TNF antagonists), and single centralized evaluator.

3.6 Efficacy

Only those efficacy outcomes identified in the review protocol (Section 2.2) are reported in this section. See 0 for additional efficacy data.

3.6.1 Clinical Remission

a) Induction Studies

A statistically significantly higher proportion of patients treated with ustekinumab (20.9% and 40.2%) than with placebo (7.3% to 19.6%; $P < 0.001$) had a remission at week 8 in UNITI-1 and UNITI-2, respectively (Table 21).

Subgroup analyses (based on those specified in the review protocol) are presented in Appendix 4 (Table 27). In general, differences between treatment groups with respect to the proportion of patients achieving remission at week 8 were consistent with those for remission in the overall study population.

The odds ratios for achieving remission were statistically significantly in favour of ustekinumab versus placebo at week 8, irrespective of baseline Crohn's disease severity (subgroups of CDAI score ≤ 300 or > 300) in both UNITI studies. Likewise, among patients in UNITI-1, a statistically significantly higher proportion of patients treated with ustekinumab, compared with placebo, were in remission at week 8, irrespective of response to previous treatment with conventional therapies. With respect to subgroups according to TNF antagonist use history in UNITI-1, remission rates were in favour of ustekinumab versus placebo, but statistically significant among those who had an initial response with, secondary nonresponse to, or intolerance to these drugs. Ustekinumab was also statistically significantly superior to placebo in achieving remission at eight weeks, regardless of whether patients enrolled in UNITI-2 had or had not received previous treatment with TNF antagonist(s).

b) Maintenance Study

Statistically significantly higher proportions of patients treated with ustekinumab every 12 weeks (48.8% and 42.6%; $P = 0.04$ and $P = 0.035$, respectively) and ustekinumab every eight weeks (53.1% and 46.9%; $P = 0.005$ and $P = 0.004$) were in clinical remission and corticosteroid-free remission, respectively, at week 44 of IM-UNITI than with placebo (35.9% and 29.8%) (Table 22). Among patients who were in remission at the start of IM-UNITI, 56.4% (44/78), 66.7% (52/78), and 45.6% (36/79) treated with the ustekinumab every 12 weeks regimen, ustekinumab every eight weeks, and placebo, respectively, were still in remission at week 44; however, the difference versus placebo was statistically significant only for the ustekinumab every eight weeks group.

Several pre-specified sensitivity analyses were performed on the remission end point. [redacted]

[redacted]

Subgroup analyses (based on those specified in the review protocol) are presented in Appendix 4 (Table 28). In general, differences between treatment groups with respect to the proportion of patients achieving remission at week 44 were consistent with those for remission in the overall study population.

[redacted] With respect to subgroups according to TNF antagonist use history, remission rates were in favour of ustekinumab versus placebo, but statistically significant only among those treated with ustekinumab every eight weeks and enrolled from UNITI-2, and those treated with ustekinumab every eight weeks who had not previously received TNF antagonist(s).

c) Phase II Study

The proportion of patients who had a remission in CERTIFI at week 6 was not statistically significantly different between the placebo (10.6%) and ustekinumab 6 mg/kg groups (12.2%) (Table 23).

3.6.2 Clinical Response

a) Induction Studies

The primary outcome for the UNITI induction studies was achieved: the proportion of patients in clinical response at week 6 was statistically significantly higher with the ustekinumab group (33.7% and 55.5%; $P = 0.003$ and $P < 0.001$, respectively) than with the placebo group (21.5% and 28.7%) in UNITI-1 and UNITI-2, respectively (Table 21).

In the sensitivity analyses performed for the primary outcome by observed case, LOCF, multiple imputation, and worst case, as well as excluding patients who were randomized but never treated, the results were consistent with those of the primary analysis.

Subgroup analyses (based on those specified in the review protocol) are presented in Appendix 4 (Table 29). In general, differences between treatment groups with respect to the proportion of patients achieving clinical response at week 6 were consistent with those for response in the overall study population.

The odds ratios for achieving clinical response were statistically significantly in favour of ustekinumab versus placebo at week 6, irrespective of baseline Crohn's disease severity (subgroups of CDAI score ≤ 300 or > 300) in both UNITI studies. Among patients in UNITI-1, a statistically significantly higher proportion of patients treated with ustekinumab compared with placebo were responders at week 6 among those who had failed to respond to previous treatment with conventional therapies; the between-group difference for response was not statistically significant among the subgroup of patients who had not failed conventional therapies. With respect to subgroups according to TNF antagonist use history in UNITI-1, remission rates were in favour of ustekinumab versus placebo, but statistically

significant among those who had an initial response with, secondary nonresponse to, or intolerance to these drugs. Ustekinumab was also statistically significantly superior to placebo in achieving remission at eight weeks regardless of whether patients enrolled in UNITI-2 had or had not received previous treatment with TNF antagonist(s).

b) Maintenance Study

Almost 60% of patients randomized to ustekinumab maintenance treatments were responders at week 44, whereas 44% of those assigned to placebo achieved clinical response. The comparison versus placebo was statistically significant for both ustekinumab regimens (*P* = 0.033 [every 12 weeks] and *P* = 0.018 [every eight weeks]) (Table 22).

c) Phase II Study

The proportion of patients who had a clinical response in CERTIFI at week 6 was statistically significantly different between the placebo (23.5%) and ustekinumab 6 mg/kg group (39.7%) (Table 23).

3.6.3 Health-Related Quality of Life, Functional and Disability Outcomes

None of the following analyses were adjusted for multiple comparisons and should be interpreted with this in mind.

a) Induction Studies

Ustekinumab-treated patients demonstrated statistically significantly greater improvements in IBDQ total score from baseline to week 8 in both UNITI-1 and UNITI-2 compared with patients receiving placebo (Table 21).

The mean scores for all four IBDQ dimension scores (bowel, emotional, systemic, and social) were statistically significantly improved from baseline with ustekinumab versus placebo through to week 8 of both induction studies (Table 30). The mean changes from baseline were higher in the ustekinumab group in UNITI-2 than in UNITI-1. Treatment differences appeared to be clinically significant only in UNITI-2 (MCID = 16 points).

A statistically significantly greater proportion of patients in the ustekinumab group (54.8%) had a ≥ 16-point improvement from baseline in the IBDQ score at week 8 compared with the placebo group (36.5%) in UNITI-1. Similarly, a statistically significantly greater proportion of patients in the ustekinumab group (68.1%) had a ≥ 16-point improvement from baseline in the IBDQ score at week 8 compared with the placebo group (41.1%) in UNITI-2 (Table 31).

As shown in Table 21, the mean change from baseline in the SF-36 PCS and MCS scores to week 8 were statistically significantly in favour of ustekinumab versus placebo, except for the PCS score change in UNITI-1. [redacted]

Statistically significantly greater proportions of ustekinumab-treated patients compared with placebo patients had at least a five-point improvement from baseline in the PCS and MCS scores of the SF-36 at week 8 in both studies, except on the PCS score in UNITI-1 (Table 31).

[redacted]

[redacted](Table 21).

b) Maintenance Study

The IBDQ total scores were statistically significantly higher at week 44 of IM-UNITI for both ustekinumab treatment regimens versus placebo (Table 22). IBDQ total scores decreased in all three treatment groups from baseline to week 44. Differences between groups were not considered clinically significant based on a MCID of 16 points.

[redacted](Table 32).

[redacted](Table 33).

As shown in Table 22, [redacted]

[redacted](Table 33).

[redacted](Table 22).

3.6.4 Mucosal Healing

The mean change from baseline in SES-CD score at week 8 of induction (endoscopy substudy primary outcome) was statistically significantly improved (decreased) in the pooled ustekinumab group (–2.8 points, standard deviation 5.7; *P* = 0.012) than in the placebo group (–0.7 points, standard deviation 5.0) (Table 35). The manufacturer reported that results of sensitivity analyses regarding approaches to handling missing data, and results across subgroup analyses by induction study and by induction dose, were consistent with the primary analysis; however, the data for these analyses were not included in the submitted materials.

The results of the first two key secondary outcomes in the endoscopy substudy — the change from induction baseline in the SES-CD score at week 44 of maintenance and the proportion of patients with mucosal healing at week 44 of maintenance — were not reported in the submitted materials. The manufacturer noted that the efficacy of ustekinumab maintenance for endoscopic outcomes could not be determined, primarily because of the very small sample size (N = 70).

The proportion of patients with mucosal healing at week 8 of induction was 9.0% and 4.1% for the pooled ustekinumab and placebo groups, respectively (*P* = 0.141).

3.6.5 Need for Surgery for Crohn's Disease

a) Induction Studies

[redacted] (Table 21).

b) Maintenance Study

[redacted] (Table 22).

TABLE 21: EFFICACY OUTCOMES FOR INDUCTION STUDIES

Parameter, n (%)	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
Clinical remission at week 8, n (%)	18 (7.3)	52 (20.9)	41 (19.6)	84 (40.2)
P value		< 0.001		< 0.001
Clinical response at week 6, n (%)	53 (21.5)	84 (33.7)	60 (28.7)	116 (55.5)
P value		0.003		< 0.001
Change in IBDQ total score at week 8	[redacted]			
Baseline score, mean (SD)				
Change from baseline, mean (SD)				
P value				
Change in SF-36 PCS score at week 8				
Baseline score, mean (SD)				
Change from baseline, mean (SD)				
P value				
Change in SF-36 MCS score at week 8				
Baseline score, mean (SD)				
Change from baseline, mean (SD)				
P value				
Time lost from work (during the previous 4 weeks) at week 8 (days)				
Mean (SD)				
P value				
Change in daily productivity (VAS)				
Baseline score, mean (SD)				
Change from baseline, mean (SD)				
P value				
Change in WLQ index at week 8				
Baseline score, mean (SD)				
Change from baseline, mean (SD)				
P value				
Number of patients with CD-related				

Parameter, n (%)	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
surgery through week 8				
n (%)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]

CD = Crohn's disease; IBDQ = Inflammatory Bowel Disease Questionnaire; MCS = mental component score; PCS = physical component score; PLA = placebo; SD = standard deviation; SF-36 = Short Form (36) Health Survey; UST = ustekinumab; VAS = visual analogue scale; WLQ = Work Limitations Questionnaire.
Source: Clinical Study Reports for UNITI-1 and UNITI-2.[7,8]

c) Maintenance Study

TABLE 22: EFFICACY OUTCOMES FOR THE MAINTENANCE STUDY

Parameter, n (%)	IM-UNITI		
	PLA (N = 131)	UST q.12.w (N = 129)	UST q.8.w. W (N = 128)
Clinical remission at week 44, n (%)	47 (35.9)	63 (48.8)	68 (53.1)
P value		0.04	0.005
CS-free clinical remission at week 44, n (%)	39 (29.8)	55 (42.6)	60 (46.9)
P value		0.035	0.004
Clinical remission at week 44 among patients in clinical remission at start of maintenance study	N = 79	N = 78	N = 78
n (%)	36 (45.6)	44 (56.4)	52 (66.7)
P value		0.189	0.007
Clinical response at week 44, n (%)	58 (44.3)	75 (58.1)	76 (59.4)
P value		0.033	0.018
Change in IBDQ total score at week 44	N = 130	N = 129	N = 128
Baseline score, median (IQR)	167.0 ([redacted])	172.0 ([redacted])	176.5 ([redacted])
Change from baseline, median (IQR)[a]	–14.5 ([redacted])	–2.5 ([redacted])	–2.0 ([redacted])
P value		[redacted]	[redacted]
Change in SF-36 PCS score at week 44	[redacted]	[redacted]	[redacted]
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD) [a]	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Change in SF-36 MCS score at week 44	[redacted]	[redacted]	[redacted]
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD) [a]	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Time lost from work (during the previous 4 weeks) at week 44 (days)	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Change in daily productivity (VAS) at week 44	[redacted]	[redacted]	[redacted]

Parameter, n (%)	IM-UNITI		
	PLA (N = 131)	UST q.12.w (N = 129)	UST q.8.w. W (N = 128)
Baseline score, median (IQR)	[redacted]	[redacted]	[redacted]
Change from baseline, median (IQR)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Change in WLQ index at week 44	[redacted]	[redacted]	[redacted]
Baseline score, median (IRQ)	[redacted]	[redacted]	[redacted]
Change from baseline, median (IQR)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Number of patients with CD-related surgery through week 44			
n (%)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]

CS = corticosteroid; IBDQ = Inflammatory Bowel Disease Questionnaire; IQR = interquartile range; MCS = mental component score; PCS = physical component score; PLA = placebo; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; SD = standard deviation; SF-36 = Short Form (36) Health Survey; UST = ustekinumab; VAS = visual analogue scale; WLQ = Work Limitations Questionnaire.
[a] Negative values indicate worsening.
Source: Clinical Study Report for IM-UNITI.[9]

TABLE 23: CLINICAL REMISSION AND RESPONSE IN THE PHASE II STUDY INDUCTION PHASE

Parameter	CERTIFI	
	PLA (N = 132)	UST 6 mg/kg (N = 131)
Clinical remission at week 6, n (%)	14 (10.6)	16 (12.2)
P value		0.682
Clinical response at week 6, n (%)	31 (23.5)	52 (39.7)
P value		0.005

PLA = placebo; UST = ustekinumab.
Source: Clinical Study Report for CERTIFI.[10]

3.7 Harms

Only those harms identified in the review protocol (Section 2.2) are reported in this section.

3.7.1 Adverse Events

a) Induction Studies

Almost two-thirds and greater than one-half of patients randomized to placebo or ustekinumab in UNITI-1 and UNITI-2, respectively, experienced an AE. In UNITI-1, the most commonly experienced AEs (in more than 5% of ustekinumab patients) were headache, arthralgia, pyrexia, nausea, and abdominal pain. In UNITI-2, these were nasopharyngitis, nausea, and pyrexia (Table 24).

b) Maintenance Study

Greater than 80% of patients in IM-UNITI reported an AE across the treatment groups. The most commonly experienced AEs (in more than 5% of patients receiving ustekinumab every eight weeks)

were arthralgia, Crohn's disease symptoms, headache, nasopharyngitis, upper respiratory infections, abdominal pain, pyrexia, cough, rash, and injection-site erythema (Table 25).

c) Phase II Study

A higher proportion of patients in the placebo group (71%) had an AE than in the ustekinumab group (61%). Gastrointestinal disorders and infections were the most common AEs, both more common in the placebo group (Table 26).

3.7.2 Serious Adverse Events

a) Induction Studies

The proportion of patients reporting SAEs in the treatment groups of both UNITI studies ranged from 3% to 7% over eight weeks (Table 24). Gastrointestinal disorders were the system-organ class in which the highest proportion of patients experienced an SAE: 3% to 4% of patients in the placebo groups and 2% to 3% of patients in the ustekinumab groups of UNITI-1 and UNITI-2. These SAEs were predominantly events of Crohn's disease or related symptoms and complications.

b) Maintenance Study

The proportion of patients reporting SAEs at week 44 in the treatment groups of IM-UNITI ranged from 10% (ustekinumab every eight weeks group) to 15% (placebo group). As in the induction studies, the most commonly experienced SAEs were gastrointestinal and occurred more frequently in the placebo group (Table 25).

c) Phase II Study

The proportion of patients reporting SAEs in CERTIFI were 8% for placebo and 7% for ustekinumab. Gastrointestinal events were the most commonly reported SAEs (Table 26).

3.7.3 Withdrawal Due to Adverse Events

a) Induction Studies

[redacted]The main AE was related to Crohn's disease (Table 24).

b) Maintenance Study

The proportion patients with a WDAE was lowest in the ustekinumab every eight weeks group (3.1%), compared with 6% for the placebo group and 7.6% for the ustekinumab every 12 weeks group. The primary AEs experienced were gastrointestinal disorders (Table 25).

c) Phase II Study

The proportion of patients with a WDAE in the induction phase of CERTIFI was similar to that in the UNITI studies (Table 26).

3.7.4 Mortality

There were no deaths in the included studies.

3.7.5 Notable Harms

a) Induction Studies

Less than 5% of patients in the UNITI studies experienced an infusion reaction, and no instances of anaphylaxis were reported. There was no clear difference between groups within the studies regarding infusion reactions (Table 24).

Infections were common, with approximately one-quarter of patients in each group in both UNITI studies reporting an infection. There were few patients with a serious infection over eight weeks of treatment; no cases of TB were reported in both studies.

None of the patients in UNITI-1 or UNITI-2 were reported as having developed cancer or experienced a major cardiovascular event.

Neurological AEs (primarily headaches) were relatively common in both studies.

b) Maintenance Study

Two per cent to 7% of patients treated with ustekinumab experienced an injection-site reaction versus less than 1% of those in the placebo group. No instances of anaphylaxis were reported (Table 25).

Infections were common, with approximately one-half of patients in each group reporting an infection. There were few patients with a serious infection, although a higher proportion was observed in the ustekinumab every 12 weeks regimen group (5.3%) versus the placebo and ustekinumab every eight weeks groups (both 2.3%). There was one suspected case of TB in a patient who received IV induction with 130 mg ustekinumab and was randomized to placebo in IM-UNITI.

None of the patients in IM-UNITI experienced a major cardiovascular event, but one patient randomized to each of placebo and ustekinumab every eight weeks group developed cancer over 44 weeks of maintenance therapy; in both cases, the diagnosis was basal cell carcinoma.

[redacted]

c) Phase II Study

Approximately 5% of patients in CERTIFI experienced an induction infusion reaction, and no instances of anaphylaxis were reported. There was no clear difference between groups regarding infusion reactions (Table 26).

Approximately one-quarter of patients in each group reported an infection, primarily nasopharyngitis. There were few patients with a serious infection during the induction phase, although there appeared to be a higher proportion of patients in the ustekinumab group (3.8%) than in the placebo group (0.8%). No one cause of infection was more frequently reported, and there were no reports of TB.

None of the patients was reported to develop cancer or experience a major cardiovascular event.

[redacted]

TABLE 24: HARMS FROM THE INDUCTION STUDIES

Parameter, n (%)	UNITI-1		UNITI-2	
	PLA (N = 245)	UST (N = 249)	PLA (N = 208)	UST (N = 207)
AEs				
Patients with > 0 AEs, n (%)	159 (64.9)	164 (65.9)	113 (54.3)	115 (55.6)
Most common AEs[a]				
Headache	22 (9.0)	20 (8.0)	14 (6.7)	10 (4.8)
Arthralgia	18 (7.3)	15 (6.0)	4 (1.9)	9 (4.3)
Pyrexia	15 (6.1)	15 (6.0)	10 (4.8)	11 (5.3)
Nausea	18 (7.3)	13 (5.2)	5 (2.4)	11 (5.3)
Abdominal pain	13 (5.3)	13 (5.2)	7 (3.4)	10 (4.8)
Nasopharyngitis	13 (5.3)	11 (4.4)	10 (4.8)	14 (6.8)
SAEs				
Patients with > 0 SAEs, n (%)	15 (6.1)	18 (7.2)	12 (5.8)	6 (2.9)
WDAEs				
WDAEs, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Most common reasons				
Crohn's disease	[redacted]	[redacted]	[redacted]	[redacted]
Deaths				
Number of deaths (%)	0	0	0	0
AEs of interest				
Infusion reactions	5 (2.0)	9 (3.6)	6 (2.9)	3 (1.4)
Anaphylaxis	0	0	0	0
Infections[b]	58 (23.7)	64 (25.7)	48 (23.1)	45 (21.7)
Serious infections[b]	3 (1.2)	7 (2.8)	3 (1.4)	1 (0.5)
Malignancy	0	0	0	0
Major cardiovascular events	0	0	0	0
Neurological	[redacted]	[redacted]	[redacted]	[redacted]
Multiple sclerosis[c]	[redacted]	[redacted]	[redacted]	[redacted]
Peripheral neuropathy[c]	[redacted]	[redacted]	[redacted]	[redacted]
Guillain–Barré syndrome[c]	[redacted]	[redacted]	[redacted]	[redacted]

AE = adverse event; PLA = placebo; SAE = serious adverse event; UST = ustekinumab; WDAE = withdrawal due to adverse event.
[a] Frequency > 5% in the UST group.
[b] Infection as assessed by the investigator.
[c] New onset or exacerbation of existing condition.
Source: Clinical Study Reports for UNITI-1 and UNITI-2.[7,8]

TABLE 25: HARMS FROM THE MAINTENANCE STUDY

Parameter, n (%)	IM-UNITI		
	PLA (N = 133)	UST q.12.w. (N = 132)	UST q.8.w. (N = 131)
AEs			
Patient with > 0 AEs, n (%)	111 (83.5)	106 (80.3)	107 (81.7)
Most common AEs[a]			
Arthralgia	19 (14.3)	22 (16.7)	18 (13.7)

Parameter, n (%)	IM-UNITI		
	PLA (N = 133)	UST q.12.w. (N = 132)	UST q.8.w. (N = 131)
Crohn's disease	19 (14.3)	16 (12.1)	16 (12.2)
Headache	15 (11.3)	15 (11.4)	16 (12.2)
Nasopharyngitis	10 (7.5)	17 (12.9)	14 (10.7)
URTI	[redacted]	[redacted]	[redacted]
Abdominal pain	16 (12.0)	13 (9.8)	11 (8.4)
Pyrexia	10 (7.5)	11 (8.3)	8 (6.1)
Cough	[redacted]	[redacted]	[redacted]
Rash	[redacted]	[redacted]	[redacted]
Injection-site erythema	1 (0.8)	1 (0.8)	7 (5.3)
SAEs			
Patients with > 0 SAEs, n (%)	20 (15.0)	16 (12.1)	13 (9.9)
WDAEs			
WDAEs, n (%)	8 (6.0)	10 (7.6)	4 (3.1)
Most common reasons			
GI disorders	[redacted]	[redacted]	[redacted]
Deaths			
Number of deaths (%)	0	0	0
AEs of Interest			
Injection-site reactions	1 (0.8)	3 (2.3)	9 (6.9)
Anaphylaxis	0	0	0
Infections[b]	66 (49.6)	61 (46.2)	63 (48.1)
Serious infections[b]	3 (2.3)	7 (5.3)	3 (2.3)
Malignancy	1	0	1
Major cardiovascular events	0	0	0
Neurological	[redacted]	[redacted]	[redacted]
Multiple sclerosis[c]	[redacted]	[redacted]	[redacted]
Peripheral neuropathy[c]	[redacted]	[redacted]	[redacted]
Guillain–Barré syndrome[c]	[redacted]	[redacted]	[redacted]

AE = adverse event; GI = gastrointestinal; PLA = placebo; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; SAE = serious adverse event; URTI = upper respiratory tract infection; UST = ustekinumab; WDAE = withdrawal due to adverse event.
[a] Frequency > 5% in the UST q.8.w. group.
[b] Infection as assessed by the investigator.
[c] New onset or exacerbation of existing condition.
Source: Clinical Study Report for IM-UNITI.[9]

TABLE 26: HARMS FROM THE PHASE II STUDY INDUCTION PHASE (BASELINE TO WEEK 8)

Parameter, n (%)	CERTIFI	
	PLA (N = 132)	UST 6 mg/kg (N = 131)
AEs		
Patient with > 0 AEs, n (%)	94 (71.2)	80 (61.1)
Most common AEs[a] SOC		
Infections and infestations		

Parameter, n (%)	CERTIFI	
	PLA (N = 132)	UST 6 mg/kg (N = 131)
Nasopharyngitis	6 (4.5)	8 (6.1)
Gastrointestinal	47 (35.6)	29 (22.1)
Abdominal pain	9 (6.8)	7 (5.3)
Nausea	11 (8.3)	8 (6.1)
Crohn's disease	13 (9.8)	5 (3.8)
Skin and subcutaneous tissue	[redacted]	[redacted]
Musculoskeletal and connective tissue	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]
Nervous system		
Headache	8 (6.1)	13 (9.9)
General and administration site	[redacted]	[redacted]
Respiratory, thoracic, mediastinal	[redacted]	[redacted]
Blood and lymphatic	[redacted]	[redacted]
Psychiatric	[redacted]	[redacted]
SAEs		
Patients with > 0 SAEs, n (%)	11 (8.3)	9 (6.9)
WDAEs		
WDAEs, n (%)	[redacted]	[redacted]
Most common reasons		
GI disorders	[redacted]	[redacted]
Deaths		
Number of deaths (%)	0	0
AEs of Interest		
Infusion reactions	6 (4.5)	7 (5.3)
Anaphylaxis	0	0
Infections	32 (24.2)	29 (22.1)
Serious infections	1 (0.8)	5 (3.8)
Malignancy	0	0
Major cardiovascular events	0	0
Neurological	[redacted]	[redacted]
Multiple sclerosis	[redacted]	[redacted]
Peripheral neuropathy	[redacted]	[redacted]
Guillain–Barré syndrome	[redacted]	[redacted]

AE = adverse event; GI = gastrointestinal; PLA = placebo; SAE = serious adverse event; SOC = system-organ class; UST = ustekinumab; WDAE = withdrawal due to adverse event.

[a] Frequency > 5% in any group.

Source: Clinical Study Report for CERTIFI.[10]

4. DISCUSSION

4.1 Summary of Available Evidence

The CDR review included four multinational, multi-centre, double-blind randomized controlled trials.[7-10] The UNITI-1 and UNITI-2 studies investigated the efficacy and safety of IV ustekinumab versus placebo for inducing clinical response in patients with moderate to severely active Crohn's disease after six weeks. Patients enrolled in UNITI-1 also had to have had an inadequate response with, intolerance to, or a contraindication to TNF antagonists, whereas the UNITI-2 population comprised patients who had had an inadequate response with, intolerance to, or a contraindication to conventional therapies (i.e., corticosteroids, AZA, 6-MP, and/or MTX). Patients who completed the UNITI studies and were still classified as responders by week 8 of could enroll in the maintenance-treatment study, IM-UNITI. IM-UNITI investigated clinical remission over 44 weeks for two regimens of SC ustekinumab versus placebo. Patients completing IM-UNITI could continue into the long-term extension study.[55] Last, a phase II dose-ranging study, CERTIFI, was included but was considered only as supportive owing to a number of important limitations (e.g., the adaptive randomization methodology). Evaluation of clinical remission and clinical response in the UNITI studies and IM-UNITI were based on changes in CDAI scores. Mucosal healing and reduction in the need for Crohn's disease–related surgery were outcomes of interest in the CDR review protocol; however, none of the studies were appropriately designed to evaluate these end points. The manufacturer conducted an endoscopy substudy[26] to evaluate mucosal healing; however, the substudy was limited methodologically and by the small sample size, particularly for the maintenance of effect (N = 70), to be able to draw any conclusions from this evidence. It was considered an exploratory analysis for the purposes of this review.

According to the clinical expert consulted by CDR, despite potential imbalances between treatment groups and study baseline characteristics, the trial populations were generally reflective of patients with moderate-to-severe Crohn's disease who are treated in Canadian clinical practice. One notable potential imbalance was in the concomitant use of oral corticosteroids at baseline in UNITI-2; 44% of ustekinumab-treated patients and 36% of placebo patients were receiving these drugs (Table 11). [redacted] The extent of corticosteroid exposure during UNITI-2 was not reported; however, the study protocol required that patients maintain their dose at a stable level throughout. Similar potential imbalance in corticosteroid use was noted for the maintenance study as well. These potential imbalances in concomitant oral corticosteroids may have led to an overestimation of the response and remission outcomes for ustekinumab versus placebo in UNITI-2 and IM-UNITI.

The definition of inadequate response to TNF antagonists may have led to misclassification of previous TNF treatment failure in some patients, based in part on the dosages of infliximab and adalimumab used to define this population. The number of patients who had received the approved dosages of infliximab and adalimumab, and therefore the number who had an adequate trial of TNF antagonist, was not specified. As well, the criteria included a TNF antagonist, certolizumab, that is not approved in Canada for the treatment of Crohn's disease.

There were no head-to-head comparisons of ustekinumab versus infliximab, adalimumab, or vedolizumab, which are the other biologics approved for the treatment of Crohn's disease in Canada. Therefore, the CDR review considered the results of three indirect comparisons that have been conducted to evaluate the comparative efficacy and safety of ustekinumab versus these drugs (Appendix 7).[12-15]

The 44-week maintenance-treatment study, IM-UNITI, included patients from the UNITI induction studies who had a clinical response using a dose that is not Health Canada–approved, ustekinumab 130 mg IV.

4.2 Interpretation of Results

4.2.1 Efficacy

Both UNITI induction studies demonstrated that ustekinumab 6 mg/kg IV is statistically significantly superior to placebo for inducing clinical response in adult patients with moderately to severely active Crohn's disease after six weeks. Likely of greater importance to patients, ustekinumab was also statistically significantly superior to placebo in inducing clinical remission by week 8. It demonstrated efficacy in two subpopulations: those with an inadequate response or intolerance to conventional therapies (corticosteroids or immunomodulators; UNITI-2 population) and in those with an inadequate response or intolerance to TNF antagonists (UNITI-1 population). The manufacturer did not report between-group percentage differences for the remission and response outcomes. Based on crude calculations, the ustekinumab versus placebo percentage difference for clinical response was approximately 27% in UNITI-2 and 12% in UNITI-1, and the estimates for clinical remission were approximately 21% in UNITI-2 and 14% in UNITI-1. Perhaps not surprisingly, the treatment effect differences in UNITI-1 are lower, given that study's population likely had more advanced Crohn's disease (i.e., had experienced a treatment failure with at least one TNF antagonist) as compared with UNITI-2 (i.e., had experienced treatment failure with conventional therapy only). The results from the induction phase of the CERTIFI study, with the same population enrolled as UNITI-1, were generally supportive of those of the UNITI studies.

The clinical expert consulted by CDR stated that these results are likely to be clinically meaningful, especially for remission among patients who had a treatment failure with TNF antagonists. According to the clinical expert, achieving response and remission after a patient experiences a failure with the initial biologic therapy is often difficult. And, because of its alternative mechanism of action, ustekinumab, like vedolizumab, may offer a biologically plausible alternative to trying another TNF antagonist among patients who experience treatment failure with the initial TNF antagonist. However, there is no trial-based comparative evidence for this and it remains hypothetical (see discussion of the manufacturer-provided network meta-analysis [NMA] in Appendix 7).

Both ustekinumab maintenance regimens were found to be statistically significantly superior in achieving clinical remission, as well as clinical response, after 44 weeks compared with placebo. However, the results are difficult to interpret. First, in accordance with the study protocol, these patients underwent a pre-specified corticosteroid-tapering regimen. Yet, the extent of exposure to oral corticosteroids in the treatment groups and its impact on the outcomes at week 44 were not clear from the manufacturer-provided clinical study reports. [redacted] Outcomes were not assessed for induction dose subgroups, because they were out of scope for this review. However, induction dose (ustekinumab 130 mg or 6 mg/kg) was a pre-specified subgroup in IM-UNITI (results for clinical remission at week 44 are presented in Appendix 4 for these subgroups). It does not appear that including this subpopulation in IM-UNITI had an important effect on clinical remission at week 44, based on similar estimates of effect for ustekinumab versus placebo in the 130 mg induction-dose subgroup and the 6 mg/kg induction-dose subgroup. Nonetheless, including this subpopulation reduces the generalizability of the maintenance study results, because not all patients were treated with the Health Canada–approved induction regimen.

An important consideration for patients and clinicians is whether the treatment can keep patients in remission over time. Approximately 78 patients per treatment group were in clinical remission at the start of IM-UNITI. Of these, 56% to 67% of ustekinumab-treated patients and 46% of placebo-treated patients were in remission at week 44; however, the difference between groups was only statistically significant for ustekinumab every eight weeks versus placebo. This outcome is difficult to interpret because it is based on a subset of patients, given that patients were not required to be in remission at the start of maintenance treatment. No pre-specified definition of durable clinical remission was provided in the study protocol; hence, it is not clear whether the study met a minimum threshold to conclude that these patients remained in remission in a clinically meaningful way.

Long-term treatment with corticosteroids is associated with an increased risk of serious AEs and is an important concern, as reported by patients. At the beginning of IM-UNITI, approximately one-half of the patients were receiving concomitant treatment with corticosteroids. A statistically significantly greater proportion of ustekinumab-treated patients (43% to 47%) achieved corticosteroid-free clinical remission at 44 weeks compared with the placebo group (30%). The clinical expert consulted by CDR considered the effect of ustekinumab to be clinically relevant for patients with Crohn's disease who are dependent on corticosteroids. Of note, corticosteroid-dependent patients with Crohn's disease are included in the Health Canada indication for ustekinumab.

In their input to CDR for this review, the patient groups noted that they hope for additional non-surgical options to treat Crohn's disease. They noted that surgery is associated with risks and should be considered the option of last resort. The clinical expert consulted by CDR also noted the importance of having additional non-surgical treatment options for patients with refractory Crohn's disease. None of the included studies was adequately designed to investigate the efficacy of ustekinumab for reducing the need for surgical intervention in patients with Crohn's disease, [redacted].

Crohn's disease has a profound negative impact on the quality of life of those living with the condition, according to the patient groups who provided input for this review. The UNITI studies and IM-UNITI used the SF-36 and IBDQ to assess the effects of ustekinumab on quality of life. Statistical analyses were conducted, but only for a subpopulation of patients for the SF-36. Analyses did not appear to be based on an intention-to-treat principle, and they were not adjusted for multiple comparisons. As well, adjusted least squares differences between groups were not reported for these outcomes. Therefore, the results for these patient-important outcomes need to be interpreted with these issues in mind. Ustekinumab was generally associated with greater improvement in these end points than placebo. [redacted][35] [redacted] It did not appear that the median difference in the IBDQ between ustekinumab and placebo groups in IM-UNITI (approximately 12 points for both groups) exceeded the published MCID for the IBDQ (i.e., an improvement of 16 points);[34] [redacted]. Of note, the European product monograph for ustekinumab states that improvements in IBDQ and SF-36 (MCS in both UNITI-1 and UNITI-2; PCS in UNITI-2), compared with placebo, were clinically meaningful.[56] The basis for this conclusion is not specified.

Patient groups also identified missed work days as a key issue related to Crohn's disease. The UNITI studies and IM-UNITI pre-specified several analyses to evaluate the effects of ustekinumab on missed work time and productivity. [redacted] (Appendix 5), [redacted].

In total, the UNITI studies plus IM-UNITI consist of one year of treatment experience with ustekinumab for Crohn's disease. Patients who completed IM-UNITI could continue in the ongoing extension study, which is planned to follow patients from week 44 of IM-UNITI to week 272; interim data up to week 92 were provided by the manufacturer and are summarized in Appendix 6. [redacted]

The Canadian product monographs for both infliximab and adalimumab provide details about dose-escalation scenarios for patients who fail to respond or who lose response to those products.[18-20] In contrast, the controlled phases of the pivotal studies for vedolizumab did not evaluate dosage escalation, and the current Canadian product monograph does not provide guidance regarding potential dosage-escalation scenarios. Dose escalation was investigated in the IM-UNITI study, in which patients who had a loss of clinical response at any scheduled visit between week 8 and week 32 were eligible for dosage adjustment (Section 3.2.3.1). A total of 57 patients treated with ustekinumab lost response and required dosage adjustment. Twenty-nine patients (22%) in the ustekinumab every 12 weeks group had a dosage adjustment to the every eight weeks regimen and, when assessed 16 weeks after their dosage was adjusted, 41.4% were reported to be in remission and 55% regained response. Twenty-eight patients (21%) in the ustekinumab every eight weeks group met the criteria for dosage adjustment due to loss of response but continued to receive ustekinumab 90 mg every eight weeks per protocol. When assessed 16 weeks afterwards, in all, 32.1% were reported to be in remission and 46% regained response. The product monograph for ustekinumab indicates that patients receiving the every 12 weeks regimen may have their dosage adjusted to every eight weeks if treatment goals are not being met. However, given the small number of patients evaluated and the absence of controlled studies examining dose escalation with ustekinumab, there is uncertainty regarding the overall clinical benefit and the comparative effectiveness of different dosage-escalation scenarios.

In the absence of head-to-head trials comparing the efficacy of ustekinumab with TNF antagonists or vedolizumab, CDR examined the results of two NMAs (Appendix 7).[12,13] The manufacturer submitted an NMA of ustekinumab versus infliximab, adalimumab, and vedolizumab.[12] The results of the

manufacturer's NMA suggested that there are no statistically significant differences in clinical response (CDAI-70), enhanced clinical response (CDAI-100), or clinical remission (CDAI < 150) between ustekinumab 6 mg/kg and adalimumab 80/40 mg, adalimumab 160/80 mg, or vedolizumab 300 mg in the induction phase in either the subpopulation that had experienced a failure of conventional therapy or of TNF antagonist therapy. Statistically significant differences in clinical response (CDAI-70) and clinical remission (CDAI < 150) in favour of infliximab 5 mg/kg compared with ustekinumab 6 mg/kg were reported in the subpopulation that had experienced a failure of conventional therapy. However, a number of variables were markedly different between the infliximab trials (namely ACCENT-1) and trials for other drugs, and these increase uncertainty concerning the indirect comparison for induction outcomes. Although the manufacturer conducted an NMA comparing maintenance therapies using treatment-sequence methodology, there is a very high degree of uncertainty as to the validity of the analysis methods, and as to trial-level clinical, design, and methodological heterogeneity. No firm conclusions on the comparative effectiveness of ustekinumab for maintenance treatment can be made from the manufacturer's NMA. Similar to the manufacturer's indirect comparison that was submitted to CDR, there is substantial clinical and methodological heterogeneity across the various studies included in the published NMA by Singh et al.,[13] as well as important limitations of the indirect comparison methods used.

Overall, there appear to be no statistically significant differences between ustekinumab, vedolizumab, and adalimumab for induction treatment, although the limitations related to the heterogeneity within the network means a conclusion of similar effects cannot be made conclusively. However, there is considerable uncertainty as to the validity of the treatment-sequence analysis, and no conclusion can be drawn regarding its results.

4.2.2 Harms

The proportions of patients who experienced at least one adverse event or serious adverse event were similar between the ustekinumab and placebo groups across all of the included studies. Nasopharyngitis and upper respiratory tract infections appeared to be more frequent with ustekinumab treatment than with placebo. As may be expected, patients treated with ustekinumab tended to report more administration-related reactions than those on placebo; however, there were no reports of anaphylaxis in any of the studies. Higher proportions of patients receiving placebo than receiving ustekinumab discontinued due to an adverse event, primarily because of gastrointestinal-related events, including worsening Crohn's disease. The interim report on the long-term extension study of IM-UNITI did not provide adverse event data.

The manufacturer provided a pooled analysis to Health Canada, which was also submitted to CDR,[26] consisting of one year of follow-up across all pooled indications for ustekinumab, with a total of 5,884 patients treated with ustekinumab (1,749 patients in the combined Crohn's disease studies, 3,117 in the combined psoriasis studies, and 1,018 in the combined psoriatic arthritis studies). Overall, the frequencies of AEs, SAEs, infections, serious infections, and WDAEs were similar between placebo-treated patients and ustekinumab-treated patients from the Crohn's disease studies compared with the adverse event profile in the approved indications of psoriasis and psoriatic arthritis. However, event proportions in the combined Crohn's disease studies were generally higher than those observed across the psoriatic indications and pooled indications, both among placebo- and ustekinumab-treated patients. The manufacturer suggested that this was a result of the underlying disease rather than ustekinumab treatment. There may be some validity to this, given that the most common event for SAEs and WDAEs in the studies included in this review were gastrointestinal disorders, including Crohn's disease–related events.

The CDR review summarized an NMA conducted by Mocko et al.,[14,15] which reported no statistically significant differences in the incidence of AEs, serious AEs, discontinuations due to AEs, or some of the more prominent AEs (e.g., infections, injection-site reactions, nausea, headache, arthralgia, etc.) among adalimumab, ustekinumab, or vedolizumab during induction therapy and adalimumab, infliximab, and vedolizumab during maintenance therapy in patients with Crohn's disease (Appendix 7). However, several major limitations associated with the conduct of this NMA introduce a very high degree of uncertainty regarding the results. Hence, caution is required when interpreting the authors' observations that there are no differences in safety among these drugs during the induction and maintenance phases of therapy for patients with CD.

Patient groups expressed an understanding of the potential risks associated with biologic treatments and noted that those living with Crohn's disease are often willing to accept these risks rather than undergo surgery, which they consider a last resort.

4.3 Potential Place in Therapy[1]

Based on current standards of practice with existing therapies, the clinical expert consulted by CDR indicated that there are several areas of unmet need where ustekinumab may play a role:

1. It may provide primary induction therapy for Crohn's disease for patients who experience primary nonresponse to either conventional therapy with immunomodulators or TNF antagonists.
2. It may also be useful in the setting of secondary nonresponse during maintenance therapy. An important proportion of patients with Crohn's disease patients lose response to TNF antagonist therapy during maintenance, either owing to formation of anti-drug antibodies or to inflammatory mechanisms that are independent of tumour necrosis factor. Evidence summarized in this review suggests that ustekinumab may provide clinically meaningful benefit in this patient group.
3. It may also provide salvage therapy for patients who respond to therapy with immunomodulators or TNF antagonists but who develop adverse effects. Immunomodulators such as AZA and MTX are generally safe medications; however, there are well-known side effects, including pancreatitis, neutropenia, hepatitis and neoplasia (e.g., skin cancers). TNF antagonists can be associated with severe allergic reactions, psoriatic skin diseases, neurological complications, congestive heart failure, lupus, and severe infections. In these situations, ustekinumab therapy may be safer and allow for continued treatment of the disease.

The clinical expert suggested that, before starting ustekinumab therapy, it is advisable to conduct the following:

1. assessment for previous TB exposure
2. hepatitis B serologic testing
3. pregnancy test in women of child-bearing age
4. immunization history and boosters for low antibody titers
5. staging of the degree of inflammatory activity (through colonoscopy and/or computed tomography enterography)

[1] This information is based on information provided in draft form by the clinical expert consulted by CDR reviewers for the purpose of this review.

5. CONCLUSIONS

Three phase III, double-blind randomized placebo-controlled trials investigated the effects of ustekinumab on treatment induction (UNITI-1 and UNITI-2) or maintenance (IM-UNITI) in patients with moderate-to-severe Crohn's disease. A single ustekinumab IV administration (approximating 6 mg/kg) appears to be significantly superior to placebo for inducing clinical response after six weeks of therapy. Likewise, both the ustekinumab 90 mg SC every 12 weeks and every eight weeks maintenance-treatment regimens were statistically significantly superior to placebo in achieving clinical remission and corticosteroid-free remission in patients who had a clinical response at week 8 of induction therapy. Moreover, these results for induction and maintenance therapy with ustekinumab were reported in subpopulations of Crohn's disease patients who had experienced a failure of conventional therapies or of TNF antagonist therapies. These findings were considered to likely be clinically meaningful by the clinician expert consulted by CDR. [redacted] These were reported as key outcomes by patients providing input to this review.

The proportion of patients who experienced at least one adverse event or serious adverse event was similar between the ustekinumab and placebo groups across all of the included studies. Nasopharyngitis and upper respiratory tract infection were more frequently reported in ustekinumab-treated patients than in placebo-treated patients, but these did not lead to discontinuation of treatment. Administration-related reactions were relatively rare.

There were no studies in which ustekinumab was compared directly with the approved TNF antagonists or vedolizumab for induction or maintenance treatment of Crohn's disease. Three indirect comparisons reviewed by CDR, including one submitted by the manufacturer, were challenging to interpret due to numerous limitations related to the source data and the NMA methods used to compare treatments. These limitations precluded any definitive conclusions regarding the efficacy and safety of ustekinumab compared with TNF antagonists and vedolizumab.

APPENDIX 1: PATIENT INPUT SUMMARY

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

1. Brief Description of Patient Groups Supplying Input

Two patient groups provided patient input for this submission: the Gastrointestinal (GI) Society and Crohn's and Colitis Canada (CCC).

The GI Society is committed to improving the lives of people with GI and liver conditions by supporting research, advocating for patient access in health care, and promoting GI and liver health. It provides evidence-based information through the BadGut basics patient information pamphlet and the Inside Tract/Du Coeur au ventre newsletter, BadGut lectures, GI support group meetings, continuing education events for health care professionals, and a website in English (www.badgut.org) and French (www.mauxdeventre.org). In the last two years, the GI Society has received funding from AbbVie Corporation, Actavis/Allergan, AstraZeneca Canada Inc., Ferring Inc., Gilead Sciences Canada Inc., GlaxoSmithKline Inc., Hoffmann-La Roche Limited, Innovative Medicines Canada, Janssen Canada, Johnson & Johnson, LifeScan, Merck Canada Inc., Pfizer Canada Inc./Hospira, Shire Canada Inc., and Takeda Canada Inc.

Crohn's and Colitis Canada (CCC) is a volunteer-based national charity dedicated to investing in education, awareness, and research for Crohn's disease (CD) and ulcerative colitis. CCC has received funding from individual donors and various pharmaceutical companies. In the 2014-2015 fiscal year, CCC received less than 11% of its total revenue from pharmaceutical companies, of which none were formally mentioned in their submission.

Both the GI Society and CCC have declared no conflict of interest with regard to the preparation of their submissions.

2. Condition-Related Information

The information in this section was collected through patient and caregiver interviews; a 2011 national survey conducted by the CCC; focus groups; a Canadian questionnaire conducted by the GI Society; patient roundtables; various interactions by telephone, email, and social media; stories obtained via patients over time; and a review of CCC published reports.

Crohn's disease is a disabling, lifelong inflammatory bowel disease (IBD), which is characterized by inflammation that can extend through the entire thickness of the bowel wall. According to the patient groups, Canada has the highest prevalence of CD in the world, with approximately 129,000 diagnosed patients. The disease can have a profound effect on patients' physical, emotional, and social well-being. In the patient input submissions, the groups expressed that the uncertainty of where and when the next flare will occur may lead to anxiety and stress, and may limit the places patients can go and/or the activities they participate in (including work and school). As one patient stated, *"It makes it difficult to leave my house, play with my son, work, etc. when I am in a flare."* This finding is supported by the CCC 2011 survey, which found that 43% of employed patients with IBD took some time off work, with an average of 7.2 missed days per year. Furthermore, 34% of respondents frequently were unable to play sports, 22% missed school trips, 40% avoided parties, and 22% did not attend special events.

Although the most commonly reported symptoms of IBD include bloody diarrhea, bloating, abdominal pain, and fatigue, the patient groups also noted that CD can lead to anemia, weight loss, fever, arthritis, ulcers of the mouth or skin, tender and inflamed nodules on the shins, and delayed development in children. The groups also reported that some patients were concerned about the increased risk of colon cancer with longstanding CD. The submissions noted two key concerns among patients with IBD. The first is the lack of control over bowel movements, including the urgent and frequent need to go to a bathroom. The CCC 2011 survey found that 73% of IBD patients reported five to 20, or even more, bowel movements per day. As one patient said, "*When you have to go to the washroom 20 times a day, it impacts everything you do*." The second major patient concern was a fear of flares and the desire for sustained remission, which has been suggested to be more important than relieving any one symptom of IBD. Concerns about future flares and uncertainty about their severity and occurrence were captured in numerous patient comments:

"*When I'm not in an active flare I live in constant fear of when the next flare will occur.*"
"The worst part is fearing the next big flare that will prevent me from being a mom to my 18-month-old."

Patient groups also reported an impact on caregivers, highlighting the inability of those who care for patients with CD to work and complete day-to day tasks, as well as the fatigue and stress associated with caregiving. Many caregivers lose their personal time to take on additional responsibilities that the person suffering from CD may no longer be able to complete. In addition, they may feel isolated and disempowered.

3. Current Therapy-Related Information

Management of CD is described as multi-faceted: it involves both symptom control and targeting of the underlying inflammation. Both submissions noted a lack of treatments available for CD. First-line therapy is aminosalicylates (e.g., 5-ASA, mesalamine) with steroids. If remission is not achieved or if the condition worsens, second-line therapy is immunomodulators (e.g., azathioprine), sometimes combined with corticosteroids (e.g., prednisone) and biologics. Patients reported few side effects with aminosalicylates, whereas some patients reported liver problems with immunomodulators. By contrast, the majority of patients reported side effects from steroids, the most common being mood swings, "moon face," and weight gain. Suppository formulations of corticosteroids are available, but patients find these inconvenient and say suppositories do not allow patients to maintain a normal routine. In addition, while these drugs may be effective in patients with mild-to-moderate disease, they often fail to maintain remission in the long-term, and are ineffective for moderate-to-severe disease. Interviews with patients suggested that these treatments help relieve some symptoms but do not offer control, as the need for constant and urgent washroom use remained. For patients who do progress to biologics, the cost and accessibility issues associated with them are important.

When the first- and second-line therapies fail to provide symptom relief, biologics are often considered effective to avoid surgery for patients with CD. The large majority of surveyed patients said they would rather receive a biologic, despite its potential risks and side effects, than undergo a colectomy. As noted by one patient, "*I have a strong desire to keep my body intact. The colon serves a myriad of beneficial functions."* According to the GI Society, surgical removal of the colon is not recommended by physicians in patients suffering from CD, as the disease can affect the entire GI tract and extend into the muscle wall. Patients further noted that surgery can be associated with later complications, including soiling, even more liquid bowel movements, poor pouch function, pouchitis, sexual dysfunction, and an increased risk of fertility loss among women. Surgery should only be considered as a last resort. Patients expressed their concern about surgery and the lack of treatment options available to them: *"Proposing*

surgery as a viable treatment option is inhumane and not fair. Surgery should be considered an option of last resort. It is a shame that there is nothing else to take."

4. Expectations About the Drug Being Reviewed

Patients hope that ustekinumab will provide them with another biologic option in their arsenal against CD, especially when other biologics either do not work or cease to work. Further, they hope that ustekinumab will provide them with a better quality of life — a life that is more normal and stable with less suffering from the effects of CD. In addition, they hope that ustekinumab will help to confer long-term remission from CD. Patients also find the route of administration (injection) appealing, as it will reduce the need to travel to infusion centres. However, the cost of ustekinumab is also a concern of patients, and many are worried that it will not be reimbursed.

One patient who responded and was currently receiving ustekinumab indicated that this was her third biologic in the last 16 years. She had first received infliximab (Remicade) for eight years before losing response and then trying adalimumab (Humira), which failed to treat CD. During the course of adalimumab, she had two additional bowel operations. Another patient who had experience with ustekinumab was able to avoid having irreversible surgery (removal of rectum) and was grateful to have received the drug. In addition, one other patient (who had both plaque psoriasis and CD) who received ustekinumab said that the drug eliminated the need for invasive and concurrent medication and greatly improved her quality of life.

Each case of CD is unique; physicians treating individual patients must take into account individual comorbidities and influences. What works for one person does not necessarily work for another. Choice among effective treatment options is essential for patients.

APPENDIX 2: LITERATURE SEARCH STRATEGY

OVERVIEW	
Interface:	Ovid
Databases:	Embase 1974 to present Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to present **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	October 7, 2016
Alerts:	Bi-weekly (twice monthly) search updates until February 15, 2017
Study Types:	No search filters were applied
Limits:	No date or language limits were used Conference abstracts were excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
MeSH	Medical Subject Heading
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
.ti	Title
.ab	Abstract
.ot	Original title
.hw	Heading word; usually includes subject headings and controlled vocabulary
.kf	Author keyword heading word (MEDLINE)
.kw	Author keyword (Embase)
.pt	Publication type
.rn	CAS registry number
.nm	Name of substance word
ppez	Ovid database code; MEDLINE Epub Ahead of Print, MEDLINE In-Process & Other Non-Indexed Citations, MEDLINE Daily and Ovid MEDLINE 1946 to Present
oemezd	Ovid database code; Embase 1974 to present, updated daily

MULTI-DATABASE STRATEGY

#	Searches
1	Ustekinumab/
2	(815610-63-0 or 949907-93-1 or FU77B4U5Z0).rn,nm.
3	(Stelara* or ustekinumab* or CNTO 1275 or CNTO1275).ti,ab,rn,nm,kf,hw,ot.
4	or/1-3
5	Crohn disease/
6	(crohn* or granulomatous colitis or granulomatous enteritis or regional enteritis or regional ileitis or

MULTI-DATABASE STRATEGY	
#	**Searches**
	regional ileitides or terminal ileitis or ileocolitis or cleron disease* or enteritis regionalis or regional enterocolitis).ti,ab,kf.
7	or/5-6
8	4 and 7
9	8 use ppez
10	*Ustekinumab/
11	(Stelara* or Ustekinumab* or CNTO 1275 or CNTO1275).ti,ab,kw.
12	or/10-11
13	exp Crohn disease/
14	(crohn* or granulomatous colitis or granulomatous enteritis or regional enteritis or regional ileitis or regional ileitides or terminal ileitis or ileocolitis or cleron disease* or enteritis regionalis or regional enterocolitis).ti,ab,kw.
15	or/13-14
16	12 and 15
17	16 use oemezd
18	conference abstract.pt.
19	17 not 18
20	9 or 19
21	remove duplicates from 20

OTHER DATABASES	
PubMed	A limited PubMed search was performed to capture records not found in MEDLINE. Same MeSH, keywords, limits, and study types used as per MEDLINE search, with appropriate syntax used.
Trial registries (Clinicaltrials.gov and others)	Same keywords, limits used as per MEDLINE search.

Grey Literature

Dates for Search:	September - October 2016
Keywords:	Stelara (ustekinumab); Crohn's disease
Limits:	No date or language limits 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
- Databases (free)
- Internet Search.

APPENDIX 3: EXCLUDED STUDIES

Reference	Reason for Exclusion
Sandborn W, Gasink C, Blank M, Lang Y, Johanns J, Gao LL, et al. A multicenter, double-blind, placebo-controlled phase3 study of ustekinumab, a human IL-12/23P40 mAB, in moderate-service Crohn's disease refractory to anti-TFN alpha: UNITI-1. Inflamm Bowel Dis. 2016 Mar;22(suppl 1):s1-O-001.	Abstract for UNITI-1
Sandborn WJ, Feagan BG, Fedorak RN, Scherl E, Fleisher MR, Katz S, et al. A randomized trial of Ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with moderate-to-severe Crohn's disease. Gastroenterology. 2008 Oct;135(4):1130-41.	Treatment regimens not Health Canada–approved

APPENDIX 4: DETAILED OUTCOME DATA

Clinical Remission Subgroup Analyses

TABLE 27: PROPORTION OF PATIENTS ACHIEVING CLINICAL REMISSION AT WEEK 8 BY DISEASE SEVERITY (BASELINE CDAI SCORE), CROHN'S DISEASE–RELATED MEDICATION HISTORY IN THE INDUCTION STUDIES

Subgroups, n (%)	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
Baseline CDAI score ≤ 300	[redacted]	[redacted]	[redacted]	[redacted]
Clinical remission at week 8	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted]		[redacted]
P value		[redacted]		[redacted]
Baseline CDAI score > 300	[redacted]	[redacted]	[redacted]	[redacted]
Clinical remission at week 8	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted]		[redacted]
P value		[redacted]		[redacted]
Inadequate response, intolerant, or contraindication to CS or immunomodulators — Yes	[redacted]	[redacted]		
Clinical remission at week 8	[redacted]	[redacted]		
OR (95% CI)		[redacted]		
P value		[redacted]		
Inadequate response, intolerant, or contraindication to CS or immunomodulators — No	[redacted]	[redacted]		
Clinical remission at week 8	[redacted]	[redacted]		
OR (95% CI)		[redacted]		
P value		[redacted]		
Initial response to TNF antagonist —Yes	[redacted]	[redacted]		
Clinical remission at week 8	[redacted]	[redacted]		
OR (95% CI)		[redacted]		
P value		[redacted]		
Initial response to TNF antagonist — No	[redacted]	[redacted]		
Clinical remission at week 8	[redacted]	[redacted]		
OR (95% CI)		[redacted]		
P value		[redacted]		
Primary nonresponder to TNF antagonist	[redacted]	[redacted]		
Clinical remission at week 8	[redacted]	[redacted]		
OR (95% CI)		[redacted]		
P value		[redacted]		
Secondary nonresponder to TNF antagonist	[redacted]	[redacted]		
Clinical remission at week 8	[redacted]	[redacted]		

Subgroups, n (%)	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
OR (95% CI)		[redacted]		
P value		[redacted]		
Intolerance to TNF antagonist	[redacted]	[redacted]		
Clinical remission at week 8	[redacted]	[redacted]		
OR (95% CI)		[redacted]		
P value		[redacted]		
Failed ≥ 2 TNF antagonists — Yes	[redacted]	[redacted]		
Clinical remission at week 8	[redacted]	[redacted]		
OR (95% CI)		[redacted]		
P value		[redacted]		
Failed ≥ 2 TNF antagonists — No	[redacted]	[redacted]		
Clinical remission at week 8	[redacted]	[redacted]		
OR (95% CI)		[redacted]		
P value		[redacted]		
Previously received TNF antagonist — Yes			[redacted]	[redacted]
Clinical remission at week 8			[redacted]	[redacted]
OR (95% CI)				[redacted]
P value				[redacted]
Previously received TNF antagonist — No			[redacted]	[redacted]
Clinical remission at week 8			[redacted]	[redacted]
OR (95% CI)				[redacted]
P value				[redacted]

CDAI = Crohn's Disease Activity Index; CI = confidence interval; CS = corticosteroid; OR = odds ratio; PLA = placebo; TNF = tumour necrosis factor; UST = ustekinumab.
Source: Clinical Study Reports for UNITI-1 and UNITI-2.[7,8]

TABLE 28: PROPORTION OF PATIENTS ACHIEVING CLINICAL REMISSION AT WEEK 44 BY DISEASE SEVERITY (BASELINE CDAI SCORE) AND CROHN'S DISEASE MEDICATION HISTORY IN THE MAINTENANCE STUDY

Subgroups, n (%)	IM-UNITI-		
	PLA (N = 131)	UST q.12.w. (N = 129)	UST q.8.w. (N = 128)
Baseline CDAI score ≤ 300	[redacted]	[redacted]	[redacted]
Clinical remission at week 44	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted]	[redacted]
P value		[redacted]	[redacted]
Baseline CDAI score > 300	[redacted]	[redacted]	[redacted]
Clinical remission at week 44	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted]	[redacted]
P value		[redacted]	[redacted]
Inadequate response, intolerant, or contraindication	[redacted]	[redacted]	[redacted]

Subgroups, n (%)	IM-UNITI-		
	PLA (N = 131)	**UST q.12.w. (N = 129)**	**UST q.8.w. (N = 128)**
to CS or immunomodulators			
Clinical remission at week 44	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted]	[redacted]
P value		[redacted]	[redacted]
TNF antagonist refractory population — Yes (UNITI-1)	N = 61	N = 57	N = 56
Clinical remission at week 44	16 (26.2)	22 (38.6)	23 (41.1)
OR (95% CI)		[redacted]	[redacted]
P value		0.140	0.102
TNF antagonist refractory population — No (UNITI-2)	N = 70	N = 72	N = 72
Clinical remission at week 44	NR (44.3)	NR (56.9)	NR (62.5)
OR (95% CI)		[redacted]	[redacted]
P value		0.146	0.020
Previously received TNF antagonist — Yes	N = 19	N = 19	N = 20
Clinical remission at week 44	NR (31.6)	NR (57.9)	NR (55.0)
OR (95% CI)		2.9 (0.7 to 11.7)	2.5 (0.7 to 9.3)
P value		0.150	0.206
Previously received TNF antagonist — No	N = 51	N = 53	N = 52
Clinical remission at week 44	NR (49.0)	NR (56.6)	NR (65.4)
OR (95% CI)		1.4 (0.6 to 3.0)	2.4 (1.0 to 5.5)
P value		0.512	0.041

CDAI = Crohn's Disease Activity Index; CI = confidence interval; CS = corticosteroid; NR = not reported; OR = odds ratio; PLA = placebo; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; TNF = tumour necrosis factor; UST = ustekinumab.
Source: Clinical Study Report for IM-UNITI.[9]

Clinical Response Subgroup Analyses

TABLE 29: PROPORTION OF PATIENTS ACHIEVING CLINICAL RESPONSE AT WEEK 6 BY DISEASE SEVERITY (BASELINE CDAI SCORE), CROHN'S DISEASE–RELATED MEDICATION HISTORY IN THE INDUCTION STUDIES

Subgroups, n (%)	UNITI-1		UNITI-2	
	PLA (N = 247)	**UST (N = 249)**	**PLA (N = 210)**	**UST (N = 209)**
Baseline CDAI score ≤ 300	N = 104	N = 94	N = 123	N = 119
Clinical response at week 6	NR (17.3)	NR (28.7)	NR (24.4)	NR (51.3)
OR (95% CI)		2.0 (1.0 to 3.9)		3.3 (1.9 to 5.7)
P value		0.048		< 0.001
Baseline CDAI score > 300	N = 143	N = 155	N = 86	N = 90
Clinical response at week 6	NR (24.5)	NR (36.8)	NR (34.9)	NR (61.1)
OR (95% CI)		1.8 (1.1 to 3.0)		3.2 (1.7 to 6.0)
P value		0.023		< 0.001
Inadequate response, intolerant, or contraindication to CS or immunomodulators — Yes	N = 215	N = 217		

Subgroups, n (%)	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
Clinical Response at week 6	NR (20.9)	NR (33.6)		
OR (95% CI)		1.9 (1.2 to 3.0)		
P value		0.005		
Inadequate response, intolerant, or contraindication to CS or immunomodulators — No	N = 32	N = 32		
Clinical Response at week 6	NR (25.0)	NR (34.4)		
OR (95% CI)		1.5 (0.5 to 4.8)		
P value		0.485		
Initial response to TNF antagonist —Yes	N = 187	N = 190		
Clinical response at week 6	NR (20.3)	NR (38.4)		
OR (95% CI)		2.5 (1.6 to 3.9)		
P value		< 0.001		
Initial response to TNF antagonist — No	N = 60	N = 59		
Clinical response at week 6	NR (25.0)	NR (18.6)		
OR (95% CI)		0.7 (0.3 to 1.7)		
P value		0.376		
Primary nonresponder to TNF antagonist	N = 74	N = 72		
Clinical Response at week 6	NR (23.0)	NR (23.6)		
OR (95% CI)		1.1 (0.5 to 2.4)		
P value		0.816		
Secondary nonresponder to TNF antagonist	N = 170	N = 171		
Clinical response at week 6	NR (20.0)	NR (36.8)		
OR (95% CI)		2.3 (1.4 to 3.8)		
P value		< 0.001		
Intolerance to TNF antagonist	N = 87	N = 105		
Clinical response at week 6	NR (24.1)	NR (34.3)		
OR (95% CI)		1.6 (0.8 to 3.0)		
P value		0.188		
Failed ≥ 2 TNF antagonists — Yes	N = 134	N = 126		
Clinical response at week 6	NR (19.4)	NR (34.9)		
OR (95% CI)		2.1 (1.2 to 3.8)		
P value		0.008		
Failed ≥ 2 TNF antagonists — No	N = 113	N = 123		
Clinical response at week 6	NR (23.9)	NR (32.5)		
OR (95% CI)		1.5 (0.8 to		

Subgroups, n (%)	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
		2.7)		
P value		0.176		
Previously received TNF antagonist — Yes			N = 74	N = 65
Clinical remission at week 8			NR (21.6)	NR (53.8)
OR (95% CI)				4.3 (2.0 to 9.2)
P value				< 0.001
Previously received TNF antagonist — No			N = 135	N = 144
Clinical remission at week 8			NR (32.6)	NR (56.3)
OR (95% CI)				2.8 (1.7 to 4.6)
P value				< 0.001

CDAI = Crohn's Disease Activity Index; CI = confidence interval; CS = corticosteroid; NR = not reported; OR = odds ratio; PLA = placebo; TNF = tumour necrosis factor; UST = ustekinumab.
Source: Clinical Study Reports for UNITI-1 and UNITI-2.[7,8]

Inflammatory Bowel Disease Questionnaire and Short Form (36) Health Survey Additional Data

TABLE 30: SUMMARY OF CHANGE FROM BASELINE IN THE IBDQ DIMENSION SCORES IN THE INDUCTION STUDIES

Parameter	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
Change in bowel score at week 8				
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]
Change in emotional score at week 8				
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]
Change in systemic score at week 8				
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]
Change in social score at week 8				
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]

PLA = placebo; SD = standard deviation; UST = ustekinumab.
Source: Clinical Study Report for UNITI-1 and UNITI-2.[7,8]

TABLE 31: RESPONDER ANALYSES FOR CHANGE FROM BASELINE IN INFLAMMATORY BOWEL DISEASE QUESTIONNAIRE AND SHORT FORM (36) HEALTH SURVEY IN THE INDUCTION STUDIES

Parameter	UNITI-1		UNITI-2	
	PLA (N = 247)	UST (N = 249)	PLA (N = 210)	UST (N = 209)
≥ 16 point improvement from baseline in IBDQ total score at week 8				
n (%)	89 (36.5)	136 (54.8)	85 (41.1)	141 (68.1)
P value		< 0.001		< 0.001
≥ 5 point improvement from baseline in SF-36 PCS score at week 8	[redacted]	[redacted]	[redacted]	[redacted]
n (%)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]
≥ 5 point improvement from baseline in SF-36 MCS score at week 8	[redacted]	[redacted]	[redacted]	[redacted]
n (%)	[redacted]	[redacted]	[redacted]	[redacted]
P value		[redacted]		[redacted]

IBDQ = Inflammatory Bowel Disease Questionnaire; MCS = mental component score; PCS = physical component score; PLA = placebo; SF-36 = Short Form (36) Health Survey; UST = ustekinumab.
Source: Clinical Study Report for UNITI-1 and UNITI-2.[7,8]

TABLE 32: SUMMARY OF CHANGE FROM BASELINE IN THE INFLAMMATORY BOWEL DISEASE QUESTIONNAIRE DIMENSION SCORES IN THE MAINTENANCE STUDY

Parameter	IM-UNITI		
	PLA (N = 131)	UST q.12.w. (N = 129)	UST q.8.w. (N = 128)
Change in bowel score at week 44			
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Change in emotional score at week 44			
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Change in systemic score at week 44			
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
Change in social score at week 44			
Baseline score, mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]

PLA = placebo; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; SD = standard deviation; UST = ustekinumab.
Source: Clinical Study Report for IM-UNITI.[9]

TABLE 33: RESPONDER ANALYSES FOR CHANGE FROM BASELINE IN INFLAMMATORY BOWEL DISEASE QUESTIONNAIRE AND SHORT FORM (36) HEALTH SURVEY IN THE MAINTENANCE STUDY

Parameter	IM-UNITI		
	PLA (N = 131)	UST q.12.w. (N = 129)	UST q.8.w. (N = 128)
≥ 16 point improvement from baseline in IBDQ total score at week 44			
n (%)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
≥ 5 point improvement from baseline in SF-36 PCS score at week 44	[redacted]	[redacted]	[redacted]
n (%)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]
≥ 5 point improvement from baseline in SF-36 MCS score at week 44	[redacted]	[redacted]	[redacted]
n (%)	[redacted]	[redacted]	[redacted]
P value		[redacted]	[redacted]

IBDQ = Inflammatory Bowel Disease Questionnaire; MCS = mental component score; PCS = physical component score; PLA = placebo; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; SF-36 = Short Form (36) Health Survey; UST = ustekinumab.
Source: Clinical Study Report for IM-UNITI.[9]

TABLE 34: CLINICAL REMISSION AT WEEK 44 OF MAINTENANCE STUDY BY USTEKINUMAB INDUCTION REGIMEN

Subgroups, n (%)	IM-UNITI		
	PLA (N = 131)	UST q.12.w. (N = 129)	UST q.8.w. (N = 128)
UST 130 mg IV induction regimen	[redacted]	[redacted]	[redacted]
Clinical remission at week 44	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted]	[redacted]
P value		[redacted]	[redacted]
UST 6 mg/kg IV induction regimen	[redacted]	[redacted]	[redacted]
Clinical remission at week 44	[redacted]	[redacted]	[redacted]
OR (95% CI)		[redacted]	[redacted]
P value		[redacted]	[redacted]

CI = confidence interval; IV = intravenous; OR = odds ratio; PLA = placebo; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; UST = ustekinumab.
Source: Clinical Study Report for IM-UNITI.[9]

TABLE 35: SUMMARY OF RESULTS FOR ENDOSCOPIC OUTCOMES AT WEEK 8 OF INDUCTION AND WEEK 44 OF MAINTENANCE STUDY

Parameter[a]	PLA (N = 97)	UST[b] (N = 155)
Change from baseline in SES-CD at week 8 of induction, mean (SD) *P* value	–0.7 (4.97)	–2.8 (5.68) 0.012
Change from induction baseline in the SES-CD score at week 44 of maintenance, mean (SD) *P* value	Not reported	Not reported
Patients with mucosal healing at week 44 of maintenance, n (%) *P* value		
Patients with mucosal healing at week 8 of induction, n (%) *P* value	4 (4.1)	14 (9.0) 0.141[c]

PLA = placebo; SD = standard deviation; SES-CD = Simplified Endoscopic Activity Score for Crohn's Disease; UST = ustekinumab.
[a] Only the primary outcome (i.e., change from baseline in the SES-CD score at week 8 of induction) and key secondary outcomes pre-specified and part of the hierarchical analysis plan in the endoscopy substudy are reported.
[b] Ustekinumab 130 mg and tiered ustekinumab doses ~6 mg/kg combined.
[c] Note that the statistical analysis hierarchy also failed at a higher-order comparison.
Source: Manufacturer's submission to CDR.[26]

APPENDIX 5: VALIDITY OF OUTCOME MEASURES

Aim

To summarize the measurement properties (e.g., reliability, validity, minimally clinically important difference [MCID]) of the following outcome measures used in the studies included in this submission:

- Crohn's Disease Activity Index (CDAI)
- Inflammatory Bowel Disease Questionnaire (IBDQ)
- Short Form (36) Health Survey (SF-36)
- Work Limitations Questionnaire (WLQ)

Findings

Crohn's Disease Activity Index

The National Cooperative Crohn's Disease Study Group developed the CDAI using prospective data gathered from 187 visits of 112 patients suffering from Crohn's disease (CD).[57] It is a disease-specific index and considered the standard for assessing CD activity. The CDAI consists of eight domains that are used to evaluate overall disease severity. The overall score is based on the sum of the weighted value of each item and ranges from 0 to 600; a score of 150 is defined as the threshold between remission and active disease. Scores ranging between 150 and 219 indicate mild-to-moderate CD and scores ranging between 220 and 450 indicate moderate-to-severe CD, whereas scores above 450 indicate very severe CD.[58,59] Item scores are derived using patient diaries, which are based on the seven days preceding each visit. Generally, the CDAI is considered impractical for use in clinical practice, as it has no MCID clearly defined.[59,60] Originally, changes of 50 points in the CDAI were associated with physician evaluation of "slightly better" and/or "slightly worse" compared with baseline.[57,59,60] However, clinical trials have commonly used changes of 50, 60, 70, or 100 points in CDAI to define clinical response.[59] More recently, the FDA and the European Medicines Agency (EMA) have suggested that a change of 100 points in CDAI is considered to be a more meaningful response (i.e., enhanced clinical response).[59]

Development of the CDAI

Gastroenterologists considered 18 parameters to inform the CDAI, including the following CD domains: subjective patient symptoms and need for symptomatic medications; objective clinical findings on physical examination; extra-intestinal manifestations of CD; complications of CD (e.g., fistulas); radiologic and endoscopic examinations; and laboratory parameters. A global assessment score was also assessed at each visit by the gastroenterologist based on the following scheme: "very well" = 1, "fair to good" = 3, "poor" = 5, "very poor" = 7.

Multiple regression and backwards stepwise deletions were utilized to assess the correlation between the 18 parameters and the physician global assessment score. Based on the results of the correlations, eight independent weighted (weighting ranges from one to 30) variables were included in the final CDAI formula.

TABLE 36: FINAL ITEMS INCLUDED IN THE CDAI AND THEIR WEIGHTS

Item (Daily Sum Per Week)	Weight
Number of liquid or very soft stools	2
Abdominal pain score in one week (rating: 0 to 3)	5
General well-being (rating: 0 to 4)	7
Sum of findings per week: • Arthritis/arthralgia • Mucocutaneous lesions (egg, erythema nodosum aphthous ulcers) • Iritis/uveitis • Anal disease (fissure, fistula, etc.) • External fistula (enterocutaneous, vesicle, vaginal, etc.) • Fever > 37.8°C	20
Antidiarrheal use (e.g., diphenoxylate hydrochloride)	30
Abdominal mass (none = 0, equivocal = 2, present = 5)	10
47 minus hematocrit (males) or 42 minus hematocrit (females)	6
100 × (1 – [body weight divided by standard weight])	1

Source: Best et al.[57]

Reliability of the CDAI

Reliability was not originally assessed during the development of the CDAI; however, the index did provide good to very good test–retest reliability, evaluated based on two successive visits for 32 patients.[57,58] The CDAI was subsequently re-evaluated and re-derived using data collected from 1,058 patients. This demonstrated little difference from the original formulation; therefore, the original version was recommended.[61]

Validity of the CDAI

Content validity: The items included in the CDAI were selected by gastroenterologists and are based on accepted features of CD, therefore demonstrating content validity.[58]

Construct validity: The CDAI appears to be able to distinguish between differing levels of CD severity.

The CDAI appears to be widely used in clinical trials, and is a measure accepted by gastroenterologists as a primary end point to assess CD activity. In contrast, the CDAI does not appear to be reflective of CD activity for pediatric patients suffering from CD, nor does the instrument address all aspects of CD such as quality of life.[58]

Criterion validity: Selecting a gold standard measure for comparison is difficult when considering CD because of the heterogeneous nature of its manifestations. Generally, the CDAI does not demonstrate any significant correlation between the overall score and objective measurements such as mucosal healing; however, the lack of correlation may not be indicative of a lack of criterion validity, given the multi-faceted nature of CD.[58] Predictive validity is another component of criterion validity. One study demonstrated that the CDAI scores increased two months preceding exacerbations of CD and decreased one month following exacerbations of CD, therefore demonstrating criterion validity.[58]

The CDAI score appears to vary depending on the observer,, even if the observers are evaluating the same case histories.[62]

Limitations of the CDAI

The overall CDAI score is derived from some subjective items such as "general well-being" and "intensity of abdominal pain" based on patient perception.

Inflammatory Bowel Disease Questionnaire

The IBDQ, developed by Guyatt et al.,[32,33] is a physician-administered questionnaire to assess health-related quality of life in patients with inflammatory bowel disease (e.g., ulcerative colitis and Crohn's disease).[63] It is a 32-item Likert-based questionnaire, divided into four dimensions (i.e., bowel symptoms [10 items], systemic symptoms [5 items], emotional function [12 items], and social function [5 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 between 32 and 224 (i.e., higher scores representing better quality of life). Scores of patients in remission typically range from 170 to 190.

This questionnaire has been validated in a variety of settings, countries, and languages.[63] A review of nine validation studies on the IBDQ in patients with inflammatory bowel disease reported that the IBDQ was able to differentiate clinically important differences between patients with disease remission and those with disease relapse. In a randomized placebo-controlled trial of patients with ulcerative colitis, the IBDQ was found to be able discriminate changes in the social and emotional state of patients.[64] The IBDQ has high test–retest reliability in all four dimensional scores (intraclass correlation coefficient = 0.96 for CD). Six studies evaluated IBDQ for sensitivity to change and all found that changes in health-related quality of life (HRQoL) were correlated with changes in clinical activity in patients with CD.[63]

A study conducted by Gregor et al.[34] noted that a clinically meaningful improvement in quality of life would be an increase of 16 points or more in the IBDQ total score or 0.5 points or more per question in patients with CD.

Short Form (36) Health Survey (SF-36)

SF-36 is a generic health assessment questionnaire that has been used in clinical trials to study the impact of chronic disease on HRQoL. SF-36 consists of eight domains: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health. SF-36 also provides two component summaries: the physical component summary (PCS) and the mental component summary (MCS), which are created by aggregating the eight domains. The SF-36 PCS, MCS, and eight domains are each measured on a scale of 0 to 100, with an increase in score indicating improvement in health status. In general use of SF-36, a change of two to four points in each domain or two to three points in each component summary indicates a clinically meaningful improvement, as determined by the patient.[36]

Validation of the survey indicates satisfactory reliability and discriminant ability for all SF-36 dimensions in patients with ulcerative colitis.[65] As symptoms increased, HRQoL scores statistically significantly reduced. In a population-based cohort in which patients were studied for 10 years, SF-36 scores of patients with ulcerative colitis were found to be comparable to those of a general population sample when adjusted for age, gender, and education.[65] The study indicated that the individual domains may present with ceiling effects in patients with less severe ulcerative colitis. Individual domain scores were also found to have less responsiveness in patients with mild ulcerative colitis, although it is unclear whether this can be generalized to the broader PCS and MCS scores.[65]

A study by Coteur et al.[35] explored MCID estimates within the CD patient population using data from multinational, multi-centre, double-blind, placebo-controlled parallel-group clinical trials in which

clinical remission of CD was assessed using the CDAI measure as the primary outcome. Secondary outcomes included the IBDQ and SF-36. All end points were measured at weeks 0, 6, 16, and 26 using standardized procedures. A total of six estimates of MCID — two analyses utilizing anchor-based methods and four analyses utilizing distribution-based methods — were evaluated for each SF-36 scale summary score to determine the most appropriate measure to use as the anchor. For the anchor-based estimates, a linear regression was performed using the two anchors, the CDAI and IBDQ. The MCID estimates for the SF-36 were then extracted from the regression equations considering a change of 16 points for the IBDQ total score or a score change of 50 points for the CDAI score as meaningful. For distribution-based estimates, measures rely on the statistical distributions of HRQoL data and include effect size (ES) measures (ES of 0.2 and 0.5 were used and suggested as small-to-moderate ESs), the standard error of measurement, and the standard error of the difference. Overall, the MCID ranged from 1.6 to 7.0 for the SF-36 PCS and from 2.3 to 8.7 for the MCS summary, depending on the approach. Because score changes in the SF-36 showed greater correlations with score changes in the IBDQ than with the CDAI, the IBDQ was selected as the "best anchor," with corresponding MCID values of 4.1 (IBDQ) and 3.9 (CDAI). The values derived by the IBDQ anchor-based method were similar to the values obtained by the distribution-based methods and were representative of small-to-moderate ESs. However, because of a number of methodological issues with this analysis, the general SF-36 MCIDs were used in this review to assess clinical significance for the SF-36.

Work Limitations Questionnaire (WLQ)

The WLQ is a self-reported tool used to assess and measure the on-the-job impact of chronic conditions and diseases and the treatment associated with them.[39] It was developed as a generic (non–disease-specific) instrument.[37-39] To develop the WLQ, focus groups were convened (for content identification), cognitive interviewing was performed (to enhance the reliability and validity of the questionnaire), and alternative forms were assessed (to assess the reliability of three different forms) in employed patients (greater than 20 hours/week) between the ages of 18 and 64 who had one of the following chronic conditions/diseases: asthma, Crohn's disease, liver disease, psychiatric disorders, and epilepsy.[39] Two studies were then performed, one study to assess the scale and recall, and the other study to assess scale reliability, construct validity, and relative validity.[39] The WLQ consists of the aforementioned four domains with 25 items; three domains ultimately examine the proportion of time over the previous two weeks that the patient had difficulties in the following four domains: time management or scheduling demands (5 items), output demands (5 items), physical demands (6 items), and mental-interpersonal demands (9 items).[37-39,66] The physical demands domain has reverse instructions and assesses the proportion of time without difficulty.[37-39,66] Scale responses and their corresponding scores are the same for all four domains: all of the time = 1 (100%); most of the time = 2; some of the time = 3 (about 50%); a slight bit of the time = 5; none of the time = 5; or does not apply to my job = 0.[38] The scores are converted from the computed mean of the non-missing responses to answers ranging from 0 (not limited) to 100 (limited all of the time); however, the response orientation is reversed for the physical demands domain.[37,38]

Although no studies identified the validity and reliability of the WLQ solely in patients with CD, these aspects have been assessed and verified in numerous indications, such as among cancer survivors[40]and in patients with rheumatoid arthritis, osteoarthritis, and other musculoskeletal conditions.[37,66,67] The WLQ has been deemed effective in assessing the responsiveness regarding work productivity changes in patients with either rheumatoid arthritis or osteoarthritis; however, Beaton et al.[66] did observe a lower-than-expected correlation in productivity-oriented constructs in this patient population. In addition, in the study by Walker et al.[67] the authors observed that, while the WLQ is reliable in assessing work productivity, it was not as strong as the Health Assessment Questionnaire and SF-36 in detecting

functional limitations in patients with rheumatoid arthritis (in part because these patients tend to select jobs that they can perform).

Tamminga et al.[40] observed that the minimal important change in improvement in their cohort of patients with cancer was 3.2 (based on the mean change method) and 4.0 (based on the receiver operating curve method); however, this was observed only at the group level and not at the individual level.

TABLE 37: SUMMARY OF OUTCOMES MEASURES

Measure	Definition	Evidence of Validity	MCID	Reference
CDAI	Physician-evaluated 8-item CD-specific index used to assess CD severity	Yes	NA	Best et al.[57]
IBDQ	Physician-administered 32-item questionnaire used to assess HRQoL in patients with IBD	Yes	16	Gregor et al.[34]
SF-36	Patient-reported generic QoL instrument	Yes	PCS 2 to 4.1 MCS 3 to 3.9	Coteur et al.[35] SF-36 Manual[36]
WLQ	Patient-reported instrument assessing on-the-job-impact of disease	Yes	3.2 (MCM) 4.0 (ROC)	Tamminga et al.[40] – in cancer survivors

CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; HRQoL = health-related quality of life; IBD = inflammatory bowel disease; IBDQ = Inflammatory Bowel Disease Questionnaire; MCS = mental component score; MCID = minimal clinically important difference; MCM = mean change method; NA = not applicable; PCS = physical component score; QoL = quality of life; ROC = receiver operating curve; SF-36 = Short Form (36) Health Survey; WLQ = Work Limitations Questionnaire.

Conclusion

The CDAI, IBDQ, and SF-36 have all been assessed within the CD population, whereas the WLQ has not. Although a minimal clinically important change in the CDAI, IBDQ, and SF-36 instruments has not been defined, some regulatory agencies rely on a reduction of 100 points in the CDAI as meaningful change, while other studies suggest an MCID of 16, 4.1, and 3.9 for the IBDQ, SF-36 PCS, and SF-36 MCS, respectively. WLQ MCIDs are lacking for patients with CD; however, there are some MCIDs (Table 37) associated with studies in cancer survivors (Tamminga et al.[40]) — 3.2 using the mean change method and 4.0 using the receiver operating curve.

APPENDIX 6: SUMMARY OF THE IM-UNITI EXTENSION STUDY

Objective

To summarize the efficacy and safety results of the long-term extension study following UNITI-1, UNITI-2, and IM-UNITI. The following is based on unpublished data.[68]

Findings

Study Design

[redacted]

[redacted]

- [redacted]

Results

[redacted]

[redacted]

Table 38 details the patient disposition.

TABLE 38: PATIENT DISPOSITION IN THE LONG-TERM EXTENSION STUDY

	PLA SC[a]	UST		
		90 mg SC q.12.w.[a]	90 mg SC q.8.w.[a]	Prior dosage adjustment[b]
Randomized patients in clinical response at week 44[c]				
N	[redacted]	[redacted]	[redacted]	[redacted]
Treatment discontinuation from week 44 to week 96, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Reasons for discontinuations, n (%)				
AEs	[redacted]	[redacted]	[redacted]	[redacted]
Lack of efficacy	[redacted]	[redacted]	[redacted]	[redacted]
Protocol violation	[redacted]	[redacted]	[redacted]	[redacted]
Study termination by sponsor	[redacted]	[redacted]	[redacted]	[redacted]
Physician decision[d]	[redacted]	[redacted]	[redacted]	[redacted]
Lost to follow-up[d]	[redacted]	[redacted]	[redacted]	[redacted]
Withdrawn consent[d]	[redacted]	[redacted]	[redacted]	[redacted]
Death, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Placebo patients who discontinued study treatment due to unblinding (after week 44 analysis completed), n (%)	[redacted]	[redacted]	[redacted]	[redacted]
	PLA SC[e]	**90 mg SC q.12.w.[f]**	**90 mg SC q.8.w.[g]**	-
Non-randomized patients in clinical response at week 44[c]				
N	[redacted]	[redacted]	[redacted]	-
Treatment discontinuation from week 44 to week 96, n (%)	[redacted]	[redacted]	[redacted]	-
Reasons for discontinuations, n (%)				
AEs	[redacted]	[redacted]	[redacted]	-
Lack of efficacy	[redacted]	[redacted]	[redacted]	-
Protocol violation	[redacted]	[redacted]	[redacted]	-
Study termination by sponsor	[redacted]	[redacted]	[redacted]	-
Physician decision[d]	[redacted]	[redacted]	[redacted]	-
Lost to follow-up[d]	[redacted]	[redacted]	[redacted]	-
Withdrawn consent[d]	[redacted]	[redacted]	[redacted]	-

	PLA SC[a]	UST		
		90 mg SC q.12.w.[a]	90 mg SC q.8.w.[a]	Prior dosage adjustment[b]
Death, n (%)	[redacted]	[redacted]	[redacted]	-
Placebo patients who discontinued study treatment due to unblinding (after week 44 analysis completed), n (%)	[redacted]	[redacted]	[redacted]	-

AE = adverse event; PLA = placebo; q.8.w. = every eight weeks; q.12.w. = every 12 weeks; SC = subcutaneous; UST = ustekinumab.

[redacted]

Source: Clinical Study Report.[68]

Clinical Remission

[redacted]

Table 39

Steroid-Free Clinical Remission

[redacted]

Table 39

Disease–Related Surgery Through Week 92

[redacted]

TABLE 39: CLINICAL OUTCOMES FOR THE LONG-TERM EXTENSION STUDY

	PLA SC[a]	UST		
		90 mg SC q.12.w.[a]	90 mg SC q.8.w.[a]	90 mg SC q.8.w. prior dosage adjustment[b]
Randomized patients in clinical response at week 44[c]				
N	[redacted]	[redacted]	[redacted]	[redacted]
Clinical remission, n (%)				
Week 44	[redacted]	[redacted]	[redacted]	[redacted]
Week 92[c,d]	[redacted]	[redacted]	[redacted]	[redacted]
Steroid-free clinical remission, n (%)				
Week 92[c,d,e]	[redacted]	[redacted]	[redacted]	[redacted]
	PLA SC[f]	**90 mg SC q.12.w.[g]**	**90 mg SC q.8.w.[h]**	
Non-randomized patients in clinical response at week 44[c]				
N	[redacted]	[redacted]	[redacted]	-
Clinical remission, n (%)				
Week 44	[redacted]	[redacted]	[redacted]	-
Week 92	[redacted]	[redacted]	[redacted]	-
Steroid-free clinical remission, n (%)				
Week 92[c,d,e]	[redacted]	[redacted]	[redacted]	-

PLA = placebo; q.8.w. = every eight weeks; q.12.w. = every 12 weeks; SC = subcutaneous; UST = ustekinumab.

[redacted]

Source: Clinical Study Report.[68]

Limitations

[redacted]

Summary

[redacted]

APPENDIX 7: SUMMARY OF INDIRECT COMPARISONS

Background

Given the lack of head-to-head studies comparing ustekinumab with other relevant biologics for moderate-to-severe Crohn's disease (CD) in this CDR review, the objective of this Appendix is to summarize and critically appraise the evidence available regarding the comparative efficacy and safety of ustekinumab versus infliximab, adalimumab, and vedolizumab through indirect comparison (IDC) using network meta-analysis (NMA) methodology. Both induction and maintenance treatment in adult patients with moderate to severely active CD were evaluated in this review. Three IDCs were assessed: the unpublished IDC submitted by the manufacturer,[12] a published IDC by Mocko et al.,[14,15], and another published IDC by Singh et al.[13]

Methods for Manufacturer's Indirect Comparison

Study Eligibility and Selection Process

The manufacturer submitted an IDC based on a systematic review that compared ustekinumab with infliximab, adalimumab, and vedolizumab in patients with moderate-to-severe CD.[12] A systematic search of randomized controlled trials was performed by searching multiple electronic databases and performing hand searches. Only biologic therapies and their respective dosages that had been approved in Canada for CD were included. Outcomes of interest for the systematic literature review included efficacy end points (Crohn's Disease Activity Index [CDAI], C-reactive protein, fecal lactoferrin and calprotectin, mucosal healing/endoscopic improvement, fistula closure), safety end points (infections/serious infections, grade 3/4 adverse events [AEs], hospitalizations/surgery, discontinuations/withdrawals, dosage escalations), and quality of life/other measures (Inflammatory Bowel Disease Questionnaire [IBDQ], Work Productivity and Activity Impairment, Short Form [36] Health Survey [SF-36], EuroQol 5-Dimensions questionnaire). Two independent reviewers assessed titles and abstracts, and potentially relevant articles were retrieved for full-text review. Eligible articles were selected based on the inclusion and exclusion criteria provided by the manufacturer, with discrepancies resolved through discussion and/or a third reviewer.

The network analysis did not include all aforementioned outcomes but included only those that were appropriate, based on a feasibility assessment conducted at the onset of the systematic review. The feasibility assessment helped to identify trials that were considered sufficiently similar; subsequently, studies that examined certolizumab and natalizumab were eliminated, as they did not contain comparable end points, nor did they report efficacy data at one year post-treatment. In addition, neither certolizumab nor natalizumab are approved for treating Crohn's disease in Canada. Safety end points were also not included in the NMA as they were deemed infeasible in the induction period owing to the differences in the AE definitions among trials. In addition, these safety end points were not included for the maintenance phase owing to multiple sources of heterogeneity in study design and lack of comparability among placebo arms in the various trials. Decision sets for the NMAs are included in Table 40.

TABLE 40: INCLUSION CRITERIA FOR THE MANUFACTURER, MOCKO ET AL., AND SINGH ET AL. INDIRECT COMPARISONS

	Manufacturer's IDC	Mocko et al. 2016	Singh et al. 2014
Patient Population	• Patients with moderate-to-severe CD **Subgroups:** • Conventional treatment failure • TNF alpha antagonist failure	• Patients (> 15 years of age) with active CD (defined by conventional clinical, radiographic, and endoscopic criteria (CDAI > 150)	**Induction:** • Biologic-naive patients with moderate-to-severe CD (CDAI > 220 but < 450) **Maintenance:** • Subset of patients with moderate-to-severe CD (CDAI > 220 but < 450) who initially responded to induction therapy with index biologic
Intervention	**Induction:** • UST (~6 mg/kg) **Maintenance:** • UST (90 mg q.8.w.) • UST (90 mg q.12.w.)	**Induction and Maintenance:** • All biologics registered for use by the EMA and FDA: ○ ADA ○ CZP ○ IFX ○ VDZ ○ UST (at the time of IDC was not yet approved but undergoing approval process)	• Anti-TNF alpha drugs ○ ADA ○ IFX ○ CZP • Anti-integrin drugs ○ NAT ○ VDZ • Anti-IL 12/23 drug ○ UST
Comparators	**Induction decision set:** • IFX (5 mg/kg) • ADA (160/80 mg) • ADA (80/40 mg) • VDZ (300 mg) **Maintenance decision set:** • IFX (5 mg/kg q.8.w.) • IFX (5 and 10 mg/kg q.8.w.) • ADA (40 mg EOW) • ADA (40 mg weekly) • VDZ (300 mg q.8.w.) • VDZ (300 mg q.4.w.) •	**Induction and maintenance:** • All biologics registered for use by the EMA and FDA: ○ ADA ○ CZP ○ IFX ○ VDZ ○ UST (at the time of IDC was not yet approved but undergoing approval process)	○ Another biologic drug ○ PLA ○ Alternate intervention with ≥ 2 biologic drugs having been compared with common intervention
Outcomes	**Efficacy outcomes:** • Clinical response (CDAI-70) • Enhanced clinical response (CDAI-100) • Clinical remission (CDAI < 150)	**Safety outcomes:** • Induction phase (6 to 10 weeks FU) ○ All end points occurring at a priori frequency of 3% • Maintenance phase (52 to 56 weeks FU) ○ All end points occurring at a priori frequency of 3%	**Efficacy outcomes:** • Clinical remission (CDAI < 150) • If clinical remission unavailable then clinical response (either CDAI-100 or CDAI-70) • Maintenance of medically induced remission (in patients with clinical response to induction therapy)

	Manufacturer's IDC	Mocko et al. 2016	Singh et al. 2014
Study Design	RCTs (induction and maintenance studies)	RCTs (induction and maintenance phase, either placebo-controlled or head-to-head trials)	RCTs (induction and maintenance)

ADA = adalimumab; CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CZP = certolizumab pegol; EMA = European Medicines Agency; EOW = every other week; FDA = US Food and Drug Administration; FU = follow-up; IDC = indirect comparison; IFX = infliximab; IL = interleukin; NAT = natalizumab; PLA = placebo; q.4.w. = every 4 weeks; q.8.w = every 8 weeks; q.12.w. = every 12 weeks; RCT =randomized controlled trial; TNF = tumour necrosis factor; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted IDC,[12] Mocko et al. 2016,[14,15] Singh et al. 2014.[13]

Quality Assessment of Included Studies

Data were extracted and quality-checked by two independent reviewers. Risk of bias was assessed using the National Institute for Health and Care Excellence (NICE) checklist, which included the following end points: adequacy of randomization and allocation concealment; similarity of prognostic factors between groups at baseline; blinding of patients, care providers, and end point assessors; similarity of dropouts; selective end point reporting; use of an intention-to-treat analysis; and adequacy of methods for handling of missing data. Results of the individual quality assessment were provided by the manufacturer.

Indirect Comparison Methods

A Bayesian hierarchical model (which helps preserve the randomization of each trial) was used for the NMA, which was performed using WinBUGS V1.4 software. The manufacturer conducted separate analyses of the subpopulation of patients who experienced treatment failure with conventional therapies, the subpopulation of patients who experienced treatment failure with anti-TNF therapies, and the overall study population (for the maintenance phase). Standard NMA methods (by Dias et al.)[69] were used for the induction phase, whereas a treatment-sequence analysis was used to assess treatment efficacy at one year (induction plus maintenance). Outputs were reported as median odds ratios with corresponding 95% credible intervals (CrI), with differences considered significant if the 95% CrIs did not include the null value (1). The deviance information criterion (DIC) was used to assess the relative goodness of fit of both the fixed-effects and random-effects models, of which the model with the lowest DIC was selected for each outcome (as the better model to use for that end point). A statistical approach that adjusted to account for the correlation between treatment effects was used for trials assessing more than two treatments of interest (e.g., multi-arm trials). No assessment of inconsistency between direct and indirect evidence was performed, as there were no closed loops in the networks. Non-informative priors were used for unknown parameters. The following priors were used for the base-case analysis (BCA): normal distributions with a mean of 0 and a variance of 10,000 for treatment effects and a uniform distribution for the between-trial standard deviation, with a range of [0,2]. Convergence was assessed, and 20,000 iterations were used as a burn-in for each of the analyses. This was followed by 20,000 or more iterations for estimation and monitoring of all parameters for the fixed and random-effects models.

For the induction phase, the time points assessed were week 4 (for infliximab and adalimumab trials) and week 6 (for vedolizumab and ustekinumab trials) in order to optimize comparability between trials. The manufacturer provided a pairwise comparison for the BCA of the induction phase in the subpopulation that had experienced treatment failure with conventional therapy and the subpopulation that had experienced treatment failure with anti-TNF therapy using a fixed-effects model. This was performed in order to ascertain that results were similar to those from the NMA. In addition, numerous

sensitivity analyses were conducted in order to test the robustness of the BCA results and included the following:

- BCA conducted under a frequentist framework (using the Bucher method)
- Different times of assessments (including week 8 results in UNITI and week 2 results from Targan et al. 1997)
- Exclusion of Targan et al. 1997 from subpopulation that had experienced treatment failure with conventional therapy network
- Exclusion of vedolizumab trials from the subpopulation that had experienced treatment failure with conventional therapy network
- Exclusion of adalimumab trials from the subpopulation that had experienced treatment failure with anti-TNF therapy network
- Exclusion of Watanabe et al. 2012 from both subpopulation networks
- Inclusion of the CERTIFI trial results (included in the main body of the report) in subpopulation that had experienced treatment failure with anti-TNF therapy, as the 6 mg/kg dose in this trial was not comparable to the 6 mg/kg dose in the UNITI-1 trial
- End points that selected time points based on re-randomization times.

Table 41 presents details regarding the outcomes and populations evaluated in the included induction trials.

TABLE 41: OUTCOMES AND POPULATIONS EVALUATED IN INDIRECT COMPARISON FOR INDUCTION STUDIES

Outcome	Definition	Drug	Trial	Time Point	Study Population	
					Conventional-Failed	TNF Failure
Ustekinumab versus infliximab						
Clinical response	CDAI-70	UST	UNITI-1	Week 6	No	Yes
			UNITI-2		Yes	No
		IFX	Targan et al. 1997	Week 4	Yes	No
Clinical remission	CDAI ≤ 150	UST	UNITI-1	Week 6	No	Yes
			UNITI-2		Yes	No
		IFX	Targan et al. 1997	Week 4	Yes	No
Ustekinumab versus adalimumab						
Clinical response	CDAI-70	UST	UNITI-1	Week 6	No	Yes
			UNITI-2		Yes	No
		ADA	CLASSIC I	Week 4	Yes	No
			GAIN		No	Yes
			Watanabe et al.		Yes	Yes
Enhanced clinical response	CDAI-100	UST	UNITI-1	Week 6	No	Yes
			UNITI-2		Yes	No
		ADA	CLASSIC I	Week 4	Yes	No
			GAIN		No	Yes
			Watanabe et al.		Yes	Yes
Clinical remission	CDAI ≤ 150	UST	UNITI-1	Week 6	No	Yes
			UNITI-2		Yes	No
		ADA	CLASSIC I	Week 4	Yes	No
			GAIN		No	Yes

Outcome	Definition	Drug	Trial	Time Point	Study Population	
					Conventional-Failed	TNF Failure
			Watanabe et al.		Yes	Yes
Ustekinumab versus vedolizumab						
Clinical response	CDAI-70	UST	UNITI-1	Week 6	No	Yes
			UNITI-2		Yes	No
		VDZ	GEMINI II	Week 6	Yes	Yes
			GEMINI III		Yes	Yes
Enhanced clinical response	CDAI-100	UST	UNITI-1	Week 6	No	Yes
			UNITI-2		Yes	No
		VDZ	GEMINI II	Week 6	Yes	Yes
			GEMINI III		Yes	Yes
Clinical remission	CDAI ≤ 150	UST	UNITI-1	Week 6	No	Yes
			UNITI-2		Yes	No
		VDZ	GEMINI II	Week 6	Yes	Yes
			GEMINI III		Yes	Yes

ADA = adalimumab; CDAI = Crohn's Disease Activity Index; IFX = infliximab; TNF = tumour necrosis factor; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted indirect comparison.[12]

For the maintenance phase, the manufacturer used treatment-sequence analyses in which outcomes from the induction and maintenance phases were used to create an overall outcome variable. The manufacturer's justification for the need to use this approach was first identified in its feasibility assessment of the maintenance-phase NMA. Several conceptual differences between the included trials were observed:

- differences in the selection criteria required for entry into the maintenance phase
- different induction active-treatment regimens for patients in the placebo arms
- different assessment times
- different clinical response definitions
- different criteria for re-randomization
- the fact that the induction phases could have been open label or double blind
- whether patients continuing on to the maintenance phase had to be in remission.

Because of this heterogeneity and to minimize bias, the manufacturer decided to consider only those trials with the most comparable maintenance phases for the analysis of efficacy after one year of treatment. In addition, the manufacturer noted that placebo response rates were different and that these arms were not truly common comparators across trials. This was tested using a chi-square test that compared the observed placebo response and remission rates with the expected placebo rates if they were truly common comparators (the *P* value was less than 0.05 in both the populations that had experienced failures of anti-TNF and conventional therapies, thereby confirming that the placebo groups were not common comparators). Therefore, in its use of the treatment-sequence analyses, the manufacturer sought to evaluate the treatment effects over the entire treatment sequence (induction followed by maintenance for each intervention) and to increase the ability to compare placebo arms across the maintenance-phase trials.

The treatment-sequence analysis involved the incorporation of induction and maintenance data for each of the interventions. This allowed the manufacturer to obtain "relative" treatment effect estimates that took into account treatment history in the induction phase. The manufacturer multiplied the "conditional probability of maintaining response until the end of the maintenance phase"[12] by the "probability of achieving response at the end of induction;"[12] hence estimating the "probability of achieving and maintaining response by the end of the maintenance phase."[12] The manufacturer determined that the placebo groups were not true comparators; therefore, the placebo arm data contained imputed maintenance data (from IM-UNITI individual patient-level data) for the treatment sequence.

The manufacturer provided a pairwise comparison for the BCA of the entire treatment-sequence analysis in the subpopulations that had experienced failures of conventional therapy and of anti-TNF therapy using a fixed-effects model. This was performed in order to ascertain that results were similar to those from the NMA. In addition, numerous sensitivity analyses were conducted in order to test the robustness of the treatment-sequence analysis results and included the following:

- BCA conducted under a frequentist framework (using the Bucher method)
- Bayesian analysis with dosage adjustments
- To replace inputs used in the analysis of the subpopulation of patients who experienced a failure with conventional therapy, inputs generated from data of patients who were truly naive to biologics from the UNITI program
- Pooled maintenance doses to assess uncertainty of IDC treatment estimate effects.

Table 42 presents details regarding the outcomes and populations evaluated in the included maintenance trials.

TABLE 42: OUTCOMES AND POPULATIONS EVALUATED IN INDIRECT COMPARISON – TREATMENT-SEQUENCE ANALYSIS

Outcome	Definition	Drug	Trial	Study Population		
				Overall Subpopulation	Conventional-Failed	TNF Failure
Ustekinumab versus infliximab						
Clinical remission	CDAI ≤ 150	UST (SC) q.12.w.	IM-UNITI	Yes	Yes	Yes
		UST (SC) q.8.w.	IM-UNITI	Yes	Yes	Yes
		IFX 5 mg/mL	ACCENT I	Yes	Yes	No
		IFX 5+10 mg/mL	ACCENT I	Yes	Yes	No
Ustekinumab versus adalimumab						
Enhanced clinical response	CDAI-100	UST (SC) q.12.w.	IM-UNITI	Yes	Yes	Yes
		UST (SC) q.8.w.	IM-UNITI	Yes	Yes	Yes
		ADA 160/80 mg; 40 mg EOW	CHARM	Yes	Yes	Yes
			EXTEND	Yes	No	No
			Watanabe et al. 2012	Yes	No	No
		ADA 160/80 mg; 40 mg weekly	CHARM	Yes	Yes	Yes

Outcome	Definition	Drug	Trial	Study Population		
				Overall Subpopulation	Conventional-Failed	TNF Failure
Clinical remission	CDAI ≤ 150	UST (SC) q.12.w.	IM-UNITI	Yes	Yes	Yes
		UST (SC) q.8.w.	IM-UNITI	Yes	Yes	Yes
		ADA 160/80 mg; 40 mg EOW	CHARM	Yes	Yes	Yes
			EXTEND	Yes	Yes	Yes
			Watanabe et al. 2012	Yes	Yes	Yes
		ADA 160/80 mg; 40 mg weekly	CHARM	Yes	Yes	Yes
Ustekinumab versus vedolizumab						
Enhanced clinical response	CDAI-100	UST (SC) q.12.w.	IM-UNITI	Yes	Yes	Yes
		UST (SC) q.8.w.	IM-UNITI	Yes	Yes	Yes
		VDZ 300 mg q.8.w.	GEMINI II	Yes	Yes	Yes
		VDZ 300 mg q.4.w.	GEMINI II	Yes	Yes	Yes
Clinical remission	CDAI ≤ 150	UST (SC) q.12.w.	IM-UNITI	Yes	Yes	Yes
		UST (SC) q.8.w	IM-UNITI	Yes	Yes	Yes
		VDZ 300 mg q.8.w	GEMINI II	Yes	Yes	Yes
		VDZ 300 mg q.4.w.	GEMINI II	Yes	Yes	Yes

ADA = adalimumab; CDAI = Crohn's Disease Activity Index; EOW = every other week; IFX = infliximab; q.4.w. = every 4 weeks; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; SC = subcutaneous; TNF = tumour necrosis factor; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted indirect comparison.[12]

Results

Study and Patient Characteristics

A total of 11 trials were included in the various NMA analyses. Only trials that assessed the efficacy of ustekinumab, adalimumab, infliximab, and vedolizumab were included. The following studies included only induction-phase results: Classic I and GAIN (adalimumab), Targan et al. 1997 (infliximab), GEMINI III (vedolizumab), and UNITI-1 and UNITI-2 (ustekinumab). The following studies included only maintenance-phase results: ACCENT I (infliximab), CHARM (adalimumab), ACCENT I (infliximab), and IM-UNITI (ustekinumab). Trials that presented both induction and maintenance results included Watanabe et al. 2012 (adalimumab) and GEMINI II (vedolizumab). Of the studies linking treatments of interest that were excluded, D'Haens 1999 (infliximab) lacked reported outcomes, Feagan 2008 (vedolizumab) used unapproved dosages, Classic II (adalimumab) included only patients in remission at weeks 4 and 8, and EXTEND (adalimumab) included a mixed population whose results were not available by subgroups. Although the CERTIFI trial (ustekinumab) was excluded from the NMA analyses (as the doses were not the same as the UNITI trials and there was a lack in reported outcomes for the 22-week analysis), it was included in a sensitivity analysis, and descriptive safety results were also reported. With regard to the included induction trials, the mean duration of CD ranged between 4.7 and 12.7 years, the mean CDAI ranged between a score of 286 and 328, the proportion of patients on corticosteroids ranged between

21.7% and 54%, and the proportion of patients on concomitant immunomodulators ranged between 16.8% and 54%. Analyses in the subpopulation who had experienced a failure of anti-TNF therapy were obtained from data in six studies (including five treatments, n = 1,433), and analyses in the subpopulation who had experienced a failure of conventional therapy were obtained from data in five studies (including six treatments, n = 1,130). Details of the induction-study characteristics are presented in Table 43.

TABLE 43: SELECT STUDY CHARACTERISTICS INCLUDED IN THE INDIRECT COMPARISON

Drug	Study	Treatment Group Included in IDC (n)	Primary End Point	Mean Years of Duration of CD (SD)	Mean CDAI (SD)	Prior Anti-TNF n (%)	Concomitant Medications, n (%)	
							CS	IM
Induction studies								
UST	UNITI-1	• UST 6 mg/kg (249) • PLA (247)	CL remission (week 6)	UST: 12.7 (9.2) PLA: 12.1 (8.4)	UST: 328 (62.0) PLA: 319 (59.7)	UST: 249 (100) PLA: 247 (100)	UST: 108 (43.4) PLA: 111 (44.9)	UST: 78 (31.3) PLA: 81 (32.8)
	UNITI-2	• UST 6 mg (NR) • PLA (NR)	CL remission (week 6)	UST: 8.7 (8.5) PLA: 8.7 (8.4)	UST: 302 (58.8) PLA: 302 (61.7)	UST: 65 (31) PLA: 75 (36)	UST: 92 (44.0) PLA: 75 (35.7)	UST: 72 (34.4) PLA: 73 (34.8)
ADA	CLASSIC I	• ADA 160 mg wk 0, 80 mg wk 2 (76) • PLA (74)	CL remission (week 4)	ADA: NR PLA: NR	ADA: 295 (52) PLA: 296 (60)	ADA: 0 (0) PLA: 0 (0)	ADA: 24 (32) PLA: 25 (34)	ADA: 22 (29) PLA: 22 (30)
	GAIN	• ADA 160 mg wk 0, 80 mg wk 2 (159) • PLA (166)	CL remission (week 4)	ADA: NR PLA: NR	ADA: 313 (58) PLA: 313 (66)	ADA: NR (100) PLA: NR (100)	ADA: 55 (35) PLA: 73 (44)	ADA: 73 (46) PLA: 85 (51)
	Watanabe et al.	• ADA 160 mg wk 0, 80 mg wk 2 (33) • PLA (23)	CL remission (week 4)	ADA: 11 (7.1) PLA: 7.9 (4.7)	ADA: 300 (66.5) PLA: 308 (63.8)	ADA: 19 (58) PLA: 13 (56)	ADA: 8 (24.2) PLA: 5 (21.7)	ADA: 10 (30.3) PLA: 8 (34.8)
IFX	Targan et al. 1997	• IFX 5 mg/kg wk 0 (NR) • PLA (NR)	CL response (week 4)	IFX: 12.5 (10.3) PLA: 10.4 (7.7)	IFX: 312 (56) PLA: 288 (54)	IFX: NR PLA:NR	IFX: 8 (30)[a] 7 (26)[b] PLA: 10 (40)[a] 6 (24)[b]	IFX: 4 (15) on 6-MP 4 (19) on AZA PLA: 4 (16) on 6-MP 7 (28) on AZA
VDZ	GEMINI II	• VDZ 300 mg wks 0, 2 (220) • PLA (148)	CL remission (week 6) Enhanced CL response (week 6)	VDZ: 9.2 (8.2) PLA: 8.2 (7.8)	VDZ:327 (71) PLA: 325 (78)	VDZ: 111 (51) PLA: 72 (49)	VDZ: 67 (30.5) PLA: 45 (30.4)	VDZ: 37 (16.8) PLA: 25 (16.9)
	GEMINI III	• VDZ 300 mg wks 0, 2, 6 (209) • PLA (207)	CL remission (week 6)	VDZ[c]: 9.4 (0.5 to 41.8)[e] VDZ[d]: 4.7 (0.3 to 4.8)[e] PLA[c]: 9.6 (1.0 to 42.9)[e] PLA[d]: NR (0.3 to 24.8)[e]	VDZ[c]: 316 (52.6) VDZ[d]: 307 (54.8) PLA[c]: 306 (55.4) PLA[d]: 286 (51.1)	VDZ[c]: 158 (100) VDZ[d]: 0 (0) PLA[c]: 157 (100) PLA[d]: 0 (0)	VDZ[c]: 86 (54) VDZ[d]: 24 (47) PLA[c]: 85 (54) PLA[d]: 23 (46)	VDZ[c]: 54 (43) VDZ[d]: 28 (55) PLA[c]: 42 (27) PLA[d]: 46 (27)

6-MP = 6-mercaptopurine; ADA = adalimumab; AZA = azathioprine; CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CL = clinical; CS = corticosteroids; IDC = indirect comparison; IFX = infliximab; IM = immunomodulators; NR = not reported; PLA = placebo; SD = standard deviation; TNF = tumour necrosis factor; UST = ustekinumab; VDZ = vedolizumab; wk = week.

[a] < 20 mg/day.

[b] > 20 mg/day.

[c] Population reported separately for failed anti-TNF.

[d] Population reported separately for failed conventional therapy.

[e] Reported as the median (range).

Source: Manufacturer submitted indirect comparison.[12]

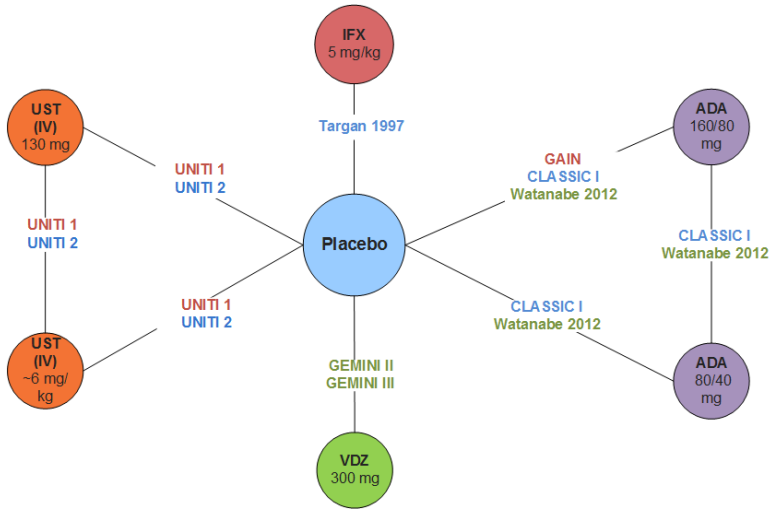
Evidence Networks

Evidence networks were provided in the manufacturer-submitted IDC. None of the networks contained any closed loops, and the networks were anchored to placebo as the only common treatment arm between studies. The following evidence networks were provided in the manufacturer's IDC.

Induction Networks

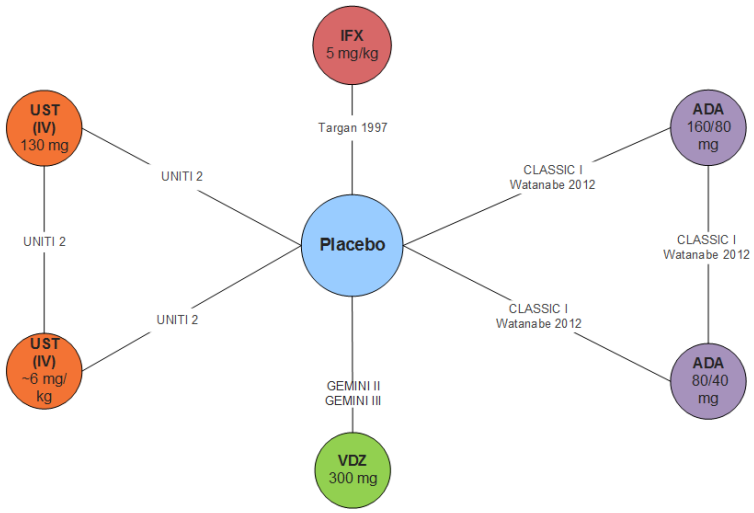
The overall network for the induction NMA is presented in Figure 5 The overall network for patients who had experienced a failure with conventional therapy is presented in Figure 6. The overall network for patients who had experienced a failure with anti-TNF therapy is presented in Figure 7.

FIGURE 5: OVERALL NETWORK FOR THE INDUCTION NETWORK META-ANALYSIS



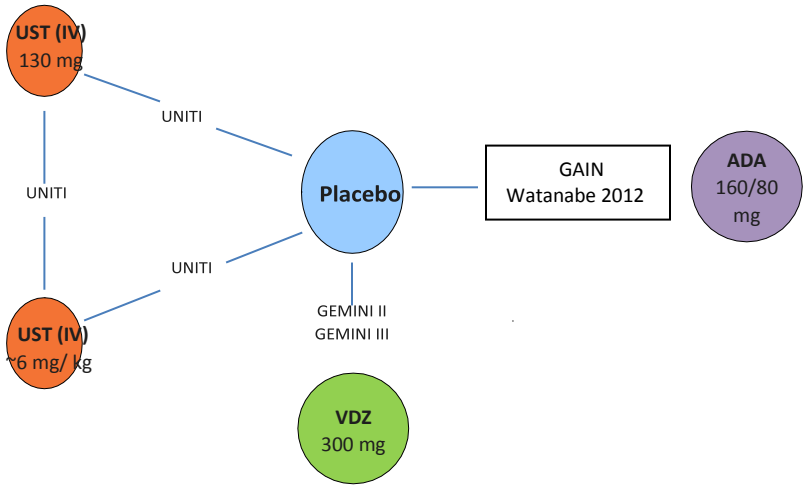
ADA = adalimumab; IFX = infliximab; IV = intravenous; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted indirect comparison.[12]

FIGURE 6: OVERALL NETWORK FOR PATIENTS WHO HAVE FAILED CONVENTIONAL THERAPY (N = 6)



ADA = adalimumab; IFX = infliximab; IV = intravenous; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted indirect comparison.[12]

FIGURE 7: OVERALL NETWORK FOR PATIENTS WHO HAVE FAILED ANTI-TUMOUR NECROSIS FACTOR THERAPY (N = 5)

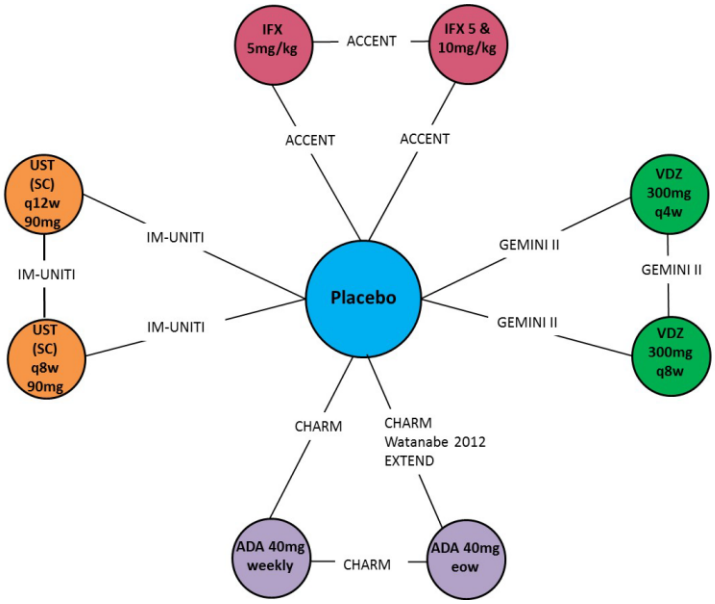


Abbreviations: ADA = adalimumab; IV = intravenous; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted IDC.[12]

Maintenance Phase Networks

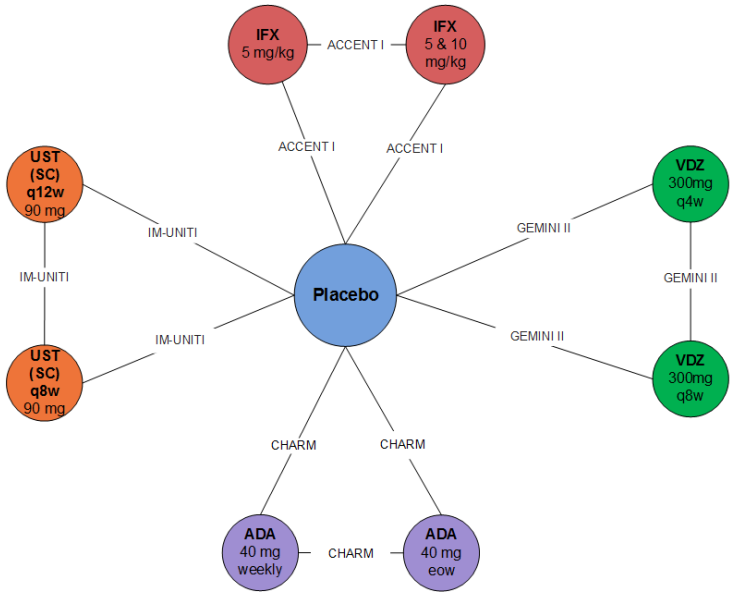
The overall network for the maintenance phase NMA is presented in Figure 8. The overall network for patients who have experienced a failure with conventional therapy is presented in Figure 9. The overall network for patients who have experienced a failure with anti-TNF therapy is presented in Figure 10.

FIGURE 8: OVERALL NETWORK FOR THE MAINTENANCE PHASE NETWORK META-ANALYSIS IN THE OVERALL POPULATION



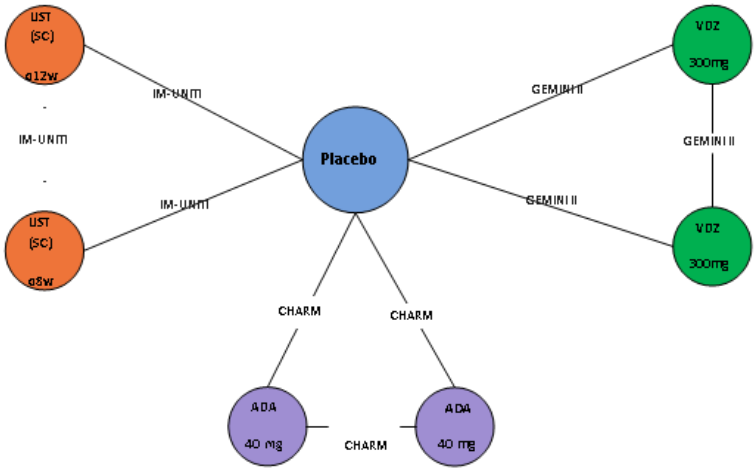
ADA = adalimumab; eow = every other week; IFX = infliximab; q4w = every 4 weeks; q8w = every 8 weeks; q12w. every 12 weeks; SC = subcutaneous; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted indirect comparison.[12]

FIGURE 9: OVERALL NETWORK FOR THE MAINTENANCE PHASE NETWORK META-ANALYSIS IN THE FAILED CONVENTIONAL THERAPY SUBPOPULATION



ADA = adalimumab; eow = every other week; IFX = infliximab; q4w = every 4 weeks; q8w = every 8 weeks; q12w = every 12 weeks; SC = subcutaneous; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted indirect comparison.[12]

FIGURE 10: OVERALL NETWORK FOR THE MAINTENANCE PHASE NETWORK META-ANALYSIS IN THE FAILED ANTI-TUMOUR NECROSIS FACTOR THERAPY SUBPOPULATION



ADA = adalimumab; SC = subcutaneous; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted indirect comparison.[12]

Efficacy Results

Induction Therapy

No statistically significant differences between ustekinumab 6 mg/kg and adalimumab (either 80/40 mg or 160/80 mg) or ustekinumab 6 mg/kg and vedolizumab 300 mg were observed in either the subpopulation who had experienced a failure with conventional or anti-TNF therapy for clinical response (CDAI-70), enhanced clinical response (CDAI-100), or clinical remission (CDAI<150). Statistically

significant differences in favour of infliximab 5 mg/kg compared with ustekinumab 6 mg/kg were observed for clinical response (CDAI-70) and clinical remission (CDAI < 150) in the subpopulation who had experienced a failure with failed conventional therapy (0.11; 95% CrI, 0.02 to 0.48; and 0.08; 95% CrI, 0.01 to 0.59; respectively.) Detailed NMA results for the induction period are presented in Table 44. The BCAs were determined to be robust, as the numerous aforementioned sensitivity analyses did not change the clinical interpretation (data not shown).

TABLE 44: NETWORK META-ANALYSIS RESULTS FOR THE INDUCTION PERIOD

Comparator	Clinical Response – CDAI-70 OR (95% CrI)[a,b]		Enhanced Clinical Response – CDAI-100 OR (95% CrI)[a,b]		Clinical Remission – CDAI < 150 OR (95% CrI)[a,b]	
	Failed Conventional	Failed Anti-TNF	Failed Conventional	Failed Anti-TNF	Failed Conventional	Failed Anti-TNF
Ustekinumab 6 mg/kg versus:						
ADA 80/40 mg	0.92 (0.46 to 2.05)	1.29 (0.38 to 4.40)	1.39 (0.64 to 2.97)	0.66 (0.18 to 2.34)	1.14 (0.44 to 2.82)	2.24 (0.36 to 20.32)
ADA 160/80 mg	0.92 (0.43 to 1.91)	0.83 (0.47 to 1.46)	1.03 (0.47 to 2.20)	0.93 (0.51 to 1.70)	0.64 (0.25; 1.53)	0.64 (0.26 to 1.51)
IFX 5 mg/kg	**0.11 (0.02 to 0.48)**	-	-	-	**0.08 (0.01 to 0.59)**	-
VDZ 300 mg	1.58 (0.85 to 2.94)	0.96 (0.57 to 1.62)	1.85 (0.96 to 3.51)	1.05 (0.59 to 1.85)	0.93 (0.39 to 2.08)	1.53 (0.69 to 3.39)

ADA = adalimumab; CDAI = Crohn's Disease Activity Index; CrI = credible interval; IFX = infliximab; OR = odds ratio; TNF = tumour necrosis factor; VDZ = vedolizumab.
Bold: Statistically significant difference.
[a] Reported as median OR with respective median credible intervals.
[b] Base-case analysis (Bayesian) based on the fixed-effects model.
Source: Manufacturer's submitted indirect comparison.[12]

Maintenance Therapy

With regard to enhanced clinical response (CDAI-100) for the maintenance phase, the only statistically significant difference was in favour of ustekinumab 90 mg every 12 weeks compared with vedolizumab 300 mg every eight weeks in the overall population (1.60; 95% CrI, 1.01 to 2.54). This statistically significant difference was not observed in either the subpopulation that had experienced a failure of conventional therapy or anti-TNF therapy. No statistically significant differences were observed in the overall population or subpopulations between the different biologic treatment regimens compared with ustekinumab 90 mg every 12 weeks with regard to clinical remission (CDAI < 150).

For the most part, no statistically significant differences between ustekinumab 90 mg every eight weeks and the other biologic regimens were observed in the overall population or subpopulations with regard to either the enhanced clinical response (CDAI-100) or clinical remission (CDAI < 150). There were two exceptions to this. First, there was a statistically significant difference in favour of ustekinumab 90 mg every eight weeks when compared with vedolizumab 300 mg every four weeks in the overall population with regard to the enhanced clinical response (CDAI-100 1.57; 95% CrI, 1.002 to 2.47). Second, there was

a statistically significance difference in favour of ustekinumab 90 mg every eight weeks when compared with vedolizumab 300 mg every eight weeks in the overall population with regard to clinical remission (CDAI < 150 1.74; 95% CrI, 1.05 to 2.88). Detailed NMA results are presented in Table 45. The BCAs were determined to be robust, as the numerous aforementioned sensitivity analyses did not change the clinical interpretation (data not shown).

TABLE 45: NETWORK META-ANALYSIS RESULTS FOR THE TREATMENT-SEQUENCE ANALYSIS

Comparator	Enhanced Clinical Response – CDAI-100 OR (95% CrI)[a,b]			Clinical Remission – CDAI < 150 OR (95% CrI)[a,b]		
	Overall	Failed Conventional	Failed Anti-TNF	Overall	Failed Conventional	Failed Anti-TNF
Ustekinumab 90 mg q.12.w. versus						
ADA 80/40 mg, ADA 40 weekly	1.27 (0.77 to 2.11)	1.16 (0.51 to 2.60)	1.07 (0.53 to 2.14)	1.10 (0.63 to 1.90)	1.01 (0.40 to 2.40)	0.94 (0.45 to 1.97)
ADA 160/80 mg, ADA 40 EOW	1.49 (0.89 to 2.48)	1.58 (0.68 to 3.62)	1.15 (0.56 to 2.32)	1.30 (0.77 to 2.17)	1.25 (0.50 to 3.07)	1.11 (0.52 to 2.35)
IFX 5 mg/kg	-	-	-	0.56 (0.10 to 2.33)	0.60 (0.07 to 3.24)	-
IFX 5 + 10 mg/kg	-	-	-	0.80 (0.14 to 3.60)	0.41 (0.05 to 2.13)	-
VDZ 300 mg q.4.w.	1.52 (0.96 to 2.39)	1.84 (0.92 to 3.65)	1.31 (0.68 to 2.50)	1.57 (0.95 to 2.59)	1.37 (0.64 to 2.91)	1.35 (0.66 to 2.70)
VDZ 300 mg q.8.w.	**1.60 (1.01 to 2.54)**	1.54 (0.77 to 3.05)	1.77 (0.91 to 3.55)	1.38 (0.83 to 2.25)	1.24 (0.58 to 2.61)	1.35 (0.66 to 2.73)
Ustekinumab 90 mg q.8.w. versus						
ADA 160/80 mg, ADA 40 weekly	1.32 (0.79 to 2.17)	1.20 (0.53 to 2.69)	1.14 (0.56 to 2.28)	1.22 (0.70 to 2.11)	1.16 (0.47 to 2.77)	1.00 (0.48 to 2.49)
ADA 160/80 mg, ADA 40 EOW	1.54 (0.92 to 2.56)	1.64 (0.71 to 3.73)	1.22 (0.60 to 2.46)	1.44 (0.86 to 2.41)	1.45 (0.58 to 3.53)	1.17 (0.56 to 2.49)
IFX 5 mg/kg	-	-	-	0.62 (0.11 to 2.59)	0.69 (0.08 to 3.73)	-
IFX 5 + 10 mg/kg	-	-	-	0.90 (0.16 to 3.97)	0.48 (0.06 to 2.46)	-
VDZ 300 mg q.4.w.	**1.57 (1.002 to 2.47)**	1.91 (0.96 to 3.78)	1.40 (0.73 to 2.66)	1.53 (0.93 to 2.50)	1.58 (0.73 to 3.35)	1.43 (0.70 to 2.87)
VDZ 300 mg	1.66	1.60	1.89	**1.74**	1.43	1.43

Comparator	Enhanced Clinical Response – CDAI-100 OR (95% CrI)[a,b]			Clinical Remission – CDAI < 150 OR (95% CrI)[a,b]		
	Overall	Failed Conventional	Failed Anti-TNF	Overall	Failed Conventional	Failed Anti-TNF
q.8.w	(1.05 to 2.62)	(0.81 to 3.15)	(0.97 to 3.67)	**(1.05 to 2.88)**	(0.66 to 2.99)	(0.70 to 2.88)

ADA = adalimumab; CDAI = Crohn's Disease Activity Index; CrI = credible interval; EOW = every other week; IFX = infliximab; OR = odds ratio; q.4.w. = every 4 weeks; q.8.w. = every 8 weeks; q.12.w. = every 12 weeks; TNF = tumour necrosis factor; VDZ = vedolizumab.

Bold: Statistically significant difference.

[a] Reported as median OR with respective median credible intervals.

[b] Base-case analysis (Bayesian) based on the fixed-effects model.

Source: Manufacturer's submitted indirect comparison.[12]

Safety Results

Induction Therapy

No NMA was performed with the safety results from the induction- or maintenance-phase trials. NMA, while statistically feasible, was determined by the manufacturer to be infeasible from a clinical standpoint. That being said, safety results were reported for the induction phase as number of patients along with their respective proportions of AEs. Total AEs in all of the induction studies identified ranged from 52% to 74%, and the incidence of infections ranged from 9% to 26%. In addition, total serious AEs ranged 1% to 9%. Caution should be used when interpreting these results, as definitions of AEs differed between trials. Detailed descriptive safety results are provided in Table 46. No maintenance-phase safety results were presented.

TABLE 46: SAFETY RESULTS FROM THE INDUCTION PERIOD

	CERTIFI		UNITI-1		UNITI-2		Watanabe et al. 2012			CLASSIC I			GAIN		GEMINI II		GEMINI III	
	UST	PLA	UST	PLA	UST	PLA	ADA 160/80 mg	ADA 80/40 mg	PLA	ADA 160/80 mg	ADA 80/40 mg	PLA	ADA 160/80 mg	PLA	VDZ	PLA	VDZ	PLA
Total AEs, n (%)[a]	131 (61)	132 (71)	249 (66)	245 (65)	207 (57)	208 (54)	66 (52)	34 (59)	23 (52)	76 (75)	75 (68)	74 (74)	159 (57)	166 (73)	220 (56)	148 (59)	209 (56)	207 (60)
Infections, n (%)[b]	131 (22)	132 (24)	249 (26)	245 (24)	207 (22)	208 (23)	33 (12)	34 (15)	23 (9)	76 (21)	75 (17)	74 (16)	159 (16)	166 (23)	220 (15)	148 (18)	209 (19)	207 (17)
SAEs, n (%)[c]	131 (7)	132 (8)	249 (7)	245 (6)	207 (3)	206 (6)	33 (3)	34 (9)	23 (9)	76 (4)	75 (1)	74 (4)	159 (1)	166 (5)	220 (9)	148 (6)	209 (6)	207 (8)

ADA = adalimumab; AE = adverse event; PLA = placebo; SAE = serious adverse event; UST = ustekinumab; VDZ = vedolizumab.
Source: Manufacturer's submitted indirect comparisons.[12]

Critical Appraisal of Manufacturer's IDC

The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Task Force on Indirect Treatment Comparisons[70] was used to guide the critical appraisal of the manufacturer's submitted IDC.

The manufacturer's rationale for conducting the IDC (i.e., absence of head-to-head studies) and the objectives of the IDC (i.e., comparisons of biologics approved for use in patients with mild-to-moderate ulcerative colitis [ustekinumab versus infliximab, adalimumab, and vedolizumab]) were clearly reported. A comprehensive systematic review was performed with a two-stage selection process, in which articles were first selected based on titles and abstracts and then full-text articles were retrieved and their inclusion criteria ascertained. In addition, data extraction was performed and quality-checked by two independent reviewers. Risk of bias was assessed using the NICE checklist (Section A1.1.2 Quality Assessment of Included Studies), and detailed results of these assessments were provided. The manufacturer provided both inclusion and exclusion criteria that were used for screening and reported lists of both included and excluded references, with accompanying reasons. Specific and detailed dosage sets were provided for both the induction and maintenance phases. The manufacturer also provided figures of all networks.

The manufacturer's IDC described the efforts taken to assess the balance of effect modifiers across studies in order to establish the feasibility of performing an NMA. Through consultation with clinical experts, the manufacturer examined the time of assessment (which varied across trials) in the induction phase and subsequently chose the four-week time point for infliximab and adalimumab and the six-week time point for vedolizumab and ustekinumab. However, heterogeneity with regard to the timing of the assessments was still apparent in the individual studies, decreasing the confidence in the induction-phase results. In an attempt to minimize heterogeneity, the manufacturer performed separate analyses for patients who had experienced a failure with conventional therapy and with anti-TNF therapy. In addition, the manufacturer recommended caution when interpreting results of ustekinumab compared with infliximab for the induction phase for all outcomes, as the only study included in the IDC containing infliximab was an older and smaller study (ACCENT I), in which there were numerous issues (e.g., missing and non-evaluable placebo arm data, smaller magnitude of effects with the higher versus lower doses of infliximab, and lack of result reproducibility in the open-label induction phase). It should also be noted that there were some differences between the baseline patient characteristics (especially in regard to the ACCENT I infliximab trial), such as the mean duration of disease, mean CDAI score, C-reactive protein at baseline, smoking status, IBDQ scores, and disease phenotype. This further increased the uncertainty, decreased confidence in the results, and decreased the generalizability of both the induction and treatment-sequence analysis results for patients with moderate-to-severe CD.

With regard to the maintenance phase, the manufacturer undertook a feasibility assessment and noted that the similarity assumption had been violated as a result of the variation among studies in the selection criteria for entry into maintenance treatment. The feasibility assessment also noted that the transitivity assumption may have been violated, as the manufacturer identified different placebo rates. This aspect led the manufacturer to believe that the placebo arms across trials were not true comparators (clinical heterogeneity). Statistical heterogeneity was assessed using a chi-square test for remission and response in the subpopulations who had experienced a failure with conventional and anti-TNF therapies, showing *P* values less than 0.05, indicating heterogeneity. The manufacturer analyzed the maintenance of efficacy after one year using only those trials with the most comparable maintenance phases and performing a treatment-sequence analysis, in which subgroup data were available for results from the induction phases of these trials. It should be noted that all the placebo

arms corresponded to patients who received only active induction therapy. In addition, the exclusion of trials that did not include the most comparable maintenance phase for the analysis of efficacy after one year of treatment (the manufacturer did not provide a list of excluded trials) may introduce a risk of bias associated with the loss of applicable information. The treatment-sequence analysis method includes both induction and maintenance data for each intervention. In addition, imputed data (from individual patient-level data from IM-UNITI) was used for the maintenance data in the placebo arms, and weighted averages for the maintenance placebo rates were used for the maintenance placebo-to-placebo arms. However, this methodology is exploratory, as this method has not been validated. Data were imputed for the placebo arms and carries with it some uncertainty. Further, the approach is problematic because it does not maintain randomization past the induction phase. Therefore, caution must be used when interpreting the results because of the potential for bias and confounding, which have not been controlled for.

In order to assess the robustness of the BCA NMA results, numerous sensitivity analyses were conducted (Section A1.1.2 Indirect Comparison Methods). The results of the sensitivity analyses for all of the outcomes and populations (or subpopulations) indicated that all of the BCAs were robust. In addition, the manufacturer performed pairwise comparisons for the BCAs in the induction phase and complete treatment-sequence analysis for the two differing subpopulations (who had experienced a failure with conventional and anti-TNF therapies) to ensure that the NMA results were robust (which they were).

In order to observe model fit, DICs were calculated using both the fixed-effects and random-effects model for each outcome (CDAI-70, CDAI-100, and CDAI < 150), for each subpopulation (who had experienced a failure with conventional therapy or with anti-TNF therapy), and for both the induction phase and the treatment-sequence analyses. The model with the lowest DIC was the one used, with the fixed-effects model consistently favoured for every analysis.

The reviewers extracted data for more outcomes than were assessed in the NMA, but only three outcomes (clinical response [CDAI-70], enhanced clinical response [CDAI-100], and clinical remission [CDAI < 150]) were included. Although the CDAI is a validated measure of disease activity in CD and is composed of many separate outcomes (number of liquid stools, abdominal pain, well-being, abdominal mass, extra-intestinal manifestations, use of antidiarrheal drugs, body weight, and hematocrit values), other outcomes also help to inform how effective treatment is. In addition, while the trials themselves were probably not powered to detect statistically significant changes in the secondary outcomes, comparisons of outcomes such as mucosal, histologic, and endoscopic healing (which often correspond with clinical remission and clinical response), anemia, and patient-reported outcomes would still have been informative, as long as methodological issues were identified and mentioned. It should also be noted that, although statistically significant differences were observed in favour of infliximab 5 mg/kg when compared with ustekinumab 6 mg/kg for clinical remission (CDAI < 150) in the induction phase, one should remain skeptical that clinical remission is possible at such an early time point, as the clinical expert who was consulted on this submission pointed out.

All of the included randomized controlled trials (RCTs) were placebo-controlled, and none contained any head-to-head comparisons of the relevant biologic drugs. For this reason, none of the networks contained closed loops with regard to the biologics, precluding the ability to assess consistency between direct and indirect comparisons. Hence, the NMA would have been stronger if there were a closed loop (e.g., active comparison). In addition, all of the trials were, at most, one year long; hence, longer-term efficacy and safety were not assessed.

Safety outcomes were not included in any NMA for either the induction or maintenance phases. The manufacturer determined that conducting analyses on subpopulations would lead to small numbers of events being assessed, a loss of statistical power, and a lack of model convergence. It also noted that using a pooled/mixed population would be unacceptable from a clinical standpoint as a result of confounding. In addition, it noted that many AEs included gastrointestinal events that are related to CD, potentially introducing uncertainty about true treatment effects. This being said, an NMA increases the power of the assessment of AEs, and performing an NMA on either AEs or SAEs in these subpopulations would still provide insight into how patients with moderate-to-severe CD are adversely affected by the use of these different biologics.

Methods for Mocko et al.

Study Eligibility and Selection Process

Mocko et al.[14,15] conducted an IDC, based on a systematic review of the literature, that compared and evaluated the safety of biologic drugs (adalimumab, infliximab, certolizumab pegol, ustekinumab, and vedolizumab) with one another or with placebo in patients with CD (defined by conventional radiographic, endoscopic, and clinical criteria [CDAI > 150]). The biologics included had to have been approved for the treatment of CD by either the European Medicines Agency (EMA) or FDA; except ustekinumab which, at the time of this IDC, was undergoing the approval process for patients with CD by the EMA and FDA. Inclusion criteria for this IDC are outlined in Table 40. A systematic search of randomized controlled trials published in English was performed using multiple databases. Studies were included if the investigators examined the induction phase (follow-up of six to 10 weeks) or maintenance phase (follow-up of 52 to 56 weeks). Two independent reviewers assessed titles, abstracts, and full-text articles for inclusion in the IDC, with any discrepancies resolved by consensus with a third reviewer. All safety outcomes with a frequency of at least 3% were included and extracted by the same two independent reviewers. Only data from studies that adhered to the approved dosage regimens (approved by the EMA or FDA) were included in the IDC, and results for the induction and maintenance phases were analyzed separately. Selected trial characteristics are provided in Table 47.

TABLE 47: SELECTED STUDY CHARACTERISTICS INCLUDED IN THE MOCKO ET AL. INDIRECT COMPARISON

Study	Trial Characteristics	Phase	Treatment Group Included in IDC (n)	Follow-Up (Weeks)
Ustekinumab Trials				
CERTIFI	• DB RCT • Multi-centre (n = 153) • Phase II • Parallel assignment	• Induction	• UST (257) • PLA (131)	• 8
Adalimumab Trials				
CHARM	• DB RCT • Multi-centre (n = 92)	• Maintenance	• ADA (517) • PLA (261)	• 56
Classic II	• DB RCT • Multi-centre (n = 53)	• Maintenance	• ADA (37) • PLA (18)	• 56
Watanabe et al.	• DB RCT • Phase II/III • Parallel assignment	• Induction • Maintenance	• ADA (25) • PLA (25)	• 8 (induction) • 52 (maintenance)
Certolizumab Pegol Trials				
Sandborn et al.	• DB RCT • Multi-centre (n = 120) • Parallel assignment	• Induction	• CZP (223) • PLA (215)	• 6

Study	Trial Characteristics	Phase	Treatment Group	Follow-Up
Infliximab Trials				
ACCENT I	• DB RCT • Multi-centre (n = 55)	• Maintenance	• IFX (193) • PLA (188)	• 54
ACCENT II	• DB RCT • Multi-centre (n = 45)	• Maintenance	• IFX (138) • PLA (144)	• 54
Regueiro et al.	• DB RCT • Single assignment	• Maintenance	• IFX (11) • PLA (13)	• 54
Vedolizumab Trials				
GEMINI II	• DB RCT • Multi-centre (n = 285) • Phase III • Parallel assignment	• Induction • Maintenance	• VDZ (814) • PLA (301)	• 8 (induction) • 52 (maintenance)
GEMINI III	• DB RCT • Multi-centre • Phase II • Parallel assignment	• Induction	• VDZ (209) • PLA (207)	• 10

ADA = adalimumab; CZP = certolizumab pegol; DB = double-blind; IDC = indirect comparison; IFX = infliximab; PLA = placebo; RCT = randomized controlled trial; UST = ustekinumab; VDZ = vedolizumab.
Source: Mocko et al. 2016.[14,15]

Quality Assessment of Included Studies

The same two independent reviewers evaluated study quality using a domain-based evaluation tool recommended by the Cochrane Collaboration, and any ambiguities were resolved through consensus with a third reviewer. The domain-based tool examined the following domains: random sequence generation, allocation concealment, blinding (participants and investigators), incomplete outcome data, selective reporting, and other bias. Results of the individual quality assessments were provided by the manufacturer.

Indirect Comparison Methods

A Bayesian random-effects model (using the Markov chain Monte Carlo simulation) was used for the NMA, which was performed using the Aggregate Data Drug Information System (ADDIS) software, version 2. A consistency model was used, and this was based on 20,000 iterations for each of the four chains with a 5,000 iteration burn-in period. The assessment of convergence was performed using the Brooks–Gelman–Rubin method. Outcomes were expressed as median odds ratios with corresponding 95% CrIs. As previously stated, results for the induction and maintenance phases were analyzed separately. Clinical heterogeneity was assessed by the extraction and examination of the following trial characteristics: study design, dosage, treatment duration, follow-up, patients achieving predefined outcomes.

A random-effects pairwise meta-analysis was performed in order to obtain direct estimates of effects relative to placebo, which were used to confirm the results obtained in the NMA framework. Estimates of effects were presented as odds ratios and 95% confidence intervals (CIs). Heterogeneity for the pairwise meta-analyses was planned to be assessed using the I^2 parameter (with values of 0% representing no heterogeneity and those higher than 50% indicating significant heterogeneity). However, no formal assessment of statistical heterogeneity was performed as a result of the limited number of RCTs for the pairwise comparisons. In the absence of an assessment of statistical

heterogeneity, the authors used a conservative random-effects model with statistical significance defined as a *P* value less than 0.05.

Results

Study and Patient Characteristics

Ten RCTs were included in the NMA analyses. Of these studies, three were on the use of adalimumab (CHARM, CLASSIC II, and Watanabe et al.), one was on the use of certolizumab pegol (Sandborn et al.), three were on the use of infliximab (ACCENT I, ACCENT II, and Regueiro et al.), one was on the use of ustekinumab (CERTIFI), and two were on the use of vedolizumab (GEMINI II and GEMINI III). With regard to the induction phase, Watanabe et al. (adalimumab), Sandborn et al. (certolizumab pegol), GEMINI II and GEMINI III (vedolizumab), and CERTIFI (ustekinumab) were considered. For the maintenance phase, Watanabe et al., CHARM, and CLASSICII (adalimumab), ACCENT I, ACCENT II, and Regueiro et al. (infliximab), and GEMINI II (vedolizumab) were considered. Using placebo as the common comparator, Mocko et al. determined that there were no significant differences among these groups (based on their characteristics); hence, they used the placebo group as the link for the NMA. No statistical heterogeneity was observed for the induction phase (I^2 = 0%) for all analyzed outcomes, and the same lack of heterogeneity was observed in the maintenance phase for the majority of the outcomes. The only exceptions to this in the maintenance phase were, in the use of infliximab, injection-site reactions and AEs leading to study drug discontinuation (I^2 = 70% to 86%) and, in the use of adalimumab, infections and nasopharyngitis (I^2 = 73% to 80%). Convergences were achieved at 20,000 simulations for all of the end points. All of the included RCTs were evaluated as having a high risk of bias, especially with regard to the "incomplete outcome data" domain. In addition, there was an unclear risk of bias regarding "allocation concealment" in six RCTs, whereas there was an apparent low risk of bias in the other domains. Specific patient characteristics for the individual trials (e.g., age, sex, mean duration of CD, proportion of patients who are treatment-naive versus -experienced, concomitant medications, etc.) were not provided.

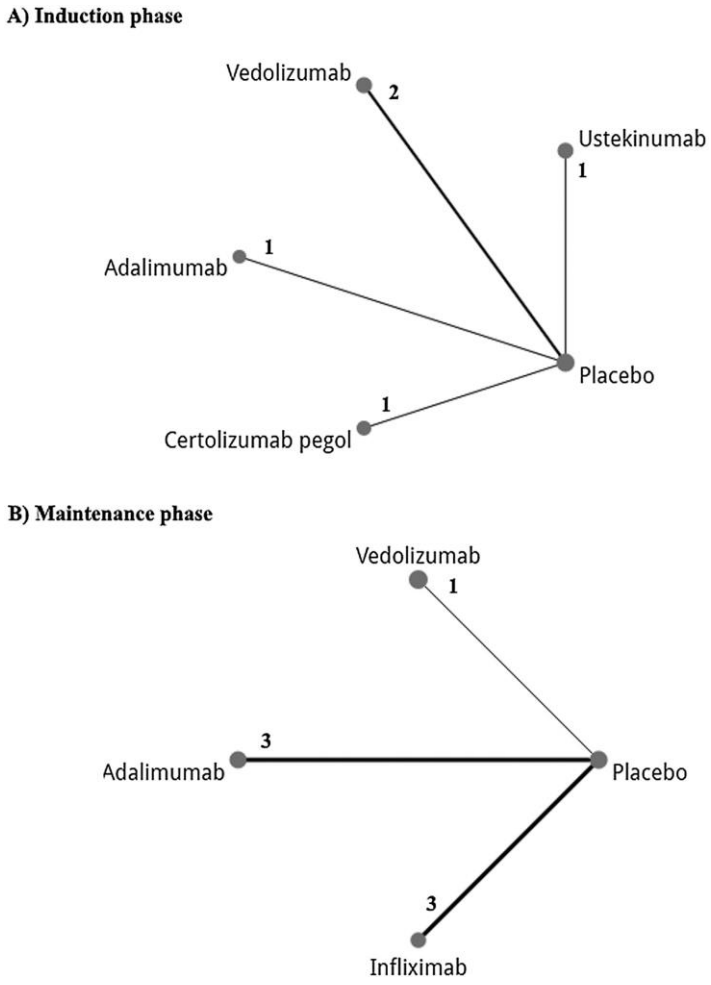
Evidence Networks

Evidence networks were provided by Mocko et al. None of the networks contained any closed loops, and networks were anchored to placebo as the only common treatment arm between studies. The following evidence networks were provided.

Induction and Maintenance Networks

The overall network for the induction NMA is presented in Figure 11A. The overall network for the maintenance NMA is presented in Figure 11B.

FIGURE 11: OVERALL NETWORKS FOR THE INDUCTION AND MAINTENANCE PHASES FOR MOCKO ET AL.



Source: Reprinted from Pharmacological Reports, 68/6, Mocko P, Kawalec P, Pilc A, Safety profile of biologic drugs in the therapy of Crohn disease: a systematic review and network meta-analysis, 1237-43 Copyright (2016), with permission from Elsevier[14]

Safety Results

Induction Therapy

No statistically significant differences were evident among any of the biologics examined in the induction-phase NMA with regard to any AEs, infections, injection-site reactions, SAEs, and drug discontinuations due to AEs. In addition, no statistically significant differences were observed among biologics for individual AEs, such as abdominal pain, arthralgia, headache, nausea, or nasopharyngitis. The pairwise meta-analyses against placebo echoed the lack of statistical significance (data not shown). Detailed NMA results for the induction phase are provided in Table 48.

TABLE 48: NETWORK META-ANALYSIS SAFETY RESULTS FOR THE INDUCTION PERIOD

Intervention[a]	Comparators		
	Adalimumab	Ustekinumab	Vedolizumab
Any AE (n = 5 RCTs), OR (95% CrI)			
Adalimumab	-	1.17 (0.36 to 4.25)	1.35 (0.43 to 4.34)
Ustekinumab	-	-	1.14 (0.54 to 2.38)
Infections (n = 5 RCTs), OR (95% CrI)			
Adalimumab	-	0.52 (0.05 to 2.69)	0.53 (0.06 to 2.82)
Ustekinumab	-	-	1.07 (0.49 to 2.29)
Injection-site reactions (n = 3 RCTs), OR (95% CrI)			
Adalimumab	-	0.52 (0.05 to 4.09)	0.29 (0.02 to 2.46)
Ustekinumab	-	-	0.56 (0.09 to 3.16)
SAEs (n = 5 RCTs), OR (95% CrI)			
Adalimumab	-	0.95 (0.09 to 7.29)	1.53 (0.16 to 10.78)
Ustekinumab	-	-	1.62 (0.48 to 5.16)
Drug discontinuations due to AE, (n = 3 RCTs), OR (95% CrI)			
Adalimumab	-	0.75 (0.04 to 8.09)	-
Individual AEs			
Abdominal pain (n = 3 RCTs), OR (95% CrI)			
Ustekinumab	-	-	2.04 (0.48 to 8.90)
Arthralgia (n = 3 RCTs), OR (95% CrI)			
Ustekinumab	-	-	0.72 (0.15 to 3.16)
Headache (n = 4 RCTs), OR (95% CrI)			
Ustekinumab	-	-	0.67 (0.23 to 1.87)
Nausea (n = 3 RCTs), OR (95% CrI)			
Ustekinumab	-	-	1.59 (0.31 to 7.91)
Nasopharyngitis (n = 2 RCTs), OR (95% CrI)			
Ustekinumab	-	-	0.74 (0.14 to 4.24)
Pyrexia (n = 2 RCTs), OR (95% CrI)			
Not applicable[b]			

AE = adverse event; CrI = credible interval; OR = odds ratio; RCT = randomized controlled trial; SAE = serious adverse event.
[a] Extracted outcomes had to have an incidence frequency of at least 3%.
[b] Data not presented because only a comparison between certolizumab and vedolizumab was reported for this outcome in the publication.
Source: Mocko et al. 2016.[14,15]

Maintenance Therapy

No statistically significant differences were evident among any of the biologics (adalimumab, infliximab, or vedolizumab) examined in the maintenance-phase NMA with regard to any AEs, infections, serious infections, injection-site reactions, SAEs, and drug discontinuations due to AEs. In addition, no statistically significant differences were observed among any of the aforementioned biologics for the individual AEs, such as abdominal pain, arthralgia, headache, nausea, nasopharyngitis, pyrexia, or upper respiratory infections. In contrast to this, a statistically significant difference in favour of adalimumab over placebo was observed for injection-site reactions and in favour of placebo over vedolizumab for any AEs and SAEs (data not shown). No comparisons were available for ustekinumab in the maintenance phase.

Critical Appraisal of Mocko et al. Indirect Comparison

While the IDC by Mocko et al. included acceptable methods in terms of the systematic review portion, some major limitations preclude definitive conclusions with regard to the results.

The rationale and objectives for conducting the IDC on the safety of these biologics for patients with CD were clearly reported, as were the methods for the systematic review and their statistical analyses. A comprehensive systematic review was performed with a two-stage selection process: articles were first selected based on titles and abstracts and then full-text articles were retrieved and their inclusion criteria ascertained. In addition, data extraction was performed and quality-checked by two independent reviewers. The two independent reviewers also assessed the risk of bias using a tool recommended by the Cochrane Collaboration based on the aforementioned domains (reported in Section A1.2.1 Quality Assessment of Included Studies). In addition, the authors were not funded externally by any pharmaceutical company, and no conflicts of interest were reported.

Mocko et al. did not report (either in the main citation or supplemental information) the individual trial or patient characteristics for the studies included in their IDC. This lack of reporting adds to the uncertainty regarding not only the severity of the disease, but also the heterogeneity of the patients across studies (although most of the studies, but not all, were included in the other IDCs). The authors also did not specify that all patients must have moderate-to-severe CD for inclusion in the IDC; instead, patients were included if they had clinical, endoscopic, and radiologic evidence of CD and CDAI greater than 150. While this is consistent with some of the other IDCs,[12,13] this score does not definitively place an individual in the moderate-to-severe category of CD. For instance, according to international definitions of CD based on the CDAI parameters, scores above 150 can be classified as mild-to-moderate (CDAI 150 to 220), moderate-to-severe (CADI 220 to 450), or severe/fulminant (CDAI > 450).[71] Other factors (such as age, mean duration of CD, concomitant medications, individual baseline CDAI, and other clinical, endoscopic, or radiologic assessments, etc.) were not reported; therefore, external readers are unable to ascertain the baseline heterogeneity of the patients in the included studies. There was also no mention of whether the patients were treatment-naive or treatment-experienced and no mention of any subgroup analyses (a priori or post hoc) that would address these potential confounders. This lack of reporting also extended to information on whether the authors assessed both the random- and fixed-effects models (and subsequent DIC calculations to ascertain the model of best fit) and which priors were used. All of this lack of reporting furthers the uncertainty surrounding the effect estimates that were obtained.

The authors made the assumption that there were no significant differences among patients in the placebo groups and therefore used the placebo groups as the common comparator and anchor for the networks. However, differences in the placebo groups, especially in those in the maintenance phase (as previously discussed in the manufacturer's submitted IDC[12]) preclude simply "using" the placebo group as a common comparator, although this is often the main anchor in the NMA process. This adds additional uncertainty with regard to the safety outcomes assessed. In addition, there were no closed loops in either the induction- or maintenance-phase networks, precluding any assessment of consistency.

The authors performed an assessment of bias in the included studies, noting a high risk of bias regarding incomplete outcome date (or attrition bias) in all 10 RCTs and an unclear risk of bias regarding allocation concealment (or selection bias) in six RCTs. Since Mocko et al. looked only at safety outcomes with an incidence of at least 3%, this high risk of bias associated with the incomplete outcome data further increases the uncertainty associated with the outcome results. Hence, results need to be interpreted

with caution. In addition, the full safety profile of these biologics was not ascertained (as AEs with an incidence of less than 3% were excluded); therefore, one must keep these results in context.

The authors noted that the limited length of the follow-up period was a significant limitation for assessment of all possible AEs, particularly with regard to the induction phase of up to 10 weeks for certolizumab pegol, infliximab, and ustekinumab. While the maintenance-phase analysis did extend the timelines, AEs or SAEs such as cancer may not have been observed during this longer follow-up period (up to 54 weeks), and a longer length of follow-up would have been more appropriate. That being said, this is a flaw of the RCTs themselves, rather than of the NMA. It should also be noted that there was no safety evidence for ustekinumab following the induction phase; therefore, no conclusions outside of the induction phase can be made.

Finally, most RCTs are designed with the power to detect efficacy end points as their primary and secondary end points; however, the same cannot be said of the safety end points. Although an NMA allows for an assessment of these types of outcomes because of the larger sample size (using multiple RCTs), the absence of a sufficiently large sample of patients to detect less commonly observed safety outcomes in the original RCTs contributes to the uncertainty surrounding the NMA results obtained.

Methods for Singh et al.

Study Eligibility and Selection Process

Singh et al.[13] conducted an IDC that compared and evaluated the relative efficacy of biologic drugs (adalimumab, certolizumab pegol, infliximab, natalizumab, ustekinumab, and vedolizumab) with one another or with placebo in treatment-naive (biologic-naive) patients with moderate-to-severe CD (defined on a CDAI greater than 220 but less than 450). Inclusion criteria for this IDC are outlined in Table 40. A systematic search of RCTs (with no language restriction) was performed using multiple databases. For trials assessing induction therapy, only those that assessed outcomes in patients not previously exposed to biologics or that reported results separately for patients not previously treated with biologics were included. Similarly for the maintenance phase, trials were excluded when outcomes were not reported separately for patients not previously treated with biologics in those who initially responded to induction therapy. Two independent reviewers assessed titles, abstracts, and full-text articles for inclusion in the IDC, with conflicts resolved by consensus only when examining the full-text articles. In addition, bibliographies of the included studies were searched, along with abstracts from major gastroenterology conferences, in order to obtain unpublished information. The primary outcome of interest was the induction of clinical remission in biologic-naive patients with active CD (up to 14 weeks) and the maintenance of remission of the patients who initially responded to biologics in the induction phase (up to 60 weeks). A hierarchical approach was used in both the induction and maintenance phases, in which CDAI < 150 was preferred, followed by CR-100 (failure to achieve a reduction of more than 100 points from baseline), followed by CR-70 (failure to achieve a reduction of 70 points from baseline). Selected trial and patient characteristics are provided in Table 49.

TABLE 49: SELECTED STUDY CHARACTERISTICS INCLUDED IN THE SINGH ET AL. INDIRECT COMPARISON

Drug	Study	Sample Size		Trial Design	Intervention and Dosage	Definition of Remission	Patients With Induction of Remission n/N (%)		Concomitant Medications %	
		Total	Biologic-Naive				Active	PLA	Active	PLA
Induction studies										
UST	Sandborn et al. 2008	104	52	• DB RCT • Multi-centre • PLA-controlled	• UST 90 mg SC at wks 0, 1, 2, and 3 • UST 4.5 mg IV at week 0	• CR-70 at wk 8	11/26 (42.3)	14/26 (53.8)	• IM: 29% • 5-ASA: 37% • CS: 30%	• IM: 38% • 5-ASA: 51% • CS: 30%
ADA	Hanauer et al. 2006	225[a]	225	• DB RCT • Multi-centre • PLA-controlled	• ADA160/80 mg SC at wks 0 and 2 • ADA 80/40 mg SC at wks 0 and 2	• CDAI < 150 at wk 4	45/151 (29.8)	9/74 (12.2)	• AZA: 13% • 6-MP: 13% • 5-ASA: 52% • CS: 37%	• AZA: 18% • 6-MP: 11% • 5-ASA: 50% • CS: 34%
	Watanabe et al. 2012	90	38	• DB RCT • Multi-centre • PLA-controlled	• ADA 160/80 mg SC at wks 0 and 2 • ADA 80/40 mg SC at wks 0 and 2	• CDAI < 150 at wk 4	10/28[b] (35.7)	2/10 (20.0)	• IM: 31% • 5-ASA: 88% • CS: 21%	• IM: 35% • 5-ASA: 100% • CS: 22%
CZP	Sandborn et al. 2007	659	469	• DB RCT • Multi-centre • PLA-controlled	• CZP 400 mg SC at wks 0, 2, and 4	• CR-100 at wk 6	91/229 (39.7)	70/240 (29.2)	• IM: 37% • CS: 39%	• IM: 38% • CS: 29%
	Sandborn et al. 2011	424	424	• DB RCT • Multi-centre • PLA-controlled	• CZP 400 mg SC at wks 0, 2, and 4	• CDAI < 150 at wk 6	68/215 (31.6)	53/209 (25.3)	• IM: 35% • CS: 44%	• IM: 31% • CS: 46%
IFX	Targan et al. 1997	108	108	• DB RCT • Multi-centre • PLA-controlled	• IFX 5, 10, or 20 mg/kg IV at wk 0	• CDAI < 150 at wk 4	27/83 (32.5)	1/25 (4.0)	• AZA: 20% • 6-MP: 14% • 5-ASA: 57% • CS: 58%	• AZA: 28% • 6-MP: 16% • 5-ASA: 68% • CS: 64%
	Lémann et al. 2006	111	111	• DB RCT • Multi-centre • PLA-controlled	• IFX 5 mg/kg IV at wks 0, 2, and6	• CDAI < 150 at wk 12	41/55 (74.5)	21/56 (37.5)	• AZA/6-MP: 100% • CS: 100%	• AZA/6-MP: 100% • CS: 100%
NAT	Ghosh et al. 2003	248[c]	180	• DB RCT • Multi-centre • PLA-controlled	• NAT 3 or 6 mg/kg IV at wks 0 and 4	• CDAI < 150 at wk 6	45/117 (38.5)	17/63 (27.0)	• IM: 22% • 5-ASA: 61% • CS: 59%	• IM: 35% • 5-ASA: 48% • CS: 49%

Drug	Study	Sample Size		Trial Design	Intervention and Dosage	Definition of Remission	Patients With Induction of Remission n/N (%)		Concomitant Medications %	
		Total	Biologic-Naive				Active	PLA	Active	PLA
	Sandborn et al. 2005	905	545	• DB RCT • Multi-centre • PLA-controlled	• NATI 300 mg IV at wks 0, 4, and 8	• CDAI < 150 at wk 10	169/433[b] (39.0)	39/112 (34.8)	• AZA: 23% • 6-MP: 7% • 5-ASA: 47% • CS: 37%	• AZA: 21% • 6-MP: 4% • 5-ASA: 44% • CS: 39%
VDZ	Feagan et al. 2008	185	185	• DB RCT • Multi-centre • PLA-controlled	• VDZ 0.5 to 2.0 mg/kg IV at wks 0 and 4	• CDAI < 150 at wk 8	43/127 (33.8)	12/58 (20.7)	• NR	• NR
	Sandborn et al. 2013	368	193	• DB RCT • Multi-centre • PLA-controlled	• VDZ 300 mg IV at wks 0 and 2	• CDAI < 150 at wk 6	21/115[a] (18.3)	7/78 (9.0)	• IM: 34% • CS: 48%	• IM: 17% • CS: 48%
Maintenance studies										
UST	Sandborn et al. 2012	-	-	• UST 1, 3, or 6 mg/kg IV at wk 0; then responders (CR-100 at wk 6) randomized to USTd or PLA • Multi-centre	• UST 90 mg SC at wks 8 and 16	• CDAI < 150 at wk 22 after initial randomization among IP responders	30/72 (41.17)	20/73 (27.4)	• NR	• NR
ADA	Sandborn et al. 2007	-	-	• Pts received either ADA 160/80, 80/40, or 40/20 mg or PLA at wks 0 and 2 (from IP CLASSIC I trial); then pts in remission (CDAI < 150 at wks 4 and 8 of original induction randomization) were randomized to PLA or ADA every wk or EOW or PLA	• ADA 40 mg SC every wk or; • ADA 40 mg SC EOW	• CDAI < 150 at wk 56 after re-randomization among IP responders	30/37 (81.1)	8/18 (44.4)	• AZA: 16% • 6-MP: 8% • 5-ASA: 70% • CS: 46%	• AZA: 6% • 6-MP: 0% • 5-ASA: 44% • CS: 56%
	Colombel et al. 2007	-	-	• Pts received OL ADA 80/40 SC at wks 0 and 2; then responders (CR-70 at wk 4) were randomized to PLA or ADA	• ADA 40 mg SC every week or; • ADA 40 mg SC EOW	• CDAI < 150 at wk 56 after randomization, among IP responders	127/329 (38.6)	20/170 (11.8)	***Overall cohort:*** • AZA: 33% • 6-MP: 8% • 5-ASA: 41% • CS: 42%	

Drug	Study	Sample Size		Trial Design	Intervention and Dosage	Definition of Remission	Patients With Induction of Remission n/N (%)		Concomitant Medications %	
		Total	Biologic-Naive				Active	PLA	Active	PLA
	Watanabe et al. 2012	-	-	• Pts received either ADA 160/80 or 80/40 mg SC or PLA at wks 0 and 2 (IP); then responders (CR-70 at wk 4) randomized to PLA or ADA	• ADA 40 mg SC EOW	• CDAI < 150 at wk 52 after re-randomization among IP responders	8/21 (38.1)	2/22 (9.1)	• IM: 44% • 5-ASA: 100% • CS: 12%	• IM: 28% • 5-ASA: 76% • CS: 20%
CZP	Schreiber et al. 2007	-	-	• Pts received OL CZP 400 mg SC at wks 0, 2, and 4; then responders (CR-100 at wk 6) were randomized to PLA or CZP	• CZP 400 mg SC q.4.w.	• CDAI < 150 at wk 26 after randomization among IP responders	103/215 (47.9)	61/210 (29.0)	• IM: 40% • CS: 35%	• IM: 41% • CS: 37%
IFX	Hanauer et al. 2002	-	-	• Pts received initial dose of IFX 5 mg/kg IV at wk 0; then responders (CR-70 at wk 2) were randomized to PLA or IFX	• IFX 5 or 10 mg/kg IV at wks 2 and 6 and q.8.w	• CDAI < 150 at wk 30 after randomization among IP responders	94/225 (41.8)	23/110 (20.9)	***Overall cohort:*** • AZA/6-MP: 24% • 5-ASA: 47% • CS: 52%	
	Rutgeerts et al. 1999	-	-	• Pts received either single dose of IFX 5, 10, or 20 mg/kg IV or PLA (from IP of Targan et al.); then responders (CR-70 at wk 8) were randomized to PLA or IFX; nonresponders at wk 4 of original IP randomization received an OL, single dose of IFX 10 mg/kg at wk 4 (considered wk 0); and 8-wk responders to these were randomized to PLA or IFX	• IFX 10 mg/kg IV q.8.w	• CDAI < 150 at wk 36 after re-randomization among IP responders	20/37 (54.0)	7/36 (19.4)	NR	NR
NAT	Sandborn et al. 2005	-	-	• Pts received either NAT 300 mg IV or PLA at wks 0,	• NAT 300 mg IV q4w	• CDAI < 150 at wk 60 after re-	92/168 (54.8)	38/171 (22.2)	• AZA: 25% • 6-MP: 7%	• AZA: 26% • 6-MP: 4%

Drug	Study	Sample Size		Trial Design	Intervention and Dosage	Definition of Remission	Patients With Induction of Remission n/N (%)		Concomitant Medications %	
		Total	Biologic-Naive				Active	PLA	Active	PLA
				4, and 8 (from IP of ENACT-I trial); then responders (CR-70 at wks 10 and 12) were randomized to PLA or NAT		randomization among IP responders			• 5-ASA: 45% • CS: 38%	• 5-ASA: 54% • CS: 44%
VDZ	Sandborn et al. 2013	-	-	• Pts received either VDZ 300 mg or PLA IV at wks 0 and 2 (from IP of GEMINI II trial); then responders (CR-70 at wk 6) were randomized to PLA or VDZ	• VDZ 300 mg IV q.4.w. or; • VDZ 300 mg IV q.8.w.	• CDAI < 150 at wk 52 after re-randomization among IP responders	116/308 (37.7)	33/153 (21.6)	• IM: 33% • CS: 53%	• IM: 32% • CS: 54%

5-ASA = 5-aminosalicylate; 6-MP = 6-mercaptopurine; ADA = adalimumab; AZA = azathioprine; CDAI = Crohn's Disease Activity Index; CR = clinical remission; CS = corticosteroid; CZP = certolizumab pegol; DB = double-blind; EOW = every other week; IFX = infliximab; IM = immunosuppressant; IP = induction phase; IV = intravenous; NAT = natalizumab; NR = not reported; OL = open label; PLA = placebo; pts = patients; q.4.w. = every 4 weeks; q.8.w. = every 8 weeks; RCT = randomized controlled trial; SC = subcutaneous; UST = ustekinumab; VDZ = vedolizumab; wk(s) = week(s).

[a] The total number of patients was 299, but the investigators excluded those patients who received induction with ADA 40/20 mg (n = 74).

[b] The efficacy data were only from the subset of biologic-naive patients.

[c] Total number of patients was 248, but the investigators excluded patients who received induction with only one dose of natalizumab IV 3 mg/kg.

[d] Data for both responders and nonresponders to induction therapy were available, but, for the purposes of analysis, the only data from the subset of responders to induction therapy were included.

Source: Singh et al. 2014.[13]

Quality Assessment of Included Studies
Two independent reviewers evaluated the quality of the included studies using criteria set out by the Evidence-Based Gastroenterology Steering Group. The criteria included random allocation concealment, patient and caregiver blinding, equal co-intervention use for both the placebo and treatment arms, follow-up of study patients (completeness), and the use of an intention-to-treat analysis.

Indirect Comparison Methods
A Bayesian random-effects model (using the Markov chain Monte Carlo simulation) was used for the NMA, which was performed using WinBUGS 1.4.3. The comparative efficacy between two treatments was assessed as a function of the active treatment relative to the reference (placebo) treatment. Non-informative priors were used, along with a burn-in of 10,000 iterations. The Markov chain Monte Carlo model used 100,000 simulations. Results were reported as a relative risk (RR) with 95% CrIs, and the authors adjusted for trials with multiple arms.

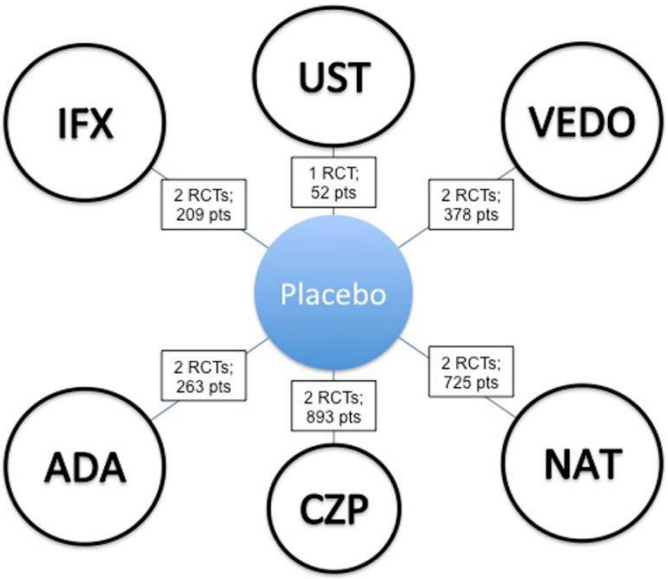
Active drugs and placebo were also compared in pairwise analyses using a random-effects model, with results presented as a pooled RR with 95% CIs. Statistical heterogeneity (although not formal assessment of clinical heterogeneity) was assessed using the I^2 statistic, and publication bias was examined using Egger's regression test to produce funnel-plot symmetry.

Results
Study and Patient Characteristics
Seventeen RCTs were included in the NMA, with 11 RCTs focused on induction of remission and nine trials focused on the maintenance of remission. With regard to the induction trials, adalimumab, certolizumab pegol, infliximab, natalizumab, and vedolizumab were each assessed in two trials, whereas only one study assessed ustekinumab. There was variability in the proportion of patients taking concomitant immunosuppressive drugs (range of 4% to 100%), corticosteroids (range of 21% to 100%), and 5-ASA drugs (range of 37% to 100%) between trials. With regard to the maintenance trials, three studies assessed adalimumab; one study assessed certolizumab pegol; two studies assessed infliximab; and one study assessed each of natalizumab, ustekinumab, and vedolizumab. The percentage of patients receiving concomitant medications ranged from 0% to 44% (immunomodulators), 12% to 56% (corticosteroids), and 41% to 100% (5-ASAs) in those trials that reported them. Based on the assessment of the risk of bias for the induction studies, one study failed to report whether the random allocation was concealed, one study did not report on the equal use of co-interventions in both treatment and placebo arms, and one study did not report on the use of the intention-to-treat analysis. With regard to the assessment of the risk of bias for the maintenance studies, four studies failed to report on the equal use of co-interventions in both treatment and placebo arms.

Evidence Networks
Evidence networks were provided by Singh et al. None of the networks contained any closed loops, and networks were anchored to placebo as the only common treatment arm between studies. The following evidence networks were provided.

Induction Network
The overall network for the induction NMA is presented in Figure 12.

FIGURE 12: OVERALL NETWORKS FOR THE INDUCTION PHASE FOR SINGH ET AL.

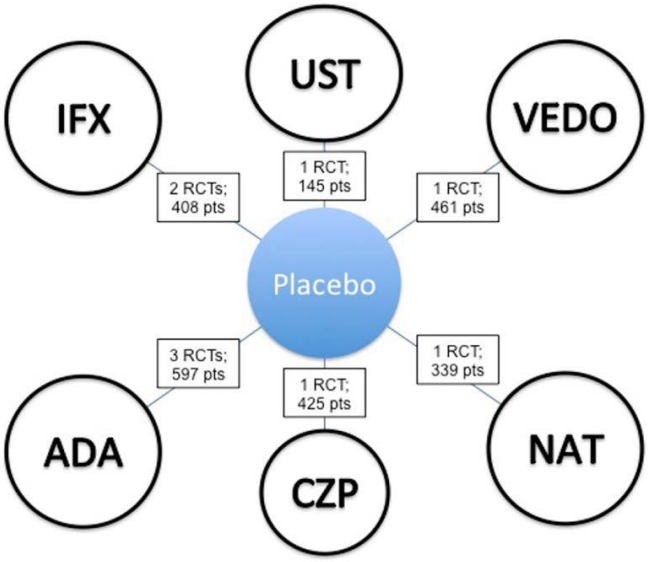


ADA = adalimumab; CZP = certolizumab pegol; IFX = infliximab; NAT = natalizumab; RCT = randomized controlled trial; VEDO = vedolizumab; UST = ustekinumab.
Source: Reprinted from Mayo Clinic Proceedings, 89/12, Singh S, Garg SK, Pardi DS, Wang Z, Murad MH, Loftus EV, Jr, Comparative efficacy of biologic therapy in biologic-naive patients with Crohn disease: a systematic review and network meta-analysis, 1621-35, Copyright (2014), with permission from Elsevier[13]

Maintenance Network

The overall network for the maintenance NMA is presented in Figure 13.

FIGURE 13: OVERALL NETWORKS FOR THE MAINTENANCE PHASE FOR SINGH ET AL.



ADA = adalimumab; CZP = certolizumab pegol; IFX = infliximab; NAT = natalizumab; RCT = randomized controlled trial; VEDO = vedolizumab; UST = ustekinumab.
Source: Reprinted from Mayo Clinic Proceedings, 89/12, Singh S, Garg SK, Pardi DS, Wang Z, Murad MH, Loftus EV, Jr, Comparative efficacy of biologic therapy in biologic-naive patients with Crohn disease: a systematic review and network meta-analysis, 1621-35, Copyright (2014), with permission from Elsevier[13]

Efficacy Results

Induction Therapy

When ustekinumab (RR 0.1; 95% CrI, 0.02 to 0.52), vedolizumab (RR 0.23; 95% CrI, 0.06 to 0.78), natalizumab (RR 0.22; 95% CrI 0.06 to 0.70), and certolizumab pegol (RR 0.24; 95% CrI, 0.07 to 0.73) were compared with infliximab, there was a statistically significant difference in favour of infliximab for the induction of remission. No other differences were observed with regard to the induction of remission. Detailed NMA results for the induction period are provided in Table 50. The anti-TNF drugs were superior to placebo in the standard pairwise meta-analysis, while the anti-integrins and the IL-12/IL-23 antogonist (ustekinumab) were not superior to placebo for the induction of remission (data not shown).

Maintenance Therapy

No statistically significant differences among any of the biologics were evident for maintaining remission in those patients who were biologic-naive upon entering the induction phase and were subsequent responders in the induction phase. Detailed results for the maintenance period are provided in Table 50. With regard to the standard pairwise meta-analysis for the maintenance of remission, anti-TNF drugs and the IL-12/IL-23 antagonist were statistically significantly superior when compared with placebo; however, the same was not observed with the anti-integrins (data not shown).

TABLE 50: NETWORK META-ANALYSIS RESULTS IN BIOLOGIC-NAIVE PATIENTS IN THE SINGH ET AL. INDIRECT COMPARISON

Comparators	Induction of Remission[a] Pooled RR (95% CrI)	Maintenance of Remission[a] Pooled RR (95% CrI)
Ustekinumab versus		
ADA	0.20 (0.04 to 1.16)	0.37 (0.04 to 3.00)
IFX	**0.10 (0.02 to 0.52)**	0.58 (0.05 to 5.20)
VDZ	0.43 (0.09 to 2.23)	0.87 (0.07 to 11.36)
Vedolizumab versus		
ADA	0.47 (0.13 to 1.75)	0.43 (0.05 to 3.36)
IFX	**0.23 (0.06 to 0.78)**	0.67 (0.06 to 5.64)
Adalimumab versus		
IFX	0.49 (0.11 to 1.85)	1.56 (0.26 to 8.92)

ADA = adalimumab; CrI = credible interval; IFX = infliximab; RR = relative risk; VDZ = vedolizumab.
Bold: Statistically significance difference.
[a] RR > 1 suggests greater efficacy when compared with reference treatment, while RR < 1 suggests lesser efficacy when compared with reference treatment.
Source: Singh et al. 2014.[13]

Critical Appraisal of Singh et al. Indirect Comparison

The IDC by Singh et al. appeared to use acceptable methods for the systematic review; however, there are some major limitations which lessen confidence in the results.

The authors' rationale and objectives for conducting the IDC on the efficacy the aforementioned biologics for inducing and maintaining remission in biologic-naive patients were clearly reported, as were the methods for the systematic review and their statistical analyses. In describing the NMA process, the authors did report that they adjusted for trials with multiple arms; however, details on this adjustment were lacking. A comprehensive systematic review was performed, in which articles were first selected based on titles and abstracts, and then full-text articles were retrieved and their inclusion criteria ascertained. However, there was uncertainty surrounding whether the dual selection pertained to both the title and abstract selection and full-text article selection, as it was described solely for the latter. If the selection pertained solely to the full-text articles (instead of being reported only for full-text articles), then certain articles may have been missed. Data extraction was performed and quality-checked by two independent reviewers, as was the assessment of bias. However, the criteria assessed were less stringent than those of the other two IDCs;[12,14,15] the authors did not assess the similarity of prognostic factors between groups at baseline (no assessment of clinical heterogeneity), selective end point reporting, or the adequacy of methods for handling missing data. Although they assessed both patient and caregiver blinding, they did not appear to assess the blinding of the end point assessor. Some of the authors had stated conflicts of interest, and the NMA was funded by a grant from the Kern Center for the Science of Health Care Delivery, which is part of the Mayo Clinic.

In order to obtain efficacy results solely for the population of biologic-naive patients, the authors included only data on this subgroup for induction therapy. In addition, they subsequently included only data from the subset of patients who responded to induction therapy for the efficacy assessment in the maintenance phase. Induction trials that did not report outcomes separately for biologic-naive patients were excluded. In addition, maintenance trials that did not report outcomes separately for the subset of patients who responded to induction therapy were also excluded. In addition, by including only the

subset of patients who responded to induction therapy in the maintenance trials, the authors artificially enriched the population of patients with those who were more likely to see an increased benefit on these biologics for the maintenance of remission. This further increases the uncertainty with regard to the results and decreases the generalizability of this treatment for those who are biologic-naive.

As with all NMAs, the primary objective was to compare drugs across similar populations of patients with predominantly the same characteristics. Although this IDC attempted to do that, there were inherent differences within the populations (as observed in Table 49); for example, some prognostic factors were different (e.g., disease duration, use of concomitant immunomodulatory medications, phenotype of disease). In addition, there were differences in the assessment timing for both the induction and maintenance phases among the included RCTs, which could have led to differences in responses. These issues, along with the fact that there were no closed loops in either the induction- or maintenance-phase networks, also call into question the generalizability of the results and further undermine confidence in them.

As in the manufacturer's submitted IDC, the lack of other end points (e.g., mucosal healing) does not allow one to obtain the full picture of efficacy for this class of drugs. As the clinical expert consulted for this review mentioned, although remission is a good overall end point, it is usually unattainable during the induction phase and does not fully encompass the entire realm of a successful treatment. Endoscopic evidence of remission (which is not always feasible to obtain in every patient), including mucosal healing, is an important efficacy end point that should be looked at.

This IDC did not attempt to perform an NMA on any safety end points. Although most RCTs are not powered to detect differences in safety end points, it is still beneficial (especially considering the classes of drugs involved with the treatment of CD) to attempt an NMA of safety end points. This would have provided some insight into the safety aspects involved in treating patients with moderate-to-severe CD with these biologics.

Conclusions

The manufacturer submitted an IDC[12] of ustekinumab versus infliximab, adalimumab, and vedolizumab using Bayesian NMA with placebo as the common comparator (for the induction phase) and using treatment-sequence analyses (which includes induction and maintenance data) for the maintenance phase. The manufacturer noted that there were no statistically significant differences in clinical response (CDAI-70), enhanced clinical response (CDAI-100), or clinical remission (CDAI < 150) among ustekinumab 6 mg/kg and either adalimumab 80/40 mg, adalimumab 160/80 mg, or vedolizumab 300 mg in the induction phase in either the subpopulation who had experienced a failure with conventional and anti-TNF therapies. Statistically significant differences in clinical response (CDAI-70) and clinical remission (CDAI < 150) in favour of infliximab 5 mg/kg compared with ustekinumab 6 mg/kg were apparent in the subpopulation that had experienced failure with conventional therapy. However, the degree of heterogeneity noted for the network comparisons, including infliximab, reduced the certainty around these results.

For treatment-sequence NMA results (based on an exploratory methodology; hence, results should be interpreted with caution), the only statistically significant differences observed were in favour of ustekinumab 90 mg every 12 weeks when compared with vedolizumab 300 mg every eight weeks when looking at enhanced clinical response (CDAI-100) in the overall population (results were not statistically significantly different among the subpopulations). All other results showed no differences between ustekinumab 90 mg every 12 weeks and the other biologic treatment regimens. With regard to

ustekinumab 90 mg every eight weeks, statistical significance in favour of ustekinumab was evident over vedolizumab 300 mg every four weeks for enhanced clinical response (CDAI-100) in the overall population (however, not in the subpopulation results) and over vedolizumab 300 mg every eight weeks for clinical remission (CDAI < 150) in the overall population (however, not in the subpopulation results). Given the uncertainty in the treatment-sequence analysis methodology and heterogeneity across studies, the comparative efficacy of these drugs in the maintenance phase of treatment is highly uncertain.

Mocko et al.[14,15] reported that there were no statistically significant differences in the incidence of AEs, serious AEs, discontinuations due to AEs, or for some of the more prominent AEs (e.g., infections, injections site reactions, nausea, headache, arthralgia, etc.) among adalimumab, ustekinumab, or vedolizumab during induction therapy and among adalimumab, infliximab, and vedolizumab during maintenance therapy in patients with CD. However, some of the major limitations associated with this IDC (e.g., lack of reporting of individual trial and patient characteristics for included studies, inclusion of patients with CDAI > 150 [with no further partitioning], lack of subgroup analyses, the bias assessment of the individual studies, lack of full safety profile [including only AEs with a frequency of 3% or greater], etc.) introduce uncertainty regarding the NMA results, decrease confidence in the results, and decrease the generalizability of the results to those patients with moderate-to-severe CD. Hence, caution is required when interpreting the authors' observations that there are no differences in safety among these drugs during the induction and maintenance phases of therapy for patients with CD.

The authors of the Singh et al.[13] IDC reported a statistically significant difference in favour of infliximab when compared with the other biologic drugs they assessed (adalimumab, ustekinumab, and vedolizumab) for the induction of remission in biologic-naive patients with moderate-to-severe CD. The results of the IDC indicated that there were no statistically significant differences among these drugs for maintaining remission in biologic-naive patients with moderate-to-severe CD. However, the limitations associated with this IDC (e.g., less stringent assessment of included study biases, the exclusion of studies in which results were not separated for the biologic-naive and biologic-experienced patients [thereby potentially losing applicable evidence], the enrichment of patients more likely to respond to biologic therapy in the maintenance phase, differences in the prognostic factors associated with the baseline patient characteristics of the individual RCTs, etc.) decrease confidence in the NMA results, decrease the generalizability of the results to those patients who are biologic-naive, and increase uncertainty. Hence, one must use caution when interpreting and using these results.

REFERENCES

1. The impact of inflammatory bowel disease in Canada: 2012 final report and recommendations [Internet]. Toronto: Crohn's and Colitis Foundation of Canada; 2012. [cited 2016 Dec 8]. Available from: http://www.isupportibd.ca/pdf/ccfc-ibd-impact-report-2012.pdf
2. Fast facts: the impact of IBD in Canada 2012 [Internet]. Toronto (ON): Crohn's and Colitis Foundation of Canada; 2012. [cited 2016 Dec 8]. Available from: http://www.isupportibd.ca/pdf/ccfc.ca-impact-report-fast-facts.pdf
3. Crohn's disease. Bethesda (MD): National Digestive Diseases Information Clearinghouse; 2014.
4. Stelara® (ustekinumab): injection 45 mg/0.5 mL, 90 mg/1.0 mL; for injection 130 mg/26 mL(5 mg/mL) [product monograph] [Final]. Toronto (ON): Janssen Inc.; 2016 Dec 12.
5. Common Drug Review. CADTH Canadian Drug Expert Review Committee (CDEC) final recommendation: Ustekinumab (Stelara - Janssen Inc.). Indication: psoriatic arthritis [Internet]. Ottawa (ON): CADTH; 2014 Oct 20. [cited 2016 Dec 8]. Available from: https://www.cadth.ca/sites/default/files/cdr/complete/cdr_complete_SR0359_Stelara_Oct-22-14.pdf
6. Common Drug Review. CEDAC final recommendation and reasons for recommendation: Ustekinumab (Stelara™ - Janssen-Ortho Inc.). Indication: chronic moderate to severe plaque psoriasis [Internet]. Ottawa (ON): CADTH; 2009 Jun 17. [cited 2016 Dec 8]. Available from: https://www.cadth.ca/sites/default/files/cdr/complete/cdr_complete_Stelara_June-17-2009.pdf
7. Clinical study report CNTO1275CRD3001. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab induction therapy in subjects with moderately to severely active Crohn's disease who have failed or are intolerant to TNF antagonist therapy [**CONFIDENTIAL** internal manufacturer's report]. Raritan (NJ): Janssen Research & Development, LLC; 2015.
8. Clinical study report CNTO1275CRD3002. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab induction therapy in subjects with moderately to severely active Crohn's disease [**CONFIDENTIAL** internal manufacturer's report]. Raritan (NJ): Janssen Research & Development, LLC; 2015.
9. Clinical study report CNTO1275CRD3003. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab maintenance therapy in subjects with moderately to severely active Crohn's disease [**CONFIDENTIAL** internal manufacturer's report]. Raritan (NJ): Janssen Research & Development, LLC; 2015.
10. Clinical study report C0743T26. A phase 2b, multicenter, randomized, double-blind, placebo-controlled, parallel-group study to evaluate the efficacy and safety of ustekinumab therapy in subjects with moderately to severely active Crohn's disease previously treated with TNF antagonist therapy: CERTIFI (Crohn's Evaluation of Response to Ustekinumab anti-IL12/23 for Induction) [**CONFIDENTIAL** internal manufacturer's report]. Raritan (NJ): Janssen Research & Development; 2011.
11. Sandborn WJ, Gasink C, Gao LL, Blank MA, Johanns J, Guzzo C, et al. Ustekinumab induction and maintenance therapy in refractory Crohn's disease. N Engl J Med. 2012 Oct 18;367(16):1519-28.
12. Cornerstone Research Group. Stelara® for the treatment of moderate to severe Crohn's disease: network meta-analysis study report [**CONFIDENTIAL** internal manufacturer's report]. Burlington (ON): Cornerstone Research Group;

13. Singh S, Garg SK, Pardi DS, Wang Z, Murad MH, Loftus EV, Jr. Comparative efficacy of biologic therapy in biologic-naive patients with Crohn disease: a systematic review and network meta-analysis. Mayo Clin Proc. 2014 Dec;89(12):1621-35.
14. Mocko P, Kawalec P, Pilc A. Safety profile of biologic drugs in the therapy of Crohn disease: a systematic review and network meta-analysis. Pharmacol Rep. 2016 Aug 1;68(6):1237-43.
15. Mocko P, Kawalec P, Pilc A. Safety profile of biologic drugs in the therapy of Crohn disease: a systematic review and network meta-analysis [supplementary material]. Pharmacol Rep. 2016 Aug 1;68(6):1237-43.
16. Lichtenstein GR, Hanauer SB, Sandborn WJ, The Practice Parameters Committee of the American College of Gastroenterology. Management of Crohn's disease in adults. Am J Gastroenterol. 2009;104:465-83.
17. Vedolizumab 300mg powder for concentrate for solution for infusion (Entyvio©). Takeda UK Ltd [Internet]. Glasgow (GB): Scottish Medicines Consortium; 2015 Jun 5. [cited 2016 Dec 8]. Available from: https://www.scottishmedicines.org.uk/files/advice/vedolizumab_Entyvio_FINAL_June_2015_for_website.pdf
18. Inflectra™ (infliximab): powder for solution, sterile, lyophilized, 100 mg/vial [product monograph]. Incheon, South Korea: Celltrion Healthcare Ltd.; 2016 Jun 10.
19. Remicade© (infliximab): powder for solution, sterile, lyophilized, 100 mg/vial [product monograph]. Toronto: Janssen Inc.; 2016 Apr 26.
20. Humira© (adalimumab): 40 mg in 0.8 mL sterile solution (50 mg/mL) subcutaneous injection [product monograph]. St. Laurent (QC): AbbVie Corporation; 2016 Apr 6.
21. Common Drug Review. CEDAC final recommendation and reason for recommendation. Adalimumab resubmission #3 (Humira© - Abbott Laboratories Ltd.) [Internet]. Ottawa (ON): CADTH; 2007 Dec 19. [cited 2016 Dec 8]. Available from: https://www.cadth.ca/sites/default/files/cdr/complete/cdr_complete_Humira-Resubmission-Crohns_Dec-19-2007.pdf
22. Common Drug Review. Adalimumab (Humira© - Abbott Laboratories Limited). Indication - Crohn's disease: overview of CDR clinical and pharmacoeconomic reports [Internet]. Ottawa: CADTH; 2008. [cited 2016 Dec 8]. Available from: https://www.cadth.ca/sites/default/files/cdr/relatedinfo/cdr_trans_humira-resub3-crohns_overview_Mar-07-08_e.pdf
23. Common Drug Review. CADTH Canadian Drug Expert Committee (CDEC) final recommendation: Infliximab (Inflectra - Hospira Healthcare Corporation). Indication: Crohn disease and ulcerative colitis [Internet]. Ottawa (ON): CADTH; 2016 Oct 25. [cited 2016 Dec 8]. Available from: https://www.cadth.ca/sites/default/files/cdr/complete/SE0483_IBD_Inflectra-Oct-28-16.pdf
24. EntyvioT (vedolizumab) powder for concentrate for solution in infusion 300 mg/vial [product monograph]. Oakville (ON): Takeda Canada; 2016 Mar 22.
25. Feagan BG, Sandborn WJ, Gasink C, Jacobstein D, Lang Y, Friedman JR, et al. Ustekinumab as induction and maintenance therapy for Crohn's disease. N Engl J Med. 2016 Nov 17;375(20):1946-60.
26. CDR submission: Stelara (ustekinumab), solution for infusion 130 mg/26 mL, 90 mg/1.0 mL solution for injection. Company: Janssen Inc. [**CONFIDENTIAL** manufacturer's submission]. Toronto (ON): Janssen Inc.; 2016 Sep 12.

27. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics. 1975 Mar;31(1):103-15.

28. Revised protocol C0743T26. A phase 2b, multicenter, randomized, double-blind, placebo-controlled, parallel-group study to evaluate the efficacy and safety of ustekinumab therapy in subjects with moderately to severely active Crohn's disease previously treated in TNF antagonist therapy. Study name: CERTIFI (Crohn's Evaluation of Response to Ustekinumab anti-IL 12/23 for Induction) [**CONFIDENTIAL** internal manufacturer's report]. Malvern (PA): Centocor Inc.; 2009.

29. Clinical protocol - Amendment 3. Protocol CNTO1275CRD3003. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab maintenance therapy in subjects wtih moderately to severely active Crohn's disease [**CONFIDENTIAL** internal manufacturer's report]. Raritan (NJ): Janssen Research & Development; 2015.

30. Clinical protocol - Amendment 3. Protocol CNTO1275CRD3001. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab induction therapy in subjects wtih moderately to severely active Crohn's disease who have failed or are intolerant to TNF antagonist therapy [**CONFIDENTIAL** internal manufacturer's report]. [Raritan (NJ)]: Janssen Research & Development; 2013.

31. Clinical protocol - Amendment 3. Protocol CNTO1275CRD3002. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab induction therapy in subjects with moderately to severely active Crohn's disease [**CONFIDENTIAL** internal manufacturer's report]. Raritan (NJ): Janssen Research & Development; 2014.

32. Guyatt G, Mitchell A, Irvine EJ, Singer J, Williams N, Goodacre R, et al. A new measure of health status for clinical trials in inflammatory bowel disease. Gastroenterology. 1989 Mar;96(3):804-10.

33. 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 Apr;28(4):S23-S27.

34. Gregor JC, McDonald JW, Klar N, Wall R, Atkinson K, Lamba B, et al. An evaluation of utility measurement in Crohn's disease. Inflamm Bowel Dis. 1997;3(4):265-76.

35. Coteur G, Feagan B, Keininger DL, Kosinski M. Evaluation of the meaningfulness of health-related quality of life improvements as assessed by the SF-36 and the EQ-5D VAS in patients with active Crohn's disease. Aliment Pharmacol Ther. 2009 May 1;29(9):1032-41.

36. User's manual for the SF-36v2 health survey. 3rd ed. Lincoln (RI): Quality Metric Incorporated; 2011.

37. Tang K, Beaton DE, Boonen A, Gignac MA, Bombardier C. Measures of work disability and productivity: Rheumatoid Arthritis Specific Work Productivity Survey (WPS-RA), Workplace Activity Limitations Scale (WALS), Work Instability Scale for Rheumatoid Arthritis (RA-WIS), Work Limitations Questionnaire (WLQ), and Work Productivity and Activity Impairment Questionnaire (WPAI). Arthritis Care Res (Hoboken) [Internet]. 2011 Nov [cited 2016 Dec 5];63 Suppl 11:S337-S349. Available from: http://onlinelibrary.wiley.com/doi/10.1002/acr.20633/pdf

38. Lerner D, Amick BC, III, Lee JC, Rooney T, Rogers WH, Chang H, et al. Relationship of employee-reported work limitations to work productivity. Med Care. 2003 May;41(5):649-59.

39. Lerner D, Amick BC, III, Rogers WH, Malspeis S, Bungay K, Cynn D. The Work Limitations Questionnaire. Med Care. 2001 Jan;39(1):72-85.

40. Tamminga SJ, Verbeek JH, Frings-Dresen MH, De Boer AG. Measurement properties of the Work Limitations Questionnaire were sufficient among cancer survivors. Qual Life Res. 2014 Mar;23(2):515-25.

41. Verhoef JA, Miedema HS, Bramsen I, Roebroeck ME. Using the work limitations questionnaire in patients with a chronic condition in the Netherlands. J Occup Environ Med. 2012 Oct;54(10):1293-9.

42. Daperno M, D'Haens G, Van AG, Baert F, Bulois P, Maunoury V, et al. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. Gastrointest Endosc. 2004 Oct;60(4):505-12.

43. Colombel JF, Sandborn WJ, Rutgeerts P, Enns R, Hanauer SB, Panaccione R, et al. Adalimumab for maintenance of clinical response and remission in patients with Crohn's disease: the CHARM trial. Gastroenterology. 2007 Jan;132(1):52-65.

44. Rutgeerts P, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, Colombel JF, et al. Comparison of scheduled and episodic treatment strategies of infliximab in Crohn's disease. Gastroenterology. 2004 Feb;126(2):402-13.

45. Dignass A, Van AG, Lindsay JO, Lemann M, Soderholm J, Colombel JF, et al. The second European evidence-based consensus on the diagnosis and management of Crohn's disease: current management. J Crohns Colitis. 2010 Feb;4(1):28-62.

46. Committee for Medicinal Products for Human Use. Guideline on the development of new medicinal products for the treatment of Crohn's disease [Internet]. London (GB): European Medicines Agency; 2008 Jul 24. [cited 2016 Dec 21]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003265.pdf

47. Guidance for industry: adaptive design clinical trials for drugs and biologics: draft guidance [Internet]. Silver Spring (MD): Food and Drug Administration; 2010. [cited 2016 Dec 8]. Available from: http://www.fda.gov/downloads/Drugs/.../Guidances/ucm201790.pdf

48. Sipponen T, Nuutinen H, Turunen U, Farkkila M. Endoscopic evaluation of Crohn's disease activity: comparison of the CDEIS and the SES-CD. Inflamm Bowel Dis. 2010 Dec;16(12):2131-6.

49. Khanna R, Nelson SA, Feagan BG, D'Haens G, Sandborn WJ, Zou GY, et al. Endoscopic scoring indices for evaluation of disease activity in Crohn's disease. Cochrane Database Syst Rev. 2016 Aug 8;(8):CD010642.

50. Sadowski DC, Bernstein CN, Bitton A, Croitoru K, Fedorak RN, Griffiths A, et al. Canadian Association of Gastroenterology clinical practice guidelines: the use of tumour necrosis factor-alpha antagonist therapy in Crohn's disease. Can J Gastroenterol [Internet]. 2009 Mar [cited 2016 Dec 21];23(3):185-202. Available from: https://www.cag-acg.org/images/publications/cag_cpg_crohnsmar2009.pdf

51. Prescribing information: Cimzia (certolizumab pegol) injection, for subcutaneous use [Internet]. Smyrna (GA): UCB Pharma S.A.; 2016. [cited 2016 Dec 8]. Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2016/125160s241lbl.pdf

52. Sandborn WJ. Crohn's disease evaluation and treatment: clinical decision tool. Gastroenterology. 2014 Sep;147(3):702-5.

53. Committee for Medicinal Products for Human Use. European public assessment report: Entyvio [Internet]. London (GB): European Medicines Agency; 2014 Mar 20. [cited 2016 Dec 21]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002782/WC500168530.pdf

54. Summary of the risk management plan (RMP) for Entyvio (vedolizumab) [Internet]. London (GB): European Medicines Agency; 2014. [cited 2016 Dec 8]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Risk-management-plan_summary/human/002782/WC500163289.pdf

55. Abbreviated health technology assessment report of ustekinumab maintenance of efficacy in Crohn's disease through 2 years. Protocol CNTO1275CRD3003. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab maintenance therapy in subjects with moderately to severely active Crohn's disease [**CONFIDENTIAL** internal manufacturer's report]. Raritan (NJ): Janssen Research & Development, LLC; 2016.

56. Stelara (ustekinumab): EPAR product information [Internet]. London (GB): European Medicines Agency; 2016 Nov 25. [cited 2017 Jan 2]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000958/WC500058513.pdf

57. Best WR, Becktel JM, Singleton JW, Kern F Jr. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. Gastroenterology. 1976 Mar;70(3):439-44.

58. Yoshida EM. The Crohn's Disease Activity Index, its derivatives and the Inflammatory Bowel Disease Questionnaire: a review of instruments to assess Crohn's disease. Can J Gastroenterol. 1999 Jan;13(1):65-73.

59. Sandborn WJ, Feagan BG, Hanauer SB, Lochs H, Lofberg R, Modigliani R, et al. A review of activity indices and efficacy endpoints for clinical trials of medical therapy in adults with Crohn's disease. Gastroenterology. 2002 Feb;122(2):512-30.

60. Mahadevan U. Clinical trial design and endpoints in biologic therapy for Crohn's disease: interpretation of the results. In: Medscape [Internet]. New York (NY): WebMD; 2006 [cited 2016 Jul 11]. Available from: http://www.medscape.com/viewarticle/543353 Free registration required.

61. Best WR, Becktel JM, Singleton JW. Rederived values of the eight coefficients of the Crohn's Disease Activity Index (CDAI). Gastroenterology. 1979 Oct;77(4 Pt 2):843-6.

62. de Dombal FT, Softley A. IOIBD report no 1: Observer variation in calculating indices of severity and activity in Crohn's disease. International Organisation for the Study of Inflammatory Bowel Disease. Gut [Internet]. 1987 Apr [cited 2017 Jan 25];28(4):474-81. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1432811

63. Pallis AG, Mouzas IA, Vlachonikolis IG. The inflammatory bowel disease questionnaire: a review of its national validation studies. Inflamm Bowel Dis. 2004 May;10(3):261-9.

64. Irvine EJ, Feagan B, Rochon J, Archambault A, Fedorak RN, Groll A, et al. Quality of life: a valid and reliable measure of therapeutic efficacy in the treatment of inflammatory bowel disease. Canadian Crohn's Relapse Prevention Trial Study Group. Gastroenterology. 1994 Feb;106(2):287-96.

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. Inflamm Bowel Dis. 2005 Oct;11(10):909-18.

66. Beaton DE, Tang K, Gignac MA, Lacaille D, Badley EM, Anis AH, et al. Reliability, validity, and responsiveness of five at-work productivity measures in patients with rheumatoid arthritis or osteoarthritis. Arthritis Care Res (Hoboken) [Internet]. 2010 Jan 15 [cited 2016 Dec 5];62(1):28-37. Available from: http://onlinelibrary.wiley.com/doi/10.1002/acr.20011/pdf

67. Walker N, Michaud K, Wolfe F. Work limitations among working persons with rheumatoid arthritis: results, reliability, and validity of the work limitations questionnaire in 836 patients. J Rheumatol. 2005 Jun;32(6):1006-12.

68. A phase 3, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the safety and efficacy of ustekinumab maintenance therapy in subjects with moderately to severely active Crohn's disease. Protocol CNTO1275CRD3003; phase 3: Stelara® (ustekinumab) [**CONFIDENTIAL** internal manufacturer's report]. Raritan (NJ): Janssen Research & Development; 2016.

69. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making [Internet]. 2013 Jul [cited 2017 Mar 29];33(5):607-17. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3704203

70. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. Value Health. 2011 Jun;14(4):417-28.

71. Peyrin-Biroulet L, Panes J, Sandborn WJ, Vermeire S, Danese S, Feagan BG, et al. Defining disease severity in inflammatory bowel diseases: current and future directions. Clin Gastroenterol Hepatol. 2016 Mar;14(3):348-54.