Common Drug Review
Clinical Review Report

April 2016

Drug	adalimumab (Humira)
Indication	For the treatment of adult patients with moderately to severely active ulcerative colitis (UC) who have had an inadequate response to conventional therapy including corticosteroids, azathioprine and/or 6-mercaptopurine (6-MP) or who are intolerant to such therapies.
Listing request	As per indication
Dosage Form(s)	40 mg/0.8 mL solution for subcutaneous injection
NOC date	November 21, 2013
Manufacturer	AbbVie Corporation

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

TABLE OF CONTENTS

Tables

Figures

ABBREVIATIONS

6-MP	6-mercaptopurine
AE	adverse event
CI	confidence interval
CrI	credible interval
DB	double-blind
GI	gastrointestinal
HRQoL	health-related quality of life
IBD	inflammatory bowel disease
IBDQ	Inflammatory Bowel Disease Questionnaire
IDC	indirect comparison
ITT	intention-to-treat
ITT-E	intention-to-treat extended
IV	intravenous
LOCF	last observation carried forward
LSMD	least squares mean difference
MCID	minimal clinically important difference
MCS	Mental Component Summary
NICE	National Institute for Health and Care Excellence
NMA	network meta-analysis
NRI	non-responder imputation
OL	open-label
OR	odds ratio
PCS	Physical Component Summary
PGA	Physician Global Assessment
PP	per-protocol
RBS	rectal bleeding subscore
RCT	randomized controlled trial
RR	relative risk
SAE	serious adverse event
SC	subcutaneous
SD	standard deviation
SFS	stool frequency subscore
TB	tuberculosis
TNF	tumour necrosis factor
UC	ulcerative colitis
WDAE	withdrawal due to adverse event
WPAI	Work Productivity and Impairment questionnaire

EXECUTIVE SUMMARY

Introduction

Ulcerative colitis (UC) and Crohn disease are both forms of inflammatory bowel disease (IBD), although the two are considered distinct from each other. UC is found only in the colon; the inflammation leads to diarrhea, pain, and bloody stools. Patients also experience extra-intestinal signs and symptoms such as fatigue and weight loss. If left untreated, the inflammation may progress, leading to mucosal damage and potentially fatal complications such as perforation and sepsis. Chronic inflammation is a recognized risk factor for malignancy, and patients with UC are at increased risk of developing colon cancer.

According to the Crohn's and Colitis Foundation of Canada, there are approximately 233,000 Canadians living with IBD, 104,000 of whom have UC. More than 10,200 new cases of IBD are diagnosed every year (5,700 with Crohn disease and 4,500 with UC), an incidence of 0.7% with 20% to 30% of people with IBD diagnosed before the age of 20.[1] Several drug classes are used in the treatment of UC, including aminosalicylates, immunosuppressants (azathioprine, 6-mercaptopurine [6-MP], and cyclosporine), corticosteroids, and the tumour necrosis factor (TNF) alpha inhibitors. Most of the drugs are associated with adverse effects that can have either short- or long-term consequences.

At the Health Canada–approved regimen, adalimumab is administered as a subcutaneous (SC) injection, 160 mg in week 0, 80 mg in week 2, then 40 mg every other week thereafter as monotherapy or in combination with conventional therapies. Conventional therapies include aminosalicylates and/or corticosteroids. Azathioprine and 6-MP may also be continued during treatment with adalimumab. Adalimumab is a human monoclonal antibody to TNF, a factor that mediates inflammation; thus, adalimumab has anti-inflammatory effects. In addition to being indicated for UC, it is indicated for rheumatoid arthritis, polyarticular juvenile arthritis, psoriatic arthritis, ankylosing spondylitis, and psoriasis, as well as Crohn disease.[2]

Indication under review
For the treatment of adult patients with moderately to severely active ulcerative colitis (UC) who have had an inadequate response to conventional therapy including corticosteroids, azathioprine and/or 6-mercaptopurine (6-MP) or who are intolerant to such therapies.
Listing criteria requested by sponsor
Per indication

The objective of this report is to perform a systematic review of the beneficial and harmful effects of adalimumab via SC injection at recommended doses for the treatment of adult patients with moderately to severely active UC who have had an inadequate response to conventional therapy or who are intolerant to such therapies.

Results and interpretation

Included studies

Three double-blind (DB), randomized controlled trials (RCTs) that enrolled patients with moderate to severe UC met the inclusion criteria for this review. Two of the studies, ULTRA 1 and ULTRA 2, were phase 3 studies for worldwide regulatory agencies, while Suzuki et al. was a phase 3 study based entirely in Japan, which was designed to meet the requirements of the Japanese regulatory process. All studies

compared adalimumab with placebo. ULTRA 1 had an 8-week DB phase, and therefore focused on inducing remission in UC patients, with a primary outcome of proportion of patients achieving clinical remission at eight weeks. ULTRA 2 focused on inducing remission at eight weeks and maintaining remission at 52 weeks, and therefore had a 52-week DB phase and co-primary outcomes of clinical remission at eight weeks and clinical remission at 52 weeks. Both of these studies allowed for dose escalation for patients with inadequate response, from every-other-week dosing to weekly dosing, after the end of the eight-week induction period, beginning at week 12 of the study. Finally, Suzuki et al. compared adalimumab with placebo for induction and maintenance over 52 weeks. The investigators stated that all analyses were exploratory, thus no primary outcome was identified. The three studies were all multi-centre and manufacturer-sponsored.

The studies with 52-week DB phases, ULTRA 2 and Suzuki et al., both had high withdrawal rates (38% for adalimumab and 47% for placebo in ULTRA 2, and 33% for adalimumab versus 24% for placebo in Suzuki et al.). In ULTRA 2, the primary analysis employed "non-responder imputation" to account for early withdrawals, where patients who withdrew were counted as non-responders. Therefore, the numerically higher proportion of withdrawals in the placebo group in ULTRA 2 may have biased results in favour of adalimumab if these were counted as treatment failures. A key limitation of the Suzuki et al. study is that it was carried out entirely in a Japanese population, and thus may have limited generalizability to Canada. It was also a relatively small study, with analysis described as "exploratory" by the manufacturer.

Efficacy

The primary outcome of both ULTRA 1 and ULTRA 2 was remission, defined as Mayo score ≤ 2 with no subscore > 1. ULTRA 1 included two treatment regimens for adalimumab, as noted above, but the focus of all data reported in this review is the Health Canada–approved regimen of adalimumab 160 mg, 80 mg, then 40 mg, as described above. Both studies tested the superiority of adalimumab over placebo for this outcome. For induction at eight weeks, adalimumab was statistically significantly superior to placebo in ULTRA 1, with 19% versus 9% of patients achieving clinical remission (difference in proportions 9.2; 95% CI, [redacted]; $P = 0.031$), and in ULTRA 2, with 17% versus 9% achieving clinical remission (difference in proportions 7.1; 95% CI, 1.2 to 12.9; $P = 0.019$). For maintenance at 52 weeks in ULTRA 2, adalimumab was again statistically significantly superior to placebo for the primary end point, with 17% versus 9% achieving clinical remission (difference in proportions 8.8; 95% CI, 2.8 to 14.5; $P = 0.004$). In Suzuki et al., 10% of adalimumab patients and 11% of placebo patients achieved remission at week 8, with no statistical comparison reported. At week 52, there was a higher proportion of adalimumab patients than placebo patients with remission, and this difference was statistically significant (23% versus 7%, $P < 0.001$). Note that all analyses were considered exploratory in this study. Across studies, although the differences in proportion of patients achieving remission between adalimumab and placebo were small, the clinical expert consulted for this review suggested they would be clinically significant in a population whose next major management option is likely to be a life-changing colectomy.

As noted above, patients were permitted to dose escalate to 40 mg weekly from 40 mg every other week in both ULTRA 1 and ULTRA 2. In ULTRA 2, 27% of adalimumab patients and 34% of placebo patients dose escalated, and because this was a 52-week study, these patients were followed through to 52 weeks. Subgroup analysis of patients who underwent dose escalation in ULTRA 2 (adalimumab 40 mg weekly instead of every other week) suggests a clinical benefit (remission in 12% of adalimumab patients) for patients whose dose interval was shortened. There were 34% of placebo patients who

achieved clinical remission after dose escalation. The product monograph for adalimumab suggests that patients who have not responded by eight weeks be discontinued from the drug.

Clinical response was defined as a decrease of Mayo score ≥ 3 points, and a decrease ≥ 30% and rectal bleeding subscore (RBS) if the Mayo score of 0 or 1, or a decrease of RBS ≥ 1. In ULTRA 1, there was no statistically significant difference between adalimumab (55%) and placebo (45%) at week 8 for clinical response (difference in proportions 10.0; 95% CI, [redacted]; *P* = 0.107). In ULTRA 2, 50% of adalimumab patients and 35% of placebo patients achieved clinical response at week 8, and this difference was statistically significant (difference in proportions 15.6; 95% CI, [redacted]; *P* < 0.001). At week 52 of ULTRA 2, 30% of adalimumab patients and 18% of placebo patients had achieved a clinical response, and this difference was also statistically significant (difference in proportions 11.7; 95% CI, [redacted]; *P* = 0.002). Clinical responses in Suzuki et al. were similar to those seen in ULTRA 2.

Quality of life was assessed using the Inflammatory Bowel Disease Questionnaire (IBDQ) and the Short Form (36) Health Survey (SF-36). Mean change from baseline was reported for both quality of life instruments, as well as the proportion of "responders" on the IBDQ (i.e., patients who achieved a certain threshold for improvement). An IBDQ responder was defined by an increase of at least 16 points from baseline, considered to be the low end of the minimal clinically important difference (MCID) for this instrument, which ranges from 16 points to 32 points. There was no statistically significant difference in the proportion of IBDQ responders at week 8 between adalimumab (61%) and placebo (58%) in ULTRA 1 (difference in proportions 3.1; 95% CI,[redacted]; *P* = 0.614). In ULTRA 2, differences in proportions between adalimumab and placebo were reported at week 8 and week 52; however, according to the statistical hierarchy, testing should not have been performed for this outcome because a previous outcome in the hierarchy had failed to reach statistical significance. Therefore, all statistical analyses reported for IBDQ should be considered exploratory. There was a statistically significant difference for adalimumab (58%) over placebo (46%) in ULTRA 2 at week 8 (difference in proportions 12.2; 95% CI, [redacted]; *P* = 0.006) and for adalimumab (26%) over placebo (16%) at week 52 (difference in proportions 9.7; 95% CI, 2.6 to 16.9]; *P* = 0.007). In Suzuki et al., there was a statistically significant difference for adalimumab over placebo at 52 weeks (25% versus 13%, respectively) but not at eight weeks (42% versus 40%, respectively) in the proportion of IBDQ responders. All analyses in Suzuki et al. were considered exploratory. SF-36 was also assessed as an exploratory outcome in ULTRA 1 and ULTRA 2. With respect to SF-36, in ULTRA 1, there was a statistically significantly greater [redacted] ([redacted]) for adalimumab over placebo in SF-36 Physical Component Summary (PCS) scores at week 8 (least squares mean difference (LSMD) [redacted]), [redacted] SF-36 Mental Component Summary (MCS) at week 8 (LSMD [redacted]). In ULTRA 2 at week 8, there was [redacted] SF-36 PCS ([redacted]) or SF-36 MCS (LSMD [redacted]) between groups, [redacted] SF-36 PCS [redacted] (LSMD [redacted]), [redacted] SF-36 MCS (LSMD [redacted]). The statistically significant improvements are [redacted], as the MCID for SF-36 component summaries is between 2.5 and 5.

No colectomies occurred during the treatment phase of ULTRA 1 or ULTRA 2, and none were reported in Suzuki et al.

There are no studies that compare adalimumab directly with other TNF inhibitors in the treatment of UC, and there is a general lack of studies that compare any of the TNF inhibitors to each other. Five

publicly available indirect comparisons (IDCs) are summarized in Appendix 7. Given the numerous limitations and inconsistent findings between the IDCs, the most conservative conclusion based on the available indirect evidence is that there is no clear evidence of a difference between the biologics, including adalimumab, with respect to inducing and maintaining remission, response, and mucosal healing in moderately to severely active UC.

Results from a long-term, open-label (OL) extension study (ULTRA 3) suggest that remission was maintained over a four-year follow-up period, while there were two deaths and one serious infection with tuberculosis (TB).

Harms

The proportions of patients with an adverse event (AE) was similar between adalimumab and placebo in each of the three included studies. In ULTRA 2, a similar proportion of adalimumab and placebo patients experienced a serious adverse event (SAE), while in ULTRA 1, 4% of adalimumab and 8% of placebo patients experienced an SAE; in Suzuki et al. the proportions were 4% versus 7%, respectively. The most common AE and the most common SAE was UC.

Withdrawals due to adverse event (WDAEs) occurred in 5% of patients in both adalimumab and placebo in ULTRA 1, 9% of adalimumab versus 13% of placebo in ULTRA 2, and 7% of adalimumab and 4% of placebo in Suzuki et al. The most common reason was UC.

Injection site reactions and infectious AEs were the most common notable harms. Injection site reactions appeared to occur in numerically more adalimumab than placebo patients across the studies; however, these studies were not powered to assess these outcomes. Infectious AEs occurred in 14% of adalimumab patients versus 16% of placebo patients in ULTRA 1, 45% versus 40% of patients, respectively, in ULTRA 2, and 19% versus 16% of patients, respectively, in Suzuki et al. Across the studies, there were three malignancies in adalimumab patients and two in placebo patients, six versus two patients with hypersensitivity reactions, and seven versus three cases of opportunistic infections excluding TB. There was one death due to TB in the adalimumab group in Suzuki et al.

Potential place in therapy[1]

The clinical expert consulted by the CADTH Common Drug Review (CDR) confirmed that there is an unmet need in patients who are acutely ill and fail to respond to corticosteroids within 72 hours, and who are not candidates for infliximab. There is also an unmet need in chronically ill patients who have failed treatment with anti-inflammatory drugs and immunosuppressants and who are intolerant of infliximab with infusion reactions, fail infliximab, do not have intravenous (IV) access, and have limited or no access to an infusion centre. The clinical expert described the patient populations in ULTRA 1 and ULTRA 2 as being consistent with those who would be targeted for adalimumab in practice. Adalimumab, therefore, may meet the needs outlined; in particular its SC administration avoids the need for IV therapy.

There remains uncertainty as to the use of adalimumab in patients who have failed or are intolerant to other biologics, and the lack of head-to-head studies versus other biologics remains a limitation in understanding the appropriate positioning of adalimumab amongst treatment options. According to the

[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.

clinical expert, infliximab is likely still the first-line therapy for acutely ill, hospitalized patients. The decision on which drug to use in ambulatory settings will likely be determined case by case.

Adalimumab may be discontinued in a certain proportion of patients who achieve remission. Generally, according to the clinical expert, this is often due to patient preference. The decision to discontinue anti-TNF therapy varies according to the severity of the index flare and discontinuation criteria are typically based on Mayo score and endoscopic score of 0.

Conclusions

Three DB RCTs met the inclusion criteria for this review, all of which compared adalimumab with placebo. ULTRA 1 was an induction study, with an eight-week, DB phase followed by an OL phase that continued out to 52 weeks. ULTRA 2 and Suzuki et al. were induction plus maintenance treatment studies, with 52-week, DB phases. Clinical remission was the primary outcome in ULTRA 1 and ULTRA 2, for which adalimumab was shown to be superior to placebo in each study. The authors of Suzuki et al. described analyses in that study as being exploratory; however, although remission was not statistically significantly improved for adalimumab over placebo at eight weeks, it was statistically significantly improved at 52 weeks. Clinical responses were not statistically significantly improved for adalimumab over placebo at eight weeks in ULTRA 1, but were statistically significantly improved for adalimumab over placebo at eight weeks and 52 weeks in ULTRA 2. Clinical responses were similar between ULTRA 2 and Suzuki et al. Quality of life was not consistently improved for adalimumab over placebo on the IBDQ and SF-36 at eight weeks, but statistically significant differences were observed for adalimumab compared to placebo on the IBDQ and SF-36 PCS at 52 weeks; however, these analyses were deemed exploratory by CDR and should be considered hypothesis-generating. There were no obvious differences in overall harms between adalimumab and placebo.

TABLE 1: SUMMARY OF RESULTS — INDUCTION STUDY (8-WEEK DOUBLE-BLIND PHASE)

Outcome	ULTRA 1	
	Adalimumab 160 mg/80 mg/40 mg N = 130	Placebo N = 130
Remission		
ITT, NRI week 8, N (%)	24 (19)	12 (9)
Difference in proportions [95% CI]	9.2 [redacted], *P* = 0.031	
Clinical Response		
Patients, 8 weeks, n (%)	71 (55)	58 (45)
Difference in proportions [95% CI]	10.0 [redacted], *P* = 0.107	
IBDQ		
Responders, week 8, n (%)	79 (61)	75 (58)
Difference in proportions [95% CI]	3.1 [redacted], *P* = 0.614[a,b]	
Mean (SD) at baseline	131.9 (NR) N = 120	125.3 (NR) N = 127
Mean (SD) change from baseline at week 8	35.9 (34.0) N = 117	26.9 (35.6) N = 124
LSMD [95% CI]	11.3 [redacted], *P* = 0.008	
SF-36 PCS		
Mean (SD) baseline	42.1 (NR) N = 125	40.2 (NR) N = 124
Mean (SD) change from baseline, week 8	[redacted]	[redacted]
LSMD [95% CI]	[redacted]	
SF-36 MCS		
Mean (SD) baseline	36.6 (NR) N = 125	36.5 (NR) N = 124
Mean (SD) change from baseline, week 8	[redacted]	[redacted]
LSMD [95% CI]	[redacted]	
Need for colectomy, n (%)		
Week 8, patients,	0 (0)	0 (0)
OL period through week 52	0 (0)	0 (0)
After last dose	4 (2)	8 (4)
AEs		
Patients with at least 1 AE, n (%)	112 (50)	108 (48)
SAEs		
Patients with at least 1 SAE, n (%)	9 (4)	17 (8)
WDAEs		
WDAEs, n (%)	12 (5)	12 (5)
Deaths		
Number of deaths, n (%)	0 (0)	0 (0)
Notable harms, n (%)		
Injection site reactions	13 (6)	7 (3)

Outcome	ULTRA 1	
	Adalimumab 160 mg/80 mg/40 mg N = 130	Placebo N = 130
Infectious AEs	32 (14)	35 (16)
Hypersensitivity reactions	2 (1)	1 (< 1)
Malignancies	0 (0)	2 (1)

AE = adverse event; CI = confidence interval; IBDQ = Inflammatory Bowel Disease Questionnaire; ITT = intention-to-treat; LOCF = last observation carried forward; LS = least squares; MCS = Mental Component Summary; MD = mean difference; NR = not reported; NRI = non-responder imputation; OL = open-label; PCS = Physical Component Summary; SAE = serious adverse event; SD = standard deviation; WDAE = withdrawal due to adverse event.

[a] According to the NRI analysis method, all missing response (or remission) values and values after dose escalation were imputed as non-response (or non-remission).

[b] Outcome was included in the hierarchical analysis plan, but testing should not have been performed for this outcome because a previous outcome in the hierarchy had failed to reach statistical significance. Therefore, all statistical test results should be considered exploratory for IBDQ responders.

Note: Remission: *P* values for adalimumab versus placebo from chi-square test (or Fisher's exact test if ≥ 20% of cells had expected cell count < 5).

Clinical Response, IBDQ: *P* value for differences between groups from chi-square test (or Fisher's exact test if ≥ 20% of the cell have an expected count < 5).

IBDQ response was defined as an increase in IBDQ ≥ 16 points from baseline.

IBDQ, mean change: *P* values for adalimumab versus placebo from one-way analysis of covariance (ANCOVA) with treatment as factor and baseline value as covariate.

SF-36: *P* values for adalimumab versus placebo from one-way ANCOVA with treatment as factor and baseline value as covariate.

Source: Clinical Study Report for ULTRA 1[3]

TABLE 2: SUMMARY OF RESULTS: INDUCTION/MAINTENANCE STUDIES (52-WEEK DOUBLE-BLIND PHASE)

Outcome	ULTRA 2		Suzuki et al. 2014	
	Adalimumab N = 248	Placebo N = 246	Adalimumab N = 90	Placebo N = 96
Remission				
Week 8, n (%)	41 (17)	23 (9)	NR (10)	NR (11)
Difference in proportions [95% CI]	7.1 [1.2 to 12.9], *P* = 0.019			
Week 52, n (%)	43 (17)	21 (9)	NR (23)	NR(7)
Difference in proportions [95% CI]	8.8 [2.8 to 14.5], *P* = 0.004		*P* = 0.001	
Clinical response				
Responders, NRI, week 8, n (%)	123 (50)	86 (35)	(50)	(35)
Difference in proportions [95% CI]	15.6 [redacted], *P* < 0.001		*P* = 0.044	
Responders, week 52, NRI, n (%)	75 (30)	47 (19)	(31)	(18)
Difference in proportions [95% CI]	11.7 [redacted], *P* = 0.002		*P* = 0.021	
IBDQ				
Responders, week 8, n (%)	144 (58)	112 (46)	NE	NE
Difference in proportions [95% CI]	12.2 [redacted], *P* = 0.006[a]		NE	NE
Responders, week 52, n (%)	65 (26)	40 (16)	NE	NE
Difference in proportions [95% CI]	9.7 [redacted], *P* = 0.007[a]		NE	NE
Mean (SD) baseline	128.0 (NR) N = 224	124.1 (NR) N = 224	NE	NE
Mean (SD) change from baseline, week 8	28.7 (35.6) N = 224	19.7 (36.0) N = 224	NE	NE
LSMD [95% CI]	[redacted]		NE	NE
Mean (SD) change from baseline, week 52	[redacted]	[redacted]	NE	NE
LSMD [95% CI]	[redacted]		NE	NE
SF-36 PCS				
Mean (SD) baseline	[redacted]	[redacted]	NE	NE
Mean (SD) change from baseline, week 8	[redacted]	[redacted]	NE	NE
LSMD [95% CI]	[redacted]		NE	NE
Mean (SD) change from baseline, week 52	[redacted]	[redacted]	NE	NE
LSMD [95% CI]	[redacted]		NE	NE
SF-36 MCS				
Mean (SD) baseline	[redacted]	[redacted]	NE	NE
Mean (SD) change from baseline, week 8	[redacted]	[redacted]	NE	NE
LSMD [95% CI]	[redacted]		NE	NE
Mean (SD) change from baseline, week 52	[redacted]	[redacted]	NE	NE
LSMD [95% CI]	[redacted]		NE	NE
Need for colectomy, n (%)				
Week 52, patients	0 (0)	0 (0)	NR	NR

Outcome	ULTRA 2		Suzuki et al. 2014	
	Adalimumab N = 248	Placebo N = 246	Adalimumab N = 90	Placebo N = 96
Follow-up phase	10 (4)	12 (5)	NR	NR
AEs				
Patients with at least 1 AE, n (%)	213 (83)	218 (84)	40 (44)	45 (47)
SAEs				
Patients with at least 1 SAE, n (%)	31 (12)	32 (12)	4 (4)	7 (7)
WDAEs				
WDAEs, n (%)	23 (9)	34 (13)	6 (7)	4 (4)
Deaths				
Number of deaths, n (%)	0 (0)	0 (0)	1	0 (0)
Notable harms, n **(%)**				
Injection site reactions	31 (12)	10 (4)	7 (8)	2 (2)
Hypersensitivity reactions	4 (2)	1 (< 1)	0 (0)	0 (0)
Malignancies	2 (1)	0 (0)	1 (1)	0 (0)
Infectious AEs	116 (45)	103 (40)	17 (19)	15 (16)

AE = adverse event; CMH = Cochran-Mantel-Haenszel; IBDQ = Inflammatory Bowel Disease Questionnaire; LS = least squares; MCS = Mental Component Summary; MD = mean difference; NE = not evaluated; NR = not reported; NRI = non-responder imputation; PCS = Physical Component Summary; SAE = serious adverse event; SD = standard deviation; TNF = tumour necrosis factor; WDAE = withdrawal due to adverse event.

[a] Outcome was included in the hierarchical analysis plan, but testing should not have been performed for this outcome because a previous outcome in the hierarchy failed to reach statistical significance. Therefore, all statistical test results should be considered exploratory for IBDQ responders.

Note: Clinical remission: *P* value to compare treatment groups was based on CMH test (stratification levels: prior anti-TNF versus anti-TNF-naive).

Clinical Response, IBDQ: *P* value for differences between groups from chi-square test (or Fisher's exact test if ≥ 20% of the cell have an expected count < 5).

IBDQ response was defined as an increase in IBDQ ≥ 16 points from baseline.

IBDQ, mean change: *P* values for adalimumab versus placebo from one-way analysis of covariance (ANCOVA) with treatment as factor and baseline value as covariate.

SF-36: *P* values for adalimumab versus placebo from one-way ANCOVA with treatment as factor and baseline value as covariate.

Source: Clinical Study Report for ULTRA 2;[4] Suzuki et al.[5]

1. INTRODUCTION

1.1 Disease prevalence and incidence

Ulcerative colitis (UC) and Crohn disease are both forms of inflammatory bowel disease (IBD), although the two are considered distinct from each other. UC is found in the colon, and the inflammation leads to diarrhea, pain, and bloody stools. Patients also experience extra-intestinal signs and symptoms such as fatigue and weight loss. If left untreated, inflammation may progress, leading to mucosal damage and potentially fatal complications such as perforation and sepsis. Chronic inflammation is a recognized risk factor for malignancy, and patients with UC are at increased risk of developing colon cancer.
According to the Crohn's and Colitis Foundation of Canada, there are approximately 233,000 Canadians living with IBD, 104,000 of whom have UC. More than 10,200 new cases of IBD are diagnosed every year (5,700 with Crohn disease and 4,500 with UC), an incidence of 0.7%; with 20% to 30% of people with IBD diagnosed before the age of 20. Canada has one of the highest incidence rates and prevalence rates of IBD in the world. There are 5,900 children in Canada living with IBD.[1]

1.2 Standards of therapy

Several drug classes are used in the treatment of UC, including aminosalicylates, immunosuppressants (azathioprine, cyclosporine), corticosteroids — which are all commonly referred to as "conventional therapies" — and tumour necrosis factor (TNF) alpha inhibitors, which are all monoclonal antibodies. Probiotics are also increasingly being recognized as useful drugs in UC management. Non-pharmacological measures include dietary and lifestyle changes, and surgery, which is the ultimate outcome in a number of patients. Most of the drugs are associated with adverse effects that can have either short- or long-term consequences. TNF inhibitors are also known for their considerable cost.

1.3 Drug

According to the Health Canada–approved regimen, adalimumab is administered as a subcutaneous (SC) injection, 160 mg in week 0, 80 mg in week 2, then 40 mg every other week thereafter as monotherapy or in combination with conventional therapies. Adalimumab is a human monoclonal antibody to TNF, a factor that mediates inflammation; thus, adalimumab has anti-inflammatory effects. In addition to being indicated for UC, adalimumab is also indicated for rheumatoid arthritis, polyarticular juvenile arthritis, psoriatic arthritis, ankylosing spondylitis, and psoriasis, as well as Crohn disease.[2]

Indication under review
For the treatment of adult patients with moderately to severely active ulcerative colitis (UC) who have had an inadequate response to conventional therapy including corticosteroids, azathioprine and/or 6-mercaptopurine (6-MP) or who are intolerant to such therapies.
Listing criteria requested by sponsor
Per indication

TABLE 3: KEY CHARACTERISTICS OF ADALIMUMAB, GOLIMUMAB, INFLIXIMAB, AND VEDOLIZUMAB

	Golimumab[6]	Infliximab[7]	Adalimumab[8]	Vedolizumab[9]
Mechanism of Action	Monoclonal antibody (human) to TNF	Monoclonal antibody (chimeric) to TNF	Monoclonal antibody (human) to TNF	Monoclonal antibody to alpha 4 beta 7 integrin
Indication[a]	Treatment of adult patients with moderate to severe UC. Patients with medical contraindications for, inadequate response to conventional therapies.	Treatment of adult patients with moderately to severely active UC who have had an inadequate response to conventional therapy.	Treatment of adult patients with moderately to severely active UC who have had an inadequate response to conventional therapy including corticosteroids, azathioprine and/or 6-MP or who are intolerant to such therapies.	Treatment of adult patients with moderately to severely active UC who have had an inadequate response, loss of response to, or were intolerant to either conventional therapy or infliximab, a TNF alpha antagonist.
Route of Administration	SC	IV	SC	IV
Recommended Dose	200 mg initially administered by SC injection at week 0, followed by 100 mg at week 2 and then 50 mg every 4 weeks, thereafter. The maintenance dose of 100 mg every 4 weeks can be considered at the discretion of the treating physician.	5 mg/kg given as an induction regimen at weeks 0, 2, and 6, followed by 5 mg/kg every 8 weeks. In some adult patients, consideration may be given to adjusting the dose up to 10 mg/kg to sustain clinical response and remission.	160 mg SC at week 0, 80 mg SC at week 2, then 40 mg SC every other week thereafter as monotherapy or in combination with conventional therapies. Adalimumab should only be continued in patients who have responded during the first 8 weeks of therapy.	300 mg administered by IV infusion at initiation, at weeks 2 and 6, and then every 8 weeks.
Serious Side Effects/Safety Issues	Infections, particularly opportunistic infections such as TB; malignancy, particularly lymphoma.	Infections, particularly TB; malignancy; allergic reactions.	Malignancy, particularly lymphoma; infections, particularly opportunistic infections such as TB.	Serious infections; infusion reactions; serious allergic reactions.

6-MP = 6-mercaptopurine; IV = intravenous; SC = subcutaneous; TB = tuberculosis; TNF = tumour necrosis factor; UC = ulcerative colitis.

[a] Health Canada indication.

2. OBJECTIVES AND METHODS

2.1 Objectives

To perform a systematic review of the beneficial and harmful effects of adalimumab (Humira) via SC injection at recommended doses for the treatment of adult patients with moderately to severely active UC who have had an inadequate response to conventional therapies or who are intolerant to such therapies.

2.2 Methods

Studies selected for inclusion in the systematic review included the pivotal studies provided in the manufacturer's submission to the CADTH Common Drug Review (CDR), as well as those meeting the selection criteria presented in Table 4.

TABLE 4: INCLUSION CRITERIA FOR THE SYSTEMATIC REVIEW

Patient Population	Adults (≥ 18 years of age) diagnosed with moderate to severe UC who have had an inadequate response to one or more conventional therapies[a] Subgroups: patients requiring dose escalation during study versus not requiring dose escalation; previous exposure to biologics versus no previous exposure to biologics; patients with fulminant disease or no fulminant disease; previous conventional therapies used	
Intervention	Adalimumab 160 mg SC at week 0, 80 mg SC at week 2, then 40 mg SC every other week thereafter as monotherapy or in combination with conventional therapies[a]	
Comparators	Conventional drugs[a] Golimumab Infliximab	Vedolizumab Placebo
Outcomes	**Key efficacy outcomes:** Clinical response (Mayo score reduction of ≥ 30%)[b] Clinical remission (Mayo score ≤ 2 with no individual subscore > 1)[b] HRQoL[b] Need for colectomy **Other efficacy outcomes:** Mucosal healing[c] Corticosteroid-free clinical remission Proportion of patients requiring dose escalation Physical function/disability[b] Days of missed work/school[b]	**Harms outcomes:** Mortality SAEs WDAEs AEs including but not limited to: • Injection site reactions • Hypersensitivity reactions • Malignancies • Infections (particularly TB and hepatitis) • Hepatotoxicity • Hematologic
Study Design	Published and unpublished RCTs	

AE = adverse event; DB = double-blind; HRQoL = health-related quality of life; RCT = randomized controlled trial; SAE = serious adverse event; SC = subcutaneous; TB = tuberculosis; UC = ulcerative colitis; WDAE = withdrawal due to adverse event.
[a] Conventional treatment: any combination of salicylates, corticosteroids (includes steroid-dependent disease with inability to taper steroids without relapse of symptoms), and immunosuppressants such as azathioprine, methotrexate, and cyclosporine.
[b] Outcomes identified as important in patient input summary.
[c] Could be determined by endoscopic or histologic investigation.

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 **Humira/adalimumab** and **colitis.**

No methodological filters were applied to limit retrieval. Where possible, retrieval was limited to the human population. Retrieval was not limited by publication year or by language. Conference abstracts were excluded from the search results. See Appendix 2: LITERATURE SEARCH STRATEGY for the detailed search strategies.

The initial search was completed on November 12, 2015. Regular alerts were established to update the search until the meeting of the Canadian Drug Expert Committee (CDEC) on March 16 2016. 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/resources/finding-evidence/grey-matters-practical-search-tool-evidence-based-medicine):

- Health Technology Assessment Agencies
- Health Economics
- Clinical Practice Guidelines
- Drug and Device Regulatory Approvals
- Advisories and Warnings
- Drug Class Reviews
- Databases (free)
- 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 5; excluded studies (with reasons) are presented in 0APPENDIX 3: EXCLUDED STUDIES.

3. RESULTS

3.1 Findings from the literature

A total of three studies were identified from the literature for inclusion in the systematic review (Figure 1). The included studies are summarized in Table 5 and Table **6** and described in section 3.2. A list of excluded studies is presented in APPENDIX 3: EXCLUDED STUDIES.

FIGURE 1: QUOROM FLOW DIAGRAM FOR INCLUSION AND EXCLUSION OF STUDIES

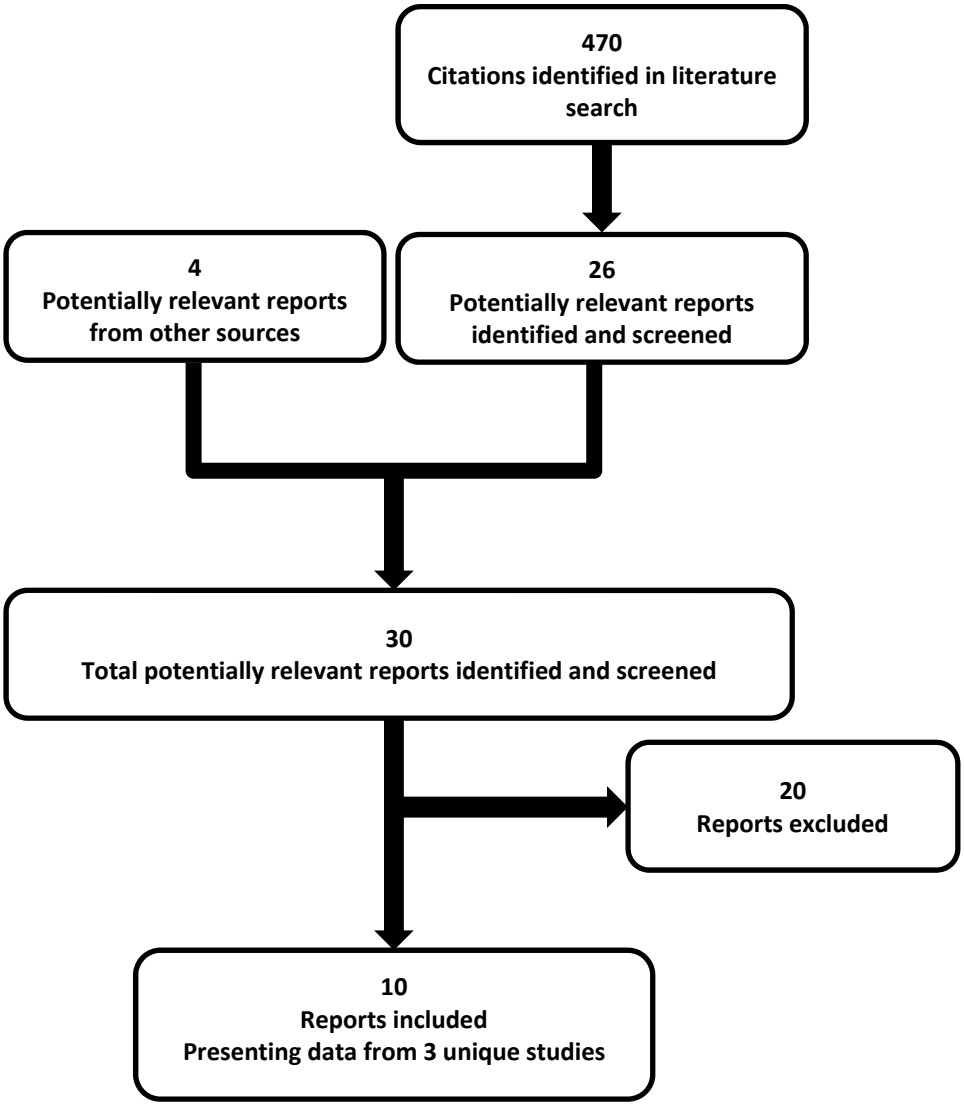


TABLE 5: DETAILS OF INCLUDED STUDIES: INDUCTION STUDY (8-WEEK DOUBLE-BLIND PHASE)

		ULTRA 1
DESIGNS & POPULATIONS	**Study Design**	DB, RCT
	Locations	80 centres: Canada, USA, Europe
	Study period	November 13, 2006 to March 5, 2010
	Randomized (N)	N = 576
	Inclusion Criteria	Male or female ≥ 18 years of age. Diagnosis of UC for > 90 days prior to baseline. Diagnosis of active UC confirmed by colonoscopy with biopsy or by flexible sigmoidoscopy with biopsy during the screening period, with exclusion of infection. Active UC with a Mayo score of 6 to 12 points and endoscopy subscore of 2 to 3 points, despite concurrent treatment with at least 1 of the following (oral corticosteroids or immunosuppressants or both as defined below): • Stable oral corticosteroid dose (prednisone dose of ≥ 20 mg/day or equivalent) for ≥ 14 days prior to baseline or stable oral corticosteroid dose (prednisone of < 20 mg/day) for ≥ 40 days prior to baseline. • At least a consecutive 90-day course of azathioprine or 6-MP prior to baseline, with a dose of azathioprine ≥ 1.5 mg/kg/day or 6-MP ≥ 1 mg/kg/day (rounded to the nearest available tablet formulation), or the highest dose tolerated by the patient (e.g., due to leukopenia, elevated liver enzymes, nausea) during that time. Patient was to be on a stable dose for at least 28 days prior to baseline. Concurrent therapy was not required for patients who had been previously treated with corticosteroids or immunosuppressants (azathioprine or 6-MP) during the previous 5 years and, in the judgment of the investigator, had failed to respond to, or could not tolerate, their treatment.
	Exclusion Criteria	History of subtotal colectomy with ileorectostomy or colectomy with ileoanal pouch, Koch pouch, or ileostomy for UC, or is planning bowel surgery. Received infliximab or any other anti-TNF drug or any biological therapy in the past.
DRUGS	**Intervention**	Adalimumab 160 mg SC at baseline, then: Week 2: 80 mg Weeks 4, 6, and 8: 40 mg
	Comparator(s)	Placebo
DURATION	Screening	21 days
	DB	8 weeks (induction treatment)
	Follow-up	70 days or extension
OUTCOMES	**Primary End Point**	Remission rate, defined as the proportion of patients with a total Mayo score ≤ 2 and no individual subscore > 1.
	Other End Points	• Clinical response per Mayo score at week 8. • Pts with mucosal healing at week 8. • RBS indicative of mild disease (≤ 1) at week 8. • PGA subscore of mild disease (≤ 1) at week 8. • Pts with SFS indicative of mild disease (≤ 1) at week 8. • Proportion of IBDQ responders at week 8. Each of above tested in a hierarchical order to account for multiplicity and for the comparison of adalimumab 160mg/80mg/40 mg versus placebo and adalimumab 80 mg/40 mg versus placebo)

		ULTRA 1
NOTES	**Publications**	Reinisch 2011,[10] Reinisch 2013[11]

6-MP = 6-mercaptopurine; DB = double-blind; IBDQ = Inflammatory Bowel Disease Questionnaire; OCS = oral corticosteroids; OL = open-label; PGA = Physician Global Assessment; pts = patients; RBS = rectal bleeding subscore; RCT = randomized controlled trial; SC = subcutaneous; SFS = stool frequency subscore; UC = ulcerative colitis.
Note: Two additional reports were included (manufacturer's submission;[12] Health Canada Reviewers Report[13]).
Source: Clinical Study Report for ULTRA 1.[3]

TABLE 6: DETAILS OF INCLUDED STUDIES: INDUCTION/MAINTENANCE STUDIES (52-WEEK DOUBLE-BLIND PHASE)

		ULTRA 2	Suzuki et al. 2014
DESIGNS & POPULATIONS	**Study Design**	DB, RCT	DB, RCT
	Locations	103 centres in Canada, US, EU, Australia, NZ, and Israel	65 centres in Japan
	Study period	November 20, 2006 to March 2, 2010	February 2009 to May 2011
	Randomized (N)	N = 518	N = 186

	Inclusion Criteria	• Male or female ≥ 18 years of age
		• Diagnosis of UC for > 90 days prior to baseline
		• Diagnosis of active UC confirmed by colonoscopy with biopsy or by flexible sigmoidoscopy with biopsy during the screening period, with exclusion of infection
		• Active UC with a Mayo score of 6 to 12 points and endoscopy subscore of 2 to 3 points, despite concurrent treatment with ≥ 1 of the following (OCS or immunosuppressants or both as defined below):
		  ○ Stable OCS dose (prednisone dose of ≥ 20 mg/day or equivalent) for ≥ 14 days prior to baseline or stable OCS dose (prednisone of < 20 mg/day) for ≥ 40 days prior to baseline
		  ○ At least a consecutive 90-day course of azathioprine or 6-MP prior to baseline, with a dose of azathioprine ≥ 1.5 mg/kg/day or 6-MP ≥ 1 mg/kg/day (rounded to the nearest available tablet formulation), or the highest dose tolerated by the patient (e.g., due to leukopenia, elevated liver enzymes, nausea) during that time. Patient was to be on a stable
			• ≥ 15 years of age with biopsy-confirmed, moderately to severely active UC (Mayo score 6 to 12 points and an endoscopy subscore of ≥ 2) despite concurrent treatment with stable doses of oral corticosteroids (prednisolone equivalent of ≥ 20 mg/day for ≥ 2 weeks or 5 to < 20 mg/day for ≥ 40 days before baseline) and/or immunomodulators
			• Patients who had been previously treated with corticosteroids or immunomodulators during the past 5 years and who, in the judgment of the investigator, had failed to respond to, or who could not tolerate, their treatment

		ULTRA 2	Suzuki et al. 2014
		dose for at least 28 days prior to baseline • Concurrent therapy was not required for patients who had been previously treated with corticosteroids or immunosuppressants (azathioprine or 6-MP) during the previous 5 years and, in the judgment of the investigator, had failed to respond to, or could not tolerate, their treatment.	
	Exclusion Criteria	• History of subtotal colectomy with ileorectostomy or colectomy with ileoanal pouch, Koch pouch, or ileostomy for UC, or planned bowel surgery • Previously used infliximab or any anti-TNF drug within 56 days of baseline • Previously used infliximab or any anti-TNF drug without clinical response at any time ("primary non-responder"), unless patient experienced a treatment-limiting reaction	• Receipt of prior treatment with anti-TNF therapies or other biologic drugs • Discontinuation of OCS within 2 weeks before baseline • Receipt of corticosteroid injection, cyclosporine, tacrolimus, or mycophenolate mofetil within 4 weeks before baseline • Receipt of therapeutic enema or suppository, other than required for endoscopy, within 2 weeks before screening endoscopy • Receipt of cytapheresis within 56 days of baseline • Receipt of total parenteral nutrition during the screening period • Receipt of any investigational drug within 4 weeks or 5 half-lives before baseline • Nonstable doses or recent discontinuation (within 4 weeks) of oral aminosalicylates or UC-related antibiotics
DRUGS	**Intervention**	Adalimumab 160 mg SC at baseline (week 0), 80 mg at week 2, and 40 mg EOW starting at week 4.	Adalimumab 160 mg at week 0, 80 mg at week 2, and then 40 mg EOW beginning at week 4 (160/80 mg group or Adalimumab 80 mg at week 0, 40 mg at week 2, and then 40 mg EOW beginning at week 4 (80/40 mg group)
	Comparator(s)	Placebo at weeks 0, 2, and EOW starting at week 4	Placebo

		ULTRA 2	Suzuki et al. 2014
DURATION	Screening	Not reported	Not reported
	DB	52 weeks (8 weeks induction; 44 weeks maintenance). At or after week 10, patients who met the criteria for inadequate response could have been switched to OL adalimumab 40 mg EOW beginning at week 12.	52 weeks (8 weeks induction; 44 weeks maintenance)
	Follow-up	OL extension for those completing 52 weeks	Not reported
OUTCOMES	**Primary End Point**	• Proportion of patients who achieved remission at week 8 • Proportion of patients who achieved remission at week 52	All statistical analyses were exploratory.
	Other End Points	• Pts with remission (sustained), at both weeks 8 and 52 • Pts who achieved clinical response per Mayo score at week 8 • Pts with clinical response per Mayo score at week 52 • Pts with clinical response per Mayo score (sustained) at both weeks 8 and 52 • Pts who achieved mucosal healing at week 8 • Pts who achieved mucosal healing at week 52 • Pts who achieved mucosal healing (sustained) at both weeks 8 and 52 • Pts who discontinued CS use before week 52 and achieved remission at week 52 • Pts with PGA subscore indicative of mild disease (≤ 1) at week 8 • Pts with SFS indicative of mild disease (≤ 1) at week 8 • Pts with RBS indicative of mild disease (≤ 1) at week 8 • Pts who discontinued CS use for at least 90 days before week 52 and achieved remission at week 52 • Pts who discontinued CS use and achieved remission (sustained) at both weeks 32 and 52 • Pts who were IBDQ responders at week 52 • Pts who were IBDQ responders at week 8	• Clinical response • Remission • Mucosal healing • Other efficacy analyses at weeks 8, 32, and 52 included • RBS, PGA, and SFS indicative of mild disease (score B1) and IBDQ response (≥16-point increase from baseline)

		ULTRA 2	Suzuki et al. 2014
NOTES	**Publications**	Sandborn 2012;[14] Sandborn 2013;[15] Wolf 2014[16]	Suzuki et al. 2014[5]

6-MP = 6-mercaptopurine; CS = corticosteroid; DB = double-blind; EOW = every other week; EU = European Union; IBDQ = Inflammatory Bowel Disease Questionnaire; NZ = New Zealand; OCS = oral corticosteroids; OL = open-label; PGA = Physician Global Assessment; pts = patients; RB = rectal bleeding score; RCT = randomized controlled trial; SC = subcutaneous; SFS = stool frequency subscore; TNF = tumour necrosis factor; UC = ulcerative colitis.
Note: Two additional reports were included (manufacturer's submission;[12] Health Canada Reviewers Report[13]).
Source: Clinical Study Report for ULTRA 2;[4] Suzuki et al. 2014.[5]

3.2 Included studies

3.2.1 Description of studies

Three DB randomized controlled trials (RCTs) met the inclusion criteria for this review. ULTRA 1 and ULTRA 2 were phase 3 studies used for registration in Canada, while Suzuki et al. was a published phase 3 study that was required for the Japanese regulatory process.

ULTRA 1 consisted of an 8-week randomized, DB, placebo-controlled period (DB period) followed by an open-label (OL) period. No randomization stratification variables were described. The primary efficacy analysis was conducted on the data set from the DB period through week 8. Patients enrolled in the study under the original protocol or Amendments 1 and 2 were randomized in a 1:1 ratio to receive 160 mg of adalimumab or placebo at baseline, 80 mg adalimumab or placebo at week 2, and 40 mg adalimumab or placebo at weeks 4 and 6. At week 8, patients randomized to placebo received 160 mg adalimumab followed by 80 mg adalimumab at week 10. Patients randomized to adalimumab continued to receive 40 mg adalimumab at weeks 8 and 10. All patients continued to receive one injection of OL adalimumab 40 mg every other week beginning at week 12 up to week 52 (or the end-of-therapy visit).

In August 2007, the ULTRA 1 study design was amended to incorporate an additional adalimumab induction dosing arm of 80 mg/40 mg. Earlier that year, both 160 mg/80 mg/40 mg and 80 mg/40 mg induction regimens had been approved in the EU as induction treatment for Crohn disease. The adalimumab induction dosing regimen of 80 mg/40 mg was therefore included so that both of these approved induction regimens would be evaluated for the induction of remission of UC.

ULTRA 2 and Suzuki et al. each had a total 52-week DB induction phase plus a maintenance treatment phase, in which adalimumab was compared with placebo. Patients were to be stratified by prior exposure to infliximab and/or other anti-TNF drugs, and randomized in a 1:1 ratio to receive adalimumab or placebo by SC injection. Stratification was not described in Suzuki et al. The primary analyses in ULTRA 2 were carried out at week 8 (induction) and at week 52 (maintenance); the co-primary outcome was clinical remission at each of these time points. The analyses in Suzuki et al. were described as exploratory by the manufacturer; thus, there was no primary outcome identified. ULTRA 1 and ULTRA 2 both assessed key efficacy outcomes in the CDR protocol, including clinical response and quality of life. Colectomies were also reported, although not as an efficacy outcome.

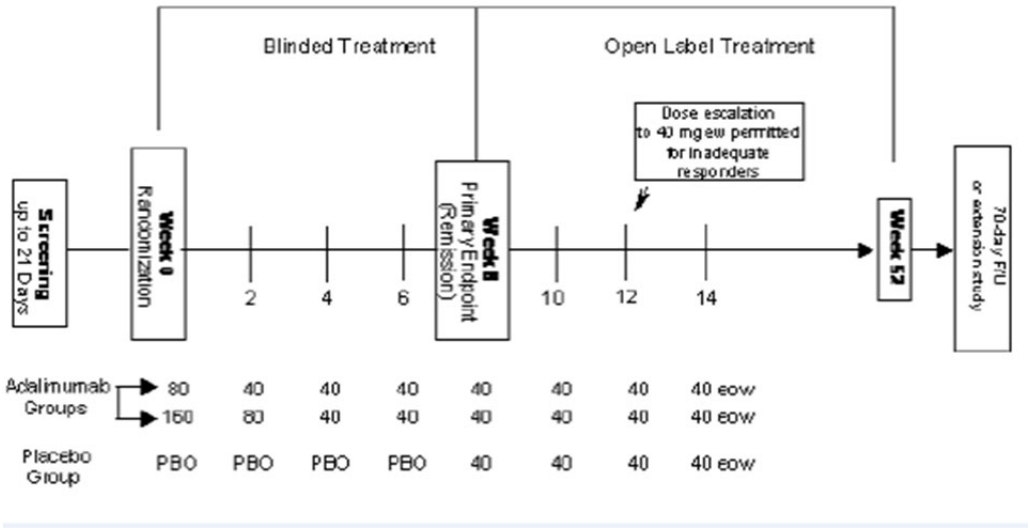
All of the included studies were manufacturer-sponsored and were multi-centre. Suzuki et al. was conducted entirely in Japan, while ULTRA 1 and ULTRA 2 were multinational, with sites in Canada.

TABLE 7: DETAILS OF PROTOCOL AMENDMENT IN ULTRA 1

Prior to Amendment 3	After Amendment 3
Two treatment arms: Placebo Adalimumab 160 mg/80 mg/40 mg	Three treatment arms: Placebo Adalimumab 80 mg/40 mg Adalimumab 160 mg/80 mg/40 mg
DB period lasting for 12 weeks.	DB period lasting for 8 weeks.
Stable (± 5 mg) corticosteroid dose (prednisone of ≥ 20 mg/day or equivalent) for at least 14 days prior to baseline or maintenance corticosteroid dose (prednisone of ≥ 10 mg/day and < 20 mg/day or equivalent) for at least 40 days prior to baseline.	Patients had to be stable on prednisone ≥ 20 mg/day or equivalent for at least 14 days prior to baseline. For doses of prednisone < 20 mg/day or equivalent, patients had to be stable for at least 40 days prior to baseline.
Prior and concurrent infliximab or anti-TNF excluded.	All prior and concurrent biologics excluded (including infliximab and anti-TNFs).
Immunosuppressants other than azathioprine or 6-MP (e.g., cyclosporine, methotrexate, or tacrolimus) prohibited within 60 days prior to baseline and during the study.	Cyclosporine, tacrolimus, mycophenolate mofetil, and investigational drugs prohibited for 30 days or 5 half-lives prior to baseline and during the study. IV corticosteroid use prohibited within 14 days prior to screening, during the screening period, and during the study.

6-MP = 6-mercaptopurine; DB = double-blind; IV = intravenous; TNF = tumour necrosis factor.

FIGURE 2: STUDY DESIGN SCHEMATIC FOLLOWING PROTOCOL AMENDMENT 3 IN ULTRA 1



3.2.2 Populations

a) Inclusion and exclusion criteria

All studies enrolled adults with active UC (Mayo score of 6 to 12, endoscopy subscore of 2 to 3), despite concurrent treatment with oral corticosteroids or immunosuppressants. Patients were not required to be on concurrent treatment if it was determined that they had had previous treatment failures with, or were unable to tolerate, these therapies (Table 5, Table 6).

b) Baseline characteristics

Patients across ULTRA 1 and ULTRA 2 were approximately 40 years of age at enrolment. The majority of patients were male (approximately 62%) and Caucasian (> 90%). A small number (< 10%) were current users of nicotine, and approximately 50% were current drinkers of ethanol. Across the two studies, patients had UC for about 8.5 years. Approximately 50% of patients across the two studies had pancolitis. In the Japanese study, the baseline characteristics were similar with respect to demographics, and patients had lived with UC for slightly less time (7.8 years). The major difference in reported data in the Japanese study was that a higher proportion of patients (approximately 65%) had pancolitis compared with patients in ULTRA 1 and ULTRA 2 (Table 8, Table 9).

With respect to differences between groups within studies, in ULTRA 1, the adalimumab group had more current nicotine users (9% versus 5%), more drinkers (48% versus 44%), and a longer duration of UC (8.1 years versus 7.5 years) than the placebo group. The largest difference between groups was in the proportion of patients with pancolitis, which was lower with adalimumab than placebo (46% versus 56%). In ULTRA 2, the groups were generally well-balanced with respect to baseline characteristics, although there was a shorter duration of UC with adalimumab versus placebo (8.1 years versus 8.5 years). Fewer baseline characteristics were reported in the Japanese study, but the largest difference between groups was the higher proportion of patients with pancolitis in the adalimumab group versus the placebo group (70% versus 62%).

TABLE 8: SUMMARY OF BASELINE CHARACTERISTICS: INDUCTION STUDY (8-WEEK DOUBLE-BLIND PHASE)

	ULTRA 1	
	Adalimumab 160 mg/80 mg/40 mg N = 130	Placebo N = 130
Age, mean (SD), years	38.2 (13.5)	38.9 (12.7)
Male gender, n (%)	83 (64)	82 (63)
Ethnicity, n (%)		
White	119 (92)	117 (90)
Black	2 (2)	5 (4)
Asian	7 (5)	5 (4)
Hispanic/Latino	4 (3)	5 (4)
Nicotine use, n (%)		
Current	12 (9)	7 (5)
Former	37 (29)	35 (27)
Non-user	81 (62)	88 (68)
Weight, mean (SD), kg	75.5 (14.2)	78.7 (17.4)
Alcohol use, n (%)		
Drinker	62 (48)	57 (44)
Former drinker	9 (7)	7 (5)
Non-drinker	59 (45)	66 (51)
Duration of UC, mean (SD), years	8.1 (7.2)	7.5 (7.2)
Site of UC, n (%)		
Pancolitis	60 (46)	73 (56)
Descending colon	61 (47)	42 (32)
Other	9 (7)	15 (12)

	ULTRA 1	
	Adalimumab 160 mg/80 mg/40 mg N = 130	Placebo N = 130
Baseline Mayo, mean (SD)	8.8 (1.6)	8.7(1.6)
Evidence of dysplasia/malignancy, n (%)		
Yes	0 (0)	0 (0)
No	130 (100)	130 (100)
Missing	0 (0)	0 (0)

SD = standard deviation; UC = ulcerative colitis.
Source: Clinical Study Report for ULTRA 1.[3]

TABLE 9: SUMMARY OF BASELINE CHARACTERISTICS: INDUCTION/MAINTENANCE STUDIES (52-WEEK DOUBLE-BLIND PHASE)

	ULTRA 2		Suzuki et al. 2014	
	Adalimumab N = 248	Placebo N = 246	Adalimumab N = 90	Placebo N = 96
Age, mean (SD), years	39.6 (12.5)	41.3 (13.2)	42.5 (14.6)	41.3 (13.6)
Male gender, n (%)	142 (57)	152 (62)	61 (68)	70 (73)
Ethnicity, n (%)				
White	236 (95)	234 (95)	NR	NR
Black	7 (3)	4 (2)	NR	NR
Asian	1 (< 1)	4 (2)	NR	NR
Other	4 (2)	4 (2)	NR	NR
Hispanic/Latino	6 (2)	7 (3)	NR	NR
Nicotine use, n (%)				
Current	20 (8)	19 (8)	NR	NR
Former	94 (38)	88 (36)	NR	NR
Non-user	134 (54)	138 (56)	50 (56)	55 (57)
Weight, mean (SD), kg	75.3 (17.7)	77.1 (17.3)	60.1 (12.3)	60.8 (14.1)
Alcohol use, n (%)				
Drinker	132 (53)	125 (51)	NR	NR
Former drinker	8 (3)	11 (5)	NR	NR
Non-drinker	108 (44)	109 (45)	43 (48)	36 (38)
Duration of UC, mean (SD), years	8.1 (7.1)	8.5 (7.4)	7.8 (7.1)	7.8 (6.6)
Site of UC, n (%)				
Pancolitis	120 (48)	120 (49)	63 (70)	59 (62)
Descending colon	96 (39)	96 (39)	27 (30)	35 (37)
Other	32 (13)	30 (12)	0 (0)	2 (2)
Baseline Mayo, mean (SD)	8.9 (1.5)	8.9 (1.8)	8.6 (1.4)	8.5 (1.6)
Evidence of dysplasia/malignancy, n (%)				
Yes	2 (1)	1 (< 1)	NR	NR
No	243 (99)	241 (100)	NR	NR

	ULTRA 2		Suzuki et al. 2014	
	Adalimumab N = 248	Placebo N = 246	Adalimumab N = 90	Placebo N = 96
Missing	3 (1.2)	4 (1.6)	NR	NR

NR = not reported; SD = standard deviation; UC = ulcerative colitis.
Source: Clinical Study Report for ULTRA 2;[4] Suzuki et al.[5]

3.2.3 Interventions

ULTRA 1 included two different dosing regimens of adalimumab as well as a placebo control. One group received adalimumab 160 mg at baseline, 80 mg at week 2, and then 40 mg at weeks 4, 6, and 8. Prior to the Protocol Amendment 3 (Table 7), there was an additional 40 mg dose administered at week 10. This eight-week regimen, 160 mg followed by 80 mg then 40 mg, is the approved dosing regimen in Canada. Another adalimumab group, which was added in Protocol Amendment 3, started with adalimumab 80 mg at baseline, then 40 mg at weeks 2, 4, 6, and 8 during the DB phase. This adalimumab 80 mg then 40 mg regimen was added after it was approved by the European Union (EU) (along with the 160 mg/80 mg/40 mg regimen). In ULTRA 1, after the end of the eight-week DB phase and the primary analysis, patients could dose escalate to 40 mg weekly at week 12 if they had an inadequate response. All placebo patients were switched to 40 mg every other week beginning with week 8, which was the end of the DB, primary analysis phase.

In ULTRA 2 and Suzuki et al., there were two groups, an adalimumab group and a placebo group. Adalimumab patients received 160 mg at baseline, 80 mg at week 2, followed by 40 mg every other week starting at week 4. In ULTRA 2, at or after week 10, patients who met the criteria for inadequate response could have been switched to OL adalimumab 40 mg every other week beginning at week 12. Patients who demonstrated inadequate response at two consecutive visits at least 14 days apart while on OL administration 40 mg every other week were permitted to dose escalate to adalimumab 40 mg weekly every week.

In Suzuki et al., patients with an inadequate response to the study drug (defined as a partial Mayo score greater than or equal to that of the baseline score on two consecutive visits at least 14 days apart for patients with a baseline partial Mayo score of 3 to 7, or partial Mayo score > 7 on two consecutive visits at least 14 days apart for patients with a baseline partial Mayo score of 8 or 9) or with a flare (defined as a partial Mayo score increase of ≥ 3 compared with the score at the last evaluation before the disease flare on two consecutive visits at least 14 days apart) at or after week 8 were entered into the ''rescue arm.'' Treatment in the ''rescue arm'' consisted of four weeks of blinded adalimumab (either 160 mg initially and 80 mg two weeks later for patients in the placebo arm, or 40 mg initially and two weeks later for patients in either adalimumab arm), followed by OL adalimumab 40 mg every other week, with a possibility to escalate to 80 mg every other week in case of inadequate response or disease flare at least eight weeks later.

In ULTRA 1 and ULTRA 2, patients on aminosalicylates, azathioprine, 6-MP, or probiotics who qualified for enrolment into the study were to continue their medication doses. Dose adjustments of UC-related concomitant treatments were not allowed, except when tapering an oral corticosteroid between week 8 and week 52 or in the event of UC–treatment-related toxicities (e.g., leukopenia or elevated liver enzymes) considered moderate to severe in the opinion of the investigator. Corticosteroids were permitted as outlined in the inclusion criteria. Patients were not allowed to adjust their corticosteroid dose during the first eight weeks of the study. At week 8 and thereafter, patients who, in the opinion of the investigator, had a satisfactory clinical response, were permitted to undergo corticosteroid tapering

according to the schedule described below, or per investigator discretion, as defined by local standards. For doses > 10 mg/day of prednisone (or equivalent), patients started with a weekly dose decrease of 5 mg until a 10 mg/day dose was reached. Thereafter, they decreased the dose weekly by 2.5 mg until discontinuation. If the patient experienced a loss of satisfactory clinical response, the patient could have his or her corticosteroid dose increased per the investigator's discretion, up to a maximum dose equivalent to the dose used at baseline. In Suzuki et al., patients were also continued on their UC therapies.

3.2.4 Outcomes

Remission was the primary outcome of both studies ULTRA 1 and ULTRA 2, defined as Mayo score ≤ 2 with no subscore > 1. Clinical response was defined as a decrease of Mayo score ≥ 3 points and a decrease ≥ 30%, and rectal bleeding subscore (RBS) of 0 or 1, or a decrease of RBS ≥ 1. The Mayo scoring system is summarized below.[12]

a) Summary of Mayo Scoring System

Stool frequency subscore*

0 = Normal number of stools for this patient
1 = 1 to 2 stools more than normal for this patient
2 = 3 to 4 stools more than normal for this patient
3 = 5 or more stools more than normal for this patient
*Each patient serves as his or her own control to establish normal stool frequency and the degree of abnormal stool frequency.

Rectal bleeding subscore**

0 = No blood seen
1 = Streaks of blood with stool less than half the time
2 = Obvious blood with stool most of the time
3 = Blood alone passed
** The daily bleeding score represents the most severe bleeding of the day.

Endoscopy subscore: Findings of flexible sigmoidoscopy

0 = Normal or inactive disease
1 = Mild disease (erythema, decreased vascular pattern, mild friability)
2 = Moderate disease (marked erythema, absent vascular pattern, friability, erosions)
3 = Severe disease (spontaneous bleeding, ulceration)

Physician's Global Assessment subscore***

0 = Normal (subscores are 0)
1 = Mild disease (subscores are mostly 1's)
2 = Moderate disease (subscores are 1 to 2)
3 = Severe disease (subscores are 2 to 3)
*** The Physician's Global Assessment (PGA) acknowledges the three other subscores, the patient's daily record of abdominal discomfort and functional assessment, other observations such as physical findings, and the patient's performance status.

The SF-36 is a generic measure of health-related quality of life (HRQoL) that consists of eight domains: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health. An increase in SF-36 score indicates an improvement in HRQoL, and a decrease in score indicates disease deterioration in HRQoL. Scores for each component range from 0 to 100, with higher scores reflecting better HRQoL. No MCID was found specific to UC although in Crohn

disease the MCID is 1.6 to 7.0 for PCS, and 2.3 to 8.7 for MCS.[17] The overall MCID (any condition) for the PCS and the MCS ranges from 2.5 to 5 points, and for individual domains ranges from 5 to 10.[18-20]

The Inflammatory Bowel Disease Questionnaire (IBDQ), a 32-item questionnaire, is a disease-specific (Crohn disease and UC) instrument that captures how the patient felt during the two weeks before the measurement time point. Questions are related to symptoms the patient might have had as a result of UC, how the patient felt in general, how the patient's mood was, and social/work problems the patient might have had resulting from UC. The total IBDQ score ranges between 32 and 224, with higher scores representing better quality of life. The scores of patients in remission usually range from 170 to 190. The MCID for the IBDQ is considered to be between 16 points and 32 points for Crohn disease; no MCID is available for UC.[21] In ULTRA 1 and in ULTRA 2, the manufacturer defined a "responder" on the IBDQ as having achieved an improvement of at least 16 points.

The Work Productivity and Impairment (WPAI) questionnaire is a six-item questionnaire that measures the effect of the patient's health problems on work and daily activities in the previous week, specifically, the number of hours worked, the number of hours missed from work due to health problems, how much the patient's health problems affected work productivity, and how much the patient's health problems affected regular activities. Low scores indicate little or no impact of health problems on work and activities, and a decrease in the WPAI score indicates improvement. No MCID has yet been defined.

3.2.5 Statistical analysis

a) ULTRA 1

The objective of the primary efficacy analysis was to demonstrate that adalimumab was superior to placebo in achieving clinical remission at week 8. The sample size was calculated assuming 15% of patients in the placebo group achieved clinical remission at week 8; therefore, a sample size of 125 in each treatment group would be adequate to detect a 15% difference using a chi-square test with 80% power at a 0.05 two-sided significance level. Thus, a total of 375 patients were to be randomized. No rationale was provided for the choice of parameters.

The primary analysis was conducted in the ITT-A3 population, defined in the next section. The following non-responder imputation (NRI) method was used to calculate the remission rate: Patients who discontinued the study for any reason prior to week 8, and patients with a missing Mayo score at week 8 were counted as not being in remission. The last observation carried forward (LOCF) method was used as a sensitivity analysis. The normal approximation to binomial distribution was used to construct the two-sided 95% confidence interval (CI) for the difference in the remission rate between each adalimumab treatment group and the placebo group.

The proportion of patients achieving a clinical response per Mayo score was presented by randomized treatment group at week 8. Patients with a missing Mayo score were not considered to have achieved a response. The difference in proportion of patients achieving response between the adalimumab group and the placebo group was assessed using the chi-square test or Fisher's exact test, as appropriate.

Twelve ranked secondary variables were tested in a hierarchical order to account for multiple testing:

- Proportion of patients with clinical response per Mayo score at week 8 (adalimumab 160 mg/80 mg/40 mg versus placebo)
- Proportion of patients with mucosal healing at week 8 (adalimumab 160 mg/80 mg/40 mg versus placebo)

- Proportion of patients with an RBS indicative of mild disease (≤ 1) at week 8 (adalimumab 160 mg/80 mg/40 mg versus placebo)
- Proportion of patients with a PGA subscore indicative of mild disease (≤ 1) at week 8 (adalimumab 160 mg/80 mg/40 mg versus placebo)
- Proportion of patients with an SFS indicative of mild disease (≤ 1) at week 8 (adalimumab 160 mg/80 mg/40 mg versus placebo)
- Proportion of IBDQ responders at week 8 (adalimumab 160 mg/80 mg/40 mg versus placebo)
- Proportion of patients with clinical response per Mayo score at week 8 (adalimumab 80 mg/40 mg versus placebo)
- Proportion of patients with mucosal healing at week 8 (adalimumab 80 mg/40 mg versus placebo)
- Proportion of patients with an RBS indicative of mild disease (≤ 1) at week 8 (adalimumab 80 mg/40 mg versus placebo)
- Proportion of patients with a PGA subscore indicative of "normal or mild disease" (or numerical score ≤ 1) at week 8 (adalimumab 80 mg/40 mg versus placebo)
- Proportion of patients with SFS indicative of mild disease (≤ 1) at week 8 (adalimumab 80 mg/40 mg versus placebo)
- Proportion of IBDQ responders at week 8 (adalimumab 80 mg/40 mg versus placebo)

Statistically significant results (*P* value ≤ 0.05) had to be achieved for a comparison in the higher rank to initiate the next comparison in the lower rank. Ranked end point number 1 (clinical response per Mayo score at week 8 in the adalimumab 160 mg/80 mg/40 mg treatment group versus placebo) did not meet the criteria for statistical significance.

Non-ranked dichotomous efficacy variables were analyzed using the same methods listed above. Change from baseline in the IBDQ scores, SF-36 scores, Mayo score, and partial Mayo score were summarized using descriptive statistics. The treatment difference in mean change was analyzed using the analysis of variance (ANOVA) model, including factors of treatment and baseline scores or non-parametric tests, as appropriate. Both the data as observed and the LOCF method could be used as appropriate. The median time to achieve response per partial Mayo score from baseline was calculated using the Kaplan-Meier method.

The primary efficacy analysis at week 8 for the intention-to-treat (ITT)-A3 set was presented for the following subgroups:

- Sex (male, female)
- Age category (< 40 years, 40 to 64 years, ≥ 65 years)
- Race (white, other)
- Weight category (< 70 kg, ≥ 70 kg)
- Duration (years) of UC (≤ median duration, > median duration)
- Baseline Mayo score (0 to 5, 6 to 9, 10 to 12)
- Baseline CRP (< 1.0 mg/dL, ≥ 1.0 mg/dL)
- Smoker, past or present (Yes, No)
- Systemic corticosteroid use at baseline (Yes, No)
- Azathioprine /6-MP therapy at baseline (Yes, No)
- Aminosalicylate therapy at baseline (Yes, No)

b) ULTRA 2

The primary efficacy analysis was to be performed on the ITT analysis set, and consisted of two ranked co-primary efficacy end points: (1) the proportion of patients achieving clinical remission at week 8, and (2) the proportion of patients achieving clinical remission at week 52.

Sample size was calculated assuming that 5% of the patients in the placebo group achieved clinical remission at week 52 or week 8; therefore, a sample size of 250 in each treatment group was adequate to detect a difference of at least percentage points from the adalimumab group using a chi-square test with 80% power at a 0.05 two-sided significance level. Thus, a total of 500 patients were to be randomized in this study. No rationale was provided for the choice of parameters.

Hypothesis testing for the ranked end points was carried out in a hierarchical order using a two-sided Cochran-Mantel-Haenszel (CMH) test adjusted for prior exposure to infliximab or other anti-TNF drugs. Additionally, the corresponding two-sided 95% CI for the difference in proportions was provided. The remission rate at week 8 was tested first. If the null hypothesis of no difference between adalimumab and placebo in remission rate at week 8 was rejected at alpha = 0.05, then the remission rate at week 52 was tested at a significance level of 0.05. However, in order to claim maintenance of remission, it was necessary to reject not only both hypotheses on the two ranked co-primary end points (clinical remission at weeks 8 and 52), but also to reject the hypothesis on the first-ordered secondary end point (proportion of patients in remission at both week 8 and week 52). The ranked secondary end points were:

- Sustained remission per Mayo score at week 8 and week 52
- Response per Mayo score at week 8
- Response per Mayo score at week 52
- Sustained response per Mayo score at week 8 and week 52
- Mucosal healing at week 8
- Mucosal healing at week 52
- Sustained mucosal healing at week 8 and week 52
- Discontinued corticosteroid use before week 52 and achieved remission at week 52
- PGA score of mild (≤ 1) at week 8
- SFS of mild (≤ 1) at week 8
- RBS of mild (≤ 1) at week 8
- Discontinued corticosteroid use for ≥ 90 days before week 52 and achieved remission at week 52
- Discontinued corticosteroid use and achieved sustained remission at both weeks 32 and 52
- Number (%) of IBDQ responders at week 52
- Number (%) of IBDQ responders at week 8

The ranked end point of PGA ≤ 1 at week 8 in the adalimumab treatment group versus placebo did not meet the criteria for statistical significance.

Non-ranked categorical secondary efficacy variables were analyzed by NRI and by LOCF as a sensitivity analysis using the CMH test. Change from baseline in Mayo score, IBDQ, SF-36, and WPAI were analyzed using an analysis of covariance (ANCOVA) model including factors of treatment, prior exposure to infliximab or other anti-TNF drugs, and baseline values. For changes, both LOCF and as-observed cases were used as imputation methods.

Non-responder imputation: The NRI approach was used for all binary variables showing some type of clinical response or clinical remission. These variables can take values of "Response" (or "Remission") or "Non-response" (or "Non-remission"), or may be missing for any reason including discontinuation from the study. According to the NRI method, all missing response (or remission) values were considered as non-responses (or non-remissions). Patients who switched to OL administration were considered non-responders (or non-remitters) at and after their switch to OL. The primary and secondary efficacy analyses used the NRI approach to impute missing values.

Last observation carried forward (LOCF): For all variables (discrete/categorical/response variables and continuous variables), the following rules were used for the LOCF approach:

1. Baseline and pre-baseline values were not used to impute the missing post-baseline values.
2. Missing values after study day 1 were imputed using the latest non-missing values after day 1 and prior to the missing value.
3. For patients who switched to OL administration, the latest non-missing value before or at the visit when the subject switched to OL administration was to be carried forward.

The LOCF approach was used for the primary and secondary efficacy analyses as a sensitivity analysis. Primary efficacy analysis and the analysis of the first-ranked secondary end point for the ITT analysis set were to be presented for the following subgroups:

- Sex (male, female)
- Age category (< 40 years, 40 to 64 years, ≥ 65 years)
- Age category (< median age in ITT analysis set, ≥ median age in ITT analysis set)
- Race (white, other)
- Weight category (< 70 kg, ≥ 70 kg)
- Tobacco use (user, ex-user, never used)
- Prior anti-TNF use (Yes, No)
- Baseline CRP (< 1.0 mg/dL, ≥ 1.0 mg/dL)
- Baseline CRP (< 0.6 mg/dL, ≥ 0.6 mg/dL)
- Baseline Mayo score category (< 10, ≥ 10)
- Concomitant use of immunosuppressants (defined as 6-MP, azathioprine) and/or non-topical corticosteroids at baseline (Yes, No)
- Week 8 remission status (Yes, No)
- Site of UC (pancolitis, descending colon, others)
- Disease duration (≤ 2 years, > 2 years)
- Baseline endoscopy subscore (< 3, 3)
- Baseline CRP ≥ 1.0 mg/dL and baseline endoscopy subscore of 3
- Geographic region

Suzuki et al. 2014

In Suzuki et al., no formal sample size calculations were performed.

c) Analysis populations

ULTRA 1

The ITT-A3 set included all patients with confirmed UC at baseline who were randomized according to Protocol Amendment 3/Amendment 4 and who received at least one injection of the following induction regimens: adalimumab 160 mg/80 mg/40 mg every other week, adalimumab 80 mg/40 mg every other week, or placebo. This analysis set was the basis for the primary efficacy analysis due to changes to the eligibility criteria in Amendment 3 and the resulting lack of homogeneity between

patients enrolled before Amendment 3 versus after Amendment 3, and for the ranked secondary efficacy analyses. In addition, non-ranked secondary efficacy analyses were also done using the ITT-A3 analysis set.

The intention-to-treat-extended (ITT-E) set included all patients with confirmed UC at baseline who were randomized according to the original protocol or any of the four protocol amendments and who received at least one injection of the following induction regimens: adalimumab 160 mg/80 mg/40 mg every other week, adalimumab 80 mg/40 mg every other week, or placebo. In the study protocol, this analysis set was referred to as ITT-A2. The focus of this review was the ITT-A3 analysis set, as this was the population used for the primary efficacy analysis.

The dose escalation analysis set included patients in the ITT-E analysis set who required dose escalation to adalimumab 40 mg every week. As per protocol, patients defined as inadequate responders were permitted to dose escalate from adalimumab 40 mg every other week to 40 mg every week at or after week 12, for subjects enrolled under Protocol Amendment 3 or 4, or at or after week 14, for subjects enrolled prior to Amendment 3.

The per-protocol (PP) set included patients in ITT-A3 analysis set after excluding the patients with major protocol deviations.
The safety analysis set included all patients enrolled in the study who received at least one dose of study medication. Analyses of adverse events (AEs), laboratory parameters, and vital signs were performed on the safety analysis set.

ULTRA 2
In ULTRA 2, the ITT analysis set consisted of patients with confirmed UC at baseline who were randomized, and excluded 24 patients from sites 22635, 36809, and 27010, which were non-compliant with good clinical practice and protocol requirements. The ITT analysis set (N = 494) was the basis for the confirmatory primary efficacy analysis and the ranked and non-ranked secondary efficacy analyses.

The PP analysis set (N = 424) consisted of patients in the ITT analysis set after excluding patients with major protocol deviations. Exploratory efficacy analyses were performed on the PP analysis set.

The safety analysis set consisted of patients who received at least one dose of study drug (including non-compliant sites) (N = 517). Analyses of AEs, laboratory parameters, and vital signs were performed on the safety analysis set.

Suzuki et al.
In Suzuki et al., all analyses were exploratory. Efficacy and safety analyses were carried out on the full analysis set (FAS), which consisted of all patients who received at least one dose of study drug under double-blind conditions.

3.3 Patient disposition

In ULTRA 1, at week 8, 7% of patients in each group had discontinued (Table 10). The most common reason for discontinuation was an AE at week 8. In ULTRA 2, 38% of adalimumab patients and 47% of placebo patients had discontinued by end of study (52 weeks), while at eight weeks, the proportions were 9% and 15%, respectively (Table 11). Lack of efficacy was the most common reason for discontinuation at both time points. In Suzuki et al., 33% of adalimumab patients versus 24% of placebo patients discontinued by end of study (52 weeks), while 4% in each group had discontinued by week 8.

The most common reason for discontinuation at week 8 was an AE, and at week 52 was lack of efficacy (Table 11).

TABLE 10: PATIENT DISPOSITION: INDUCTION STUDY (8-WEEK DOUBLE-BLIND PHASE)

	ULTRA 1	
	Adalimumab 160 mg/80 mg/40 mg	**Placebo**
Screened, N	NR	
Randomized, N (%)	130	130
Completed week 8	121 (93)	121 (93)
Discontinued study prior to week 8	9 (7)	9 (7)
AE	4 (3)	6 (5)
Withdrew consent	1 (1)	0 (0)
Lost to follow-up	0 (0)	0 (0)
Lack of efficacy	2 (2)	5 (4)
Protocol violation	2 (2)	0 (0)
Other[a]	1 (1)	0 (0)
ITT-A3, N	130	130
PP, N	124	120
Safety[b], N	223	223
ITT-E[b], N	223	222

AE = adverse event; ITT = intention-to-treat; ITT-E = intention-to-treat-extended; NR = not rated; PP = per-protocol; TB = tuberculosis; UC = ulcerative colitis.
[a] Reasons for discontinuation recorded as "other" included diagnosis of Crohn disease, loss of response, primary non-responder, UC symptoms not improving, investigator decision, patient noncompliance, positive TB skin test, patient wanted to start family, or total colectomy surgery within the 70-day follow-up period.
[b] The safety and ITT-E populations included all patients originally enrolled into the study as well as those enrolled under any of the subsequent protocol amendments.
Source: Clinical Study Report for ULTRA 1.[3]

TABLE 11: PATIENT DISPOSITION: INDUCTION/MAINTENANCE STUDIES (52-WEEK DOUBLE-BLIND PHASE)

	ULTRA 2		Suzuki et al. 2014[b]	
	Adalimumab	**Placebo**	**Adalimumab**	**Placebo**
Screened, N	NR		339	
Randomized, N	258	260	90	96
Treated, n (%)	257 (92)	260 (100)	NR	NR
Completed, n (%)	163 (62)	145 (53)	60 (NR)	73 (NR)
Discontinued, n (%)	94 (38)	115 (47)	30 (33)	23 (24)
AE	12 (5)	25 (10)	13 (14)	7 (7)
Lack of efficacy	63 (25)	70 (29)	16 (18)	14 (14)
Protocol deviation	1 (< 1)	5 (2)	0 (0)	0 (0)
Withdrew consent	8 (3)	4 (2)	0 (0)	2 (2)
Lost to follow-up	1 (< 1)	0 (0)	0 (0)	0 (0)
Other	9 (4)[a]	11 (5)[a]	1 (1)	0 (0)
Discontinued prior to week 8	23 (9)	36 (15)	4 (4)	4 (4)
AE	5 (2)	10 (4)	3 (3)	2 (2)

	ULTRA 2		Suzuki et al. 2014[b]	
	Adalimumab	**Placebo**	**Adalimumab**	**Placebo**
Withdrew consent	1 (< 1)	2 (1)	0 (0)	0 (0)
Lost to follow-up	0 (0)	0 (0)	0 (0)	0 (0)
Lack of efficacy	13 (5)	15 (6)	1 (1)	2 (2)
Protocol violation	1 (< 1)	3 (1)	0 (0)	0 (0)
Other[a]	3 (1)	6 (2)	0 (0)	0 (0)
ITT, N	248	246	NR	NR
PP, N	212	212	NR	NR
Safety, N	257	260	NR	NR

AE = adverse event; ITT = intention-to-treat; NR = not reported; PP = per-protocol.
[a] Reasons for discontinuation recorded as "other" included: diagnosis of Crohn disease, loss of response, primary non-responder, UC symptoms not improving, investigator decision, patient noncompliance, positive TB skin test, patient wanted to start family, or total colectomy surgery within the 70-day follow-up period.
[b] Only the arm with the approved Health Canada dose (160/80 mg) is included in the Adalimumab arm.
Source: Clinical Study Report for ULTRA 2;[4] Suzuki et al.[5]

3.4 Exposure to study treatments

The mean treatment duration in ULTRA 1 was [redacted] days for each of adalimumab 160 mg/80 mg/40 mg group and placebo group, and each group had an average of [redacted] injections per patient. In ULTRA 2, the mean treatment duration for adalimumab was [redacted] days and for placebo was [redacted] days. The mean number of injections was [redacted] per patient for adalimumab and [redacted] per patient for placebo.

3.5 Critical appraisal

3.5.1 Internal validity

In ULTRA 1 and ULTRA 2, the manufacturer employed a hierarchical testing procedure as a means for adjusting for multiplicity. Secondary outcomes were presented in rank order, in a way that made it clear whether each subsequent outcome was being tested in an appropriate manner. However, not all outcomes were ranked, and for unranked outcomes no adjustment for multiplicity appears to have been made. This applied to some key efficacy outcomes, particularly in ULTRA 1, where the quality of life analysis was not tested with control of multiplicity. There were some quality of life outcomes that were not statistically significant, regardless of adjustment, but there were others, including mean changes on the IBDQ, that were reported as statistically significant despite the fact that, according to the hierarchy, they should not have been tested. Other outcomes outside of the hierarchy, such as the SF-36, should be considered only as exploratory. It is not clear how the ranking was determined, including what outcomes would and would not be ranked.

A major protocol amendment was undertaken in ULTRA 1, whereby an additional dosing arm of 80 mg/40 mg adalimumab was added after the study was underway. According to the manufacturer, this arm was added because the EU had approved both the 160 mg/80 mg/40 mg and the 80 mg/40 mg regimens as induction for Crohn disease. Note that in Canada, adalimumab was only approved at the 160 mg/80 mg/40 mg regimen for UC. As a result of this amendment, changes were also made that applied to all arms in the study, most notably the shortening of the treatment period from 12 weeks to eight weeks. As a result, ULTRA 1 was essentially re-initiated from the beginning, and the population in this re-initiated study was referred to as the “A3” population, named for Protocol Amendment 3. The A3 population became the population for the primary analysis, with the other populations also presented. Data for these other populations are provided in Appendix 3, Table 23.

Early withdrawals were counted as treatment failures ("non-responder imputation") as part of the primary analysis. This is a simple and straightforward method for accounting for patients who withdrew from the study; however, it might bias results in cases of differential rates of withdrawal. In ULTRA 1, 7% of patients in each of the adalimumab and placebo groups discontinued by week 8; therefore, the primary analysis in this study is unlikely to have been biased by a differential rate of withdrawal. However, in ULTRA 2 at week 8, 9% of adalimumab patients and 15% of placebo patients had withdrawn, and at end of study the withdrawal rates were 38% and 47%, respectively. In these cases, where higher rates of withdrawal were seen with placebo, a bias may have been introduced favouring adalimumab, as more placebo patients may have been counted as non-responders because they withdrew. At week 52, 25% of adalimumab patients and 29% of placebo patients withdrew due to lack of efficacy. Counting these patients as treatment failures may be appropriate, as this reflects their reason for withdrawing. These data suggest, however, that there may have been patients who withdrew for reasons other than treatment failure who were counted as treatment failures, and in this group, there was still a numerically higher proportion of placebo patients who withdrew compared with adalimumab patients. Given the small differences in treatment effect between adalimumab and placebo in the study, even a small difference in the number of patients imputed as treatment failures who may not have been treatment failures has the potential to impact the analysis in a way that may favour adalimumab.

All studies were DB, and appropriate measures appear to have been taken in order to maintain blinding throughout the study. The investigator, patients, as well as personnel at the study site were to remain blinded throughout the DB phase of the study. Nevertheless, one of the concerns in maintaining adequate blinding is whether adverse effects that are associated with a particular drug may allow patients to speculate on what intervention they have been assigned to. Examples relevant to adalimumab are injection site reactions and hypersensitivity responses, adverse effects that have been classically associated with the therapeutic use of monoclonal antibodies. Although injection site reactions were infrequent AEs, patients who experienced them may have been able to accurately guess which group they had been assigned to; the same possibility applies to the investigators who observed the reactions in their patients. Knowledge of treatment assignment is more likely to bias key patient-reported outcomes such as quality of life, and less likely to bias objective outcomes such as the need for colectomy. Clinical remission and clinical response both rely on the Mayo scoring system, an instrument that has subjective components that may be influenced by investigator knowledge of treatment allocation, and is therefore prone to such bias.

The manufacturer included a number of pre-planned subgroup analyses in both ULTRA 1 and ULTRA 2. Common limitations of such analyses include limited power, not maintaining randomization (unless a stratification variable was used), and not adjusting for multiplicity. The manufacturer did not state whether the analyses were planned or post hoc analyses (for those where statistical tests were performed).

3.5.2 External validity

Suzuki et al. (2014) was conducted in Japan as part of that country's regulatory process. The fact that the population was entirely Japanese limits the generalizability of study findings to Canada. There were also a relatively large proportion of patients who were screened out of the study (339 screened, 180 randomized), which may further limit generalizability. The number of patients screened could not be found for ULTRA 1 and ULTRA 2.

All of the included studies were limited in their ability to assess key safety issues such as malignancy because of a relatively short duration of treatment (maximum of 52 weeks in the DB phase of ULTRA 2, for example). Malignancies are a potential safety issue associated with the use of TNF inhibitors, although the nature and extent of the risk have yet to be established. One would not expect to see differences in risk of malignancy between adalimumab and placebo in studies with a duration of one year or less. Although long-term extensions provide some data, they do not adequately assess risk because they lack a control group.

The indication and listing criteria for adalimumab require patients to have had an inadequate response to, or be intolerant of, corticosteroids, azathioprine, or 6-MP. However, not all patients in the included studies met this criteria (Table 18). For example, in ULTRA 2, [redacted] of patients had used UC-related medications within 90 days of baseline. If these patients were to have contributed disproportionately to the beneficial effects of adalimumab, this may have led to conclusions being drawn about adalimumab that are not applicable to the population in which the drug will be used. Additionally, according to the clinical expert involved in the review, the required dose of azathioprine was lower than that typically used in clinical practice. If patients previously treated with azathioprine were underdosed, that may mean that these patients were potential responders but were counted as non-responders, again impacting generalizability.

All included studies enrolled patients with moderate to severe UC, based largely on their Mayo scores. Many clinicians do not use Mayo scores to categorize patients, although they may use components of the scoring system, such as the SFS, to assess severity of disease in their patients. The suggested listing criteria for adalimumab is for use in moderate to severe UC patients, presumably based on the phase 3 studies; however, it is not clear whether clinicians in practice will use Mayo scores to identify patients who are eligible to receive the drug.

The clinical expert involved in the review noted that patients who are attending school/university were not specifically assessed as a subgroup in any of the included studies. The clinical expert noted that this was an important omission, as a university student whose condition worsens may be at risk of losing an entire term of their program, and this can have significant consequences for students who are in programs that do not have the flexibility to make up for that loss in a subsequent term. Many patients with UC are in their 20s, and therefore of university age.

3.6 Efficacy

Only those efficacy outcomes identified in the review protocol are reported below (section 2.2, Table 4). See 0APPENDIX 4: DETAILED OUTCOME DATA for detailed efficacy data.

3.6.1 Remission

Remission was defined as Mayo score ≤ 2 with no subscore > 1.

In ULTRA 1, the proportion of patients achieving remission at eight weeks was 19% with adalimumab and 9% with placebo, and this difference was statistically significant (difference in proportions 9.2; 95% CI, [redacted]; *P* = 0.031) (Table 12). This was the primary outcome of this superiority study; therefore, adalimumab achieved its primary outcome for this study.

In ULTRA 2, the proportion of patients achieving remission at eight weeks was 17% with adalimumab and 9% with placebo, and this difference was statistically significant (difference in proportions 7.1; 95% CI, 1.2 to 12.9; *P* = 0.019). The proportion achieving remission at 52 weeks was also 17% versus 9%

(difference in proportions 8.6; 95% CI, 2.8 to 14.5; *P* = 0.004) (Table 13). These two outcomes, remission at 8 weeks and at 52 weeks, were the co-primary outcomes of this superiority study; therefore, adalimumab achieved both of its co-primary outcomes in ULTRA 2.

In Suzuki et al., 10% of adalimumab patients and 11% of placebo patients achieved remission at week 8, and statistical analyses were not reported for this outcome (Table 13). At week 52, a higher proportion of adalimumab patients than placebo patients had achieved remission, and this difference was statistically significant (23% versus 7%, *P* < 0.001). Note that all analyses were considered exploratory in this study.

a) Subgroups

In ULTRA 1, 23% of adalimumab and 31% of placebo patients underwent dose escalation. Of those, the proportion of patients achieving remission was 14% with adalimumab and 12% with placebo/adalimumab at week 52, compared with the overall proportions of 25% and 26% at week 52, respectively (Table 20). No analysis appears to have been performed for subgroups based on prior anti-TNF use. Subgroup analyses based on baseline use of conventional therapies were also reported. In patients who were users of azathioprine or 6-MP at baseline, the proportion achieving remission at week 52 was 16% with adalimumab and 4% with placebo, while in patients who were not users of these drugs at baseline, the proportion of responders was 20% versus 13%, adalimumab versus placebo, respectively.

In ULTRA 2, the proportion of patients who underwent dose escalation was 27% with adalimumab and 34% with placebo. Of those, the proportion who achieved remission was 12% with adalimumab and 34% with adalimumab/placebo (Table 21). In patients with no prior anti-TNF, remission occurred in 21% of adalimumab patients and 11% of placebo patients at week 8 (P = 0.017), and in patients with prior anti-TNF, the proportions were 9% and 7%, respectively (P = 0.559). At week 52, in patients with no prior anti-TNF, 22% of adalimumab patients and 12% of placebo achieved remission (P = 0.029), while in patients with prior anti-TNF the proportions were 10% and 3%, respectively (P = 0.039). Subgroup analyses based on baseline use of conventional therapies were also reported in ULTRA 2. In patients who had taken azathioprine or 6-MP at baseline, the proportion achieving remission at week 8 was 13% with adalimumab and 15% with placebo. However, in patients who had not received azathioprine or 6-MP, the proportions were 19% for adalimumab and 7% for placebo. After 52 weeks, these differences in remission were no longer evident between prior azathioprine/6-MP users (remission in 18% of adalimumab and 10% of placebo) and non-users (remission in 17% of adalimumab and 8% placebo). These differences in remission between previous conventional therapy users/non-users were not apparent when the conventional therapies included corticosteroids in addition to azathioprine/6-MP. Subgroup analyses were not evaluated for Suzuki et al.

3.6.2 Clinical response

Clinical response was defined as a decrease of Mayo score ≥ 3 points and a decrease ≥ 30%, and RBS of 0 or 1, or a decrease of RBS ≥ 1.

In ULTRA 1, 55% of adalimumab patients and 45% of placebo patients had a clinical response at week 8, and this difference was not statistically significant (Table 12).

In ULTRA 2, 50% of adalimumab patients and 35% of placebo patients achieved clinical response at week 8, and this difference was statistically significant (difference in proportions 15.6; 95% CI, 7.0 to 24.2; $P < 0.001$) (Table 13). At week 52, 30% of adalimumab patients and 18% of placebo patients had achieved a clinical response, and this difference was also statistically significant (difference in proportions 11.7; 95% CI, 4.3 to 19.2; $P = 0.002$).

In Suzuki et al., at week 8, the proportion of patients with a clinical response was higher with adalimumab than with placebo, and this difference was statistically significant (50% versus 35%, $P = 0.044$) (Table 13). At week 52, adalimumab also had a statistically significantly higher proportion of responders than placebo (31% versus 18%, $P = 0.021$).

3.6.3 Health-related quality of life

HRQoL was assessed using the IBDQ and SF-36. The SF-36 score was an exploratory outcome in all studies. An IBDQ responder was defined by an increase of at least 16 points from baseline, and there was no difference in the proportion of responders at week 8 between adalimumab (61%) and placebo (58%) in ULTRA 1 (Table 12). Of note, based on the statistical analysis hierarchy plan, testing should not have been performed for this outcome because a higher-ranked outcome in the hierarchy failed to reach statistical significance. Change from baseline in IBDQ was not part of the hierarchy. Therefore, all statistical analyses reported for IBDQ should be considered exploratory. The MCID for the IBDQ has not been established for UC, but is considered to be between 16 and 32 points for Crohn disease (See Appendix 5 for review). The mean change from baseline to week 8 was statistically significantly greater (improved) for adalimumab versus placebo at week 8 (least squares mean difference [LSMD] 11.3; 95% CI, 2.9 to 19.6; $P = 0.008$). There was also a statistically significantly greater increase (improvement) for adalimumab over placebo in SF-36 PCS scores at week 8 (LSMD 3.4; 95% CI, 1.8 to 5.0; $P < 0.001$), but no statistically significant difference between groups on the SF-36 MCS at week 8 (Table 12). While the improvement in the SF-36 PCS was statistically significant, the clinical significance was less clear, given that the MCID for SF-36 component summaries is between 2.5 and 5.

In ULTRA 2, differences in proportions of IBDQ responders between adalimumab and placebo were reported at week 8 and week 52. However, based on the hierarchical statistical testing plan, testing should not have been performed for this outcome because a previous outcome in the hierarchy failed to reach statistical significance. Therefore, all statistical analyses reported for IBDQ should be considered exploratory. In ULTRA 2, the proportion of IBDQ responders at week 8 was higher with adalimumab (58%) than with placebo (46%), and this difference was statistically significant (difference in proportions 12.2; 95% CI, 3.6 to 20.9; $P = 0.006$) (Table 13). The proportion of IBDQ responders was also higher with adalimumab (26%) than with placebo (16%) at week 52 (difference in proportions 9.7; 95% CI, 2.6 to 16.9; $P = 0.007$). At week 8, there was no statistically significant difference in SF-36 PCS or SF-36 MCS between groups (Table 13), but at week 52, there was a statistically significantly greater improvement in SF-36 PCS scores for adalimumab versus placebo (LSMD 1.9; 95% CI, 0.4 to 3.3; $P = 0.011$), but no statistically significant difference between groups on the SF-36 MCS. This statistically significant difference is unlikely to be clinically significant.

In Suzuki et al., the proportion of IBDQ responders was not reported as statistically significantly different between adalimumab and placebo (42% versus 40%, respectively) at week 8 (Table 13). There was a statistically significantly higher proportion of adalimumab responders versus placebo responders at week 52 (25% versus 13%, $P \leq 0.01$). All analyses in Suzuki et al. were considered exploratory.

3.6.4 Need for colectomy

In ULTRA 1, no patients underwent colectomy during the treatment phase (DB and OL) phase of the study. There were 2% of adalimumab patients and 4% of placebo patients who underwent colectomy after they had received their final dose of study drug (Table 12).
In ULTRA 2, no patients underwent colectomy during the treatment phase of the study; however, during the follow-up phase, 4% of adalimumab patients and 5% of placebo patients underwent colectomy (Table 13). No colectomies were reported in Suzuki et al.

3.6.5 Other efficacy outcomes

Missed days of work were reported as an exploratory outcome within the WPAI in ULTRA 2, and were not reported in ULTRA 1. There was no statistically significant difference between adalimumab and placebo in change in work time missed from baseline to 52 weeks. Physical function and disability were not reported as separate outcomes, but were likely captured within the IBDQ, reported above.

In ULTRA 1, there was no statistically significant difference between adalimumab and placebo in the proportion of patients with mucosal healing (Table 16). In ULTRA 2, there was a larger proportion of adalimumab patients versus placebo patients who experienced mucosal healing, and these differences were statistically significant at both eight weeks (41% versus 32%, *P* = 0.032) and at 52 weeks (25% versus 15%, *P* = 0.009). In Suzuki et al., the proportion of patients with mucosal healing was also higher with adalimumab versus placebo at both week 8 (44% versus 30%, *P* = 0.045) and at week 52 (29% versus 16%, *P* = 0.015).

In ULTRA 1 and in Suzuki et al., no *P* values were reported for steroid-free remission (Table 16, Table 17).In ULTRA 2, there were statistically significantly higher proportions of adalimumab patients versus placebo patients who discontinued corticosteroids before week 52 and achieved remission at week 52 (13% versus 6%, *P* = 0.035), discontinued corticosteroids for at least 90 days before week 52 and achieved remission at week 52 (13% versus 6%, *P* = 0.035), and discontinued corticosteroid use and achieved sustained remission at both weeks 32 and 52 (10% versus 1%, *P* = 0.002).

In ULTRA 1, 23% of adalimumab patients and 31% of placebo patients underwent dose escalation, while in ULTRA 2, the proportions were 27% and 34%, respectively.

TABLE 12: KEY EFFICACY OUTCOMES — INDUCTION STUDY (8-WEEK DOUBLE-BLIND PHASE)

	ULTRA 1	
Remission	**Adalimumab N = 130**	**Placebo N = 130**
ITT, NRI, week 8, n/N (%)	24/130 (19)	12/130 (9)
Difference in proportions [95% CI]	9.2 [redacted], *P* = 0.031	
ITT, LOCF, week 8, n/N (%)	24/124 (19)	12/123 (10)
Difference in proportions [95% CI]	9.6 [redacted], *P* = 0.033	
clinical response		
Patients, week 8, n (%)	71/130 (55)	58/130 (45)
Difference in proportions [95% CI]	10.0 [redacted], *P* = 0.107	
IBDQ		
Responders, week 8, n (%)	79/130 (61)	75/130 (58)
Difference in proportions [95% CI]	3.1 [redacted], *P* = 0.614[a,b]	

	ULTRA 1	
Remission	**Adalimumab N = 130**	**Placebo N = 130**
Mean (SD) baseline	131.9 n = 120	125.3 n = 127
Mean (SD) change at week 8	35.9 (34.0) n = 117	26.9 (35.6) n = 124
LSMD [95% CI]	[redacted]	
SF-36 PCS		
Mean (SD) baseline	[redacted]	[redacted]
Mean (SD) change from baseline, week 8	[redacted]	[redacted]
LSMD [95% CI]	[redacted]	
SF-36 MCS		
Mean (SD) baseline	[redacted]	[redacted]
Mean (SD) change from baseline, week 8	[redacted]	[redacted]
LSMD [95% CI]	[redacted]	
Need for colectomy, patients, n (%)		
Week 8	0 (0)	0 (0)
OL period through week 52	0 (0)	0 (0)
After last dose	4 (2)	8 (4)

IBDQ = Inflammatory Bowel Disease Questionnaire; ITT = intention-to-treat; NRI = non-responder imputation; LOCF = last observation carried forward; LS = least squares; MCS = Mental Component Summary; MD = mean difference; OL = open-label; PCS = Physical Component Summary; SD = standard deviation; SF-36 = Short Form (36) Health Survey.

[a] According to the NRI analysis method, all missing response (or remission) values and values after dose escalation were imputed as non-response (or non-remission).

[b] Outcome was included in the hierarchical analysis plan, but testing should not have been performed for this outcome because a previous outcome in the hierarchy failed to reach statistical significance. Therefore all statistical test results should be considered exploratory for IBDQ responders.

Note: Remission: *P* values for adalimumab versus placebo in ITT-A3 set (NRI and LOCF analyses) and placebo set from chi-square test (or Fisher's exact test if ≥ 20% of cells had expected cell count < 5).

Clinical Response, IBDQ : *P* value for differences between active treatment group and placebo from chi-square test (or Fisher's exact test if ≥ 20% of the cell have an expected count < 5).

IBDQ response was defined as an increase in IBDQ ≥ 16 points from baseline.

IBDQ, mean change: *P* values for adalimumab versus placebo from one-way analysis of covariance (ANCOVA) with treatment as factor and baseline value as covariate.

SF-36: *P* values for adalimumab versus placebo from one-way ANCOVA with treatment as factor and baseline value as covariate.

Source: Clinical Study Report for ULTRA 1.[3]

TABLE 13: KEY EFFICACY OUTCOMES — INDUCTION/MAINTENANCE STUDIES (52-WEEK DOUBLE-BLIND PHASE)

	ULTRA 2		Suzuki et al. 2014	
Remission	**Adalimumab N = 248**	**Placebo N = 246**	**Adalimumab N = 90**	**Placebo N = 96**
Week 8, NRI, n (%)	41/248 (17)	23/246 (9)	NR (10)	NR (11)
Difference in proportions [95% CI]	7.1 [1.2, 12.9], *P* = 0.019		NR	

	ULTRA 2		Suzuki et al. 2014	
Remission	**Adalimumab N = 248**	**Placebo N = 246**	**Adalimumab N = 90**	**Placebo N = 96**
Week 8 LOCF	41/248 (18)	23/246 (11)	NR	NR
Difference in proportions [95% CI]	7.5 [1.0, 14.0], *P* = 0.024			
Week 52, NRI	43/248 (17)	21/246 (9)	NR (23)	NR (7)
Difference in proportions [95% CI]	8.8 [2.8, 14.5], *P* = 0.004		*P* = 0.001[b]	
Week 52 LOCF	46/248 (20)	23/246 (11)	NR	NR
Difference in proportions [95% CI]	9.2 [redacted], *P* = 0.006			
Clinical Response				
Week 8, Patients, n (%)	123/248 (50)	86/246 (35)	NR (50)	NR (35)
Difference in proportions [95% CI]	15.6 [redacted], *P* < 0.001		*P* = 0.044	
Week 52, Patients, n (%)	75/248 (30)	45/246 (19)	NR (31)	NR (18)
Difference in proportions [95% CI]	11.7 [redacted], *P* = 0.002		*P* = 0.021	
IBDQ				
Responders, NRI, week 8, n (%)	144/248 (58)	112/246 (46)	38 (42)	38 (40)
Difference in proportions [95% CI]	12.2 [redacted], *P* = 0.006[a]		NR	
Responders, week 52, NRI, n (%)	65/248 (26)	40/246 (16)	45 (25)	12 (13)
Difference in proportions [95% CI]	9.7 [redacted], *P* = 0.007[a]		*P* ≤ 0.01	
Mean (SD) baseline	[redacted]	[redacted]	NR	NR
Mean (SD) change at week 8	[redacted]	[redacted]	NR	NR
LSMD [95% CI]	[redacted]			
Mean (SD) change at week 52	[redacted]	[redacted]	NR	NR
LSMD [95% CI]	[redacted]			
SF-36 PCS				
Mean (SD) baseline	[redacted]	[redacted]	NE	NE
Mean (SD) change from baseline, week 8	[redacted]	[redacted]	NE	NE
LSMD [95% CI]	[redacted]			
Mean (SD) change from baseline, week 52	[redacted]	[redacted]	NE	NE
LSMD [95% CI]	[redacted]			
SF-36 MCS				
Mean (SD) baseline	[redacted]	[redacted]	NE	NE
Mean (SD) change from baseline, week 8	[redacted]	[redacted]	NE	NE
LSMD [95% CI]	[redacted]			
Mean (SD) change from baseline, week 52	[redacted]	[redacted]	NE	NE
LSMD [95% CI]	[redacted]			
Need for colectomy, patients, n (%)				
Week 8	0 (0)	0 (0)	0 (0)	0 (0)

	ULTRA 2		Suzuki et al. 2014	
Remission	**Adalimumab N = 248**	**Placebo N = 246**	**Adalimumab N = 90**	**Placebo N = 96**
Follow-up phase	10 (4)	12 (5)	NR	NR

IBDQ = Inflammatory Bowel Disease Questionnaire; ITT = intention-to-treat; LOCF = last observation carried forward; MCS = Mental Component Summary; NE = not evaluated; NR = not reported; NRI = non-responder imputation; OL = open-label; PCS = Physical Component Summary; SD = standard deviation; SF-36 = Short Form (36) Health Survey.
[a] Outcome was included in the hierarchical analysis plan, but testing should not have been performed for this outcome because a previous outcome in the hierarchy failed to reach statistical significance. Therefore, all statistical test results should be considered exploratory for IBDQ responders.
[b] Confidence intervals on the difference were not available; on the p-values.
Note: Remission was defined as Mayo score ≤ 2 with no subscore > 1. Clinical response was defined as a decrease of Mayo score ≥ 3 points and a decrease ≥ 30%, and rectal bleeding subscore (RBS) of 0 or 1, or a decrease of RBS ≥ 1.
Clinical remission: *P* value to compare treatment groups was based on Cochran-Mantel-Haenszel test (stratification levels: prior anti-tumour necrosis factor (TNF) versus anti-TNF-naive).
Source: Clinical Study Report for ULTRA 2;[4] Suzuki et al.[5]

3.7 Harms

Only those harms identified in the review protocol are reported below (see 2.2.1, Protocol). See 0APPENDIX 4: DETAILED OUTCOME DATA for detailed harms data.

3.7.1 Adverse events

In ULTRA 1, 50% of adalimumab patients and 48% of placebo patients experienced an AE (Table 14), and in ULTRA 2, the proportions were 83% versus 84%, respectively (Table 15). In Suzuki et al., 44% of adalimumab patients and 47% of placebo patients experienced an AE (Table 15). The most common AE was UC.

3.7.2 Serious adverse events

SAEs were reported by 4% of adalimumab patients and 8% of placebo patients in ULTRA 1 (Table 14), 12% of patients in each of the adalimumab and placebo groups in ULTRA 2, and 4% versus 7% of adalimumab versus placebo patients in Suzuki et al. (Table 15). The most common SAE in ULTRA 1 and ULTRA 2 was UC.

3.7.3 Withdrawals due to adverse events

WDAEs occurred in 5% of patients in both adalimumab and placebo in ULTRA 1 (Table 14), 9% of adalimumab versus 13% of placebo in ULTRA 2, and 7% of adalimumab versus 4% of placebo in Suzuki et al. (Table 15). The most common reason for withdrawal was UC.

3.7.4 Mortality

There was one death in the adalimumab group in Suzuki et al. due to TB, and no deaths in either ULTRA 1 or ULTRA 2.

3.7.5 Notable harms

Injection site reactions and infectious AEs were the most common notable harms across all studies. Injection site reactions occurred in 6% of adalimumab patients and 3% of placebo patients in ULTRA 1 (Table 14), 12% of adalimumab and 4% of placebo patients in ULTRA 2, and 8% of adalimumab 2% of placebo patients, respectively, in Suzuki et al. (Table 15). Infectious AEs occurred in 14% of adalimumab patients versus 16% of placebo patients in ULTRA 1 (Table 14), 45% of adalimumab versus 40% of placebo patients in ULTRA 2, and 19% of adalimumab versus 16% of placebo patients, respectively, in

Suzuki et al. (Table 15). Across the studies, there were three patients with malignancies in the adalimumab group versus two in placebo group, six versus two patients with hypersensitivity reactions, respectively, and seven versus three cases of opportunistic infections excluding TB, respectively. As noted above, there was one death due to TB in the adalimumab group in Suzuki et al.

TABLE 14: HARMS: INDUCTION STUDY

	ULTRA 1 (8 weeks)		ULTRA 1-OL (52 weeks)	
Harms	**Adalimumab N = 130**	**Placebo N = 130**	**Adalimumab/ OL Adalimumab**	**Placebo/ OL Adalimumab**
AEs, n (%)				
Patients with > 0 AEs	112 (50)	108 (48)	218 (70)	147 (74)
Most common AEs				
Anemia	1 (< 1)	5(2)	-	-
Abdominal pain	3 (1)	9 (4)	-	-
Ulcerative colitis	13 (6)	21 (9)	50 (16)	42 (21)
Fatigue	9 (4)	5 (2)	-	-
Nasopharyngitis	6 (3)	4 (2)	37 (12)	8 (4)
URTI	3 (1)	6 (3)	14 (5)	14 (7)
Arthralgia	4 (2)	1 (< 1)	24 (8)	13 (7)
Headache	7 (3)	16 (7)	15 (5)	16 (8)
Rash	3 (1)	1 (< 1)	-	-
SAEs, n (%)				
Patients with > 0 SAEs	9 (4)	17 (8)	31 (10)	30 (15)
Most common SAEs[a]				
Ulcerative colitis	4 (2)	13 (6)	14 (5)	12 (6)
WDAEs, n (%)				
WDAEs,	12 (5)	12 (5)	30 (10)	25 (13)
Most common reasons				
Ulcerative colitis	8 (4)	9 (4)	12 (4)	17 (9)
Deaths, n (%)				
Number of deaths	0 (0)	0 (0)	0 (0)	0 (0)
Notable harms, n (%)				
Injection site reactions	13 (6)	7 (3)	19 (6)	11 (6)
Hypersensitivity reactions	2 (1)	1 (< 1)	3 (1)	0 (0)
Malignancies	0 (0)	2 (1)	2 (1)	1 (1)
NMSC	0 (0)	1 (< 1)	0 (0)	0 (0)
Lymphoma	0 (0)	0 (0)	0 (0)	0 (0)
Infectious AEs	32 (14)	35 (16)	116 (37)	58 (29)
Opportunistic infections excluding TB	1 (< 1)	0 (0)	3 (1)	1 (1)
TB	0 (0)	0 (0)	0 (0)	0 (0)
Hepatic-related	5 (2)	1 (< 1)	8 (3)	2 (1)

AE = adverse event; OL = open-label; NMSC = non-melanoma skin cancer; SAE = serious adverse event; TB = tuberculosis; URTI = upper respiratory tract infection; WDAE = withdrawal due to adverse event.
Source: Clinical Study Report for ULTRA 1.[3]

TABLE 15: HARMS: INDUCTION/MAINTENANCE STUDIES (52-WEEK DOUBLE-BLIND PHASE)

	ULTRA 2		Suzuki et al. 2014	
Harms	**Adalimumab N = 248**	**Placebo N = 246**	**Adalimumab N = 90**	**Placebo N = 96**
AEs, n (%)				
Patients with > 0 AEs, N (%)	213 (83)	218 (84)	40 (44)	45 (47)
Most common AEs				
Anemia	10 (4)	15 (6)	NR	NR
Abdominal pain	20 (8)	16 (6)	NR	NR
Ulcerative colitis	58 (23)	76 (29)	NR	NR
Fatigue	16 (6)	15 (6)	NR	NR
Nasopharyngitis	45 (18)	27 (10)	NR	NR
URTI	11 (4)	14 (5)	NR	NR
Arthralgia	20 (8)	16 (6)	NR	NR
Headache	22 (9)	37 (14)	NR	NR
Nausea	15 (6)	22 (9)	NR	NR
Pyrexia	11 (4)	14 (5)	NR	NR
Oropharyngeal pain	15 (6)	7 (3)	NR	NR
SAEs, n (%)				
Patients with > 0 SAEs,	31 (12)	32 (12)	4 (4)	7 (7)
Most common SAEs[a]				
Ulcerative colitis	16 (6)	18 (7)	NR	NR
Anemia	2 (1)	1 (< 1)	NR	NR
Pyoderma gangrenosum	0 (0)	4 (2)	NR	NR
Deep vein thrombosis	2 (1)	1 (< 1)	NR	NR
Serious infection			3 (3)	0 (0)
WDAEs, n (%)				
WDAEs	23 (9)	34 (13)	6 (7)	4 (4)
Most common reasons, n (%)				
Ulcerative colitis	18 (7)	20 (8)	NR	NR
Pyoderma gangrenosum	0 (0)	3 (1)	NR	NR
Deaths, n (%)				
Number of deaths, n (%)	0 (0)	0 (0)	1 (1)	0 (0)
Reasons			TB	
Notable harms, n (%)				
Injection site reactions	31 (12)	10 (4)	7 (8)	2 (2)
Hypersensitivity reactions	4 (2)	1 (< 1)	0 (0)	0 (0)
Malignancies	2 (1)	0 (0)	1 (1)	0 (0)
NMSC	1 (< 1)	0 (0)	NR	NR
Lymphoma	0 (0)	0 (0)	NR	NR
Infectious AEs	116 (45)	103 (40)	17 (19)	15 (16)
Opportunistic infections excluding TB	5 (2)	3 (1)	1 (1)	0 (0)
TB	0 (0)	0 (0)	1 (1)	0 (0)

	ULTRA 2		Suzuki et al. 2014	
Harms	**Adalimumab N = 248**	**Placebo N = 246**	**Adalimumab N = 90**	**Placebo N = 96**
Hepatic-related	10 (4)	7 (3)	1 (1)	1 (1)

AE = adverse event; NMSC = non-melanoma skin cancer; NR = not reported; SAE = serious adverse event; TB = tuberculosis; URTI = upper respiratory tract infection; WDAE = withdrawal due to adverse event.
Source: Clinical Study Report for ULTRA 2;[4] Suzuki et al.[5]

4. DISCUSSION

4.1 Summary of available evidence

Three DB RCTs, all comparing adalimumab with placebo, met the inclusion criteria for this review. ULTRA 1 and ULTRA 2 were pivotal studies for international registration, while Suzuki et al. was a published study that was conducted entirely in Japan as part of that country's regulatory process. The DB phase of ULTRA 1 was eight weeks (induction treatment), with an OL phase that extended to 52 weeks, while the DB phases of ULTRA 2 and Suzuki et al. were 52 weeks in total (induction plus maintenance treatments). The primary outcome of ULTRA 1 and ULTRA 2 was clinical remission, while Suzuki et al. was described as an exploratory analysis. In both ULTRA 1 at eight weeks and in ULTRA 2 at eight weeks and 52 weeks, adalimumab was superior to placebo with respect to the primary outcome of clinical remission. In Suzuki et al., there was no statistically significant difference between adalimumab and placebo at eight weeks with respect to remission, but there was a statistically significant difference at 52 weeks. The proportion of patients with clinical response was not statistically significantly different between adalimumab and placebo at eight weeks in ULTRA 1; however, there were statistically significantly greater proportions of adalimumab versus placebo patients with clinical responses at weeks 8 and 52 in ULTRA 2. Sustained clinical responses (clinical response at weeks 8, 32, and 52) were achieved by a statistically significantly greater proportion of adalimumab patients than placebo patients in ULTRA 2. Both clinical remission and clinical response rely on the Mayo scoring system. Although this instrument itself has been validated, it is not clear whether the thresholds used to determine remission and response have been clearly validated.

No colectomies occurred during the treatment phases of ULTRA 1 and ULTRA 2; however, there were colectomies performed during follow-up, and there was no obvious difference between groups. Statistically significant improvements in quality of life were seen on the IBDQ in ULTRA 1 at eight weeks and ULTRA 2 at eight weeks and 52 weeks, but results of the SF-36 were inconsistent, with no statistically significant improvement in either the PCS or MCS at eight weeks and a statistically significant improvement on only the PCS at 52 weeks. However, while IBDQ responder analyses were included in the hierarchical analysis plan, no statistical test should have been performed because a higher-ranked comparison failed. Change from baseline in IBDQ scores and SF-36 were not part of the hierarchical statistical testing plan, and therefore may be subject to false-positive findings due to multiple comparisons. All of these results were treated as exploratory analyses by CDR, and thus should be considered hypothesis-generating. The most common AE, SAE, and reason for WDAE in both studies was exacerbation of UC. The most common notable harms were injection site reactions and infections.

4.2 Interpretation of results

4.2.1 Efficacy

The manufacturer's listing criteria for adalimumab reflects the indication for adalimumab. The indication requires that patients have failed or are intolerant to prior standard therapy. This criterion appears to be

addressed with the inclusion criteria for both ULTRA 1 and ULTRA 2, as patients had to have Mayo scores in the moderate to severe range despite a trial of an oral corticosteroid, or thiopurines (azathioprine or 6-MP). Subgroup analyses did not report interaction *P* values, so it is not known whether there were statistically significant differences between groups based on baseline use of conventional therapies. In ULTRA 2, there was no difference in the proportion of patients achieving remission at eight weeks between adalimumab and placebo in patients who had previously used thiopurines, while there was a difference in patients who had not previously used thiopurines. However, this numerical difference in remission based on prior thiopurine use was no longer evident after 52 weeks, and, in ULTRA 1, there were no obvious differences in response based on prior thiopurine use. It is also not clear exactly how treatment failure is going to be defined in the clinical setting. Similarly, the product monograph states that patients who have not responded by eight weeks of therapy should not be continued on adalimumab. Once again, it is not clear how non-response will be defined in this case. If response is defined as clinical response, then in ULTRA 1 there was no statistically significant difference in clinical response between adalimumab and placebo at eight weeks.

Patients were excluded from ULTRA 1 if they had used prior biologics for UC; however, prior use was allowed in ULTRA 2, and approximately 40% of patients had received anti-TNF therapy at one time. Based on subgroup analyses, remission responses appeared to have been better in TNF-naive patients than in patients who had prior treatment with TNF inhibitors; however, no interaction *P* values were provided and no adjustments were made for multiple comparisons, so this data must be interpreted with caution. The listing criteria for adalimumab does not restrict its use to bio-naive patients, although according to the clinical expert it is unlikely that a patient who failed on one anti-TNF drug would be tried on another. However, the approval of vedolizumab in 2015 has added another biologic option for patients with UC. Vedolizumab is an integrin inhibitor, and therefore presents an alternative mechanism of action to the TNF inhibitors. Vedolizumab has been reviewed by CDR, and is indicated for patients with moderate to severe UC who have had an inadequate response, or loss of response, to either conventional therapies or infliximab. Therefore, under the current indication, vedolizumab would be considered an option after TNF inhibitors. According to the clinical expert, vedolizumab may soon be considered as a second-line option to conventional therapies, ahead of the TNF inhibitors.

Clinical remission was the primary outcome of both ULTRA 1 and ULTRA 2. This outcome, as well as clinical response, a key efficacy outcome of this review, both rely on the Mayo scoring system (see Appendix 6 for review). The Mayo scoring system takes into account a number of the symptoms identified by patients as being important in their input to CDR, including frequency of bowel movements. Remission is defined by an improvement across a certain threshold in Mayo score, while response is defined by a specific change in Mayo score; however, it is not clear whether these definitions of remission and response have been validated. Although Mayo itself is widely used and has been validated, there are limitations associated with its use. For instance, the physician assessment component is subjective, and the PGA double-counts some of the symptoms in the scale. Additionally, the SFS might not necessarily be an accurate reflection of disease activity, as the number of stools per day that would be considered “normal” can vary widely.

All of the studies included in this review were placebo-controlled. Adalimumab is the third anti-TNF monoclonal antibody to be approved for use in UC, following infliximab and golimumab, yet there is a lack of trials comparing these drugs directly. Of the three anti-TNF biologics, infliximab is the oldest and uses a chimeric monoclonal antibody, while golimumab and adalimumab are human sourced. Accordingly, infliximab might be expected to have a higher risk of immune reactions; however, this has not been established in a direct, head-to-head comparison. Many of the other adverse effects

associated with anti-TNF may be related to the blockade of TNF, most notably malignancy; therefore, there is no clear mechanistic rationale for an advantage of one drug over the other with respect to key harms such as this. According to the clinical expert, an important distinction between the TNF inhibitors is route of administration. Infliximab is delivered via IV infusion, and is almost always the anti-TNF used for patients who are in hospital, and thus are likely in a severe stage of their disease. Because it is delivered by IV infusion, infliximab may be less suitable for patients who live in a rural setting, or who have mobility issues; therefore, adalimumab and golimumab would likely be more appropriate in this setting. Accordingly, the clinical expert believes that the most relevant comparator for adalimumab is golimumab, as these two drugs would tend to be used by the same patients, i.e., by those for whom infliximab is deemed inappropriate. A summary of indirect comparisons of TNF inhibitors is found in Appendix 7 of this report. Given the numerous limitations and inconsistent findings between the indirect comparisons (IDCs), the most conservative conclusion based on the available indirect evidence is that there is no clear evidence that there is a difference between the biologics, including adalimumab, with respect to inducing and maintaining remission, response, and mucosal healing in patients with moderately to severely active UC.

4.2.2 Harms

There were no clear or consistent differences in the risks of notable harms such as malignancies or opportunistic infections such as tuberculosis (TB). As noted earlier, none of the included studies were likely of sufficient duration to assess risk of malignancies; therefore, a finding of no difference in risk is not surprising. There has been a long-standing debate over the potential connection between TNF inhibitors and development of malignancy. TNF is believed to play a role in cancer surveillance by the immune system; therefore, it has long been theorized that long-term inhibition might lead to development of cancer. Over many years of experience with these drugs, the malignancy that is most commonly noted as carrying elevated risk is lymphoma. Hepatosplenic T-cell lymphoma is singled out under serious warnings and precautions in the product monograph for adalimumab, and lymphoma and “other malignancies” is listed under risks of pediatric malignancies. Data from ULTRA 3, a long-term OL extension of studies ULTRA 1 and ULTRA 2, are summarized in Appendix 6.There were no obvious issues with respect to malignancies noted.

Infections are also listed as a safety issue in the adalimumab product monograph. Of the many opportunistic infections, TB is often singled out as a key risk of therapy. In the included studies, an adalimumab patient died of TB in the Suzuki et al. study. The case was reported 44 days after the last dose of study drug in that patient; it is not clear whether the patient developed TB while on adalimumab, but the patient had tested negative with respect to induration and chest X-ray at baseline.[5] In the product monograph, it is recommended that patients who have been exposed to TB or have travelled to an endemic area carefully assess the risk versus benefit of therapy. The product monograph also recommends that any patient who develops an infection while on therapy should be monitored closely, and that treatment should be discontinued in any patient who develops a serious infection.[2]

Other notable harms include injection site reactions and hypersensitivity reactions. Not surprisingly, injection site reactions were more common with adalimumab than with placebo. Adalimumab is a monoclonal antibody, and injection site reactions and hypersensitivity are known risks associated with these drugs. Injection site reactions did appear to occur more often with adalimumab than placebo, although no definitive comparisons can be made as the studies were not powered to assess this outcome. Patients are likely at highest risk for hypersensitivity reactions as they initiate therapy; thus, analysis of this AE is not likely to have benefited from a longer follow-up. However, because a

hypersensitivity reaction is a relatively rare AE, a larger sample might have been better able to characterize the extent of the risk with adalimumab, if any.

4.3 Potential place in therapy[2]

The clinical expert involved in the review confirmed that anti-TNF therapy is an important component in the treatment of UC patients. There is currently an unmet need in both acute and chronic care, particularly in patients who are:

- acutely ill and fail to respond within 72 hours to corticosteroids. The clinical expert stated that currently IV infliximab is typically used in this situation. There is a small unmet need for an anti-TNF that can be administered via the SC route when IV access is not available. The clinical expert noted that this is uncommon, but does occur. It is possible that a patient may have previously responded to infliximab, but subsequently fails and flares, therefore, another option for anti-TNF therapy would be useful for such patients.
- chronically ill and have failed treatment with anti-inflammatory and immunosuppressant drugs and who:
    - are intolerant of infliximab with infusion reactions such as headache or arthralgia
    - have failed on infliximab
    - do not have IV access
    - live far from an infusion centre

The clinical expert consulted by CDR described the patient populations in ULTRA 1 and ULTRA 2 as having moderate to severe active UC defined both clinically and endoscopically. Hence, they represent a group of ill patients with UC, especially given that their disease is active despite steroid use. Adalimumab, according to the clinical expert, may meet the needs outlined: it is administered by SC injection and thus avoids the need for IV therapy; and in less acutely ill patients, adalimumab may be useful if the patient has failed on infliximab (defined by symptoms, blood testing, and endoscopy), does not have IV access, or lives in a remote community with limited or no access to an infusion centre. Access to an infusion centre is not required for adalimumab, as the drug is self-administered, making it suitable for use by patients living in remote areas.

Of the two studies, only ULTRA 2 allowed inclusion of patients with prior anti-TNF use. There remains uncertainty as to the use of adalimumab in this subgroup of patients. According to the clinical expert, infliximab is likely still the first-line therapy for acutely ill, hospitalized patients. The decision on which drug to use in ambulatory settings will likely be determined case by case.

The lack of head-to-head studies versus other biologics, including vedolizumab, remains a limitation in understanding the exact position of adalimumab among treatment options.

Adalimumab may be discontinued in a certain proportion of patients who achieve remission. Generally, according to the clinical expert, this is often due to patient preference. The decision to discontinue anti-TNF therapy varies according to the severity of the index flare and discontinuation criteria are typically based on Mayo score and endoscopic score of 0.

[2] 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 DB RCTs met the inclusion criteria for this review, all of which compared adalimumab with placebo. ULTRA 1 was an induction study, with an eight-week, DB phase followed by an OL phase that continued out to 52 weeks. ULTRA 2 and Suzuki et al. were induction plus maintenance treatment studies, with 52-week, DB phases. Clinical remission was the primary outcome in ULTRA 1 and ULTRA 2, for which adalimumab was shown to be superior to placebo in each study. The authors of Suzuki et al. described analyses in that study as being exploratory; however, although remission was not statistically significantly improved for adalimumab over placebo at eight weeks, it was statistically significantly improved at 52 weeks. Clinical responses were not statistically significantly improved for adalimumab over placebo at eight weeks in ULTRA 1, but were statistically significantly improved for adalimumab over placebo at eight weeks and 52 weeks in ULTRA 2. Clinical responses were similar between ULTRA 2 and Suzuki et al. Quality of life was not consistently improved for adalimumab over placebo on the IBDQ and SF-36 at eight weeks, but statistically significant differences were observed for adalimumab compared to placebo on the IBDQ and SF-36 PCS at 52 weeks; however, these analyses were deemed exploratory by CDR and should be considered hypothesis-generating. There were no obvious differences in overall harms between adalimumab and placebo.

APPENDIX 1: PATIENT INPUT SUMMARY

This section was summarized by CDR staff based on the input provided by patient groups. It has not been systematically reviewed. It has been reviewed by the submitting patient groups.

1. Brief description of patient groups supplying input

Crohn's and Colitis Canada is a volunteer-based national charity that aims to both find a cure for Crohn disease and ulcerative colitis (UC) and improve the lives of those affected by these diseases. In the fiscal year 2013-2014, Crohn's and Colitis Canada received less than 10% of total revenues from the following manufacturers: AbbVie, Aptalis, Celltrion, Ferring, Janssen, Shire, Takeda, Vertex, and Warner Chilcott. The funds are used to help sponsor patient education events, research and medical conferences, educational brochures, kids' camps, and post-secondary scholarships for inflammatory bowel disease (IBD) patients.

The Gastrointestinal (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. The GI Society has two websites, one in English (www.badgut.org) and one in French (www.mauxdeventre.org). In the last two years, the GI Society has received funding from AbbVie Corporation, Actavis/Allergan Canada Inc., AstraZeneca Canada Inc., Canada's Research-Based Pharmaceutical Companies (Rx&D), Ferring Inc., Gilead Sciences Canada Inc., GlaxoSmithKline Inc., Janssen Canada, Merck Canada Inc., Pfizer Canada Inc., and Takeda Canada Inc.

Crohn's and Colitis Canada and The GI Society declared no conflicts of interest with regard to the preparation of their submissions.

2. Condition-related information

Information was obtained through telephone interviews, a 2011 national online survey, a questionnaire, one-to-one conversations, round tables, discussions with health care professionals, and Crohn's and Colitis Canada published reports.

Ulcerative colitis (UC) is a serious IBD with no cure that is characterized by fine ulcerations in the inner mucosal lining of the large intestine. These ulcerations subsequently cause inflammation that extends varying distances upward from the anus. The highest occurrence of UC is in young children, and then peaks again around 40 to 50 years of age. If left untreated, long-standing UC can lead to colon cancer.

Patients with UC experience urgent and frequent bowel movements (usually between five and 20 per day, but sometimes even more) during flare-ups in the active stage of the disease. Rectal bleeding and bloody diarrhea (which can lead to anemia in severe cases), cramping, abdominal pain, fatigue, and fluctuations in weight are other symptoms. In addition, patients may experience extra-intestinal manifestations of UC, such as fever, arthritis, mouth or skin ulcers, tender and inflamed nodules on the shins, and other systemic disease symptoms. All of these symptoms can profoundly affect a person's physical, emotional, and social well-being by causing anxiety and stress (at having to face the uncertainty of where and when they will experience a bowel movement or regarding the associated financial burdens), limiting the places they can go and the activities (including work) they can participate

in. The stigma and symptoms are particularly problematic for children, as they rarely get to experience a "normal life." Patients noted some of the following effects of UC:

- *"There is no quality of life. Virtually any activity that would take an individual away from the bathroom cannot be done."*
- *"I have low energy. I'm tired often. My employer does not understand and it affects my attendance."*

3. Current therapy-related information

Managing UC is multi-faceted, and involves managing the symptoms of UC along with targeting the underlying inflammation. At present, there are a limited number of treatment options for people with UC, even fewer than for Crohn disease. Aminosalicylates (e.g., mesalamine) are first-line drugs, followed by immunosuppressants (e.g., azathioprine) and/or steroids (e.g., prednisone or other corticosteroids) if remission is not achieved or if the condition becomes worse. While these drugs can be effective in patients with mild to moderate disease, they often do not maintain remission in patients with severe forms of UC. Patients have indicated that sustained remission/treatment response is more important than relieving any one symptom of UC.

Biologics (infliximab, golimumab, and now adalimumab) are approved in Canada for treatment of UC. While they come with a number of potential side effects and risk factors, they provide an option before surgery when first- and second-line therapies have failed. The majority of surveyed patients said they would rather receive a biologic medication, despite the potential risks and side effects, than get a colectomy. Drug coverage is a concern, given the inequalities of access to treatment across Canada. Many UC patients do not have private insurance, and costs of the prescribed biologics are prohibitive. One patient noted, *"While it is true that these drugs are exorbitantly expensive, in the long run they are less costly for society than the alternative, which is other numerous health care expenses, surgeries, and hospital stays, as well as lost work productivity and long-term disability funding."* The use of adalimumab increases remission rates, enabling patients to return to work more quickly and to be productive, tax-paying citizens. Some interviewees claimed that they are deliberately being misdiagnosed as living with Crohn disease in order to get access to the greater variety of biologic treatments that are currently unavailable in their province or territory.

4. Expectations about the drug being reviewed

Patients living with UC feel that the introduction of biologics will potentially normalize and improve their quality of life. Patients expect a decrease in the number of surgeries required as well as a fall in hospitalization rates. They see Humira as a valuable next step when other medications do not work well or when patients cannot tolerate them. Women living with IBD hope that with the introduction of new biologic treatments, they can have the same fertility rates as women in the general population when their disease is in remission.

Of the patients who have experience with Humira, many noticed improved health and well-being. Home, school, and social lives were all improved along with their general emotional well-being; however, anxiety over the cost of this treatment remained. Successful treatment was echoed through the following quote, *"Wish I had had this option eight years earlier. I have been able to have a life."* Patients also remarked that Humira is more administration-friendly as it is a self-injection that can be administered at home and does not require travel to a hospital or clinic. Adherence issues are, and will continue to be, less of a problem.

While patients expect more freedom and independence while on Humira, there is apprehension about its costs and the access to it. They hope that it will be available on the public drug formularies so that there are fewer barriers to the treatment they may require, should other types of treatment be unsuccessful.

APPENDIX 2: LITERATURE SEARCH STRATEGY

OVERVIEW	
Interface:	Ovid
Databases:	Embase 1974 to present MEDLINE Daily and MEDLINE 1946 to present MEDLINE In-Process & Other Non-Indexed Citations **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	November 12, 2015
Alerts:	Weekly search updates until March 16 2016
Study Types:	No search filters were applied
Limits:	No date or language limits were used Human filter was applied Conference abstracts were excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
.sh	At the end of a phrase, searches the phrase as a subject heading
MeSH	Medical Subject Heading
fs	Floating subheading
exp	Explode a subject heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
#	Truncation symbol for one character
?	Truncation symbol for one or no characters only
adj	Requires words are adjacent to each other (in any order)
adj#	Adjacency within # number of words (in any order)
.ti	Title
.ab	Abstract
.ot	Original title
.hw	Heading word; usually includes subject headings and controlled vocabulary
.pt	Publication type
.po	Population group [PsycInfo only]
.rn	CAS registry number
.nm	Name of substance word
pmez	Ovid database code; 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	
Embase, Ovid MEDLINE(R)	
#	**Searches**
1	(humira* or adalimumab* or D2E7 or trudexa* or hsdb 7851 or hsdb7851 or lu200134).ti,ab,ot,hw,rn,nm,kf.
2	(331731-18-1 or FYS6T7F842).rn,nm.
3	1 or 2
4	Colitis, ulcerative/
5	(colitis or proctocolitis or rectocolitis or proctitis or colorectitis or rectosigmoiditis or proctosigmoiditis).ti,ab.
6	4 or 5
7	3 and 6
8	7 use pmez
9	*adalimumab/
10	(humira* or adalimumab* or D2E7 or trudexa* or hsdb 7851 or hsdb7851 or lu200134).ti,ab.
11	9 or 10
12	*ulcerative colitis/
13	5 or 12
14	11 and 13
15	14 use oemezd
16	8 or 15
17	conference abstract.pt.
18	16 not 17
19	remove duplicates from 18

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:	November 2015
Keywords:	Humira/adalimumab and colitis
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 evidence-based searching" (https://www.cadth.ca/resources/finding-evidence/grey-matters-practical-search-tool-evidence-based-medicine) 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
Clinical Study Report: M10-223[22]	Not a RCT

APPENDIX 4: DETAILED OUTCOME DATA

TABLE 16: OTHER OUTCOMES: INDUCTION STUDY

	ULTRA 1	
Outcome	**Adalimumab N = 130**	**Placebo N = 130**
Mucosal healing		
ITT, n (%)	61 (47)	54 (42)
P value	*P* = 0.382	
Steroid-free remission, n (%)		
Week 8	25 (22) n = 113	14 (11) n = 130
Week 52	52 (46)	55 (42)
Weeks 8 and 52	20 (18)	10 (8)
Corticosteroid-free for ≥ 90 days at week 52	50 (47) n = 106	52 (43) n = 122
P value	NR	
Dose escalation		
Patients requiring, n (%)	NR (22.9)	NR (31.1)
WPAI–Change in work time missed		
Not investigated		

ITT = intention-to-treat; NR = not reported; WPAI = Work Productivity and Impairment Questionnaire.
Note: Mucosal healing: *P* value for differences between active treatment group and placebo from chi-square test (or Fisher's exact test if ≥ 20% of the cell have an expected count < 5).
Source: Clinical Study Report for ULTRA 1.[3]

TABLE 17: OTHER OUTCOMES: INDUCTION/MAINTENANCE STUDIES

	ULTRA 2		Suzuki et al. 2014	
Mucosal healing, n (%)	**Adalimumab N = 248**	**Placebo N = 246**	**Adalimumab N = 90**	**Placebo N = 96**
Week 8,	102 (41)	78 (32)	NR (44)	NR (30)
	P = 0.032		*P* = 0.045	
Week 52	62 (25)	38 (15)	NR (29)	NR (16)
	P = 0.009		*P* = 0.015	
Steroid-free remission				
DC CS before week 52 and achieved remission at week 52, n (%)	20 (13)	8 (6)		
P value	*P* = 0.035			
DC CS use for ≥ 90 days before week 52 and achieved remission at week 52, n (%)	20 (13)	8 (6)		
P value	*P* = 0.033			
DC CS use and achieved sustained remission at both weeks 32 and 52, n (%)	15 (10)	2 (1)		
P value	*P* = 0.002			
Patients, n/N (%)			17/120 (14)[a]	4/58 (7)[a]
P value				
Dose escalation				
Patients requiring, n (%)	68 (27)	84 (34)		
WPAI–Change in work time missed				
% of work time missed, Mean (SD) baseline	21.1 (NR) N = 153	22.5 (NR) N = 140		
Mean (SD) change from baseline to week 52	–9.2 (40)	–4.8 (35)		
LSM [95% CI] between group change vs. placebo	–5.28 (–11.54, 0.98), *P* = 0.098			

CS = corticosteroid; DC = discontinued; LSM = least squares mean; NR = not reported; SD = standard deviation.
[a] The percentages were calculated based on the baseline number of patients taking steroids.
Note: *P* value to compare mean change from baseline between adalimumab and placebo from analysis of covariance (ANCOVA) with treatment and prior anti-TNF status as factors and baseline value as covariate.
Source: Clinical Study Report for ULTRA 2;[4] Suzuki et al.[5]

TABLE 18: MEDICATIONS TAKEN AT/PRIOR TO BASELINE

ULTRA 1		
Medication	**Adalimumab N = 130**	**Placebo N = 130**
Medications taken at baseline, n (%)		
UC-related	121 (93)	125 (96)
Any CS	71 (55)	89 (69)
Any AZA, 6-MP	51 (39)	52 (40)
Any aminosalicylate	105 (81)	98 (75)
Current medications (within 90 days of baseline), n (%)		
UC-related	[redacted]	[redacted]
Any CS	[redacted]	[redacted]
Any AZA, 6-MP	[redacted]	[redacted]
Any aminosalicylate	[redacted]	[redacted]
Any anti-TNF	[redacted]	[redacted]
ULTRA 2		
Medication	**Adalimumab N = 248**	**Placebo N = 246**
Medications taken at baseline, n (%)		
UC-related	224 (90)	218 (89)
Any CS	150 (61)	140 (57)
Any AZA, 6-MP	93 (38)	80 (33)
Any aminosalicylate	146 (59)	155 (63)
Any anti-TNF	NR	NR
Current medications (within 90 days of screening), n (%)		
UC-related	[redacted]	[redacted]
Any CS	[redacted]	[redacted]
Any AZA, 6-MP	[redacted]	[redacted]
Any aminosalicylate	[redacted]	[redacted]
Any anti-TNF	[redacted]	[redacted]
Suzuki et al. 2014		
Medication	**Adalimumab N = 90**	**Placebo N = 96**
Medications taken at baseline, n (%)		
UC-related		
Any SCS	57 (63.3)	58 (60.4)
Any AZA, 6-MP	41 (45.6)	52 (54.2)
5-ASAs	83 (92.2)	89 (92.7)

5-ASA = 5-aminosalicylic acid; 6-MP = 6-mercaptopurine; AZA = azathioprine; CS = corticosteroid; SCS = systemic corticosteroids; UC = ulcerative colitis.
Source: Clinical Study Report for ULTRA 1;[3] Clinical Study Report for ULTRA 2;[4] Suzuki et al.[5]

TABLE 19: COMPLIANCE

	ULTRA 1	
	Adalimumab N = 130	**Placebo N = 130**
Mean (SD), %	[redacted]	[redacted]
	ULTRA 2	
	Adalimumab N = 247	**Placebo N = 246**
Mean (SD), %	[redacted]	[redacted]

SD = standard deviation.
Source: Clinical Study Report for ULTRA 1;[3] Clinical Study Report for ULTRA 2.[4]

Subgroups

The dose escalation analysis set comprised patients in the intention-to-treat extended (ITT-E) analysis set who required dose escalation to adalimumab 40 mg every week. Per protocol, patients defined as inadequate responders were permitted to dose escalate from adalimumab 40 mg every other week to 40 mg weekly at or after week 12 (A3 population). To describe the total clinical response (or remission) at week 52 including the dose escalation portion of the study, the following post hoc modified non-responder imputation (mNRI) imputation method was also used. The mNRI imputation method considered all missing response (or remission) values as non-response (or non-remission). However, patients who dose escalated to adalimumab 40 mg weekly were considered as responders (or remitters) or non-responders (or non-remitters) according to their response (or remission) status after their dose escalation.

TABLE 20: SUBGROUPS: INDUCTION STUDY — PATIENTS ACHIEVING REMISSION

	ULTRA 1	
Dose escalation, n (%)	**Adalimumab 160 mg/80 mg/40 mg**	**Placebo**
Dose Escalators, week 52,	[redacted]	[redacted]
Overall, week 52	[redacted]	[redacted]
Anti-TNF	[redacted]	[redacted]
AZA/6-MP use at baseline, n (%)		
Yes	[redacted]	[redacted]
No	[redacted]	[redacted]
Aminosalicylate use at baseline, n (%)		
Yes	[redacted]	[redacted]
No	[redacted]	[redacted]

6-MP = 6-mercaptopurine; AZA = azathioprine; ITT = intention-to-treat; NR = not reported; TNF = tumour necrosis factor.
Source: Clinical Study Report for ULTRA 1.[3]

TABLE 21: SUBGROUPS: INDUCTION/MAINTENANCE STUDY

	ULTRA 2	
Remission based on prior TNF	**Adalimumab 160 mg/80 mg/40 mg**	**Placebo**
No prior anti-TNF — week 8	32/150 (21) *P* = 0.017	16/145 (11)
— week 52	33/150 (22) *P* = 0.029	18/145 (12)
Prior anti-TNF — week 8	9/98 (9) *P* = 0.559	7/101 (7)
— week 52	10/98 (10) *P* = 0.039	3/101 (3)
Remission based on dose escalation, n (%)		
Dose escalators, week 52	[redacted]	[redacted]
Overall, week 52	43 (17)	21 (9)
Remission based on AZA/6-MP use at baseline, n (%)		
Yes, week 8	[redacted]	[redacted]
No, week 8	[redacted]	[redacted]
Yes, week 52	[redacted]	[redacted]
No, week 52	[redacted]	[redacted]
Remission based on AZA/6-MP or CS use at baseline, n (%)		
Yes, week 8	[redacted]	[redacted]
No, week 8	[redacted]	[redacted]
Yes, week 52	[redacted]	[redacted]
No, week 52	[redacted]	[redacted]
Remission based on SCS use at baseline, n (%)		
Yes, week 8	[redacted]	[redacted]
No, week 8	[redacted]	[redacted]
Yes, week 52	[redacted]	[redacted]
No, week 52	[redacted]	[redacted]

6-MP = 6-mercaptopurine; AZA = azathioprine; CS = corticosteroids; ITT = intention-to-treat; SCS = systemic corticosteroids; TNF = tumour necrosis factor.
Source: Clinical Study Report for ULTRA 2.[4]

TABLE 22: COMPARISON OF POPULATIONS IN ULTRA 1

	ULTRA 1-A3		ULTRA 1-ITT-E	
	Adalimumab 160 mg/80 mg/40 mg N = 130	**Placebo N = 130**	**Adalimumab 160/80 mg/40 mg N = 223**	**Placebo N = 222**
Mean (SD) age, years	38.2 (13.5)	38.9 (12.7)	38.5 (13.1)	39.7 (12.7)
Male gender, n (%)	83 (64)	82 (63)	138 (62)	139 (63)
Ethnicity, n (%)				
White	119 (92)	117 (90)	206 (92)	202 (91)
Black	2 (2)	5 (4)	4 (2)	9 (4)
Asian	7 (5)	5 (4)	10 (5)	7 (3)
Hispanic/Latino	4 (3)	5 (4)	13 (6)	7(3)
Other	2 (2)	3 (2)	3 (1)	4 (2)
Nicotine use, n (%)				
Current	12 (9)	7 (5)	23 (10)	10 (5)
Former	37 (29)	35 (27)	63 (28)	65 (29)
Non-user	81 (62)	88 (68)	137 (61)	147 (66)
Weight, mean (SD) kg	75.5 (14.2)	78.7 (17.4)	73.9 (13.8)	78.4 (18.1)
Alcohol use, n (%)				
Drinker	62 (48)	57 (44)	108 (48)	102 (46)
Former drinker	9 (7)	7 (5)	12 (5)	14 (6)
Non-drinker	59 (45)	66 (51)	103 (46)	106(48)
Duration of UC, mean (SD) years	8.1 (7.2)	7.5 (7.2)	8.4 (7.3)	7.9 (7.5)
Site of UC, n (%)				
Pancolitis	60 (46)	73 (56)	114 (51)	132 (60)
Descending colon	61 (47)	42 (32)	87 (39)	67 (30)
Other	9 (7)	15 (12)	22 (10)	23 (10)
Baseline Mayo score, mean (SD)	8.8 (1.6)	8.7(1.6)	8.9 (1.7)	8.8 (1.6)
Evidence of dysplasia/malignancy, n (%)				
Yes	0 (0)	0 (0)	0 (0)	1 (1)
No	130 (100)	130 (100)	223 (100)	218 (> 99)
Missing	0 (0)	0 (0)	0 (0)	3 (1.4)

SD = standard deviation; UC = ulcerative colitis.
Source: Clinical Study Report for ULTRA 1.[3]

TABLE 23: COMPARISON OF EFFICACY DATA FROM POPULATIONS IN ULTRA 1

	ULTRA 1-A3		ULTRA 1-ITT-E	
Remission	**Adalimumab 160 mg/ 80 mg/40 mg**	**Placebo**	**Adalimumab 160 mg/80 mg/40 mg N = 223**	**Placebo N = 222**
ITT, Week 8, N (%)	24 (19)	12 (9)	35 (16) N = 223	16 (7) N = 222
P value	*P* = 0.031		*P* = 0.005	
ITT-A3, LOCF	24 (19) N = 124	12 (10) N = 123	-	-
Clinical Response				
Patients, week 8, n (%)	71 (55)	58 (45)	116 (52)	95 (43)
P value	*P* = 0.107		NR	
IBDQ				
Responders, week 8, N (%)	79 (61)	75 (58)	130 (58)	128 (58)
P value	*P* = 0.614		NR	
Mean (SD) baseline	131.9	125.3	NR	NR
Mean (SD) change at week 8	35.9 (34.0) N = 117	26.9 (35.6) N = 124	NR	NR
P value	*P* = 0.008		NR	
SF-36 PCS				
Mean (SD) baseline	42.1 N = 125	40.2 N = 124	NR	NR
Mean (SD) change from baseline, week 8	6.5 (7.3)	3.8 (7.2)	NR	NR
P value	*P* < 0.001		NR	
SF-36 MCS				
Mean (SD) baseline	36.6 N = 125	36.5 N = 124	NR	NR
Mean (SD) change from baseline, week 8	7.6 (12.2)	5.8 (11.4)	NR	NR
P value	*P* = 0.145		NR	
Need for colectomy, n (%)				
Week 8, patients	0 (0)	0 (0)	NR	NR
OL period through week 52	0 (0)	0 (0)	NR	NR
After last dose	4 (2)	8 (4)	NR	NR

IBDQ = Inflammatory Bowel Disease Questionnaire; ITT = intention-to-treat; LOCF = last observation carried forward; MCS = Mental Component Summary; NR = not reported; OL = open-label; PCS = Physical Component Summary; SD = standard deviation.

Notes: Remission: *P* values for adalimumab versus placebo in ITT-A3 set (non-responder imputation [NRI] and LOCF analyses) and placebo set from chi-square test (or Fisher's exact test if ≥ 20% of cells had expected cell count < 5). For patients in the ITT-E set, the *P* value to compare adalimumab 160 mg/80 mg/40 mg versus placebo is from Cochran-Mantel-Haenszel test with patients in/not in the ITT-A3 set as the stratification factor; and the *P* value to compare adalimumab 80 mg/40 mg versus placebo is from chi-square test (or Fisher's exact test if ≥ 20% of cells had expected cell count < 5).

Clinical response, IBDQ: *P* value for differences between active treatment group and placebo from chi-square test (or Fisher's exact test if ≥ 20% of the cell have an expected count < 5).

IBDQ response was defined as an increase in IBDQ ≥ 16 points from baseline IBDQ, mean change: *P* values for adalimumab vs. placebo from one-way analysis of covariance (ANCOVA) with treatment as factor and baseline value as covariate.

SF-36: *P* values for adalimumab vs. placebo from one-way ANCOVA with treatment as factor and baseline value as covariate.

Source: Clinical Study Report (CSR) 238, CSR 264, CSR 334, and CSR 304.

APPENDIX 5: VALIDITY OF OUTCOME MEASURES

Aim

To summarize evidence concerning the reliability, validity, scoring, and minimal clinically important difference (MCID) of the following scales used to assess changes in ulcerative colitis (UC) disease activity, and outcome measurement in the clinical trials:

- Mayo scoring system
- Inflammatory Bowel Disease Questionnaire (IBDQ)
- Short Form (36) Health Survey (SF-36): Medical Outcomes Study

Findings

Mayo scoring system

The Mayo score is one of the most commonly used disease activity indices in placebo-controlled trials in UC. In its complete form, it is composed of four parts: rectal bleeding, stool frequency, physician assessment, and endoscopy appearance. Each part is rated from 0 to 3, giving a total score of 0 to 12. A score of 3 to 5 points indicates mildly active disease, a score of 6 to 10 points indicates moderately active disease, and a score of 11 to 12 points indicates severely active disease. Two abridged versions, the partial Mayo score that excludes the endoscopy subscore and the non-invasive six-point score comprising only the rectal bleeding and stool frequency portions, have been developed and validated.[23] The Mayo score and the partial Mayo score have been demonstrated to correlate with patient assessment of change in UC activity.[23] Lewis et al. reported a reduction of ≥ 3 points on the Mayo score and the partial Mayo score to constitute a clinically meaningful change.[23] Lewis et al. also recommended that clinical remission of UC be defined using a Mayo score of ≤ 2 points.[23]

Although the Mayo score is a widely recognized UC activity index and is accepted by regulatory bodies, including Health Canada and the US FDA, it may not be optimal. Cooney et al. argue that two components of the Mayo score, the Physician Global Assessment (PGA) and the sigmoidoscopy subscore, are subjective and introduce variability and a lack of precision into the index. The PGA also includes a sigmoidoscopy score that introduces double-counts of some elements.[24]

Inflammatory Bowel Disease Questionnaire

The IBDQ was developed by Guyatt et al.[25] as a physician-administered questionnaire and it is widely used to assess health-related quality of life (HRQoL) in patients with inflammatory bowel disease (IBD) (UC and Crohn disease).[26] It is a 32-item Likert-based questionnaire divided into four dimensions: bowel symptoms (10 items), systemic symptoms (5 items), emotional function (12 items), and social function (5 items). Responses to each question are graded from 1 to 7 (1 being the worst situation and 7 the best). Therefore, the total IBDQ score ranges between 32 and 224, with higher scores representing better quality of life. The scores of patients in remission usually range from 170 to 190. An increase in IBDQ score of 16 to 32 points constitutes the upper and lower bounds of the clinically meaningful improvement in HRQoL in patients with Crohn disease.[21] Information on whether this correlation between score and levels of clinical improvement translates directly to UC was not available through the literature search for this summary.

A systematic review[21] of nine validation studies on the IBDQ for UC reported that, in seven of the studies, the IBDQ was able to differentiate clinically important differences between patients with disease remission and patients with disease relapse, by demonstrating significant differences in score.[26] The IBDQ can also discriminate changes in the social and emotional state of patients; however, the correlation of this dimension with disease activity is not as high as the correlation with remission of

bowel symptoms.[26] The IBDQ also demonstrated high test-retest reliability in all the four IBDQ dimensional scores. Six studies evaluated IBDQ for sensitivity to change, and all suggested it to be a sensitive instrument to quantify changes in HRQoL relative to clinical activity changes in UC.[26]

Short Form (36) Health Survey: Medical Outcomes Study

The SF-36 is a 36-item, general health status instrument that has been used extensively in clinical trials in many disease areas.[27] The SF-36 consists of eight health domains: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health.[28] For each of the eight categories, a subscale score can be calculated. The SF-36 also provides two component summaries, the Physical Component Summary (PCS) and the Mental Component Summary (MCS). The PCS and MCS scores range from 0 to 100, with higher scores indicating better health status. The summary scales are scored using norm-based methods, with regression weights and constants derived from the general US population. Both the PCS and MCS scales are transformed to have a mean of 50 and a standard deviation (SD) of 10 in the general US population. Therefore, all scores above/below 50 are considered above/below average for the general US population. In patients with either UC or Crohn disease, the SF-36 showed good discriminant ability and had satisfactory reliability.[29] While reliability was satisfactory, the authors did observe substantial floor effects within the role-physical and role-emotional dimensions, underscoring the lack of sensitivity of the scale to detect small changes in certain groups of patients.[29] For patients with UC and patients with Crohn disease, high ceiling effects along with low responsiveness scores (obtained using the Guyatt statistic) indicate some limitations associated with the ability of SF-36 to detect either deterioration or improvement over periods of time, particularly in longitudinal studies.[29]

The MCID for either the PCS or MCS of the SF-36 is typically between 2.5 and 5 points,[18-20] while in Crohn disease the PCS and MCS MCIDs were estimated to range between 1.6 and 7.0 and 2.3 to 8.7, respectively, using various distribution- and anchor-based approaches.[17] No MCID was identified in UC.

Summary

The Mayo score and the partial Mayo score are commonly used disease activity indices in placebo-controlled trials in UC. Both have demonstrated correlation with patient assessment of change in UC activity. Mild, moderate, and severe disease activity are indicated by score ranges of 3 to 5 points, 6 to 10 points, and 11 to 12 points, respectively. Lewis et al. reported that a reduction of ≥ 3 points on the Mayo score and the partial Mayo score reflect a clinically meaningful change.[23]

The IBDQ is a physician-administered, 32-item questionnaire used to assess HRQoL in patients with IBD (UC and Crohn disease).[26] It evaluates bowel and systemic symptoms, as well as emotional and social functions. Responses to each question are graded from 1 to 7, with the overall score ranging from 32 (very poor HRQoL) to 224 (perfect HRQoL). Patients in symptomatic remission usually have a score of 170 or greater. An increase in IBDQ score of 16 to 32 points constitutes the upper and lower bounds of the clinically meaningful improvement in HRQoL in patients with Crohn disease.

The SF-36 is a 36-item, general health status instrument that has been used extensively in clinical trials in many disease areas,[27] with an MCID generally ranging between 2.5 and 5 points.[18-20] With regard to Crohn disease, however, the estimates of MCID for the PCS and MCS ranged between 1.6 and 8.7, while no MCID was identified in UC. The SF-36 was found to have good discriminant ability and satisfactory reliability; however, due to high floor and ceiling effects and low scores obtained for responsiveness using the Guyatt statistic, the SF-36 might be limited in its ability to detect either deterioration or improvement over time in patients with UC and Crohn disease.

APPENDIX 6: SUMMARY OF OTHER STUDIES

Objective

To summarize the study design and results of the open-label (OL) extension study ULTRA 3 (M10-223). The following summary is based on published data.[30]

Findings

Study design

ULTRA 3 (M10-223) is an ongoing extension study that continues to assess the treatment of adalimumab in patients with ulcerative colitis (UC) for up to approximately four years (208 weeks of treatment) who were enrolled in and completed ULTRA 1 (M06-826) and ULTRA 2 (M06-827). All patients who entered the extension study were on adalimumab 40 mg every other week or every week (for inadequate responders), depending on their treatment response or lack thereof. Inadequate response was defined as:

- "Subject with a Baseline Partial Mayo Score of 4 to 7 who presents with a Partial Mayo Score greater than or equal to their Baseline score on two consecutive visits at least 14 days apart."[30]
- "Subject with a Baseline Partial Mayo Score of 8 or 9 who presents with a Partial Mayo Score ≥ 7 on two consecutive visits at least 14 days apart."[30]

Patients from the blinded cohort or those who received every-other-week dosing were permitted to escalate the adalimumab dose to 40 mg every week after week 12 in ULTRA 3 if their response was deemed inadequate or if they had a disease flare (defined as partial Mayo score difference ≥ 3 compared with the baseline partial Mayo score of ULTRA 3 on two consecutive visits with least 14 days between visits). In addition, dose escalation to 40 mg every week was permitted in the second week of ULTRA 3 in those patients who had been in the OL cohort receiving every-other-week dosing. If patients were experiencing clinical response, corticosteroid tapering was permitted after week 12 of ULTRA 3. Additionally, patients could continue corticosteroid tapering in ULTRA 3 if they had already begun this process in ULTRA 1 and ULTRA 2.[30]

Long-term efficacy of adalimumab was evaluated in the all-adalimumab population (all-adalimumab set, N = 993), which consisted of patients in ULTRA 1, ULTRA 2, or ULTRA 3 who had at least one dose of blinded or OL adalimumab (this excluded 17 patients from non-compliant sites). Efficacy was reported from the first dose through 208 weeks of treatment; however, some patients (those randomized to placebo, or those who discontinued from or did not enter ULTRA 3, or those who could not be reached) did not have the full 208 weeks of treatment. Efficacy results were obtained in 321 patients of the all-adalimumab set at the full week 208 follow-up. Maintenance efficacy of adalimumab was also analyzed in patients who entered ULTRA 3 from ULTRA 1 or ULTRA 2 (adalimumab extension, N = 588) from weeks 0 through 156 of ULTRA 3 (corresponding to week 208 from baseline of ULTRA 1 and ULTRA 2). There remained 360 patients still receiving adalimumab at week 156 at the study visit. Previous anti-tumour necrosis factor (TNF) use was also used to perform a subgroup analysis.[30]

Efficacy end points that were assessed in the adalimumab randomized set included remissions in partial Mayo scores and Inflammatory Bowel Disease Questionnaire (IBDQ) scores, mucosal healing, discontinuation of corticosteroids, and corticosteroid-free remission. Efficacy end points assessed in the adalimumab extension set through year 3 were maintenance of remission (per partial Mayo score; those who entered ULTRA 3 in remission per full Mayo score, N = 242), and maintenance of mucosal healing (those who entered ULTRA 3 in remission, N = 409). Due to the timing of the endoscopy in ULTRA 3,

remission per full Mayo score and mucosal healing were reported up to week 196 from baseline of lead-in studies. Adverse events (AEs), serious adverse events (SAE)s, hospitalizations, and mortality were also reported.[30]

Results

Baseline characteristics from the lead-in studies (ULTRA 1 and ULTRA 2) can be observed in detail in the main body of the report. In brief, most of the patients were white males, with a mean disease duration of approximately six years. In addition, the mean Mayo score was approximately 9, with more than half the patients using oral corticosteroids and half having pancolitis.[30] Patient cohorts, as they progressed through ULTRA 1, ULTRA 2, and ULTRA 3 are presented in Table 24.

TABLE 24: PATIENT COHORTS IN ULTRA 1, ULTRA 2, AND ULTRA 3

	ULTRA 1 (N = 575)			ULTRA 2 (N = 518)	
Patients Cohort, n	**PL (N = 222)**	**ADA 80 mg/ 40 mg/40 (N = 130)**	**ADA 160/ 80 mg/40 mg (N = 223)**	**PL (N = 260)**	**ADA 160 mg/ 80 mg/40 mg (N = 258)**
Completed lead-in study	382			285	
Entered ULTRA 3 on:	334			254	
ADA OL EOW	264			58	
ADA OL EW				71	
PL DB	NA			54	
ADA DB EOW	NA			71	
Entered ULTRA 3 on:					
ADA EW	70			71	
ADA EOW	264			183	
Moved to ADA EW in ULTRA 3	120				

ADA = adalimumab; DB = double-blind; EOW = every other week; EW = every week; OL = open-label; NA = not applicable; PL = placebo.
Source: Colombel et al.[30]

Long-term efficacy outcomes were reported for patients up until 208 weeks from baseline in both ULTRA 1 and ULTRA 2, and up to week 156 from the start of ULTRA 3. Remission and mucosal healing rates were maintained throughout the four years of adalimumab treatment, with the four-year rate at 24.7% and 27.7% per partial Mayo score and mucosal healing, respectively (including non-responder imputation [NRI] and excluding dose escalators). When including the patients who dose escalated to receive weekly adalimumab, the rate of remission per partial Mayo score was 27.3% and 18.2% at weeks 52 and 208, respectively. At the three-year mark, 78.5% (last observation carried forward [LOCF]) and 63.6% (NRI) of those who entered ULTRA 3 in remission (per full Mayo score) remained in remission. Mucosal healing was maintained in 81.7% (LOCF) and 59.9 % (NRI) through week 144 of ULTRA 3 in patients who entered the extension phase with mucosal healing. In addition, anti-TNF-naive patients (at the commencement of the lead-in studies) were observed to have better mucosal healing than those who had experience with anti-TNFs. Weekly percentages of remission and mucosal healing are presented in Table 25.

TABLE 25: PERCENTAGE OF PATIENTS EXPERIENCING CLINICAL REMISSION AND MUCOSAL HEALING THROUGHOUT THE LEAD-IN STUDIES AND ULTRA 3

	Remission[a] (%)		Mucosal Healing (%)	
	All, NRI	No dose escalation, mNRI		
Treatment Week[b]				
8	28.2	28.2	41.7	
52	32.2	27.3	42.3	
100	30.3	24.3	32.3	
148	28.3	21.2	30.3	
196	25.5	18.5	27.7	
208	24.7	18.2	NR	
	LOCF	**NRI**	**LOCF**	**NRI**
Week in ULTRA-3				
0	100	100	100	100
48	78.5	75.6	81.4	73.3
96	81.8	73.1	80.4	66.3
144	78.1	64.9	81.7	59.9
156	78.5	63.6	NR	NR

LOCF = last observation carried forward; mNRI = modified non-responder imputation; NRI = non-responder imputation; NR = not reported.
[a] Per partial Mayo score.
[b] From start of lead-in studies.
Source: Colombel et al.[30]

Corticosteroid-free remission rates increased from week 52 (27.4%) to week 196 (39.7%) of adalimumab use in patients in the adalimumab randomized set who received corticosteroids at baseline of ULTRA 1 and ULTRA 2. An increased proportion of patients discontinued corticosteroid use from week 16 through week 196 of ULTRA 3. There were 40.3% of patients in the adalimumab randomized set who achieved remission per IBDQ score at one year of adalimumab therapy, while the average IBDQ score throughout ULTRA 3 ranged from 26% to 36%. Throughout the entire duration of follow-up, anti-TNF-naive patients had numerically higher remission rates compared with those with anti-TNF experience. Percentages of weekly corticosteroid discontinuations and remissions per IBDQ are presented in Table 26.

TABLE 26: PERCENTAGE OF PATIENTS WHO DISCONTINUED CORTICOSTEROIDS OR EXPERIENCED IBDQ REMISSION

	Discontinued Corticosteroid Use (%)[a]	Remission — IBDQ[bc] (%)
Treatment Week		
8	NR	38.7
16	28.6	NR
52	53.7	40.3
100	65.4	32.8
148	60.0	29.3
208	59.2	26.3

IBDQ = Inflammatory Bowel Disease Questionnaire; NR = not reported.
[a] Adalimumab randomized set; as-observed analysis.
[b] IBDQ remission score ± 170.
[c] Adalimumab randomized set; non-responder imputation (NRI) analysis.
Source: Colombel et al.[30]

For patients receiving adalimumab 160 mg/80 mg/40 mg, incidence rates of hospitalizations and colectomy (including all-cause and UC-related) were observed to be lower in the ULTRA 3 extension study than in the ULTRA 1 and ULTRA 2 studies. The most frequent SAEs were worsening or flare of UC and injection site reactions. Two deaths and one serious tuberculosis (TB) infection were observed after the double-blind (DB) period, and there were three cases of B-cell lymphoma during ULTRA 3. Details of the harms data during the 0 to 52 week DB period and ULTRA 3 are presented in Table 27.

TABLE 27: HARMS DATA THROUGH ULTRA 1, ULTRA 2, AND ULTRA 3

	ULTRA 1 and 2[a] Weeks 0–52		ULTRA 3 (N = 592)	All-ADA[b] (N = 1,010)
	ADA 160 mg/80 mg/40 mg (N = 480)	PL (N = 483)		
Any AE, n[c]	1,412	1,318	-	8,057
AE leading to discontinuation, n[c]	39	63	-	249
SAE, n[c]	55	69	-	414
Opportunistic infection (excluding TB)	2	1	-	6
Active TB	0	0	-	1
Injection site reaction	84	25	-	246
Any malignancies	2	2	-	23
Lymphoma	0	0	-	3[d]
Congestive HF	1	0	-	4
Demyelinating disease	0	0	-	3
Hepatic event	0	0	-	12
UC worsening or flare	82	106	-	588
Death	0	0	-	2
Hospitalizations and colectomy,[e] n[f] (IR)				
All-cause	69 (0.18)	-	135 (0.09)	-
UC-related	47 (0.12)	-	59 (0.04)	-
Colectomy	15 (0.04)	-	16 (0.01)	-
Hospitalizations,[g] n[h] (IR)				
All-cause	85 (0.21)	-	204 (0.12)	-
UC-related	56 (0.14)	-	86 (0.05)	-

ADA = adalimumab; AE = adverse event; DB = double-blind; HF = heart failure; IR = incidence rate; TB = tuberculosis; UC = ulcerative colitis.
[a] DB studies.
[b] Exposure as of April 15, 2013.
[d] One malignant event was reported twice in one patient.
[e] Exposure-adjusted patient-based analysis.
[f] Number of patients with event.
[g] Exposure-adjusted event-based analysis.
[h] Number of events.
Source: Colombel et al.[30]

Summary

The goal of the extension study ULTRA 3 was to acquire long-term data on the efficacy and safety associated with adalimumab use. Overall, both remission and mucosal healing were maintained over the four-year observation period, along with an increase in corticosteroid-free remission during ULTRA 3. The most frequent SAEs were worsening or flare of UC and injection site reactions. In addition, two deaths and one serious TB infection were noted during ULTRA 3.

APPENDIX 7: SUMMARY OF INDIRECT COMPARISONS

Introduction

Background

To provide a summary and critical appraisal of the 2014 National Institute for Health and Care Excellence (NICE),[31] Thorlund et al.,[32] Danese et al.,[33] Stidham et al.,[34] and Galvan-Banqueri et al.[35] indirect comparisons (IDCs). No trials have been identified that directly compare the tumour necrosis factor (TNF) alpha inhibitors adalimumab, infliximab, and golimumab or the integrin inhibitor vedolizumab for the treatment of patients with moderate to severe ulcerative colitis (UC) who have had an inadequate response to, were intolerant to, conventional treatment, or for whom conventional treatment was contraindicated. Therefore, IDCs (network meta-analyses [NMAs]) have been performed to ascertain the clinical effectiveness and safety of the aforementioned treatments.

Methods

One IDC was provided by the manufacturer (NICE[31]), with four other IDCs identified through an electronic literature search (Galvan-Banqueri, 2015,[35] Thorlund 2015,[32] Stidham, 2014,[34] and Danese 2014[33]). IDCs by NICE,[31] Thorlund et al.,[32] and Danese et al.[33] have been summarized and critically appraised in this section, while the other identified IDCs (Stidham et al.[34] and Galvan-Banqueri et al.[35]) will be discussed in the context of how their results pertain to those obtained by NICE.[31] This is because they had similar methods, population, comparators, outcomes, and results to the aforementioned IDCs.

Description of indirect comparisons identified

All of the IDCs assessed both the induction and maintenance of biologic treatment with TNF alpha inhibitors (adalimumab, golimumab, and infliximab)[31-35] and one integrin inhibitor (vedolizumab)[33] in patients with moderately to severely active UC that was uncontrolled (or in patients intolerant accepted first-line therapies including any combination of salicylates, corticosteroids, or immunosuppressants, or for whom these therapies were contraindicated).[31-35] Comparators of interest included the TNF alpha inhibitor treatments,[31-35] vedolizumab,[33] and surgery.[31] Primary efficacy outcomes included clinical response and remission,[31-35] mucosal healing,[31-33,35] adverse events (AEs), and serious adverse events (SAEs).[31,33] Other outcomes of interest included health-related quality of life (HRQoL) and rates of hospitalization or surgical intervention.[31] Details of the inclusion criteria for each of the five identified IDCs are provided in Table 28.

TABLE 28: INCLUSION CRITERIA FOR THE NETWORK META-ANALYSES

Inclusion Criteria	NICE 2015[31]	Galvan-Banqueri 2015[35]	Thorlund 2015[32]	Danese 2014[33]	Stidham 2014[34]
Patient population	Adult (≥ 18 years) and pediatric (6 to 17 years) patients with moderately to severely active UC after the failure of conventional therapy[b] or who are intolerant of or have medical contraindications to such therapies	Adult patients naive to biological drugs with moderate to severe UC	Adults diagnosed with moderately to severely active UC who had not previously received treatment with an anti-TNF	Adults diagnosed with moderately to severely active UC (defined as a Mayo score of 6 to 12 points, with an endoscopic subscore of 2 or 3)	Adults diagnosed with UC
Interventions and comparators	Induction and maintenance treatment with: • ADA • GOL • IFX At the approved dose[a] or unlicensed dose as monotherapy or in combination with conventional therapies,[b] calcineurin inhibitors, or elective surgical intervention	Induction (6 to 8 weeks) and/or maintenance treatment (52 to 54 weeks) with: • ADA • GOL • IFX	Induction (6 to 8 weeks), maintenance (1 year) treatment, as well as sustained remission and response (induction through maintenance) with: • ADA • GOL • IFX At the approved dose[a] as monotherapy or in combination with conventional therapies[b]	Induction and/or maintenance treatment with: • ADA • GOL • IFX • VED At the approved dose[a] as monotherapy or in combination with conventional therapies[b]	Induction and/or maintenance treatment with: • ADA • GOL • IFX At the approved dose as monotherapy or in combination with conventional therapies[b]
Outcomes	***Efficacy outcomes:*** • Measures of disease activity • Rates of and duration of response, relapse, and remission[c] • HRQoL ***Harms outcomes:*** • Mortality	***Efficacy outcomes:*** • Clinical response (decrease from baseline in the total Mayo score by ≥ 3 points and at least 30% with an accompanying decrease in RBS of ≥ 1 point or an absolute	***Efficacy outcomes:*** • Clinical response (decrease from baseline in the total Mayo score by ≥ 3 points and at least 30% with an accompanying decrease in RBS of ≥ 1 point or an	***Efficacy outcomes:*** • Clinical response (Mayo score reduction of ≥ 3 points and/or ≥ 30%, plus decrease in RBS of ≥ 1 point or an absolute RBS of 0 or 1) • Clinical remission	***Efficacy outcomes:*** • Clinical remission (Mayo score or UCSS within 8 wks of treatment start; from induction to ≥ 52 wks for maintenance) • Clinical response (Mayo score reduction of ≥ 3

Inclusion Criteria	NICE 2015[31]	Galvan-Banqueri 2015[35]	Thorlund 2015[32]	Danese 2014[33]	Stidham 2014[34]
	• AEs (including leakage and infections following surgery) • Rates of hospitalization • Rates of surgical intervention (both elective and emergency) • Time to surgical intervention (both elective and emergency)	RBS of 0 or 1) • Clinical remission (total Mayo score ≤ 2 with no individual subscore > 1) • Mucosal healing (absolute endoscopy subscore of 0 or 1) ***Harms outcomes:*** • Not evaluated	absolute RBS of 0 or 1) • Clinical remission (Mayo score ≤ 2 with no individual subscore > 1) • Mucosal healing (endoscopy subscore of 0 or 1) • Sustained remission (remission after 1 year of maintenance therapy conditional on remission after induction therapy) • Sustained response (response after 1 year of maintenance therapy conditional on response after induction therapy) ***Harms outcomes:*** Not evaluated	(Mayo score ≤ 2 with no individual subscore > 1) • Mucosal healing (absolute subscore for endoscopy of 0 or 1) ***Harms outcomes:*** • SAEs • WDAEs • AEs including: o Total infections o Serious infections o TB o Heart failure	points and/or ≥ 30%, plus decrease in RBS of ≥ 1 point or an absolute RBS of 0 or 1; within 8 weeks of treatment start; from induction to ≥ 52 weeks for maintenance) ***Harms outcomes:*** • Not evaluated
Study designs	RCTs and their OL extension studies	RCTs			

ADA = adalimumab; AE = adverse event; GOL = golimumab; HRQoL = health-related quality of life; IFX = infliximab; NICE = National Institute for Health and Care Excellence (UK); NMA = network meta-analysis; OL = open-label; RBS = rectal bleeding subscore; RCT = randomized controlled trial; SAE = serious adverse event; TB = tuberculosis; TNF = tumour necrosis factor; UC = ulcerative colitis; UCSS = ulcerative colitis symptom score; VED = vedolizumab; WDAE = withdrawal due to adverse event.

[a] Different doses of the same treatment were treated as separate interventions.

[b] Conventional treatment: any combination of salicylates, corticosteroids, and immunosuppressants such as azathioprine, 6-mercaptopurine, cyclosporine, or methotrexate.

[c] Definitions of each outcome not provided.

Sources: Danese et al.,[33] Stidham et al.,[34] Galvan-Banqueri et al.,[35] Thorlund et al.,[32] and NICE.[31]

Review and appraisal of NICE[31] IDC
Review of NICE[31] IDC
Due to the paucity of direct active comparator evidence, NICE IDC performed an NMA in order to ascertain comparative clinical effectiveness and AE evidence of three TNF alpha inhibitors (adalimumab, infliximab, and golimumab) approved in the UK for the treatment of adult patients with moderate to severely active UC uncontrolled on, or in patients intolerant to, conventional therapies, or for whom those therapies are contraindicated.

Methods for NICE[31] indirect comparison
Study eligibility and selection process
A systematic review was performed with specific inclusion and extensive exclusion criteria provided and clearly outlined.

Data extraction
One reviewer who was not blinded to either authors or journals extracted the data, which were double-checked by another reviewer.

Trial characteristics
Nine trials were included in the NICE IDC: three for adalimumab (ULTRA 1, ULTRA 2, and Suzuki et al.), two for golimumab (PURSUIT-SC and PURSUIT-M), and four for infliximab (ACT-1, ACT-2, Probert et al., and UC-SUCCESS). The extension study data from ULTRA 3, ACT-1, and ACT-2 were also included in the IDC; however, these studies were not included as individual studies. All trials included adult patients with moderately to severely active UC who were not controlled on, had failed on, or had intolerances or contraindications to one or more pharmacological standard of care treatments (including a combination of corticosteroids, aminosalicylates, calcineurin inhibitors, and thiopurines), or had surgical intervention. In addition, all trials included the same or very similar definitions of clinical response and clinical remission, with the exception of Probert et al., who did not provide any definition for clinical response, and their definition of clinical remission was based on the UC symptom score (as opposed to the other studies that used Mayo scores). Outcomes of interest included the clinical effectiveness (measures of disease activity, rates of response, relapse, remission, and HRQoL) and AEs (mortality, rates of hospitalization, rates of surgical intervention, time to surgical intervention, common AEs) of each intervention, along with the effect of disease duration on the clinical effectiveness of each intervention (if available). The NICE IDC did not provide any information on early escape for the included RCTs. Details of the trial design characteristics are provided in Table 29.

TABLE 29: OVERVIEW OF INCLUDED RANDOMIZED CONTROLLED TRIALS IN NICE INDIRECT COMPARISON

Trial, Year	Trial Type	Treatment	Sample Size	Primary Outcome(s)	Induction Assessment	Maintenance Assessment
Adalimumab Trials						
ULTRA 1	• Phase 3 • Multi-centre/ multi-country	• ADA 80 mg/40 mg • ADA 160 mg/80 mg • PL	390	Clinical remission (at week 8 with ITT-A3 amendment population)	Week 8	NA
ULTRA 2	• Phase 3 • Multi-centre/multi-country	• ADA 160 mg at week 0, 80 mg at week 2, then 40 mg EOW beginning at week 4 • PL	494	Proportion of patients achieving clinical remission at weeks 8 and 52	Week 8	Week 52
ULTRA 3	• OLE • Multi-centre/ multi-country	• ADA 40 mg EOW or EW	588	NR	NA	Week 104
Suzuki et al.	• Phase 2/3 • Japan	• ADA 80 mg/ 40 mg • ADA 160 mg/80 mg • PL	274	NR	Week 8	Week 32 Week 52
Golimumab Trials						
PURSUIT-SC	• Phase 2/3 dose ranging study • Multi-centre/ multi-country	• GOL 200 mg/100 mg • PL	845	Clinical response at week 6	Week 6	NA
PURSUIT-M	• Phase 3 • Multi-centre/ multi-country	• GOL 50 mg • GOL 100 mg • PL	1,228	Clinical response maintained through week 54; however, only in GOL induction responders	NA	Week 30 Week 52
Infliximab Trials						
ACT-1	• Phase 3 • Multi-centre	• IFX 5 mg/kg • IFX 10 mg/kg • PL	363	Clinical response at week 8	Week 8	Week 30 Week 54
ACT-2	• Phase 3 • Multi-centre	• IFX 5 mg/kg • IFX 10 mg/kg • PL	364	Clinical response at week 8	Week 8	Week 30 Week 54

Trial, Year	Trial Type	Treatment	Sample Size	Primary Outcome(s)	Induction Assessment	Maintenance Assessment
ACT-1 and ACT-2 ES	• Phase 3 • Multi-centre	• IFX 5 mg/kg • IFX 10 mg/kg • PL	229	NR	NA	To week 152
Probert et al.	• Phase not reported • Multi-centre (US and Germany)	• IFX 5 mg/kg • PL	43	Clinical remission at week 6	Week 6	NA
UC-SUCCESS	• Multi-centre	• IFX 5 mg/kg • AZA 2.5 mg/kg • Combo (IFX + AZA)	239	Corticosteroid-free remission at week 16	Week 8	Week 16

A3 = Amendment 3; ADA = adalimumab; AZA = azathioprine; EOW = every other week; ES = extension studies; EW = every week; GOL = golimumab; IDC = indirect comparison; IFX = infliximab; ITT = intention-to-treat; NA = not applicable; NICE = National Institute for Health and Care Excellence (UK); NR = not reported; OLE = open-label extension; PL = placebo; RCT = randomized controlled trial; SD = standard deviation.
Source: NICE.[31]

Population characteristics

Mean and median ages of the patients across trials were consistent, ranging between 37 years to 42.5 years, while white men made up the majority of the patients (range 79.5% to 95.9%), with the exception of the Suzuki et al. study, in which all patients were Japanese. Baseline Mayo scores were consistent between trials, ranging from 8.1 to 8.9, while baseline medications were variable between studies. Disease duration of the patients at baseline ranged from 4.9 years to 8.5 years; however, weight and smoking status were poorly reported. While all trials included patients naive to anti-TNF alpha treatments, ULTRA 2 (adalimumab) also included patients with anti-TNF alpha experience and PURSUIT-M (golimumab) re-randomized only patients who responded to golimumab in the induction phase. The NICE IDC[31] authors requested, and ultimately received, data for anti-TNF alpha naive patients for the maintenance period from AbbVie for the adalimumab trial.

Comparators

The comparators of interest included anti-TNF alpha drugs (adalimumab, golimumab, and infliximab) approved for patients with moderately to severely active UC uncontrolled on prior treatment. In addition, standard of care was also a comparator. Of note, at the time of the NICE[31] IDC, vedolizumab had not been approved for this indication; hence, this was not included as a comparator.

Outcomes

Outcomes of interest included the clinical effectiveness (measures of disease activity, rates of response, relapse, and remission, and HRQoL) and AEs (mortality, rates of hospitalization, rates of surgical intervention, time to surgical intervention, common AEs) of each intervention, as well as the effect of disease duration on the clinical effectiveness of each intervention (if available).

Quality assessment of included studies

The Cochrane Risk of Bias tool was used by one reviewer to assess the methodological quality of the included studies, and a second reviewer double-checked these findings.

Evidence network

Evidence networks were provided in the NICE[31] study and were complete. None of the networks contained any closed loops and were anchored to placebo as the only common treatment arm between studies.

Indirect comparison methods

Inclusion in the NMA required studies to have both clinical response and remission outcomes in either an induction period of six to eight weeks or a maintenance period of either 30 to 52 or 54 weeks. Two adult populations from the infliximab studies (Probert et al. and UC-SUCCESS) were excluded from the induction treatment NMA analysis. With regard to the base-case analysis, the anti-TNF alpha naive patients from ULTRA 2 were used rather than the intention-to-treat (ITT) population, as the ITT population included both treatment-naive and treatment-experienced patients. In addition, the Suzuki et al. patients were excluded from the base-case analysis due to their being exclusively Japanese. These populations (adding the Suzuki et al. patients, replacing the ULTRA 2 anti-TNF alpha naive population with the ITT population, and replacing the anti-TNF alpha naive population with the ITT population as well as adding the Suzuki et al. patients) were assessed in subsequent sensitivity analyses. With regard to the induction base-case analysis, both the PURSUIT-SC phase 2 and phase 3 data were utilized.

The authors considered clinical response and remission as ordinal categorized data with the following mutually exclusive categories of no response, response, and remission, of which the model assumed the

treatment effect (of which the type of treatment effect was not specifically stated) was the same regardless of these categories. Six- and eight-week, 30- and 32-week, and 52- and 54-week data were combined, respectively.

Bayesian NMA analyses were conducted using the OpenBUGS software package. The authors used a random effects model that was completed with conventional reference prior distributions when there were sufficient sample data. However, there were relatively few studies for both the induction and maintenance phases; therefore, weakly informative priors were used. Specifically, prior distributions used for "trial-specific baselines were N(0, 1000), N(0, 1000) for treatment effects relative to the reference, U(0,2) for between-study standard deviation of treatments effect, U(0,5) for population cut-offs, U(0,2) for between-study deviation of cut-offs."[31] A burn-in of 50,000 iterations of the Markov chain (plus an additional 10,000 iterations for parameter estimation) were used for both the baseline and relative treatment effects model.

Between-study heterogeneity was assessed by finding the between-study standard deviation. Model fit was also assessed by calculating the total residual deviance, which was compared with the total number of data points used (e.g., if close to the number of data points then the model fit the data well).

In addition to the aforementioned NMA methods employed, secondary outcomes (colectomy, ostomy, serious infections, and injection site reactions) were assessed using classical meta-analysis methodology with Cochrane Review Manager software (version 5.2). Dichotomous outcomes (colectomy, ostomy, serious infections, injection site reactions) were reported as risk ratios with associated 95% CIs. Data were pooled where appropriate; however, clinical heterogeneity (including differences in patients, intervention types, intervention duration, and outcome types) was considered prior to pooling; hence, a random effects model was employed.

Results

Clinical efficacy

Induction phase

No statistically significant differences in treatment effects were observed when comparing any of the biologics with one another (Table 30).

Maintenance phase — eight to 32 weeks for patients starting in response

No statistically significant differences in treatment effects (in either induction or maintenance phase therapy) were observed when comparing any of the biologics with one another (Table 30).

Maintenance phase — eight to 32 weeks for patients starting in remission

No statistically significant differences in treatment effects were observed when comparing any of the biologics with one another (Table 30).

Maintenance phase — 32 to 52 weeks for patients starting in response

No statistically significant differences in treatment effects were observed when comparing any of the biologics with one another (Table 30).

Maintenance phase — 32 to 52 weeks for patients starting in remission

No statistically significant differences in treatment effects were observed when comparing any of the biologics with one another or with placebo, with the exception of adalimumab, which had a statistically

significantly beneficial effect (clinical response/clinical remission) versus placebo (–1.04; 95% credible interval [CrI],–1.93 to –0.12).

TABLE 30: EFFICACY RESULTS FOR BIOLOGIC DRUGS IN PATIENTS WITH MODERATE TO SEVERE ULCERATIVE COLITIS UNCONTROLLED ON CONVENTIONAL THERAPY FROM ALL REVIEWED INDIRECT COMPARISONS

Comparison	Clinical Response	Clinical Remission	Mucosal Healing
NICE[31]	**Effect**		**Effect**
Induction Therapy			
GOL vs. ADA	–0.10 (–0.69 to 0.50)		-
IFX vs. ADA	–0.52 (–1.03 to 0.00)		-
IFX vs. GOL	–0.42 (–1.00 to 0.17)		-
Maintenance Therapy (8 to 32 weeks) for Patients Starting in Response			
GOL 50 mg vs. ADA	–0.29 (–1.24 to 0.67)		-
GOL 100 mg vs. ADA	–0.38 (–1.36 to 0.59)		-
IFX vs. ADA	–0.20 (–1.09 to 0.69)		-
GOL 100 mg vs. GOL 50 mg	–0.10 (–0.72 to 0.55)		-
IFX vs. GOL 50 mg	0.07 (–0.75 to 0.91)		-
IFX vs. GOL 100 mg	0.18 (–0.68 to 1.01)		-
Maintenance Therapy (8 to 32 weeks) for Patients Starting in Remission			
GOL 50 mg vs. ADA	–0.82 (–1.96 to 0.39)		-
GOL 100 mg vs. ADA	–0.79 (–1.96 to 0.42)		-
IFX vs. ADA	–0.29 (–1.41 to 0.85)		-
GOL 100 mg vs. GOL 50 mg	0.02 (–0.69 to 0.75)		-
IFX vs. GOL 50 mg	0.52 (–0.46 to 1.49)		-
IFX vs. GOL 100 mg	0.51 (–0.48 to 1.45)		-
Maintenance Therapy (32 to 52 weeks) for Patients Starting in Response			
GOL 50 mg vs. ADA	–0.49 (–1.77 to 0.77)		-
GOL 100 mg vs. ADA	–0.12 (–1.36 to 1.11)		-
IFX vs. ADA	–0.67 (–2.04 to 0.66)		-
GOL 100 mg vs. GOL 50 mg	0.38 (–0.47 to 1.20)		-
IFX vs. GOL 50 mg	–0.18 (–1.51 to 1.10)		-
IFX vs. GOL 100 mg	–0.56 (–1.85 to 0.73)		-
Maintenance Therapy (32 to 52 weeks) for Patients Starting in Remission			
GOL 50 mg vs. ADA	1.08 (–0.18 to 2.31)		-
GOL 100 mg vs. ADA	0.87 (–0.38 to 2.11)		-
IFX vs. ADA	0.78 (–0.53 to 2.14)		-
GOL 100 mg vs. GOL 50 mg	–0.20 (–1.01 to 0.61		-
IFX vs. GOL 50 mg	–0.29 (–1.57 to 1.02)		-
IFX vs. GOL 100 mg	–0.08 (–1.37 to 1.24)		-
***Danese et al.*[33]**	**OR (95% CrI)**	**OR (95% CrI)**	**OR (95% CrI)**
Induction Therapy[a]			
GOL vs. ADA	1.20 (0.60 to 2.12)	1.52 (0.50 to 4.28)	1.12 (0.64 to 1.92)
IFX vs. ADA	**2.36 (1.22 to 4.63)**	2.79 (0.95 to 8.83)	**2.02 (1.13 to 3.59)**
VED vs. ADA	1.84 (0.74 to 4.66)	2.39 (0.51 to 12.40)	NA
IFX vs. GOL	1.96 (0.99 to 4.48)	1.84 (0.58 to 6.92)	1.80 (0.96 to 3.46)

Comparison	Clinical Response	Clinical Remission	Mucosal Healing
VED vs. GOL	1.53 (0.62 to 4.32)	1.56 (0.32 to 9.19)	NA
IFX vs. VED	1.28 (0.48 to 3.45)	1.18 (0.21 to 6.32)	NA
Galvan-Banqueri et al.[35]	**RR (95% CrI)**	**RR (95% CrI)**	**RR (95% CrI)**
Induction Therapy[b]			
ADA vs. GOL	0.79 (0.59 to 1.04)	0.66 (0.33 to 1.30)	0.83 (0.61 to 1.14)
IFX vs. ADA	**1.46 (1.12 to 1.90)**	1.68 (0.94 to 3.03)	**1.49 (1.12 to 1.98)**
IFX vs. GOL	1.15 (0.85 to 1.55)	1.10 (0.56 to 2.17)	1.25 (0.91 to 1.71)
Maintenance Therapy[c]			
ADA vs. GOL 50 mg	NA	1.18 (0.62 to 2.25)	NA
ADA vs. GOL 100 mg	NA	1.16 (0.61 to 2.20)	NA
IFX vs. ADA	1.51 (0.87 to 2.60)	1.19 (0.59 to 2.40)	1.54 (0.86 to 2.79)
IFX vs. GOL 50 mg	NA	1.40 (0.77 to 2.56)	NA
IFX vs. GOL 100 mg	NA	1.37 (0.75 to 2.50)	NA
Stidham et al.[34]	**RR (95% CrI)**	**RR (95% CrI)**	**RR (95% CrI)**
Induction Therapy (6-8 weeks)			
GOL vs. ADA	1.46 (0.42 to 5.38)	1.75 (0.17 to 16.86)	NA
IFX vs. ADA	2.15 (0.73 to 5.80)	2.08 (0.32 to 12.03)	NA
IFX vs. GOL	1.48 (0.38 to 4.69)	1.18 (0.13 to 10.63)	NA
Maintenance Therapy (54 weeks)			
GOL vs. ADA[d]	1.14 (0.11 to 1092)	1.04 (0.16 to 6.96)	NA
IFX vs. ADA	1.70 (0.17 to 16.59)	1.18 (0.19 to 8.02)	NA
IFX vs. GOL	1.47 (0.15 to 14.43)	1.22 (0.18 to 8.43)	NA
Thorlund et al.[32]	**OR (95% CrI)**	**OR (95% CrI)**	**OR (95% CrI)**
Induction Therapy			
GOL 100 mg vs. ADA	1.30 (0.56 to 3.10)	1.48 (0.54 to 4.18)	1.24 (0.60 to 2.56)
IFX vs. ADA	**2.19 (1.03 to 4.68)**	**2.36 (1.02 to 5.57)**	**2.20 (1.14 to 4.26)**
GOL 100 mg vs. IFX	0.60 (0.25 to 1.44)	0.63 (0.23 to 1.73)	0.57 (0.27 to 1.19)
Maintenance Therapy			
GOL 50 mg vs. ADA 40 mg	**2.31 (1.23 to 4.27)**	1.17 (0.54 to 2.46)	1.91 (0.98 to 3.61)
GOL 100 mg vs. ADA 40 mg	**2.63 (1.41 to 4.86)**	1.47 (0.69 to 3.08)	**2.03 (1.05 to 3.86)**
IFX vs. ADA 40 mg	1.87 (0.88 to 4.03)	1.34 (0.55 to 3.23)	2.09 (0.96 to 4.69)
GOL 50 mg vs. GOL 100 mg	0.88 (0.63 to 1.22)	0.80 (0.53 to 1.19)	0.94 (0.66 to 1.33)
GOL 50 mg vs. IFX	1.24 (0.62 to 2.41)	0.86 (0.41 to 1.80)	0.96 (0.48 to 1.89)
GOL 100 mg vs. IFX	1.41 (0.70 to 2.74)	1.09 (0.52 to 2.25)	1.02 (0.51 to 2.01)

ADA = adalimumab; CrI = credible interval; GOL = golimumab; IFX = infliximab; NA = not available; NICE = National Institute for Health and Care Excellence (UK); PL = placebo; RR = relative risk; UC = ulcerative colitis; VED = vedolizumab; vs. = versus.
[a] No comparative efficacy results were available for the maintenance therapy at 52 weeks except for comparisons with placebo.
[b] Induction phase was 6 weeks for golimumab studies and 8 weeks for both adalimumab and infliximab studies.
[c] Maintenance phase was 52 weeks for adalimumab studies and 54 weeks for both golimumab and infliximab studies.
[d] Comparison reported in Stidham et al. as adalimumab versus golimumab.
Note: Bolding indicates statistical significance.
Sources: Danese et al.,[33] Stidham et al.,[34] Galvan-Banqueri et al.,[35] Thorlund et al.,[32] and NICE.[31]

Sensitivity analyses

The a priori sensitivity analyses (adding Suzuki et al. patients, replacing the ULTRA 2 anti-TNF alpha naive population with the ITT population, and replacing the anti-TNF alpha naive population with the ITT population and adding the Suzuki et al. patients) did not change the direction of the treatment effect observed in the main analyses; the differences between treatments were not statistically significant.

Safety

Adverse events, serious adverse events, and withdrawals due to adverse events

Rates of AEs, SAEs, or withdrawals due to adverse events (WDAEs) were not formally meta-analyzed or pooled; rather, they were descriptively reported. One or more AEs were experienced in similar proportions between patients treated with placebo and adalimumab (ULTRA 1) or golimumab (PURSUIT-SC) in the induction phase of treatment; however, more patients experienced AEs in the azathioprine group (52%) than in either the infliximab (33%) or azathioprine plus infliximab (38%) groups in the induction phase of UC-SUCCESS. In the maintenance phase of treatment, the proportions of patients experiencing AEs were similar when comparing adalimumab with placebo (ULTRA 1) and infliximab with placebo (ACT-1). A numerically higher proportion of patients experienced AEs compared with placebo in the golimumab trial (PURSUIT-M; 72.7% for 50 mg, 73.4% for 100 mg golimumab, and 66% for placebo) and the second infliximab trial (ACT-2; 81.8% and 73.2%, respectively). Details are provided in Table 31.

The only results that were meta-analyzed included those for serious infections and injection site reactions. In the induction phase, serious infection meta-analysis revealed a risk ratio of 0.14 (95% confidence interval [CI], 0.01 to 2.75) for adalimumab versus placebo (less than 1 indicating more events in placebo arm) in ULTRA 1, and a risk ratio of 7.46 (95% CI, 0.39 to 142.47) in the Suzuki et al. trial, while the risk ratio for golimumab versus placebo in PURSUIT-SC was 0.17 (95% CI, 0.02 to 1.37). In the maintenance phase, the risk ratio for serious infections of adalimumab versus placebo was 0.81 (95% CI, 0.22 to 2.98); for golimumab 50 mg versus placebo in PURSUIT-M, the risk ratio was 1.67 (95% CI, 0.41 to 6.85); for golimumab 100 mg versus placebo the risk ratio was 1.69 (95% CI, 0.41 to 6.94); the risk ratio for infliximab versus placebo in ACT 1 (week 54) was 0.60 (95% CI, 0.15 to 2.46); and for ACT-2 (week 30) the risk ratio was 2.03 (95% CI, 0.19 to 22.13). In the induction phase, injection site reaction meta-analysis revealed a risk ratio of 1.86 (95% CI, 0.76 to 4.57) for adalimumab versus placebo in ULTRA 1 and 3.73 (95% CI, 0.80 to 17.50) in Suzuki et al., while the risk ratio was 2.19 (95% CI, 0.77 to 6.24) for golimumab versus placebo in PURSUIT-SC. In the maintenance phase, the injection site reaction meta-analysis revealed a risk ratio of 3.14 (95% CI, 1.57 to 6.26) for adalimumab versus placebo in ULTRA 2, a risk ratio of 1.00 (95% CI, 0.20 to 4.88) for golimumab 50 mg versus placebo in PURSUIT-M, and a risk ratio of 3.71 (95% CI, 1.06 to 13.06) for golimumab 100 mg versus placebo in PURSUIT-M.

NICE noted that SAEs were reported poorly across the included trials. The proportions of patients experiencing SAEs were similar in the adalimumab trial (ULTRA 1) in the induction phase; however, a higher proportion of patients experienced SAEs in the placebo group of the golimumab trial (PURSUIT-SC; 6.1% placebo and 2.7% golimumab) and in the azathioprine group of the infliximab trial (UC-SUCCESS; 8% compared with 4% for infliximab and azathioprine, respectively and 0% for infliximab) in the induction phase. For the maintenance phase, the proportions of patients experiencing SAEs were similar in ULTRA 2, Suzuki et al., and ACT-1, while those taking the active treatments experienced more AEs compared with placebo in PURSUIT-M, with the reciprocal occurring in ACT-1. Details are provided in Table 31.

WDAEs occurred in slightly higher proportions in the placebo groups compared with active treatment groups in all of the induction and maintenance trials, with the exception of the ULTRA trials (in which there were equal proportions) and the PURSUIT-M trials (in which those patients taking 100 mg of golimumab had more WDAEs than either the 50 mg golimumab group or placebo groups). Details are provided in Table 31.

TABLE 31: SAFETY ASSOCIATED WITH TREATMENTS IN THE INDUCTION AND MAINTENANCE PHASES OF TREATMENT: NICE AND DANESE ET AL. INDIRECT COMPARISONS[A]

	Placebo	Adalimumab	Infliximab	Golimumab
NICE[31]				
Patients experiencing ≥ 1 AE — Induction				
ULTRA 1, **Dose** n/N (%)	- 108/223 (48.4)	**160 mg/80 mg** 112/223 (50.2)	-	-
PURSUIT-SC, **Dose** n/N (%)	- 216/330 (38.2)	-	-	**200 mg/100 mg** 124/331 (37.5)
UC-SUCCESS, n/N (%)	-	-	**AZA:** 41/79 (52) **IFX:**26/78 (33) **AZA/IFX:** 30/80 (38)	-
Patients experiencing ≥ 1 AE — Maintenance				
ULTRA 2, **Dose** n/N (%)	- 218/260 (83.8)	**160/80 mg** 213/257 (82.9)	-	-
Suzuki et al., **Dose** n/PY	- 609.4/100	**40 mg EOW** 547.9/100	-	-
PURSUIT-M, **Dose** n/N (%)	- 103/156 (66.0)	-	-	**50 mg** 112/154 (72.7) **100 mg** 113/154 (73.4)
ACT-1, **Dose** n/N (%)	- 103/121 (85.1)	-	**5 mg/kg** 106/121 (87.6)	-
ACT-2, **Dose** n/N (%)	- 90/123 (73.2)	-	**5 mg/kg** 99/121 (81.8)	-
Patients experiencing ≥ 1 SAE[a] — Induction				
ULTRA 1, **Dose** n/N (%)	- 12/223 (5.4)	**160/80 mg** 12/223 (5.4)	-	-
PURSUIT-SC, **Dose** n/N (%)	- 20/330 (6.1)	-	-	**200 mg/100 mg** 9/331 (2.7)
UC-SUCCESS, n/N (%)	-	-	**AZA:** 6/79 (8) **IFX:** 0/78 (0) **AZA/IFX:** 3/80 (4)	-
Patients experiencing ≥ 1 SAE[b] — Maintenance				
ULTRA 2, **Dose** n/N (%)	- 32/260 (12.3)	**160/80 mg** 31/257 (12.1)	-	-
Suzuki et al., **Dose** n/PY	- 31.3/100	**40 mg EOW** 33.6/100	-	-
PURSUIT-M, **Dose** n/N (%)	- 12/156 (7.7)	-	-	**50 mg** 13/154 (8.4) **100 mg** 22/154 (14.3)
ACT-1, **Dose** n/N (%)	- 31/121 (25.6)	-	**5 mg/kg** 26/121 (21.5)	-
ACT-2, **Dose** n/N (%)	- 24/123 (19.5)	-	**5 mg/kg** 13/121 (10.7)	-
WDAE — Induction				
ULTRA 1, **Dose** (%)	- 5.4	**160 mg/80 mg** 5.4	-	-

	Placebo	Adalimumab	Infliximab	Golimumab
PURSUIT-SC, **Dose** n/N (%)	**-** 3/330 (0.9)	-	-	**200 mg/100 mg** 1/331 (0.3)
UC-SUCCESS, n/N (%)	-	-	**AZA:** 6/79 (8) **IFX:** 2/78 (3) **AZA/IFX:** 3/80 (4)	-
WDAE – Maintenance				
ULTRA 2, **Dose** n/N (%)	**-** 34/260 (13.1)	**160 mg/80 mg** 23/257 (8.9)	-	-
Suzuki et al., **Dose** N n/PY	**-** 6 13.4/100	**40 mg EOW** 22 22.4/100	-	-
PURSUIT-M, **Dose** n/N (%)	**-** 10/156 (6.4)	-	-	**50 mg** 8/154 (5.2) **100 mg** 14/154 (9.1)
ACT-1, **Dose** n/N (%)	**-** 11/121 (9.1)	-	**5 mg/kg** 10/121 (8.3)	-
ACT-2, **Dose** n/N (%)	**-** 12/123 (9.8)	-	**5 mg/kg** 2/121 (1.7)	-
Danese et al.[33]				
	Median Event Rates, % (Range)			
Any AEs, n = 3,354	69.8 (38.2 to 85.1)	81.4 (37.5 to 87.6)		
SAEs, n = 3,354	12.3 (6.1 to 25.6)	10.7 (2.2 to 21.5)		
AEs leading to drug discontinuation, n = 2,607	6.4 (0.9 to 13.1)	8.3 (0.3 to 13.2)		
Infectious AEs, n = 3,354	28.2 (12.1 to 70.6)	39.0 (11.8 to 71.7)		
Serious infection, n = 3,354	1.9 (0.8 to 4.1)	1.7 (0.0 to 5.5)		
TB, n = 2,607	0.0 (0.0 to 0.6)	0.0 (0.0 to 1.9)		

AE = adverse event; AZA = azathioprine; EOW = every other week; IFX = infliximab; NICE = National Institute for Health and Care Excellence (UK); PY = patient-years; RCT = randomized controlled trial; SAE = serious adverse event; TB = tuberculosis; WDAE = withdrawal due to adverse event.

[a] Definitions of serious adverse event were poorly reported across included RCTs.

[b] Only NICE and Danese et al. IDCs looked at harms data.

Health-related quality of life

No IDCs were performed for any of the HRQoL data.

Critical appraisal

The NICE IDC[31] was, in general, well conducted and well reported. A comprehensive systematic review was performed with a two-stage selection process, whereby articles were first selected based on titles and abstracts and then full-texts articles were retrieved and ascertained for their inclusion criteria. However, while two reviewers selected articles based on titles and abstracts, only one reviewer subsequently examined full-text articles for inclusion, thereby reducing the redundancy associated with this step. In addition, data extraction was performed by only one reviewer (not in duplicate); however, another reviewer checked the data.

When selecting evidence for inclusion, the NICE IDC included only those studies that reported both response and remission for the induction and maintenance phases. Selection in such a manner could have potentially led to the possibility that some relevant literature may have been excluded. In addition, no specific rationale was provided regarding the reasoning for the exclusion of the adult populations from the induction treatment NMA analysis for the Probert et al. and UC-SUCCESS infliximab studies. Instead, NICE specified only that these populations were excluded for "other reasons" and directed the reader to the table of study characteristics.

With regard to clinical definitions, Probert et al. did not provide any definition of clinical response and based their definition of clinical remission on the UC symptom score rather than on the Mayo score, which was used by the other eight included trials. This could have potentially increased the heterogeneity in the overall population and could have affected the results, as no adjustments for this discrepancy was apparent. In addition, neither the NICE IDC[31] nor any of the other IDCs[32-35] included any information on whether the trials included the potential for early escape and, if the trials did look at this (but failed to report on it), how those patients were dealt with in the trial analyses.

Using the Cochrane Risk of Bias instrument, only three trials included in the NICE IDC were considered to be at low risk of bias with regard to allocation concealment, completeness of outcome data, and blinding. In addition, the PURSUIT-M trial included an alternate trial design, whereby patients who were responders in the induction phase were re-randomized to treatment groups for the maintenance phase.

While the target population to be included in the NICE IDC is relevant, the clinical expert involved in the CDR review noted that treatments with biologics cannot be extrapolated to certain populations of patients with UC as the trials tend to exclude them. These patients include those with acute UC, proctitis, localized UC (not involving the entire colon, indicating that these patients are not as ill), and those patients who are extremely ill with UC (and therefore would not ethically be entered into these types of RCTs).

All of the included 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; hence, the strength of the NMA was not as strong as it would be if there was a closed loop (e.g., active comparison). In addition, all of the trials were, at most, one year long; hence, long-term efficacy and safety were not assessed. Another important aspect to consider regarding the trials were that these biologics were assessed in mostly treatment-naive patients; therefore, the ability to extrapolate these results to patients who were unsuccessful on previous biologic treatment remains uncertain.

The NMA itself was comprehensively performed, with discussion and reasoning provided for the type of model used (random or fixed effects; ultimately using random effects model), the addition of supplementary meta-analyses for comparison (for secondary outcomes, using appropriate measures for continuous and dichotomous outcomes), the assessment of heterogeneity, the network diagrams, appropriate base cases used and subsequent a priori sensitivity analyses to confirm these, descriptions of the priors used, burn-ins, and iterations to obtain convergence. In-depth discussions supporting the authors' use of the aforementioned components were provided to explain the relevance and appropriateness of each. In the Results section, the authors noted when heterogeneity was high, thereby highlighting potential areas of concern within the IDC. However, while the random effects model was the main model for the analysis and results for the fixed effects model were obtained, but were not provided. NICE also did not specify which treatment effect was used for the assessment of the

clinical response/remission end point (e.g., OR or RR); therefore, there remains uncertainty surrounding how to interpret this information (Table 30). In addition, the credible intervals surrounding these estimates were wide, implying imprecision and uncertainty.

Upon calculating the between-study treatment effects for the induction phase, the eight- to 32-week maintenance phase for patients starting in remission, the 32- to 52-week maintenance phase for patients starting in response, and the 32- to 52-week maintenance phase for patients starting in remission, the authors ascertained there to be mild to moderate between-study heterogeneity. However, for the eight- to 32-week maintenance phase for patients starting in response, the authors noted that mild to moderate between-study heterogeneity was implied. A suggestion that the model did not fit the data well was postulated by the authors, as the total residual deviance was 11.73, which was smaller than that of the total number of data points for this analysis (18). All of the included studies (ULTRA 2, ACT-1, ACT-2, PURSUIT-M) had smaller-than-anticipated residual deviances.

The outcomes of clinical response, clinical remission, and mucosal healing were, for the most part, assessed in all of the RCTs, thereby providing a rationale for using these outcomes in the IDC (as these were the primary areas of concern along with HRQoL for patients with UC, according to the clinical expert associated with this review). The doses are comparable to those available in Canada and hence can be generalized to patients in Canada.

The conclusions appear congruent with the direct evidence that these biologics are superior to placebo; however, for the maintenance treatment of UC in patients with moderately to severely active UC who cannot tolerate standard treatment or for whom standard treatment is contraindicated, it appears that there are no statistically significant differences between the biologics adalimumab, infliximab, and golimumab. In addition, the estimates are all in the same direction and are of approximately the same magnitude; these results are similar among the different IDCs listed in this summary with the exception of Thorlund et al., in which golimumab 100 mg was favoured (through statistical significance) over adalimumab in both clinical response and mucosal healing, and golimumab 50 mg was favoured over adalimumab for clinical response.

Review of Thorlund et al. indirect comparison

Objectives and rationale for Thorlund et al. indirect comparison[32]

Like that of the NICE IDC,[31] Thorlund et al.[32] performed an IDC on the comparative efficacy of the anti-TNF alpha drugs (adalimumab, golimumab, and infliximab) in patients with moderately to severely active UC uncontrolled on conventional treatment. However, only one previously completed IDC included one-year outcomes that included golimumab comparisons. This IDC did not account for the re-randomization that occurred at the end of the induction phase; hence, (according to the authors) there was a possibility of an overestimation of the overall placebo response rates, as patients were randomized again to placebo, golimumab 50 mg, or golimumab 100 mg, or were crossed over into golimumab 100 mg if they were not responders. Therefore, the authors conducted mathematical conversions in an attempt to account for this discrepancy and to try to make the results more comparable to trials with a parallel-group design.

Methods for Thorlund et al.[32]

Systematic review

Multiple databases were searched by two independent reviewers, and a full and extensive search strategy was provided; however, there was no specific mention of whether the two reviewers were

involved in both the article selection based on titles and abstracts, and again in determining the inclusion of the full-text articles.

Data extraction

Data were extracted by two independent reviewers, with a third reviewer verifying the information.

Trial characteristics

Five RCTs were included in the IDC: two assessed adalimumab (ULTRA 1 and ULTRA 2), one assessed golimumab (PURSUIT in which both the PURSUIT-SC and PURSUIT-M trials were considered one RCT), and two assessed infliximab (ACT-1 and ACT-2). Only the subgroups that had no previous anti-TNF alpha exposure from ULTRA 1 and ULTRA 2 were included in the analysis. The baseline characteristics and design features of the studies included in the Thorlund et al. analysis were described previously for the description of the NICE IDC (also see Table 28).

Population, interventions, comparators, and outcomes

Populations, interventions, comparisons, and outcomes of interest (and their definitions) were similar to those in the NICE IDC,[31] with the main outcomes of interest being clinical response, clinical remission, mucosal healing in both the induction and maintenance phases, along with sustained response and remission (which was specified as either response or remission "after one year of maintenance therapy conditional on response/remission after induction therapy"[32]). Thorlund et al. did not compare biologics for UC with respect to AEs or SAEs. While the induction phase had similar timelines to the NICE IDC,[31] the maintenance phase was extended to include 52 to 60 weeks. Details of the inclusion criteria for this IDC are summarized in Table 28.

Quality assessment

No description regarding the quality assessment of the included studies was provided.

Evidence network

Evidence networks were provided by Thorlund et al.[32] and were complete. None of the networks contained any closed loops and were anchored to placebo as the only common treatment arm between studies.

Meta-analysis and indirect comparison for Thorlund et al.[32]

As mentioned previously, a key source of heterogeneity is the difference in study design for the golimumab PURSUIT studies (re-randomization to treatment groups moving from the induction phase to the maintenance phase) versus the infliximab and adalimumab studies (which did not re-randomize patients following the induction phase). In order to include the golimumab PURSUIT study results in the analysis, the authors used all of the available data from both the re-randomized arms and the crossover arms as if the trial was a conventional three-arm trial. In obtaining overall maintenance proportions, the number of responders in the induction non-responders group were reduced by a factor of three, as the re-randomization process reduced the golimumab 100 mg patient population size by a factor of three; therefore, the proportion for this group was 41.8%. In addition, numbers were estimated for induction response for patients on golimumab 50 mg, as there were no data available for this (the induction study did not have a golimumab 50 mg treatment group, which was a group in the maintenance study). Estimations were obtained under the assumption that, regardless of the dose, there was a similar relationship between induction response and non-response for the maintenance period; hence, using the previous method for the golimumab 100 mg group, the estimated overall response was 38.8% for golimumab 50 mg. As only the data for those patients who were responders in the placebo group were

available, the authors estimated the overall response proportion of 13.2% (95% CrI, 10.2% to 16.8%) using Bayesian methodology and a priori distribution.

Bayesian NMA was used for all outcomes and modelled in a logistic regression; hence, odds ratios (ORs) were presented with their corresponding 95% CrI. A fixed effects model was performed when only one randomized controlled trial (RCT) contributed to the outcome. To stabilize heterogeneity, a random effects model was employed using a moderately informative prior obtained from an empirical study in which 14,000 Cochrane meta-analyses examined heterogeneity values.

Results of Thorlund et al.[32]

Induction phase

Adalimumab, infliximab, and golimumab were all observed to be statistically significantly superior to placebo for clinical response, remission, and mucosal healing in the induction phase, with the only exception being adalimumab for mucosal healing (OR 1.51; 95% CrI, 0.96 to 2.39). No statistically significant differences were evident between infliximab and golimumab for any of the outcome measures, while there were statistically significant differences in favour of infliximab over adalimumab for clinical response, remission, and mucosal healing (OR 2.19; 95% CrI, 1.03 to 4.68; OR 2.36; 95% CrI, 1.02 to 5.57; and OR 2.20; 95% CrI, 1.14 to 4.26, respectively).

Maintenance phase

The only statistically significant differences evident between the anti-TNF alpha drugs was in favour of golimumab 50 mg over adalimumab for clinical response (OR 2.31; 95% CrI, 1.23 to 4.27) and in favour of golimumab 100 mg over adalimumab for clinical response and mucosal healing (OR 2.36; 95% CrI, 1.14 to 4.86; and OR 2.03; 95% CrI, 1.05 to 3.86, respectively). Sustained response and remission were statistically significantly in favour of the anti-TNF alpha drugs compared with placebo, with the exception of adalimumab, where there was no statistically significant difference observed for sustained remission (OR 1.85; 95% CrI, 0.79 to 4.54). No statistically significant differences were observed when comparing anti-TNF alpha drugs, with the exception of sustained response, in which statistically significant differences favoured golimumab 50 mg and golimumab 100 mg over adalimumab (OR 2.11; 95% CrI, 1.00 to 4.36; and OR 2.26; 95% CrI, 1.07 to 4.17, respectively).

Detailed results for both the induction and maintenance phases are provided in Table 30.

Critical appraisal of Thorlund et al.[32]

The authors of this NMA received funding support from Janssen Canada, the manufacturers of golimumab. The manufacturer was involved in the study design; however, the authors maintain their independence with the analysis and subsequent interpretation.

Thorlund et al. did not report conducting a critical appraisal of the included RCTs; therefore, there is no indication of any potential issues with randomization, allocation concealment, blinding, or dropouts (e.g., potential biases and how they were adjusted for [or if indeed they were]). The authors did provide baseline patient and disease characteristics of the included RCTs, which were consistent for the most part across trials; however, individual study results were not reported. Like that of the other IDCs,[31,33-35] the population is generally relevant; however, there is the possibility that these results may not pertain to the sickest patients, or to those with acute UC, proctitis, or with UC localized only to the rectum (as was evidenced in the NICE IDC[31] critical appraisal and echoed by the clinical expert who participated in this review).

In order to include the PURSUIT trial in the IDC, treatment effects from this study were re-estimated as if the results of the trial had been a conventional three-arm study with a parallel group. To obtain the estimated treatment effects, a number of assumptions were made. The validity and accuracy of the estimates, based upon these assumptions, were uncertain. It did not appear that there were any sensitivity analyses performed to determine if the IDC was robust to variation in these estimates. Moreover, Thorlund et al. did not perform sensitivity analyses for the IDC with or without the PURSUIT trial included to determine its impact on the results. As such, the validity of the approach used and its potential to bias the results of the IDC is uncertain. Another potential issue with this IDC (echoed in all of the other IDCs identified) is that there were no closed loops in the network; thereby potentially decreasing the confidence in the effect estimates. Discussion and reasons pertaining to the models used (random effects versus fixed effects) were included and appropriate.

Review of Danese et al. indirect comparison

Methods for Danese et al. indirect comparison[33]

Systematic review

Multiple databases were searched, and a full and extensive search strategy was provided; however, there was no specific mention of the numbers of reviewers involved in either the primary article selection or the selection of the full-text articles for inclusion.

Data extraction

Data were extracted by two independent reviewers. Eight trials were included in the IDC: three for adalimumab (ULTRA 1, ULTRA 2, and Suzuki et al.), two for golimumab (PURSUIT-SC and PURSUIT-M), two for infliximab (ACT-1 and ACT-2), and one for vedolizumab (GEMINI 1). Only data for those patients who were naive to biologic treatment were extracted along with data only for approved doses and administrations (with the only exception being vedolizumab as both dosing regimens [300 mg IV every eight weeks or four weeks] were included and had not received regulatory approval). In addition, different doses of the same biologic drug were treated as separate interventions. Clinical response, remission, and mucosal healing and their corresponding definitions were the primary outcomes of interest, were similar between trials, and were extracted for both the induction and maintenance phases, if available.

AEs, infectious AEs, AEs leading to study drug discontinuation, and SAEs were also extracted.

Population, interventions, comparators, and outcomes

Populations, interventions, comparisons, and outcomes of interest were similar to the other identified IDCs,[31,32,34,35] with the main exception that Danese et al. also included vedolizumab as an intervention and comparator. The baseline characteristics between trials were similar, as were the trial designs (with the exception of PURSUIT-M and GEMINI 1, in which only patients who responded in the induction phase were eligible for re-randomization into the maintenance phase). Clinical response, remission, and mucosal healing and their corresponding definitions were similar between trials and were extracted for both the induction and maintenance phases, if available. They were presented as ORs (calculated using ITT principals). Details of the inclusion criteria for this IDC are summarized in Table 28.

Quality assessment

The Cochrane Risk of Bias tool was used by an unspecified number of multiple reviewers in order to assess the risk of bias of the included RCTs. Consensus was reached through discussion if there were disagreements.

Evidence network

The induction evidence network was provided by Danese et al.;[33] however, no separate evidence network was provided for the maintenance phase. The network did not contain any closed loops and was anchored to placebo as the only common treatment arm between studies.

Meta-analysis and indirect comparison for Danese et al.[33]

The NMA was conducted in WINBUGS using a Bayesian framework and Markov chain Monte Carlo methods. Non-informative priors for relative-effect parameters and between-study standard deviation were used for the analyses. A burn-in phase of 5,000 simulations was performed before convergence, and the lack of auto correlation was checked. This was followed by a 20,000 burn-in iteration phase and a 50,000 burn-in iteration phase for estimate parameters. Only those estimates for the default model were reported, as no differences in model fit, according to the deviance information criterion (DIC), were observed. Fixed effects ORs and their corresponding 95% CrIs were also calculated. No formal assessments of publication bias or statistical heterogeneity were completed, as a limited number of RCTs informed the pairwise comparisons.

Results of Danese et al.[33]

Clinical efficacy

Induction phase

Across induction trials (ACT-1, ACT-2, ULTRA 1, ULTRA 2, Suzuki et al., PURSUIT-SC, and GEMINI 1), the risk of bias was assessed to be low.

Statistical significance in favour of the biologics over placebo was reported for clinical response, remission, and mucosal healing in the induction phase, with the exception of adalimumab, which was not statistically different from placebo for clinical remission (OR 1.91; 95% CrI, 0.98 to 3.72). There was no statistically significant difference between golimumab, infliximab, or vedolizumab; however, ORs in favour of infliximab compared with adalimumab were evident for clinical response and mucosal healing in the induction phase (OR 2.36; 95% CrI, 1.22 to 4.63; and OR 2.02; 95% CrI, 1.13 to 3.59, respectively). Details are provided in Table 30.

Maintenance phase

The risk of bias was assessed as high in the incomplete outcome data domain across all of the maintenance trials (ACT-1, ACT-2, ULTRA 2, Suzuki et al., PURSUIT-M, and GEMINI 1), mainly due to the lower follow-up (55% to 76%) in this phase (compared with 91% to 97% in the induction phase).

Only results when compared to placebo were reported for the maintenance phase. Statistically significant ORs in favour of adalimumab, golimumab, and infliximab over placebo were reported. No comparative efficacy results were reported for any of the anti-TNF alpha drugs or vedolizumab for the maintenance phase.

Safety

Median event ranges for AEs in the Danese et al. IDC[33] indicated that more AEs occurred in the biologics group than in the placebo group. In addition, there were more infectious AEs in the biologics group than in the placebo group.

Critical appraisal of Danese et al.[33]

Key limitations included a lack of description of how different treatment groups based on dose were handled in the NMA, a lack of presentation of results from traditional pairwise meta-analyses, and a poor description of how fixed versus random effects models were chosen. In addition, there was minimal description with regard to how assumptions related to potential confounding factors and effect modifiers were assessed and handled.

As with the other IDCs,[31,32,34,35] another limitation is the relatively sparse data and considerable limitations of the included RCTs, particularly the lack of head-to-head studies between anti-TNF drugs. Placebo was the only link between the studies in the NMA. This was especially an issue for assessing the relative treatment effects of these drugs on maintenance treatment of UC, for which only three RCTs were included in this review. Danese et al. reported they were unable to perform NMA for maintenance treatment because of the methodological heterogeneity among these studies.

A substantial proportion of patients in the longer-term RCTs withdrew from the studies. For example, the two large manufacturer-sponsored trials for infliximab were marred by large withdrawals and loss to follow-up in the placebo group at week 30 (ACT-1 and ACT-2) and at week 54 in ACT-1, thereby making these data unreliable. The most common reason for discontinuation was lack of efficacy, which in ACT-1, for example, occurred approximately two times more frequently in the placebo group than in either of the two infliximab groups.

Another key issue is the design heterogeneity between PURSUIT-M and the other large maintenance treatment studies for adalimumab (ULTRA 2) and infliximab (ACT-1). In PURSUIT-M, patients who responded to treatment during the induction phase studies were re-randomized for the maintenance phase, as previously described in Thorlund et al.[32] Of note, in their NMA, Stidham et al. conducted a sensitivity analysis by excluding PURSUIT-M from the analysis, but this did not affect either the direction or the magnitude of the treatment effects in the model.

Safety outcomes were poorly reported, and therefore limited the ability for an NMA of AEs. Indeed, although Danese et al. planned to conduct such a comparison, they did not, as the data were sparse.

Conclusion

Five IDCs were included in this summary and critical appraisal, one of which was provided by the manufacturer (NICE[31]), and four others that were identified from an electronic database search (Thorlund et al.,[32] Danese et al.,[33] Stidham et al.,[34] and Galvan-Banqueri et al.[35]). The NICE, [31] Stidham et al.,[34] and Galvan-Banqueri et al.[35] IDCs were similar in their methods, populations, comparators, outcomes, and results, while the IDC by Danese et al.[33] also included vedolizumab as a treatment comparator, and the IDC by Thorlund et al.[32] provided an additional mathematical model to take into account the difference in the trial design of the golimumab maintenance study, PURSUIT-M.

While most of the IDCs provided evidence to support the superiority of the anti-TNF alpha biologics over placebo in both the induction[31-35] and maintenance phases[32-35] of treatment for clinical response, clinical remission, and mucosal healing, there were small differences observed between IDCs with regard to the NMA biologic comparisons. With regard to clinical response and mucosal healing in the induction phase, the IDCs by Danese et al.,[33] Galvan-Banqueri et al.,[35] and Thorlund et al.[32] produced statistically significant results favouring infliximab over adalimumab. In addition, Thorlund et al.[32] also produced statistically significant results favouring infliximab over adalimumab for clinical remission in the induction phase. In contrast, however, no statistically significant results favouring either infliximab, adalimumab, or golimumab for either clinical response or clinical remission in the induction or

maintenance phases were observed in the IDC by Stidham et al.[34] Finally, comparisons of golimumab versus adalimumab, infliximab versus golimumab, vedolizumab versus adalimumab, vedolizumab versus golimumab, and infliximab versus vedolizumab in the induction phase produced no statistically significant results favouring any of the aforementioned biologic interventions over another.

With regard to results from the IDC for maintenance phase treatment, most IDCs produced non-statistically significant results when comparing biologics treatments against one another for clinical response,[31,34,35] clinical remission,[31,34,35] or mucosal healing.[31,35] Only Thorlund et al. reported statistically significant differences favouring golimumab 100 mg over adalimumab for clinical response and mucosal healing, and for golimumab 50 mg over adalimumab for clinical response; however, questionable statistical methods were used to account for the different design of the PURSUIT-M trial.[32]

The Danese et al. IDC[33] suggested that the occurrence of AEs was not different between placebo and the biologic drugs (including vedolizumab). The NICE IDC[31] identified similar trends to that of Danese et al. with regard to AEs. However, these assessments were descriptive and no NMA analyses were conducted for AEs; hence, the comparative safety of the biologic drugs in UC remains uncertain.

The most conservative conclusion based on the available indirect evidence is that there is no clear evidence that there is a difference between the biologics, including adalimumab, with respect to inducing and maintaining remission, response, and mucosal healing in moderately to severely active UC. The numerous limitations identified preclude a more definitive conclusion, particularly based on the small number of studies with no head-to-head comparisons. In particular, the lack of consistent results across the IDCs makes it difficult to draw firm conclusions regarding the comparative efficacy of adalimumab versus other biologic treatments for inducing response and remission in UC. The methodological heterogeneity between studies means there is a high degree of uncertainty about whether these drugs are similarly efficacious in maintaining remission and response.

REFERENCES

1. Fast facts: the impact of IBD in Canada 2012 [Internet]. Toronto: Crohn's and Colitis Foundation of Canada; 2012. [cited 2016 Jan 25]. Available from: http://www.isupportibd.ca/pdf/ccfc.ca-impact-report-fast-facts.pdf
2. [Pr]Humira® (adalimumab): 40 mg in 0.8 mL sterile solution (50 mg/mL) subcutaneous injection [product monograph]. St-Laurent (QC): AbbVie Corporation; 2015 Jul 3.
3. Clinical study report: M06-826. A multicenter, randomized, double-blind, placebo controlled study of the human anti-TNF monoclonal antibody adalimumab for the induction of clinical remission in subjects with moderately to severely active ulcerative colitis [**CONFIDENTIAL** internal manufacturer's report]. Abbott Park (IL): Abbott Laboratories; 2012 Dec.
4. Clinical study report: M06-827. A multicenter, randomized, double-blind, placebo controlled study of the human anti-TNF monoclonal antibody adalimumab for the induction and maintenance of clinical remission in subjects with moderately to severely active ulcerative colitis [**CONFIDENTIAL** internal manufacturer's report]. Abbott Park (IL): Abbott Laboratories; 2012 Dec.
5. Suzuki Y, Motoya S, Hanai H, Matsumoto T, Hibi T, Robinson AM, et al. Efficacy and safety of adalimumab in Japanese patients with moderately to severely active ulcerative colitis. J Gastroenterol [Internet]. 2014 Feb [cited 2015 Nov 16];49(2):283-94. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3925299
6. e-CPS [Internet]. Ottawa: Canadian Pharmacists Association; 2009 -. Simponi, golimumab; 2013 Sep 19 [cited 2016 Jan 28]. Available from: https://www.e-therapeutics.ca Subscription required.
7. e-CPS [Internet]. Ottawa: Canadian Pharmacists Association; 2009 -. Remicade, infliximab; 2013 Aug 26 [cited 2016 Jan 26]. Available from: https://www.e-therapeutics.ca Subscription required.
8. e-CPS [Internet]. Ottawa: Canadian Pharmacists Association; 2009 -. Humira, adalimumab; 2014 Jan 17 [cited 2016 Jan 26]. Available from: https://www.e-therapeutics.ca Subscription required.
9. e-CPS [Internet]. Ottawa: Canadian Pharmacists Association; 2009 -. Entyvio, vedolizumab; 2015 Jan 25 [cited 2016 Jan 28]. Available from: https://www.e-therapeutics.ca Subscription required.
10. Reinisch W, Sandborn WJ, Hommes DW, D'Haens G, Hanauer S, Schreiber S, et al. Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial. Gut. 2011 Jun;60(6):780-7.
11. Reinisch W, Sandborn WJ, Panaccione R, Huang B, Pollack PF, Lazar A, et al. 52-week efficacy of adalimumab in patients with moderately to severely active ulcerative colitis who failed corticosteroids and/or immunosuppressants. Inflamm Bowel Dis. 2013 Jul;19(8):1700-9.
12. CDR submission: Humira (adalimumab) for ulcerative colitis, 40 mg in 0.8 mL sterile solution (50 mg/mL). Company: AbbVie Corporation [**CONFIDENTIAL** manufacturer's submission]. Dorval (QC): AbbVie Corporation; 2015 Oct 27.
13. Health Canada reviewer's report: Humira (adalimumab) [**CONFIDENTIAL** internal report]. Ottawa: Therapeutics Products Directorate, Health Canada; 2013 Nov 21.
14. Sandborn WJ, van Assche G, Reinisch W, Colombel JF, D'Haens G, Wolf DC, et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2012 Feb;142(2):257-65.
15. Sandborn WJ, Colombel JF, D'Haens G, van Assche G, Wolf D, Kron M, et al. One-year maintenance outcomes among patients with moderately-to-severely active ulcerative colitis who responded to induction therapy with adalimumab: subgroup analyses from ULTRA 2. Aliment Pharmacol Ther. 2013 Jan;37(2):204-13.

16. Wolf D, D'Haens G, Sandborn WJ, Colombel JF, van Assche G, Robinson AM, et al. Escalation to weekly dosing recaptures response in adalimumab-treated patients with moderately to severely active ulcerative colitis. Aliment Pharmacol Ther. 2014 Sep;40(5):486-97.

17. 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 [Internet]. 2009 May 1 [cited 2016 Jan 26];29(9):1032-41. Available from: http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2036.2009.03966.x/epdf

18. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. Ann Med. 2001 Jul;33(5):350-7.

19. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. PharmacoEconomics. 1999 Feb;15(2):141-55.

20. Strand V, Singh JA. Improved health-related quality of life with effective disease-modifying antirheumatic drugs: evidence from randomized controlled trials. Am J Manag Care. 2008 Apr;14(4):239-53.

21. 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.

22. Clinical Study Report: M10-223. A multicenter,open-label study of the human anti-TNF monoclonal antibody adalimumab to evaluate the long-term safety and tolerability of repeated administration of adalimumab in subjects with ulcerative colitis [**CONFIDENTIAL** internal manufacturer's report]. Abbott Park (IL): Abbott Laboratories; 2009 Dec 31.

23. Lewis JD, Chuai S, Nessel L, Lichtenstein GR, Aberra FN, Ellenberg JH. Use of the non-invasive components of the Mayo score to assess clinical response in ulcerative colitis. Inflamm Bowel Dis [Internet]. 2008 Dec [cited 2015 Nov 19];14(12):1660-6. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2597552/pdf/nihms65682.pdf

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

25. 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.

26. 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.

27. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. conceptual framework and item selection. Med Care. 1992 Jun;30(6):473-83.

28. Mease PJ, Menter MA. Quality-of-life issues in psoriasis and psoriatic arthritis: outcome measures and therapies from a dermatological perspective. J Am Acad Dermatol. 2006 Apr;54(4):685-704.

29. 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.

30. Colombel JF, Sandborn WJ, Ghosh S, Wolf DC, Panaccione R, Feagan B, et al. Four-year maintenance treatment with adalimumab in patients with moderately to severely active ulcerative colitis: Data from ULTRA 1, 2, and 3. Am J Gastroenterol. 2014 Nov [cited 2016 Jan 26];109(11):1771-80. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4223868

31. Archer R, Tappenden P, Ren S, Martyn-St James M, Harvey R, Besarir H, et al. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262) [ID695]. Sheffield, UK: University of Sheffield; 2014.

32. Thorlund K, Druyts E, Toor K, Mills EJ. Comparative efficacy of golimumab, infliximab, and adalimumab for moderately to severely active ulcerative colitis: a network meta-analysis accounting for differences in trial designs. Expert Rev Gastroenterol Hepatol. 2015 May;9(5):693-700.

33. Danese S, Fiorino G, Peyrin-Biroulet L, Lucenteforte E, Virgili G, Moja L, et al. Biological agents for moderately to severely active ulcerative colitis: a systematic review and network meta-analysis. Ann Intern Med. 2014 May 20;160(10):704-11.

34. Stidham RW, Lee TC, Higgins PD, Deshpande AR, Sussman DA, Singal AG, et al. Systematic review with network meta-analysis: the efficacy of anti-tumour necrosis factor-alpha agents for the treatment of ulcerative colitis. Aliment Pharmacol Ther. 2014 Apr;39(7):660-71.

35. Galvan-Banqueri M, Vega-Coca MD, Castillo-Munoz MA, Beltran CC, Molina LT. Indirect comparison for Anti-TNF drugs in moderate to severe ulcerative colitis. Farm Hosp [Internet]. 2015 Mar [cited 2015 Nov 16];39(2):80-91. Available from: http://www.aulamedica.es/fh/pdf/8218.pdf