CADTH

CADTH COMMON DRUG REVIEW

Clinical Review Report

Upadacitinib (Rinvoq)

(AbbVie)

Indication: For the treatment of adults with moderately to severely active rheumatoid arthritis who have had an inadequate response or intolerance to methotrexate.

Service Line:	CADTH Common Drug Review
Version:	Final (with redactions)
Publication Date:	March 2020
Report Length:	116 Pages

CADTH

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

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

CADTH

Table of Contents

Abbreviations 6
Executive Summary 7
Introduction 7
Stakeholder Engagement 8
Clinical Evidence 9
Conclusions 19
Introduction 20
Disease Background 20
Standards of Therapy 20
Drug 21
Stakeholder Engagement 23
Patient Group Input 23
Clinician Input 26
Clinical Evidence 29
Systematic Review (Pivotal and Protocol-Selected Studies) 29
Findings From the Literature 32
Results 45
Indirect Evidence 65
Other Relevant Studies 77
Discussion 88
Summary of Available Evidence 88
Interpretation of Results 90
Conclusions 92
Appendix 1: Literature Search Strategy 93
Appendix 2: Excluded Studies 96
Appendix 3: Detailed Outcome Data 97
Appendix 4: Description and Appraisal of Outcome Measures 103
References 113

Tables

Table 1: Summary of Key Results From Pivotal and Protocol-Selected Studies 11
Table 2: Key Characteristics of JAK Inhibitors, IL-6 Inhibitors, T-Cell Co-Stimulation Modulators, CD20 Inhibitors, IL-1 Inhibitors, and TNF Inhibitors 22
Table 3: Inclusion Criteria for the Systematic Review 30
Table 4: Details of Included Studies 33
Table 5: Summary of Baseline Characteristics (NEXT, BEYOND) 39
Table 6: Baseline Characteristics (COMPARE, MONOTHERAPY, EARLY) 40
Table 7: EULAR Improvement Response Criteria (DAS28) 42
Table 8: Multiple Testing Hierarchical Testing Procedure 45
Table 9: Patient Disposition up to Week 14 (COMPARE, MONOTHERAPY) 46
Table 10: Patient Disposition up to Week 12 (NEXT, BEYOND) 46
Table 11: Patient Disposition up to Week 24 (EARLY) 47
Table 12: Exposure to Study Treatments (Safety Analysis Set) 48
Table 13: Clinical Response (ACR20/50/70) at Week 12 or Week 14 49
Table 14: Radiographic Response: mTSS at Week 24 or 26 (COMPARE, EARLY) 51
Table 15: Radiographic Response: No Radiographic Progression at Week 24 or 26 (COMPARE, EARLY) 51
Table 16: Disease Activity: DAS28-CRP at Week 12 or 14 52
Table 17: Disease Activity Outcomes: LDA, CR 53
Table 18: Quality of Life Outcomes at Week 12 or 14 55
Table 19: Patient Global Assessment of Pain at Week 12 or 14 56
Table 20: Functional and Disability Outcomes at Week 12 or 14 57
Table 21: Summary of Harms (COMPARE up to Week 26, Censored at Treatment Switching) 58
Table 22: Summary of Harms (MONOTHERAPY up to Week 14, EARLY up to Week 24) 60
Table 23: Summary of Harms (NEXT up to Week 12, BEYOND up to Week 12) 62
Table 24: Study Selection Criteria and Methods for ITCs 65
Table 25: ITC Analysis Methods 68
Table 26: Probability of ACR Response at 12 Weeks: csDMARD Inadequate Responders 69
Table 27: ACR Response at 24 Weeks: csDMARD Inadequate Responders 71
Table 28: ACR Response at 12 Weeks: bDMARD Inadequate Responders 73
Table 29: League Tables Showing Odds Ratios and 95% Credible Intervals: (A) Efficacy and (B) Safety 76
Table 30: Long-Term Efficacy Results (SELECT-COMPARE): Dichotomous Outcomes 80

Table 31: Long-Term Efficacy Results (SELECT-COMPARE): Continuous Outcomes 81
Table 32: Long-Term Efficacy Results (SELECT-MONOTHERAPY): Dichotomous Outcomes 82
Table 33: Long-Term Efficacy Results (SELECT-MONOTHERAPY): Continuous Outcomes 82
Table 34: Long-Term Efficacy Results (SELECT-NEXT): Dichotomous Outcomes 83
Table 35: Long-Term Efficacy Results (SELECT-NEXT): Continuous Outcomes 83
Table 36: Long-Term Efficacy Results (SELECT-BEYOND): Dichotomous Outcomes 84
Table 37: Long-Term Efficacy Results (SELECT-BEYOND): Continuous Outcomes 84
Table 38: Long-Term Efficacy Results (SELECT-EARLY): Dichotomous Outcomes 85
Table 39: Long-Term Efficacy Results (SELECT-EARLY): Continuous Outcomes 85
Table 40: Long-Term Safety Outcomes: Methotrexate-Controlled Upadacitinib 15 mg, at Six Months (Data From SELECT-EARLY and SELECT-MONOTHERAPY) 86
Table 41: Long-Term Safety Outcomes: Any Phase III Upadacitinib 15 mg Analysis Set, One-Year Exposure 87
Table 42: Excluded Studies 96
Table 43: Clinical Response (ACR20/50/70) at Week 2 97
Table 44: Functional and Disability Outcomes: Fatigue 98
Table 45: Subgroup Data: ACR20 at Week 12 (Primary Outcome) by Disease Activity 99
Table 46: Subgroup Data: ACR20 at Week 12 (Primary Outcome) by Prior Treatment Failure 100
Table 47: Summary of Harms (COMPARE up to Week 14, Prior to Availability of Rescue Therapy) 100
Table 48: Summary of Harms (COMPARE up to Week 26, After Treatment Switching) 102
Table 49: Outcome Measures Included in Each Study 103
Table 50: Summary of Outcome Measures and Their Measurement Properties 103
Table 51: EULAR Improvement Response Criteria (DAS28) 107
Table 52: Scoring for the Modified Total Sharp Score 109

Figure

Figure 1: Flow Diagram for Inclusion and Exclusion of Studies 32

Abbreviations

ACE	Arthritis Consumer Experts
ACR	American College of Rheumatology
ACR20	20% improvement in the American College of Rheumatology criteria
ACR50	50% improvement in the American College of Rheumatology criteria
ACR70	70% improvement in the American College of Rheumatology criteria
bDMARD	biologic disease-modifying antirheumatic drug
CAPA	Canadian Arthritis Patient Alliance
CCP	cyclic citrullinated peptide
CDAI	Clinical Disease Activity Index
cDMARD	conventional disease-modifying antirheumatic drug
CI	confidence interval
CrI	credible interval
CRP	C-reactive protein
csDMARD	conventional synthetic disease-modifying antirheumatic drug
DAS	Disease Activity Score
DMARD	disease-modifying antirheumatic drug
EULAR	European League Against Rheumatism
ESR	erythrocyte sedimentation rate
FACIT	Functional Assessment of Chronic Illness Therapy
HAQ-DI	Health Assessment Questionnaire–Disability Index
HRQoL	health-related quality of life
ITC	indirect treatment comparison
JAK	Janus kinase
LDA	low disease activity
MCID	minimal clinically important difference
MCS	mental component summary
mTSS	modified total Sharp score
NMA	network meta-analysis
NSAID	nonsteroidal anti-inflammatory drug
OR	odds ratio
PCS	physical component summary
RA	rheumatoid arthritis
RCT	randomized controlled trial
SD	standard deviation
SDAI	Simplified Disease Activity Index
SF-36	Short Form (36) Health Survey
VAS	visual analogue scale

Drug	Upadacitinib (Rinvoq)
Indication	For the treatment of adults with moderately to severely active rheumatoid arthritis who have had an inadequate response or intolerance to methotrexate. Upadacitinib may be used as monotherapy or in combination with methotrexate or other nonbiologic disease-modifying antirheumatic drugs (DMARDs).
Reimbursement request	As per indication; reimburse in a similar manner to biologic DMARDs and targeted synthetic DMARDs for the treatment of moderate to severe rheumatoid arthritis.
Dosage form(s) and route of administration and strength(s)	15 mg extended-release tablets, oral
NOC date	December 23, 2019
Sponsor	AbbVie

Executive Summary

Introduction

Rheumatoid arthritis (RA) is an autoimmune inflammatory disease that primarily affects the joints of the body. Characterized by acute and chronic inflammation of the synovium, or soft tissue surrounding the joints, patients are subject to severe pain, stiffness, and fatigue, all of which can affect a patient's ability to perform activities of daily living and overall health-related quality of life (HRQoL). Prolonged inflammation may lead to damage and destruction of the joints through erosion of the cartilage and bone and, consequently, disability and premature mortality. Other areas of the body may be affected as well, including the eyes, lungs, heart, or skin. It is estimated that about 1% of Canadians have the disease.

Upadacitinib is a Janus kinase (JAK) inhibitor. JAK mediates the effects of cytokines and their production, thus JAK inhibitors may have more of a global effect on various cytokine production than do biologics, which tend to target specific cytokines. Upadacitinib is the third JAK inhibitor approved in Canada, the first being tofacitinib, followed by baricitinib, both of which were previously reviewed and issued recommendations by the CADTH Canadian Drug Expert Committee.

The systematic review protocol for the current review was established before the granting of Notice of Compliance from Health Canada for upadacitinib. The objective is to perform a systematic review of the beneficial and harmful effects of upadacitinib 15 mg extended-release tablets for once daily administration for the treatment of moderate to severe RA in adult patients who have responded inadequately to, or who are intolerant to, one or more disease-modifying antirheumatic drugs (DMARDs). Upadacitinib may be used as monotherapy or in combination with methotrexate or other conventional synthetic DMARDs (csDMARDs).

Stakeholder Engagement

Patient Input

CADTH received a joint submission from the Arthritis Society and the Canadian Arthritis Patient Alliance (CAPA) and a second submission from the Arthritis Consumer Experts (ACE). The patient groups described RA as a chronic autoimmune disease in which the body's immune system attacks and causes inflammation in joints. When uncontrolled, inflammation can result in irreversible damage to the affected joints. Some patients also reported living with fibromyalgia, myasthenia gravis, and asthma. Patients with RA reported that they may experience periods of active disease (flares or flare-ups) and periods of decreased activity of the disease (remission). Flares can be unpredictable in their onset, frequency, and length. The patient groups also reported that symptoms and manifestation of the disease can vary from patient to patient. Some patients reported having to deal with flares reactively, feeling that they do not have control over their RA.

Patients explained that relief of pain, fatigue, and stiffness, and the ability to maintain mobility, are important outcomes for them. The patient groups indicated that for many patients with RA, treatment is determined by trial and error. Several patients reported that they are taking or have tried up to four different RA treatments. While a few patients indicated that their present treatment is controlling their disease, the remainder described a continuing struggle to find both an effective and tolerable treatment. According to the survey responses, treatments are difficult to tolerate because of the side effects. Furthermore, some of the side effects require additional treatment with other medications, such as folic acid or antiemetics. Generally, patients expect upadacitinib to be easy to take, to treat symptoms of pain and stiffness, and to increase their mood, sense of independence, and overall quality of life.

Clinician Input

The following input is a summary of information provided by one clinical specialist with expertise in the diagnosis and management of RA.

The JAK inhibitors represent the latest therapeutic advance in RA. Unlike the biologics, they are small molecules for oral administration. A biologic drug is specific in affecting an immunologic mechanism. The JAK inhibitors are not specific and affect several immunologic mechanisms. Upadacitinib is the third JAK inhibitor and is the most selective JAK inhibitor to date. Tofacitinib is a pan-JAK inhibitor. Baricitinib is more selective against JAK 1 and 2, and upadacitinib is a selective JAK 1 inhibitor. JAK inhibitors have been shown to be DMARDs. In addition to controlling signs and symptoms of disease and improving functional status, JAK inhibitors inhibit radiographic progression. An ideal treatment would result in remission, a drug-free immunologic remission, or cure. More pragmatically, effective RA treatment is intended to delay or stop disease progression, although requiring long-term drug use. However, each new treatment has been associated with benefit in a percentage of patients heretofore unsuccessfully treated. This is the rationale for ongoing therapy development.

Clinically, any patient who is diagnosed with RA can be considered for treatment; classification criteria include number of inflamed joints and the level of disease activity being at least moderate and most likely high. Rheumatologists use a treat-to-target strategy. The goal of therapy is to achieve remission, and if that is not possible, the goal is

to achieve low disease activity (LDA). Several definitions of remission are used in clinical practice. Patients should be assessed every three to six months.

Clinical Evidence

Pivotal Studies and Protocol-Selected Studies

Description of Studies

Five pivotal multinational double-blind randomized controlled trials (RCTs) met the criteria for this systematic review: SELECT-COMPARE, SELECT-MONOTHERAPY, SELECT-NEXT, SELECT-BEYOND, and SELECT-EARLY. All five studies enrolled adults with adult-onset RA, and in all but the EARLY trial (where patients were methotrexate naive), patients' symptoms had been inadequately controlled with a DMARD. In COMPARE (N = 1,629) patients had to be inadequately controlled on methotrexate, in NEXT (N = 661) patients had to be inadequately controlled on any csDMARD, and in BEYOND (N = 499) patients had to be inadequately controlled on a biologic DMARD (bDMARD). The five studies reflected patients with RA who had received various prior treatments and had switched to different regimens. In EARLY (N = 947), patients could not have been considered as methotrexate inadequate responders and could not have been on any DMARD beside methotrexate for longer than three weeks; the patients were randomized to receive upadacitinib 15 mg once daily, upadacitinib 30 mg once daily, or methotrexate (7.5 mg or 10 mg) once weekly. In MONOTHERAPY, patients had had a csDMARD experience and patients were randomized to upadacitinib 15 mg once daily, upadacitinib 30 mg once daily, or methotrexate (to continue on a prior stable dose) with no other added therapy in the background. In COMPARE, patients had been considered as csDMARD inadequate responders but not bDMARD inadequate responders and moved into the trial with a background therapy of methotrexate; they were randomized (2:1:2) into upadacitinib 15 mg once daily, adalimumab 40 mg injection every other week, or placebo. In NEXT, patients had been considered csDMARD inadequate responders but not bDMARD inadequate responders and moved into the study with csDMARD background therapy; they were randomized into upadacitinib 15 mg once daily, upadacitinib 30 mg once daily, or placebo. In BEYOND, patients had been considered bDMARD inadequate responders and moved into the study with csDMARD background therapy; they were randomized into upadacitinib 15 mg once daily, upadacitinib 30 mg once daily, or placebo. The primary outcome in the COMPARE, NEXT, and BEYOND studies was the proportion of patients achieving an American College of Rheumatology (ACR) 20 (20% improvement in ACR criteria) response at 12 weeks. The primary outcome in MONOTHERAPY was achieving ACR20 at 14 weeks, and the primary outcome in EARLY was achieving 50% improvement in ACR criteria (ACR50) at 12 weeks. Key secondary outcomes that were accounted for type I errorincluded HRQoL on the Health Assessment Questionnaire–Disability Index (HAQ-DI), the Disease Activity Score-28 and C-reactive protein (DAS28-CRP), and the modified total Sharp score (mTSS). The primary outcome was reported at 12 weeks in all studies except MONOTHERAPY, in which it was reported at week 14. Each of these studies has an ongoing long-term extension study.

Efficacy Results

Efficacy results are summarized in Table 1. The primary outcome in EARLY of ACR50 at week 12 showed a response rate of 52.1% (95% confidence interval [CI], 46.6 to 57.5) in the upadacitinib and 28.3% (95% CI, 23.4 to 33.3) in the methotrexate arms, with a rate difference of 23.7 (95% CI, 16.3 to 31.1). The primary outcome in MONOTHERAPY of

ACR20 at week 14 showed a response rate of 67.7% (95% CI, 61.5 to 74.0) in the upadacitinib and 41.2% (95% CI, 34.6 to 47.8) in the methotrexate arms, with a rate difference of 26.5 (95% CI, 17.5 to 35.6). The primary outcome in COMPARE of ACR20 at week 12 showed a response of 70.5% (95% CI, 67.0 to 74.0) in the upadacitinib, 63.0% (95% CI, 57.8 to 68.2) in the adalimumab, and 36.4% (95% CI, 32.7 to 40.1) in the placebo arms, with a rate difference of 7.5 (95% CI, 1.2 to 13.8) versus adalimumab and 34.1 (95% CI, 29.0 to 39.2) versus placebo. The primary outcome in NEXT of ACR20 at week 12 was 63.8% (95% CI, 57.5 to 70.1) in the upadacitinib and 35.7% (95% CI, 29.4 to 42.1) in the placebo groups, with a rate difference of 28.1 (95% CI, 19.1 to 37.0). In BEYOND, the primary outcome of ACR20 at week 12 was 64.6% (95% CI, 57.3 to 72.0) in the upadacitinib and 28.4% (95% CI, 21.6 to 35.2) in the placebo groups, with a rate difference of 36.2 (95% CI, 26.2 to 46.2).

The HAQ-DI was a secondary outcome in all studies. Results of upadacitinib 15 mg versus the placebo and methotrexate groups showed a statistically significant magnitude of difference greater than the estimated minimal important difference of 0.22 in all the studies at week 12 (week 14 in MONOTHERAPY). The DAS28-CRP was a secondary outcome in all studies. Results of upadacitinib 15 mg versus the placebo and methotrexate groups showed a statistically significant mean difference in the results in all the studies. The mTSS was reported in COMPARE (week 24) and EARLY (week 26) as a secondary outcome, with a responder analysis where a responder was defined as having no change in mTSS. In both studies, the mean difference was statistically significantly in favour of upadacitinib and, similarly, the response rate in upadacitinib-treated patients was statistically significantly higher than placebo (COMPARE, rate difference 7.5 [95% CI, 3.0 to 12.1]) and than methotrexate (EARLY, rate difference 9.8 [95% CI, 3.5 to 16.2]). Other secondary outcomes included the proportion of patients with LDA, the Short Form (36) Health Survey (SF-36), and morning stiffness. Overall, these results are consistent in showing the benefit of upadacitinib 15 mg once daily over the placebo and methotrexate control arms.

One of the comparisons provided in the studies was that of upadacitinib versus adalimumab in the COMPARE study. The sponsor's initial outcome for this comparison was to achieve a noninferiority on ACR50 at week 12; the result showed upadacitinib to be statistically superior to adalimumab (rate difference 16.1 [95% CI, 9.9 to 22.3]). In addition, upadacitinib was superior to adalimumab in the HAQ-DI measure, but the treatment difference did not exceed the identified minimal important difference of 0.22 points (least squares mean difference –0.11 [95% CI, –0.184 to –0.036]). Other outcomes beyond ACR50, HAQ-DI, and patients' assessments of pain, in comparing upadacitinib 15 mg to adalimumab, were outside the statistical testing hierarchy. However, upadacitinib showed better results than adalimumab in all examined outcomes except in mTSS.

Harms Results

In COMPARE, 64.2% of upadacitinib patients, 60.2% of adalimumab patients, and 53.2% of placebo patients experienced an adverse event. In MONOTHERAPY, the percentages were 47.5% in upadacitinib and 47.2% in placebo groups. In NEXT, the percentages were 56.6% in upadacitinib and 48.9% in placebo groups. In BEYOND, the percentages were 55.5% in upadacitinib and 56.2% in placebo groups. In EARLY, the percentages were 64.0% in upadacitinib and 65.3% in placebo groups. Respiratory tract infections were the most common adverse events in all the included studies. Serious adverse generally occurred in less than 5% of patients across the studies. In COMPARE, 3.7% of upadacitinib patients, 4.3% of adalimumab patients, and 2.9% of placebo patients experienced a serious adverse event. In MONOTHERAPY, the percentages were 5.1% in upadacitinib and 2.8%

in methotrexate groups. In NEXT, the percentages were 4.1% in upadacitinib and 2.3% in placebo groups. In BEYOND, the percentages were 4.9% in upadacitinib and 0% in placebo groups. In EARLY, the percentages were 4.7% in upadacitinib and 4.1% in methotrexate groups. No single serious adverse event was most common across the five included studies.

According to the available data, there was a numerically higher incidence of herpes zoster infection in the upadacitinib treatment groups when contrasted to non-upadacitinib treatment groups. Over the course of the efficacy and extension phases of the studies, malignancies were reported in 1.1% of all patients who started and stayed on upadacitinib and 1.2% of patients who switched over to upadacitinib for the extension phase. Overall, notable harms identified for this review did not show explicit imbalance between groups, with the exception of a numerically higher proportion of neutropenia in COMPARE, BEYOND, and EARLY. Also, there was no explicit imbalance in the number of thromboembolic events between upadacitinib-treated patients and other groups.

Table 1: Summary of Key Results From Pivotal and Protocol-Selected Studies

	Total N	Responder	Response rate (95% CI)	Response rate difference versus control	
		n		Point estimate (95% CI)	P value
ACR20 response at week 12 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	459	70.5 (67.0 to 74.0)	7.5 (1.2 to 13.8) [UPA vs. ADA] 34.1 (29.0 to 39.2) [UPA vs. PBO]	0.018[a] [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	206	63.0 (57.8 to 68.2)		
Placebo	651	237	36.4 (32.7 to 40.1)		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	147	67.7 (61.5 to 74.0)	26.5 (17.5 to 35.6)	< 0.001
MTX	216	89	41.2 (34.6 to 47.8)		
SELECT-NEXT					
UPA 15 mg q.d.	221	141	63.8 (57.5 to 70.1)	28.1 (19.1 to 37.0)	< 0.001
Placebo	221	79	35.7 (29.4 to 42.1)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	106	64.6 (57.3 to 72.0)	36.2 (26.2 to 46.2)	< 0.001
Placebo	169	48	28.4 (21.6 to 35.2)		

	Total N	Responder	Response rate (95% CI)	Response rate difference versus control	
		n		Point estimate (95% CI)	P value
ACR50 response at week 12 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	294	45.2 (41.3 to 49.0)	16.1 (9.9 to 22.3) [UPA vs. ADA] 30.3 (25.6 to 35.0) [UPA vs PBO]	NI met; Sup.: < 0.001 [UPA vs. ADA] < 0.001[a] [UPA vs. PBO]
ADA 40 mg e.o.w.	327	95	29.1 (24.1 to 34.0)		
Placebo	651	97	14.9 (12.2 to 17.6)		
SELECT-EARLY					
UPA 15 mg q.d.	317	165	52.1 (46.6 to 57.5)	23.7 (16.3 to 31.1)	< 0.001
MTX	314	89	28.3 (23.4 to 33.3)		
	Total N	**Baseline**	**Week 24/26**	**Treatment group difference versus control**	
		Mean	**Mean**	**LS mean difference (95% CI)**	**P value**
Modified total Sharp scores, change from baseline at week 24 or 26 (linear extrapolation, FAS)					
SELECT-COMPARE (at week 26)					
UPA 15 mg q.d.	651	34.73	34.93	0.14 (–0.23 to 0.51) [UPA vs. ADA] –0.67 (–0.97 to –0.37) [UPA vs. PBO]	0.448 [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	35.09	35.15		
Placebo	651	35.47	36.35		
SELECT-EARLY					
UPA 15 mg q.d.	317	17.03	17.16	– 0.53 (–0.85 to –0.20)	0.001[a]
MTX	315	13.89	14.55		
Proportion of patients with no radiographic progression (change from baseline mTSS ≤ 0), at week 24 or 26[a] (linear extrapolation, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	495	83.5 (80.5 to 86.5)	–3.4 (–8.2 to 1.5) [UPA vs. ADA] 7.5 (3.0 to 12.1) [UPA vs. PBO]	0.187 [UPA vs. ADA] 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	257	86.8 (83.0 to 90.7)		
Placebo	651	455	76.0 (72.5 to 79.4)		
SELECT-EARLY					
UPA 15 mg q.d.	317	244	87.5 (83.6 to 91.3)	9.8 (3.5 to 16.2)	0.002
MTX	315	205	77.7 (72.6 to 82.7)		

	Total N	Baseline	Week 24/26	Treatment group difference versus control	
		Mean	Mean	LS mean difference (95% CI)	P value
Disease activity: DAS28-CRP at week 12					
SELECT-COMPARE					
UPA 15 mg q.d.	651	5.78	3.33	–0.47 (–0.638 to –0.295) [UPA vs. ADA] –1.33 (–1.47 to –1.19) [UPA vs. PBO]	< 0.001[a] [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	5.87	3.84		
Placebo	651	5.83	4.69		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	5.61	3.36	–1.08 (–1.319 to –0.848)	< 0.001
MTX	216	5.59	4.43		
SELECT-NEXT					
UPA 15 mg q.d.	221	5.65	3.39	–1.18 (–1.420 to –0.939)	< 0.001
Placebo	221	5.55	4.51		
SELECT-BEYOND					
UPA 15 mg q.d.	164	5.87	3.48	–1.29 (–1.574 to –1.008)	< 0.001
Placebo	169	5.83	4.75		
SELECT-EARLY					
UPA 15 mg q.d.	317	5.90	3.17	–0.88 (–1.09 to –0.67)	< 0.001
MTX	315	5.89	4.05		
	Total N	**Responder**	**Response rate (95% CI)**	**Response rate difference versus control**	
		n		**Point estimate (95% CI)**	**P value**
Proportion of patients achieving LDA based on CDAI ≤ 10, at week 12 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	263	40.4 (36.6 to 44.2)	10.4 (4.2 to 16.7) [UPA vs. ADA] 24.1 (19.4 to 28.8) [UPA vs. PBO]	0.001 [UPA vs. ADA] < 0.001[UPA vs. PBO]
ADA 40 mg e.o.w.	327	98	30.0 (25.0 to 34.9)		
Placebo	651	106	16.3 (13.4 to 19.1)		
SELECT-NEXT					
UPA 15 mg q.d.	221	89	40.3 (33.8 to 46.7)	21.3 (13.0 to 29.5)	< 0.001
Placebo	221	42	19.0 (13.8 to 24.2)		

	Total N	Baseline	Week 24/26	Treatment group difference versus control	
		Mean (SD)	Mean (SD)	LS mean difference (95% CI)	P value
HAQ-DI, change from baseline at week 12 (MI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	1.63	0.99	–0.11 (–0.184 to –0.036) [UPA vs. ADA] –0.31 (–0.372 to –0.253) [UPA vs. PBO]	0.004 [UPA vs. ADA] < 0.001[UPA vs. PBO]
ADA 40 mg e.o.w.	327	1.65	1.11		
Placebo	651	1.61	1.28		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	1.47	0.87	–0.33 (–0.431 to –0.220)	< 0.001
MTX	216	1.47	1.19		
SELECT-NEXT					
UPA 15 mg q.d.	221	1.48	0.85	–0.33 (–0.432 to –0.236)	< 0.001
Placebo	221	1.43	1.15		
SELECT-BEYOND					
UPA 15 mg q.d.	164	1.67	1.24	–0.22 (–0.343 to –0.100)	< 0.001
Placebo	169	1.57	1.38		
SELECT-EARLY					
UPA 15 mg q.d.	317	1.60	0.77	–0.34 (–0.44 to –0.25)	< 0.001
MTX	315	1.60	1.11		
SF-36 (physical component score), change from baseline at week 12 (MMRM, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	32.46	40.92	1.62 (0.62 to 2.62) [UPA vs. ADA] 4.33 (3.52 to 5.15) [UPA vs. PBO]	0.002[a] [UPA vs. ADA] < 0.001[UPA vs. PBO]
ADA 40 mg e.o.w.	327	32.15	39.07		
Placebo	651	32.46	36.57		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	33.22	41.30	3.97 (2.52 to 5.42)	< 0.001
MTX	216	33.24	37.08		
SELECT-NEXT					
UPA 15 mg q.d.	221	33.26	41.34	4.55 (3.13 to 5.98)	< 0.001
Placebo	221	33.18	36.85		
SELECT-BEYOND					
UPA 15 mg q.d.	164	30.71	36.99	3.44 (1.72 to 5.15)	< 0.001
Placebo	169	31.84	34.60		
SELECT-EARLY (MI)					
UPA 15 mg q.d.	317	32.74	43.13	4.25 (3.00 to 5.50)	< 0.001
MTX	315	33.11	39.10		

	Total N	Baseline	Week 24/26	Treatment group difference versus control	
		Mean (SD)	Mean (SD)	LS mean difference (95% CI)	P value
Morning stiffness duration (minutes), change from baseline at week 12 (MMRM, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	141.08	48.16	–9.92 (–23.89 to 4.05) [UPA vs. ADA] –44.04 (–55.39 to –32.69) [UPA vs. PBO]	0.164[a] [UPA vs. ADA] < 0.001[UPA vs. PBO]
ADA 40 mg e.o.w.	327	149.06	59.75		
Placebo	651	144.21	91.84		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	147.65	55.79	–41.53 (–66.56 to –16.50)	0.001
MTX	216	155.70	102.26		
SELECT-NEXT					
UPA 15 mg q.d.	221	147.52	54.27	–51.01 (–78.14 to –23.87)	< 0.001
Placebo	221	141.53	95.67		

Safety at end of double-blind period					
	SAE n (%)	**WDAE n (%)**	**Thrombosis n (%)**	**Herpes zoster infection n (%)**	**Neutropenia n (%)**
SELECT-COMPARE					
UPA 15 mg q.d. (n = 650)	24 (3.7)	23 (3.5)	2 (0.3)	5 (0.8)	18 (2.8)
ADA 40 mg e.o.w. (n = 327)	14 (4.3)	20 (6.1)	3 (0.9)	1 (0.3)	3 (0.9)
Placebo (n = 652)	19 (2.9)	15 (2.3)	1 (0.2)	3 (0.5)	2 (0.3)
SELECT-MONOTHERAPY					
UPA 15 mg q.d. (n = 217)	11 (5.1)	8 (3.7)	1 (0.5)	3 (1.4)	2 (0.9)
MTX (n = 216)	6 (2.8)	6 (2.8)	0	1 (0.5)	1 (0.5)
SELECT-EARLY					
UPA 15 mg q.d. (n = 317)	15 (4.7)	14 (4.4)	0	7 (2.2)	10 (3.2)
MTX (n = 314)	13 (4.1)	16 (5.1)	1 (0.3)	1 (0.3)	2 (0.6)
SELECT-NEXT					
UPA 15 mg q.d. (n = 221)	9 (4.1)	7 (3.2)	0	1 (0.5)	4 (1.8)
Placebo (n = 221)	5 (2.3)	7 (3.2)	0	1 (0.5)	1 (0.5)

Safety at end of double-blind period					
	SAE n (%)	WDAE n (%)	Thrombosis n (%)	Herpes zoster infection n (%)	Neutropenia n (%)
SELECT-BEYOND					
UPA 15 mg q.d. (n = 164)	8 (4.9)	4 (2.4)	3 (1.8)	1 (0.6)	5 (3.0)
Placebo (n = 169)	0	9 (5.3)	1 (0.6)	1 (0.6)	0

ACR = American College of Rheumatology; ADA = adalimumab; CDAI = Clinical Disease Activity Index; CI = confidence interval; e.o.w. = every other week; FAS = full analysis set; LDA = low disease activity; LS = least squares; MI = multiple imputation; MMRM = mixed models for repeated measures; MTX = methotrexate; NI = non-inferiority; NRI = nonresponder imputation; PBO = placebo; q.d. = once daily; SAE = serious adverse event; SD = standard deviation; sup = superiority; UPA = upadacitinib; vs. = versus; WDAE = withdrawal due to adverse event.

[a] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

Critical Appraisal

The main limitations of the pivotal studies include an imbalance in the discontinuation rate between groups in the EARLY study (14.6% withdrawals in methotrexate and 8.5% in upadacitinib) as well as a disproportional discontinuation rate between adalimumab (8.3%) and upadacitinib (4.8%) in the COMPARE study. It is not clear if these imbalances in discontinuation could have led to any sort of bias in the outcomes. In addition, the lack of direct comparison against other existing JAK inhibitors reduces the ability to determine the benefit of upadacitinib versus other existing JAK inhibitors (e.g., baricitinib). Also, several outcomes that were identified in our protocol and reported in the studies fell outside the statistical testing hierarchy and thus need to be interpreted with consideration of type I error.

As well, the EARLY study included treatment-naive patients, which does not align with the Health Canada indication for the treatment of adults with moderately to severely active RA who have had an inadequate response or intolerance to methotrexate. Additionally, although the trials' inclusion and exclusion criteria as well as the baseline demographic characteristics of enrolled patients are generally in line with other RA clinical trials, according to the clinical expert, the enrolled patients in these trials do not represent the majority of patients present in clinical practice. Specifically, the high ratio of female, white, and rheumatoid factor–positive patients may not be representative of Canadian patients. However, there is no clear evidence that these factors will lead to variation in the response to the treatment.

Indirect Comparisons

Description of Studies

Two indirect treatment comparisons (ITCs) are discussed in this review: one submitted by the sponsor and one identified in the literature search conducted by CADTH. Both used a Bayesian network meta-analysis (NMA) approach. The sponsor-submitted ITC is a systematic review of upadacitinib and existing csDMARD, bDMARD, and JAK inhibitors. The published ITC identified by CADTH, by Song et al., only compared upadacitinib to tofacitinib.

Efficacy Results

In the csDMARD-experienced population, the authors of the sponsor's ITC reported that upadacitinib achieved the highest numerical probability of achieving ACR20, 50, and 70 (70% improvement in the ACR criteria) at both the 12- and 24-week end points when contrasted with the probability of other bDMARD and csDMARD interventions. Wide credible intervals (CrIs) were reported across all calculations. However, no comparative result versus bDMARD or JAK inhibitors was provided to assess the magnitude of potential difference in treatment response between upadacitinib and other bDMARD or JAK inhibitors. In the bDMARD-experienced population, the results for upadacitinib were available for 12 weeks only, showing CrIs that are wider than those reported in the csDMARD-experienced patients and suggesting a lack of statistical robustness in the data.

Song et al. reported that upadacitinib had a higher odds ratio (OR) of achieving the efficacy outcome than did tofacitinib. However, the CrI was wide and included the null (OR 1.52 [95% CrI, 0.64 to 3.26]). A recent draft publication by the Institute for Clinical and Economic Review examining upadacitinib reported conclusions similar to the sponsor's submitted ITC, but the report had insufficient details and outcome results to be included in this review.

Harms Results

The sponsor's submitted ITC did not report a safety outcomes analysis. Song et al. provided an analysis of serious adverse events, where the results suggest that the OR of serious adverse events is lower in upadacitinib than in other comparators with a CrI that includes the null.

Critical Appraisal

Several limitations increase the uncertainty in the results provided in the ITC discussed in this review. The sponsor's ITC did not provide an indirect comparison result versus comparators beyond csDMARD. Also, CrIs were wide across the reported outcomes, suggesting considerable statistical heterogeneity in the included studies. In addition, the sponsor's ITC did not provide the results of inconsistency modelling but reported that it was conducted and that no inconsistency was observed. The ITC by Song et al. does not provide sufficient information regarding the included studies' characteristics, the baseline demographics of enrolled patients, or the methods of combining different routes of administration and different definitions of inadequate responders and potential outcomes. An informed judgment of potential clinical or methodological heterogeneity cannot be made in the absence of this information.

Other Relevant Evidence

Description of Studies

Each of the five pivotal studies consisted of two periods, with the first ranging from 12 weeks to 48 weeks in duration and the second ranging from 192 weeks to five years. At the time of this review, data up to 48 weeks were available for SELECT-COMPARE, SELECT-MONOTHERAPY, and SELECT-EARLY. SELECT-NEXT and SELECT-BEYOND included data up until week 60. The five included studies were double-blind RCTs, followed by an open-label extension (SELECT-COMPARE and SELECT-EARLY) or a blinded long-term extension (SELECT-MONOTHERAPY, SELECT-NEXT, and SELECT-BEYOND).

Efficacy Results

The proportion of patients meeting the ACR20, 50, and 70 response criteria at week 48 was 64.7% (95% CI, 61.0 to 68.3), 49.5% (95% CI, 45.6 to 53.3), and 36.1% (95% CI, 32.4 to 39.8), respectively, in SELECT-COMPARE. In the SELECT-EARLY study, at week 48, 91.9% (95% CI, 88.5 to 95.3) of patients met the ACR20 criteria, 79.3% (95% CI, 74.2 to 84.3) met the ACR50 criteria, and 63.3% (95% CI, 57.2 to 69.3) met the ACR70 criteria.

The three remaining studies used the as-observed dataset to describe their results and reported that 87.2% (95% CI, 82.2 to 92.2) met the ACR20 criteria at week 48 (SELECT-MONOTHERAPY) and 76.7% (95% CI, 69.5 to 83.9) and 85.0% (95% CI, 79.6 to 90.3) met it at week 60 (SELECT-NEXT and SELECT-BEYOND). Further, 69.5% (95% CI, 62.7 to 76.4) of patients met the ACR50 criteria at week 48 in SELECT-MONOTHERAPY, and 52.2% (95% CI, 43.8 to 60.7) and 72.6% (95% CI, 65.9 to 79.4) did at week 60 in SELECT-BEYOND and SELECT-NEXT, respectively. The ACR70 response rate at week 48 was 45.5% (95% CI, 38.1 to 52.8) in SELECT-MONOTHERAPY; at week 60, it was between 33.3% (95% CI, 25.4 to 41.3) and 51.5% (95% CI, 44.0 to 59.0) in SELECT-BEYOND and SELECT-NEXT.

Harms Results

The pooled harms data included data from patients treated with upadacitinib 15 mg for up to one year in any of the five included studies. In total, 2,630 patients were included. Of the 2,630 patients, [redacted] reported experiencing an adverse event, [redacted] experienced a serious adverse event, [redacted] stopped treatment owing to adverse events [redacted]

Critical Appraisal

The results of the long-term efficacy and safety outcomes are limited by a lack of comparator in SELECT-NEXT, SELECT-BEYOND, and SELECT-MONOTHERAPY. In addition, only descriptive statistics were provided and any statistical testing that was performed was not included in the statistical hierarchy; thus, there is a risk of type I error. Missing data were only accounted for in SELECT-COMPARE, which used a nonresponder imputation method to label missing data as nonresponders. Consequently, the efficacy of upadacitinib in patients in SELECT-COMPARE was lower than in the other four studies, which used nonimputed, observed data. It is uncertain whether this is due to the bias introduced by classifying missing data as nonresponders or by an overestimation of the results caused by not imputing for missing data.

Conclusions

The included studies showed that upadacitinib at 15 mg, orally once daily, after 12 to 14 weeks of treatment, improved clinical response in terms of the ACR20 outcome compared to placebo (two studies) and methotrexate (two studies) and in terms of the ACR50 outcome compared to adalimumab (one study) in a population of patients with RA who had either an inadequate response to csDMARDs (three studies), an inadequate response to bDMARDs (one study), or an undetermined response to either (one study). Three out of five studies had patients on either methotrexate or other csDMARD background therapy, alone or in combination. One study, BEYOND, specifically recruited patients who had failed bDMARD therapy. In all five studies, there was a statistically significant improvement in HRQoL and in disease activity with the upadacitinib versus the methotrexate or placebo groups, with a difference larger than the minimal important difference in the HAQ-DI outcome. Upadacitinib showed better treatment outcomes than adalimumab in disease activity measures and HRQoL measures but not in radiographic progression. The benefit of upadacitinib versus other JAK inhibitors remains uncertain owing to the lack of direct or indirect comparative estimates versus baricitinib and the indirect evidence versus tofacitinib, which carries a large degree of uncertainty. According to the indirect evidence, upadacitinib is likely at least as efficacious as other bDMARDs, but owing to the lack of reported comparative estimates versus bDMARDs, no magnitude of treatment difference between upadacitinib and other bDMARDs could be reported. The risk of notable harms such as serious infections, malignancies, cardiovascular events, dyslipidemia, and elevated hepatic enzymes did not appear to differ between upadacitinib and placebo, although the included studies were not designed to assess outcomes such as these. Long-term extension studies are ongoing.

Introduction

Disease Background

RA is an autoimmune inflammatory disease that primarily affects the joints of the body. Characterized by acute and chronic inflammation of the synovium, or soft tissue surrounding the joints, patients are subject to severe pain, stiffness, and fatigue, all of which can affect a patient's ability to perform activities of daily living and overall HRQoL. Prolonged inflammation may lead to damage and destruction of the joints through erosion of the cartilage and bone and, consequently, to disability and premature mortality. Other areas of the body may be affected as well, including the eyes, lungs, heart, and skin. The cause of RA is unknown, and onset can occur at any time, but the risk of developing RA increases with age. It is almost three times more common among women than among men, and certain genetic factors put patients at higher risk of RA as well.[6,7] Data about the prevalence and incidence of RA in Canada are limited; however, it is estimated that about 1% of Canadians have the disease.[7,8] The diagnosis of RA is made clinically. Prolonged joint swelling and inflammatory join pain may raise suspicion of RA, which should be further investigated through a clinical exam or imaging.[6] Serology can be used to test for rheumatoid factor, but this is limited by poor specificity.[9] Similarly, anti–cyclic citrullinated peptide (CCP) antibodies may support a diagnosis of RA as well.[9,10]

Upadacitinib is a JAK inhibitor. JAK mediates the effects of cytokines and their production; thus, JAK inhibitors may have more of a global effect on various cytokine production than do biologics, which tend to target specific cytokines. Upadacitinib is the third JAK inhibitor approved in Canada, the first being tofacitinib, followed by baricitinib, both of which were previously reviewed and issued a recommendation by the CADTH Canadian Drug Expert Committee. Although all are JAK inhibitors, upadacitinib inhibits JAKs with a high degree of selectivity against other kinases in the human genome. Specifically, upadacitinib-inhibited, cytokine-induced STAT phosphorylation mediated JAK1 and JAK1/JAK3 more potently than JAK2/JAK2 mediated STAT phosphorylation. The clinical significance of these differences in pharmacodynamics is yet to be determined. Although JAK inhibitors target cytokines and thus have much in common with biologics, and are often lumped in with the biologics, they are in fact small molecules. Unlike the bDMARDs, JAK inhibitors are administered orally.

Standards of Therapy

Treatment of RA consists both of acute therapies used to address intense flares of the disease and more chronic therapies that are aimed at the underlying disease process itself, known as DMARDs. These DMARDs consist of small molecules that address various pathways involved in inflammatory or immune processes and include a diverse array of drugs, such as the antimalarials, sulfasalazine, leflunomide, and — the most commonly used — methotrexate. As a group, these drugs are referred to as the conventional DMARDs (cDMARDs). More recently, these cDMARDs have been joined by the bDMARDs, a group of drugs with a shared design, being either monoclonal antibodies or fusion proteins. Common limitations of all approaches are increased risk of infection and, possibly, an increased risk, albeit rare, of certain cancers.

According to the 2015 ACR guideline for the treatment of RA, it is recommended that for patients with symptomatic early RA who are DMARD naive, DMARD monotherapy should be initiated. If disease activity remains moderate or high, then a combination of DMARDs or

a tumour necrosis factor inhibitor or a non–tumour necrosis factor biologic (all choices with or without methotrexate), in no particular order of preference, could be used.[11] The recommendations for established RA state that a treat-to-target strategy should be used, regardless of the disease activity level. For LDA, DMARD monotherapy should be used, preferably using methotrexate. For moderate to high disease activity, combination DMARDs or the addition of a tumour necrosis factor inhibitor, a non–tumour necrosis factor biologic, or tofacitinib is recommended.[11] Nonpharmacological therapies are also used alongside pharmacological options, such as physiotherapy, occupational therapy, and surgery.[10,12]

Drug

Upadacitinib is an oral selective and reversible JAK inhibitor developed for the treatment of moderate to severe active RA. It is available as a 15 mg extended-release tablet and is recommended as a single 15 mg dose once daily.

Upadacitinib has a Health Canada indication for the treatment of adults with moderately to severely active RA who have had an inadequate response or intolerance to methotrexate. Upadacitinib may be used as monotherapy or in combination with methotrexate or other nonbiologic DMARDs.

The sponsor is requesting reimbursement of upadacitinib per the indication, in a similar manner to bDMARDs and targeted synthetic DMARDs for the treatment of moderate to severe RA.

Table 2: Key Characteristics of JAK Inhibitors, IL-6 Inhibitors, T-Cell Co-Stimulation Modulators, CD20 Inhibitors, IL-1 Inhibitors, and TNF Inhibitors

	Mechanism	Indication[a]	Monotherapy	Combination	Route
		Inadequate response required			
Upadacitinib	JAK inhibitor	MTX	Yes[b]	Monotherapy or MTX	Oral
Baricitinib	JAK inhibitor	≥ 1 DMARD	Yes[c]	+MTX	Oral
Tofacitinib	JAK inhibitor	MTX	Yes[c]	+MTX	Oral
Tocilizumab	IL-6 inhibitor	Not required	Yes[c]	MTX or DMARD	SC or IV
Abatacept	T-cell co-stimulation modulator	≥ 1 DMARD or TNF inhibitor	Yes	+DMARD[d]	SC or IV
Rituximab	CD20 inhibitor	≥ 1 TNF inhibitor	No	+MTX	IV
Anakinra	IL-1 inhibitor	Not required	Yes	+DMARD[e]	SC
Adalimumab	TNF inhibitor	Not required	Yes	+MTX[f]	SC
Etanercept			Yes	+MTX	SC
Golimumab			No	+MTX	SC or IV
Certolizumab pegol			Yes[c]	+MTX	SC
Infliximab			No	+MTX	IV

DMARD = disease-modifying antirheumatic drug; IL = interleukin; IV = intravenous; JAK = Janus kinase; MTX = methotrexate; SC = subcutaneous; TNF = tumour necrosis factor.

[a] Health Canada–approved indication (all approved for adults with moderately to severely active RA, except anakinra, which is approved for active RA, severity not specified).

[b] If patient has had an inadequate response or is intolerant to MTX.

[c] If patient is intolerant to MTX.

[d] If first-line treatment, give with MTX.

[e] The DMARD used is usually MTX.

[f] Other DMARDs may also be used.

Stakeholder Engagement

Patient Group Input

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

1. Brief Description of Patient Groups Supplying Input

Two responses to CADTH's call for patient input for the upadacitinib submission were received: a joint submission from the Arthritis Society and CAPA and a second submission from ACE.

The Arthritis Society is the largest nongovernment funder of arthritis research in Canada, investing more than $200 million in projects that have led to breakthroughs in the diagnosis, treatment, and care of individuals with arthritis. The Arthritis Society has been established for 70 years, supporting more than six million Canadians with arthritis. It is dedicated to a vision of living in a world where individuals are free from the devastating effects of arthritis. CAPA is a virtual, patient-driven, independent, national education and advocacy organization that facilitates Canadians living with arthritis to become effective advocates, as well as to improve their quality of life.

ACE is Canada's largest, longest-running national arthritis patient organization; it provides free, science-based information and education programs to individuals with arthritis. It strives to help individuals with arthritis take control of their disease and improve their quality of life through education and empowerment. It was founded and is led by individuals with arthritis and is involved in advocating for arthritis health policy issues through its JointHealth family of programs and the Arthritis Broadcast Network.

The patient input submissions were prepared independently without influence from any outside party. An AbbVie contact provided CAPA with the names and contact information of the Canadian health care professionals who had patients who participated in the upadacitinib clinical trial; however, AbbVie was not involved in the preparation of the submission. For a complete summary of the conflict of interest declarations, please refer to the patient input summary for upadacitinib on the CADTH website.

2. Condition-Related Information

In response to CADTH's call for patient input, CAPA and the Arthritis Society collaboratively distributed a survey, with questions that were informed by the CAPA board members, who all have experience living with various forms of arthritis. The survey was shared via email and social media and was open from June 7, 2019, to July 8, 2019. Fifty-one online survey responses were received from individuals with no experience taking upadacitinib, and one survey response was received from a rheumatology nurse on behalf of an upadacitinib clinical trial participant. Of the demographic data collected (n = 36), the majority (75%) of the respondents were between 45 years old and 75 years old (range: 8 years old to 77 years old) and indicated that their RA was moderate in severity. ACE conducted an online survey using Survey Monkey from June 2019 to July 2019, which was shared through social media and ACE's list of subscribers. Respondents who lived outside Canada were removed from the survey, leaving a total of six responses from patients living in Canada to inform the patient input submission. ACE also interviewed one upadacitinib clinical trial participant.

The patient groups described RA as a chronic autoimmune disease in which the body's immune system attacks and causes inflammation in joints. When uncontrolled, inflammation can result in irreversible damage to the affected joints. RA can also affect vital organs, such as the eyes, lungs, and heart, as well as lead to depression and other mental health issues. Some patients also report living with fibromyalgia, myasthenia gravis, and asthma. Patients with RA report that they may experience periods of active disease (flares or flare-ups) and periods of decreased activity of the disease (remission). Flares can be unpredictable in their onset, frequency, and length. The patient groups also reported that symptoms and manifestation of the disease can vary from patient to patient. Some patients report having to deal with flares reactively, feeling that they do not have control over their RA; for example, one patient reported "the episodic nature makes planning ahead difficult and uncertain, and has an impact on my social life."

The patient input received for this submission highlighted that RA impacts patients' day-to-day lives in many ways, such as completing daily tasks, participating in leisure activities, and caring for — and spending time with — loved ones. According to the written comments received, several patients described being limited in daily tasks due to joint stiffness, pain, brain fog, and overall fatigue. For example, one patient reported, "A good bra day means I can get a bra on within 15 minutes." A second patient described "pain, stiffness, swelling, loss of mobility and fatigue" and stated that "RA affects my every day by limiting what I am able to do." Furthermore, many patients reported negative consequences of their disease on both their employment and financial status, whether it was due to no longer being able to work, having to take a demotion, having to go on Canada Pension Plan disability, or having to abandon postgraduate education. RA also impacts the patient's circle, including spouses, partners, and children. Often, the patient's circle must take on additional household chores, such as cleaning, cooking, shopping, and accompanying the patients to and from medical appointments. While some patients reported no specific challenges for their caregivers, others reported that it was difficult for caregivers to provide both mental and physical support to the patients given that RA is "silent" and "unpredictable." One patient reported their caregivers feeling depressed. Patients also reported that RA affects intimacy.

3. Current Therapy-Related Information

Current treatments for RA aim to control inflammation and decrease disease activity and joint damage. These treatments include nonsteroidal anti-inflammatory drugs (NSAIDs); corticosteroids; DMARDs such as methotrexate, sulfasalazine, and hydroxychloroquine; and biologics such as Humira (adalimumab), Remicade (infliximab), Simponi (golimumab), and Rituxan (rituximab). More recent treatment options include JAK inhibitors tofacitinib and baricitinib, as well as upadacitinib.

Currently, there is no way to predict who will best respond to which therapies. The patient groups indicated that for many patients with RA, treatment is determined by trial and error. Several patients reported that they are taking or have tried up to four different RA treatments. While a few patients indicated that a present treatment is controlling their disease, the remainder described a continuing struggle to find a both effective and tolerable treatment. For example, one patient reported that they were being treated with methotrexate/sulfasalazine and hydroquinine and, more recently, Remicade: "With the addition of Remicade, my symptoms appear to be under control." Another patient reported that:

> I've tried 2 DMARDS, prednisone, and have just started on my second biologic. My first DMARD (methotrexate) worked quite well, but due to elevated liver enzymes, I

> was unable to continue it. Prednisone helped, but only at higher doses (above 15 mg), and was not an acceptable long-term solution (my main side-effect was a noticeable increase in appetite). Plaquenil didn't seem to improve my RA symptoms at all, and I had headaches almost daily.

Another patient reported that they were being treated with multiple drugs at once, including a DMARD, an anti-inflammatory, an opioid, and a steroid, as well as their fourth biologic, which seemed to be losing efficacy. Overall, patients reported switching treatments owing to treatment toxicity, lack of efficacy, or a reduction in efficacy over time. Patients also reported that treatments were costly.

Both patient groups indicated that some patients pursue medical cannabis and/or nonpharmacological approaches to manage RA symptoms; the nonpharmacological approaches include physiotherapy, massage therapy, acupuncture, and counselling, which can help control some symptoms of pain or fatigue. However, these treatments are often not reimbursed through provincial health care systems.

According to the survey responses, treatments are difficult to tolerate because of the side effects. Furthermore, some of the side effects require additional treatment with other medications, such as folic acid or antiemetics. The patient groups describe minimizing these side effects as an important outcome that should be considered when evaluating new therapies. Additional treatment outcomes that are important to patients with RA include reduction in pain and fatigue, reduction in RA complications, increased mobility, ability to work and be productive, ability to carry out daily activities and social roles, and ability to effectively carry out caregiver and parenting tasks. Patients reported that unmet needs include homecare assistance, counselling, and quicker access to specialists and occupational therapists.

4. Expectations About the Drug Being Reviewed

CAPA and the Arthritis Society received a response to their survey from one upadacitinib clinical trial participant. The patient reported that:

> I started the trial drug in pill form in May of 2017, and within days I noticed some relief from pain and swelling. My ankle and wrist joints were my biggest problem, and after a few weeks I was able to do some yard work like using the lawn mower to cut grass. I don't recall any negative effects at all.

The patient also reported that the trial drug "reduced the feeling of uselessness and reliance on others" and that the "pill form was easier to take than the Humira injections." ACE interviewed an upadacitinib clinical trial participant who also described the treatment as effective, with no adverse effects. The participant added that upadacitinib "manages all of my symptoms by limiting inflammation, no pain, no morning stiffness, more energy, better mood, want to complete tasks and do more, able to work."

Generally, patients expect upadacitinib to be easy to take; treat symptoms of pain and stiffness; and increase their mood, sense of independence, and overall quality of life.

5. Additional Information

Additional information was provided in the ACE submission, which includes a recommendation for a well-rounded treatment plan for RA that includes education, appropriate therapeutic and recreational exercise, physical therapy, and an overall healthy lifestyle. The patient group also stressed the timely initiation of the most suitable medications, chosen by the patient in consultation with their rheumatologist.

Lastly, ACE indicated that the following questions were frequently asked by patients:

- Is there one or a few advanced therapies that you feel may work the best for me? If yes, why?
- What are the different ways to take the ones you think might work best for me at this point in my disease course?
- What are the most common and the most serious side effects for the advanced therapy you are recommending?
- Do I have to stop what I'm taking now to clear it from my body before starting on the advanced therapy you recommend?
- Can I stop any of the other medications I'm currently taking when I start on the advanced therapy you recommend? If so, when?
- How long do I have to be off one or all of my current medications before I can start on the advanced therapy you recommend?
- Can I get pregnant while taking the advanced therapy you recommend?
- How quickly do I need to decide?

Clinician Input

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

Description of the Current Treatment Paradigm for the Disease

The treatment paradigm for active RA starts with cDMARDs such as methotrexate, sulfasalazine, Plaquenil (hydroxychloroquine), or leflunomide. The goal of treatment is, at a minimum, to reach a state of LDA and, at a maximum, to achieve remission. Failure to reach an acceptable treatment target in three to four months would lead to the addition of a biologic drug. The number of DMARDs used during the initial stages of treatment can be influenced by the need for payers to establish lack of response before reimbursing the more expensive biologics. For example, Ontario requires failure of three months of combination DMARDs (methotrexate/Plaquenil (hydroxychloroquine) /sulfasalazine or methotrexate/leflunomide), Quebec is satisfied with failure after three months of methotrexate/Plaquenil (hydroxychloroquine) combination therapy, while British Columbia requires failure of four DMARDs, to include parenteral-administered methotrexate. All these strategies are supported by results of RCTs and/or expert opinion and are advocated by guidelines developed by, among others, the Canadian Rheumatology Association, the ACR, and the European League Against Rheumatism (EULAR). The aforementioned treatments inhibit underlying disease mechanisms; disease progression is modified, radiographic progression is inhibited or arrested, and functional status is maintained.

The role of nonmedicinal treatment (e.g., physical therapy, splints, ice, rest) is adjunctive and cannot substitute for the use of medications. Pain-relieving modalities (NSAIDs, analgesics, cannabinoids) do not modify disease and cannot substitute for DMARDs and

biologics. High-dose prednisone is not a safe option, and low-dose prednisone is inadequate as a DMARD.

The JAK inhibitors represent the latest therapeutic advance. Unlike the biologics, they are oral small molecules. A biologic drug is specific in affecting an immunologic mechanism. The JAK inhibitors are not specific and affect several immunologic mechanisms. JAK inhibitors have been shown to be DMARDs. In addition to controlling signs and symptoms of disease and improving functional status, JAK inhibitors inhibit radiographic progression.

Treatment Goals

An ideal treatment would result in remission, a drug-free immunologic remission, or cure. More pragmatically, effective RA treatment is intended to delay or stop disease progression, although requiring long-term drug use. Effective treatment means prevention of crippling and disability, no loss of days from work, and maintenance of physical and mental function. Numerous observational studies have shown that successful treatment, by inhibiting radiographic progression, has been associated with a reduction in hip and knee replacements and, through reduction of inflammation, has been associated with diminished cardiovascular morbidity.

Unmet Needs

An ideal therapy would induce total remission in 100% of patients and have no safety issues. No such treatment exists. Furthermore, the addition of all available treatments is not 100% effective. However, each new treatment has been associated with benefit in a percentage of patients heretofore unsuccessfully treated. This is the rationale for ongoing therapy development.

Even “good” therapies have limitations. Loss of effect occurs. Issues with convenience, safety, tolerability, and adherence to therapy occur. The identification of patient characteristics predicting response and safety would improve therapeutic algorithms dramatically, but the lack of such biomarkers represents a large unmet need.

Place in Therapy

Upadacitinib is the third JAK inhibitor and the most selective JAK inhibitor to date. Tofacitinib is a pan-JAK inhibitor. Baricitinib is more selective against JAK1 and 2, and upadacitinib is a selective JAK1 inhibitor. Other JAK inhibitors (e.g., filgotinib) are under development. The aim of selectivity is enhanced safety. For the most part, the JAK inhibitors share the same safety profile as the biologics, but the use of JAK inhibitors is associated with a greater risk of reactivation of herpes zoster infections.

The efficacy of the JAK inhibitors appears to be as good as that of the current biologics, and for the most part, the same can be said for safety. The use of JAK inhibitors is expected to occur post-DMARD inadequate response, and owing to the convenience of once daily oral dosing, their use is expected to replace that of biologics as the first line of treatment post-DMARD inadequate response. JAK inhibitors can be used as monotherapy or in combination with methotrexate and other cDMARDs, but not in combination with biologics.

Given the data provided for the efficacy and safety of upadacitinib, it is reasonable to consider it as the first treatment for a patient with an inadequate response to cDMARDs.

Patient Population

Clinically, any patient who is diagnosed with RA can be considered for treatment; classification criteria include number of inflamed joints and the level of disease activity being at least moderate and most likely high. There are no specific disease characteristics that argue against the initiation of some treatment in a patient with active RA. Characteristics associated with a poorer prognosis, such as high levels of anti-CCP antibodies, elevated erythrocyte sedimentation rate (ESR) or CRP, and baseline erosions, prompt urgency for aggressive treatment, but therapeutic algorithms demand failure of cDMARDs before eligibility for a biologic or a JAK inhibitor. As we cannot predict response, patients are subjected to three to six months of potentially ineffective DMARD therapy before access to potentially ineffective biologic or JAK inhibitor therapy.

Current RCT design in RA research has limited the generalizability of results to real-life patients with RA. Patients in RCTs are selected usually for seropositivity and/or having radiographic erosions at baseline, having elevated levels of acute phase reactants, or having fewer comorbidities. According to the clinical expert's estimate, less than half of patients in cohorts and registers are eligible for RCTs, but they still require therapeutic decisions. Because provincial access to biologics and JAK inhibitors is usually derived from the inclusion criteria of the RCTs, many patients are denied access to biologics, for example patients with seronegative non-erosive disease who meet the classification criteria for RA. The decision not to use a particular treatment for a patient is dependent on circumstances that can change. Certainly, patients with active tuberculosis, active herpes, and/or other active infection require treatment for these conditions first. Patients with active heart failure, hepatitis, recent bowel perforation, leukopenia, and other comorbidities have more pressing needs, requiring delay in initiation of RA therapy.

Assessing Response to Treatment

Rheumatologists use a treat-to-target strategy. The goal of therapy is to achieve remission and, if that is not possible, to achieve LDA. There are several definitions of remission and, as expected, the stricter the definition, the lower the percentage of patients achieving remission. Patients should be assessed every three to six months.

As opposed to using the disease status defined by DAS, the Clinical Disease Activity Index (CDAI), or the Simplified Disease Activity Index (SDAI), provincial formularies define a level of disease activity for the initiation and switching of biologics. In Ontario, the number of swollen joints is a good compromise. Someone with more than five swollen joints after appropriate cDMARD therapy is eligible for a biologic (provided the person is seropositive or erosive), and failure to reduce the swollen joint count below five makes the person eligible for switching.

An ACR20 is a meaningful response in RCTs and clinical practice. Patients achieving an ACR20 are, on average, about 35% better. Despite the use of the ACR response for more that 25 years, a minimal clinically important difference (MCID) has not been defined. The ACR measures a change in disease activity, and a patient with an ACR20 response can have a considerable amount of ongoing disease activity. Therefore, achieving a state of LDA is considered a more relevant end point.

There is debate about the advantages and disadvantages of using patient-reported outcomes in the making of therapeutic decisions in clinical practice. Patients often rate themselves as more active than their physician does. Issues of fibromyalgia, fatigue, and

depression confound the decrease in inflammation measured by the swollen joint count or lack of X-ray progression. In clinical practice, many rheumatologists use the swollen joint count as a measure of success or failure and a basis for therapeutic decisions.

Discontinuing Treatment

Discontinuation of treatment can be decided upon lack of response and/or safety concerns. This can be decided at each assessment visit.

Prescribing Conditions

Although a specialist is required to establish diagnosis and initiate a treatment plan, family physicians may be expected to monitor patients who are clinically stable.

Clinical Evidence

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

Systematic Review (Pivotal and Protocol-Selected Studies)

Objectives

To perform a systematic review of the beneficial and harmful effects of upadacitinib 15 mg once daily orally, used as monotherapy or in combination with methotrexate or other csDMARDs, for the treatment of moderate to severe active RA in adult patients who have responded inadequately to, or who are intolerant to, one or more DMARDs.

This systematic review protocol was established before the granting of a Notice of Compliance from Health Canada for upadacitinib.

Methods

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

Table 3: Inclusion Criteria for the Systematic Review

Patient population	Adult patients with moderate to severe rheumatoid arthritis who have responded inadequately to, or who are intolerant to, one or more DMARDs. Subgroups: • patients who have responded inadequately to, or who are intolerant to, methotrexate • patients by disease severity • patients who have responded inadequately to, or who are intolerant to, conventional DMARDs • patients who have responded inadequately to, or who are intolerant to, biologic DMARDs
Intervention	Upadacitinib 15 mg orally once daily, used as monotherapy or in combination with methotrexate or other conventional synthetic DMARDs
Comparators	• TNF-alpha inhibitors (infliximab, adalimumab, certolizumab pegol, golimumab, etanercept) • T-cell stimulation inhibitor (abatacept) • CD20 inhibitor (rituximab) • IL-6 inhibitors (tocilizumab, sarilumab) • JAK inhibitor (tofacitinib, baricitinib) • Conventional synthetic (nonbiologic) DMARDs
Outcomes	**Efficacy outcomes:** • clinical response (ACR20, ACR50, ACR70)[a] • radiographic response • health-related quality of life[a] • functional and disability outcomes[a] • disease activity[a] • health care resource use **Harms outcomes:** • AEs,[a] SAEs,[a] WDAEs • mortality • AEs of special interest (e.g., serious infection [including herpes zoster],[a] neutropenia, lymphopenia, thrombocytopenia, malignancies, thrombosis [including increased platelets], major cardiovascular events, gastrointestinal perforations and other gastrointestinal SAEs,[a] liver toxicity, dyslipidemia)
Study design	Published and unpublished phase III and IV RCTs

ACR = American College of Rheumatology; AE = adverse event; DMARD = disease-modifying antirheumatic drug; IL = interleukin; JAK = Janus kinase; RCT = randomized controlled trial; SAE = serious adverse event; TNF = tumour necrosis factor; WDAE = withdrawal due to adverse event.

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

The literature search for clinical studies was performed by an information specialist using a peer-reviewed search strategy according to the PRESS (Peer Review of Electronic Search Strategies) checklist (https://www.cadth.ca/resources/finding-evidence/press).[13]

Published literature was identified by searching the following bibliographic databases: MEDLINE All (1946–) 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 upadacitinib. Clinical trial registries were searched: the US National Institutes of Health's clinicaltrials.gov and the WHO's International Clinical Trials Registry Platform search portal.

No filters were applied to limit retrieval by study type. Retrieval was not limited by publication date or by language. Conference abstracts were excluded from the search results. See Appendix 1 for the detailed search strategies.

The initial search was completed on August 1, 2019. Regular alerts updated the search until the meeting of the CADTH Canadian Drug Expert Committee on November 20, 2019.

Grey literature (literature that is not commercially published) was identified by searching relevant websites from the following sections of the *Grey Matters: A Practical Tool for Searching Health-Related Grey Literature* checklist (https://www.cadth.ca/grey-matters):[14] health technology assessment agencies, health economics, clinical practice guidelines, drug and device regulatory approvals, advisories and warnings, drug class reviews, clinical trials registries, and databases (free). Google was used to search for additional internet-based materials. These searches were supplemented by reviewing bibliographies of key papers and through contacts with appropriate experts. In addition, the sponsor of the drug was contacted for information regarding unpublished studies. See Appendix 1 for more information on the grey literature search strategy.

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

Findings From the Literature

Fifty-one studies were identified from the literature for inclusion in the systematic review (Figure 1). The included studies are summarized in Table 4. A list of excluded studies is presented in Appendix 2.

Figure 1: Flow Diagram for Inclusion and Exclusion of Studies

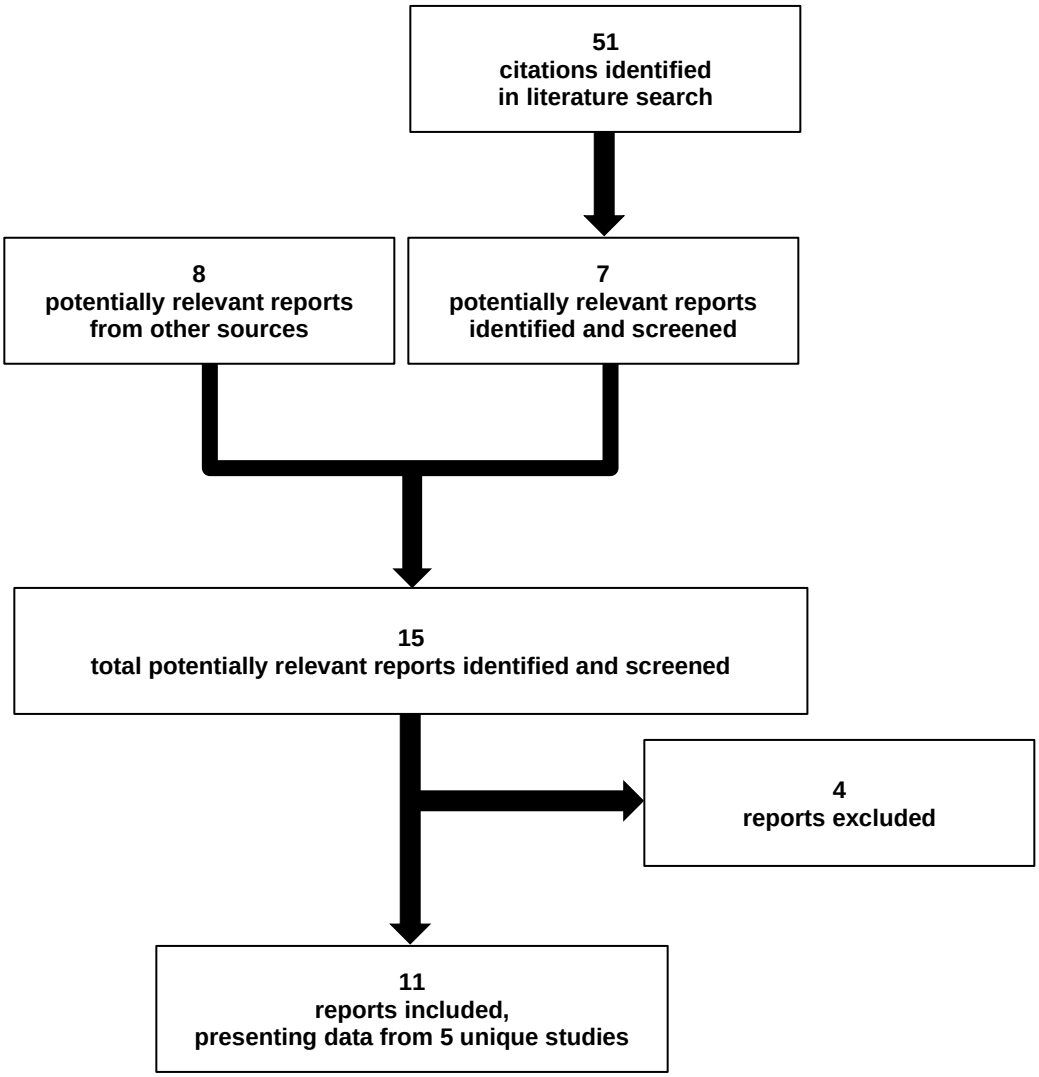

Table 4: Details of Included Studies

		COMPARE (M14-465)	MONOTHERAPY (M15-555)	NEXT (M13-549)	BEYOND (M13-542)	EARLY (M13-545)
DESIGNS AND POPULATIONS	**Study design**	DB RCT followed by OLE	DB RCT followed by blinded LTE	DB RCT followed by blinded LTE	DB RCT followed by blinded LTE	DB RCT followed by OLE
	Locations	286 sites in 41 countries Canada, USA, Mexico, Australia, New Zealand, Republic of Korea, Malaysia, South Africa, Europe, South America	138 sites in 24 countries USA, Mexico, Japan, South Africa, Europe, South America	150 sites in 35 countries Canada, USA, Mexico, Australia, New Zealand, Republic of Korea, South Africa, Europe	152 sites in 26 countries Canada, USA, Australia, Europe	229 sites in 43 countries Canada, USA, Mexico, China, Japan, Europe, South America, Australia, New Zealand
	Randomized (N)	1,629	648	661	499	947
	Inclusion criteria	• ≥ 18 years old • Diagnosis of RA for ≥ 3 months and fulfills the 2010 ACR/EULAR classification criteria for RA • Must have been on oral or parenteral MTX therapy for ≥ 3 months and on a stable prescription of 15 to 25 mg/week (or if intolerant, MTX at ≥ 10 mg/week) for ≥ 4 weeks before first dose of study drug • Disease activity: ≥ 6 swollen joints (based on 66 joint counts) and ≥ 6 tender joints (based on 68 joint counts) at screening and baseline visits; hsCRP ≥ 5 mg/L at screening visit • At screening, has ≥ 3 bone erosions on X-ray, or	• ≥ 18 years old • Diagnosis of RA for ≥ 3 months and fulfills the 2010 ACR/EULAR classification criteria for RA • Must have been on oral or parenteral MTX therapy for ≥ 3 months and on a stable prescription of 15 to 25 mg/week (or if intolerant, MTX at ≥ 10 mg/week) for ≥ 4 weeks before first dose of study drug • Must have discontinued csDMARDs (except MTX) ≥ 4 weeks before first dose of study drug, with a washout period of at least 5 times the	• ≥ 18 years old • Diagnosis of RA for ≥ 3 months and fulfills either the 1987 revised ACR classification or the 2010 ACR/EULAR classification criteria for RA • Must have been receiving csDMARD therapy for ≥ 3 months and on a stable dose for ≥ 4 weeks before first dose of study drug; failed at least one of MTX, sulfasalazine, or leflunomide; had inadequate response to hydroxychloroquine and/or chloroquine and failed MTX, sulfasalazine, or leflunomide • Disease activity: ≥ 6 swollen joints	• ≥ 18 years old • Diagnosis of RA for ≥ 3 months and fulfills the 2010 ACR/EULAR classification criteria for RA • Received previous treatment with bDMARD for RA and failed ≥ 1 bDMARD therapy before first does of study drug • Must have discontinued bDMARDs before first dose of study drug, with a washout period of at least 5 times the mean terminal elimination half-life of the drug • Receiving csDMARD therapy for ≥ 3 months and on a stable dose for ≥ 4 weeks before first dose of study drug	• ≥ 18 years old • RA symptoms for ≥ 6 weeks and fulfills the 2010 ACR/EULAR classification criteria for RA • MTX naive, or had received ≤ 3 weekly MTX doses with a required 4-week washout before first dose of study drug • Disease activity: ≥ 6 swollen joints (based on 66 joint counts) and ≥ 6 tender joints (based on 68 joint counts) at screening and baseline visits; hsCRP ≥ 5 mg/L at screening visit

		COMPARE (M14-465)	MONOTHERAPY (M15-555)	NEXT (M13-549)	BEYOND (M13-542)	EARLY (M13-545)
		≥ 1 bone erosion and positive RF, or ≥ 1 bone erosion and positive anti-CCP autoantibody • Must have discontinued bDMARD therapy and all csDMARDs (except MTX) before first dose of study drug, with a washout period of at least 5 times the mean terminal elimination half-life of the drug	mean terminal elimination half-life of the drug • Disease activity: ≥ 6 swollen joints (based on 66 joint counts) and ≥ 6 tender joints (based on 68 joint counts) at screening and baseline visits; hsCRP ≥ 3 mg/L at screening visit	(based on 66 joint counts) and ≥ 6 tender joints (based on 68 joint counts) at screening and baseline visits; hsCRP ≥ 3 mg/L at screening visit	• Disease activity: ≥ 6 swollen joints (based on 66 joint counts) and ≥ 6 tender joints (based on 68 joint counts) at screening and baseline visits; hsCRP ≥ 3 mg/L at screening visit	
	Exclusion criteria	• Prior exposure to any JAK inhibitor • Exposure to adalimumab, or treatment with other bDMARD therapy for ≥ 3 months or considered inadequate responder to bDMARD therapy • History of arthritis with onset before age 17 years or current diagnosis of inflammatory joint disease other than RA	• Prior exposure to JAK inhibitor or any bDMARDs • History of arthritis with onset before age 17 years or current diagnosis of inflammatory joint disease other than RA • Has been treated with intra-articular, intramuscular, IV, trigger point or tender point, intra-bursa, or intra-tendon sheath corticosteroids in the 8 weeks preceding first dose of study drug	• Prior exposure to JAK inhibitor • Inadequate responder to bDMARD therapy • History of arthritis with onset before age 17 years or current diagnosis of inflammatory joint disease other than RA • Has been treated with intra-articular, intramuscular, IV, trigger point or tender point, intra-bursa, or intra-tendon sheath corticosteroids in the 8 weeks preceding first dose of study drug	• Prior exposure to JAK inhibitor • History of arthritis with onset before age 17 years or current diagnosis of inflammatory joint disease other than RA • Has been treated with intra-articular, intramuscular, IV, trigger point or tender point, intra-bursa, or intra-tendon sheath corticosteroids in the 8 weeks preceding first dose of study drug	• Intolerant to MTX • Prior exposure to JAK inhibitor or any bDMARD • History of arthritis with onset before age 17 years or current diagnosis of inflammatory joint disease other than RA
Drugs	**Intervention**	15 mg upadacitinib orally, once daily	15 mg and 30 mg upadacitinib orally, once daily	15 mg and 30 mg upadacitinib orally, once daily	15 mg and 30 mg upadacitinib orally, once daily	7.5 mg[a], 15 mg, and 30 mg upadacitinib orally, once daily

		COMPARE (M14-465)	MONOTHERAPY (M15-555)	NEXT (M13-549)	BEYOND (M13-542)	EARLY (M13-545)
	Comparator(s)	40 mg adalimumab SC, biweekly; matching placebo, either SC or oral	MTX orally, once weekly, dose based on patient's prior stable dose	Placebo, oral	Placebo, oral	10 mg MTX orally, once weekly, titrated up to 20 mg[b]
DURATION	**Phase**					
	Screening period	35 days	35 days	35 days	35 days	35 days
	Double-blind	48 weeks	14 weeks	12 weeks	24 weeks	48 weeks
	Long-term extension	Up to 5 years	226 weeks	Up to 5 years	216 weeks	Up to 192 weeks
	Follow-up	30 days (call or visit) and 70 days (call)	30 days	30 days (call or visit)	30 days (call or visit)	30 days
OUTCOMES	**Primary end point**	Proportion of patients achieving ACR20 response at week 12	Proportion of patients achieving ACR20 response at week 14	Proportion of patients achieving ACR20 response at week 12	Proportion of patients achieving ACR20 response at week 12	Proportion of patients achieving ACR50 response at week 12
	Secondary end points	Change from baseline at week 12 in: • DAS28-CRP • HAQ-DI • SF-36 PCS • mTSS at week 26 • morning stiffness (duration) • FACIT-F • patient's global assessment of pain Proportion of patients achieving: • LDA (DAS28-CRP ≤ 3.2) at week 12 • LDA (CDAI ≤ 10) at week 12 • CR based on DAS28-CRP at week 12	Change from baseline at week 14 in: • DAS28-CRP • HAQ-DI • SF-36 PCS • morning stiffness (duration) Proportion of patients at week 14 achieving: • LDA (DAS28-CRP ≤ 3.2) • CR based on DAS28-CRP • ACR50 response rate • ACR70 response rate	Change from baseline at week 12 in: • DAS28-CRP • HAQ-DI • SF-36 PCS • morning stiffness (duration) • FACIT-F Proportion of patients achieving: • LDA (DAS28-CRP ≤ 3.2) at week 12 • CR based on DAS28-CRP at week 12 • LDA (CDAI ≤ 10) • ACR20 response rate at week 1	Change from baseline at week 12 in: • DAS28-CRP • HAQ-DI • SF-36 PCS Proportion of patients achieving: • LDA (DAS28-CRP ≤ 3.2) at week 12 • ACR50 response rate at week 12 • ACR70 response rate at week 12 • ACR20 response rate at week 1	Change from baseline at week 12 in: • DAS28-CRP • HAQ-DI • SF-36 PCS • mTSS at week 24 Proportion of patients achieving: • LDA (DAS28-CRP ≤ 3.2) at week 12 • CR (DAS28-CRP ≤ 2.6) at week 24 • no radiographic progression (change from baseline mTSS ≤ 0) at week 24 • ACR20 response at week 12

		COMPARE (M14-465)	MONOTHERAPY (M15-555)	NEXT (M13-549)	BEYOND (M13-542)	EARLY (M13-545)
		• with no radiographic progression (change from baseline in mTSS ≤ 0) at week 26 • ACR50 response rate at week 12 • ACR70 response rate at week 12		• ACR50 response rate at week 12 • ACR70 response rate at week 12		• ACR70 response at week 12
	Exploratory end points	**Exploratory:** ACR 20/50/70 response rates Change from baseline in: • individual components of ACR response • DAS28-CRP and DAS28-ESR • CDAI and SDAI • morning stiffness (severity and duration) Proportion of patients: • achieving LDA or CR by DAS28-CRP, DAS28-ESR, SDAI, CDAI • with ≤ –0.3 and ≤ –0.22 change from baseline in HAQ-DI ACR/EULAR Boolean remission End points at weeks 12, 26, and 48: • change from baseline in SF-36	**Exploratory:** Change from baseline at weeks 2, 4, 8, and 14 in: • individual components of ACR response • CDAI and SDAI • DAS28-CRP and DAS28-ESR • morning stiffness (severity and duration) • EQ-5D-5L at weeks 4 and 14 • SF-36 at weeks 4 and 14 Proportion of patients: • achieving LDA and CR based on DAS28-CRP, DAS28-ESR, SDAI and CDAI criteria • with change from baseline in HAQ-DI ≤ –0.22 and ≤ –0.3	**Additional end points at all visits in period 1:** Change from baseline in: • individual components of ACR response • DAS28-CRP and DAS28-ESR • CDAI and SDAI • morning stiffness (severity and duration) • EQ-5D-5L • SF-36 • FACIT-F • RA-WIS Proportion of patients achieving: • LDA and CR based on DAS28-CRP, DAS28-ESR, SDAI and CDAI criteria • MCID in change from baseline in HAQ-DI (i.e., ≤ –0.3) among those with baseline HAQ-DI ≥ 0.3 ACR20/50/70 response rates	**Exploratory:** Change from baseline in: • individual components of ACR response • DAS28-CRP and DAS28-ESR • CDAI and SDAI • morning stiffness (severity and duration) Proportion of patients achieving: • LDA and CR based on DAS28-CRP, DAS28-ESR, SDAI and CDAI criteria • MCID in change from baseline in HAQ-DI (i.e., ≤ –0.3) among those with baseline HAQ-DI ≥ 0.3 ACR20/50/70 response rates ACR/EULAR Boolean remission	**Exploratory:** ACR 20/50/70 response rates Change from baseline in: • individual components of ACR response • DAS28-CRP and DAS28-ESR • CDAI and SDAI • morning stiffness (severity and duration) • EQ-5D-5L • FACIT-F • WPAI • SF-36 • mTSS at weeks 24 and 48 • Radiographic joint space narrowing and erosion scores at weeks 24 and 48 Proportion of patients: • achieving LDA or CR by DAS28-CRP, DAS28-ESR, SDAI, CDAI

		COMPARE (M14-465)	MONOTHERAPY (M15-555)	NEXT (M13-549)	BEYOND (M13-542)	EARLY (M13-545)
		• change from baseline in FACIT-F • change from baseline in RA-WIS • change from baseline in EQ-5D-5L **Additional end points at weeks 26 and 48:** • change from baseline in mTSS • proportion of patients with no radiographic progression (defined as change from baseline in mTSS ≤ 0) Change from baseline in joint space narrowing score and joint erosion score	ACR20/50/70 response rates at weeks 2, 4, 8, and 14 ACR/EULAR Boolean remission at weeks 2, 4, 8, and 14	ACR/EULAR Boolean remission	Proportion of patients with no concomitant corticosteroid use (among patients with corticosteroid use at baseline) up until week 48 **Additional end points at weeks 4 and 12:** Change from baseline in: • EQ-5D-5L • ISI (sleep) • SF-36	• with ≤ –0.3 and ≤ –0.22 change from baseline in HAQ-DI • with no radiographic progression (change from baseline in mTSS ≤ 0) at weeks 24 and 48 Proportion of patients with no concomitant corticosteroid use (among patients with corticosteroid use at baseline) up until completion of study
NOTES	**Publications**	Fleischmann et al. (2019)[15] Fleischmann et al. (2019)[16]	Smolen et al. (2019)[17]	Burmester et al. (2018)[18]	Genovese et al. (2018)[19]	None

ACR = American College of Rheumatology; bDMARD = biological DMARD; CCP = cyclic citrullinated peptide; CDAI = Clinical Disease Activity Index; CR = clinical remission; CRP = C-reactive protein; csDMARD = conventional synthetic DMARD; DAS28 = Disease Activity Score 28; DB = double-blind; DMARD = disease-modifying antirheumatic drug; EQ-5D-5L = EuroQol 5-Dimensions 5-Levels; ESR = erythrocyte sedimentation rate; EULAR = European League Against Rheumatism; FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue; HAQ-DI = Health Assessment Questionnaire–Disability Index; hsCRP = high-sensitivity CRP; ISI = Insomnia Severity Index; IV = intravenous; JAK = Janus kinase; LDA = low disease activity; LTE = long-term extension; MCID = minimal clinically important difference; mTSS = modified total Sharp score; MTX = methotrexate; OLE = open-label extension; PCS = physical component summary; RA = rheumatoid arthritis; RA-WIS = Work Instability Scale for Rheumatoid Arthritis; RCT = randomized controlled trial; RF = rheumatoid factor; SC = subcutaneous; SDAI = Simplified Disease Activity Index; SF-36 = Short Form (36) Health Survey; WPAI = work productivity and activity impairment.

[a] Japan only.

[b] In China and Japan 7.5 mg MTX was administered orally, once weekly, titrated up to 15 mg.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

Description of Studies

Five pivotal multinational DB RCTs met the criteria for this systematic review. All five studies enrolled adults with adult-onset RA, and in all but EARLY (where patients were methotrexate naive), patients' symptoms had been inadequately controlled on DMARDs. In COMPARE, patients had to be inadequately controlled on methotrexate, in NEXT patients had to be inadequately controlled on any csDMARD, and in BEYOND patients had to be inadequately controlled on bDMARDs. The primary outcome in the COMPARE, NEXT, and BEYOND studies was the proportion of patients achieving an ACR20 response at 12 weeks. The primary outcome in MONOTHERAPY was the proportion achieving an ACR20 response at 14 weeks, and the primary outcome in EARLY (N = 947) was the proportion achieving an ACR50 response at 12 weeks. The primary outcome was reported at week 12 in all studies except MONOTHERAPY, in which it was reported at week 14. All the included studies compared the approved upadacitinib 15 mg dose to placebo (NEXT and BEYOND), methotrexate (EARLY and MONOTHERAPY), or placebo and adalimumab with methotrexate background (COMPARE). In all studies, key secondary outcomes with control of the type I error rate included HRQoL on HAQ-DI, the DAS28-CRP, and the mTSS. There was a screening period in each study that ran up to 35 days. The double-blind period ranged from 12 weeks to 48 weeks, and each study had an extension period (ongoing at the time of writing this report), which are planned for up to five years.

Randomization was carried out using an interactive response system in all studies. Randomization in COMPARE and NEXT was stratified by geographic region and prior exposure to bDMARDs. Randomization in MONOTHERAPY and EARLY was stratified by geographic region, and randomization in BEYOND was stratified by the number of failed bDMARDs a patient had experienced. The randomization ratio in COMPARE was at a 2:2:1 ratio of oral upadacitinib 15 mg once daily, placebo, and adalimumab injection every other week, respectively. In EARLY, randomization was at a 1:1:1 ratio of oral upadacitinib 15 mg once daily, oral upadacitinib 30 mg once daily, and methotrexate oral daily. In the rest of the studies, randomization was at a 2:2:1:1 ratio of oral upadacitinib 15 mg once daily, oral upadacitinib 30 mg once daily, placebo (upadacitinib 15 mg in the extension phase), and placebo (upadacitinib 30 mg in the extension phase), respectively.

Populations

Inclusion and Exclusion Criteria

All the examined studies included adults diagnosed with adult-onset RA for three or more months (except EARLY, in which patients had to have had symptoms for six or more weeks) according to the ACR/EULAR 2010 definition and have a disease activity involving six or more swollen joints and six or more tender joints, with the exception of NEXT, which also allowed the use of the 1987 revised ACR classification as a diagnosis of RA. All studies excluded patients who were exposed to any JAK inhibitor. However, the population inclusion and exclusion criteria differed among the studies in terms of exposure to treatment before enrolment: COMPARE included patients who had been on methotrexate for three months or more, but they could not have been on bDMARDs for more than three months; MONOTHERAPY included patients who had been on methotrexate for three months or more but had no bDMARD exposure; NEXT included patients who were on any csDMARD for three or more months, but they could not have been diagnosed as bDMARD inadequate responders; BEYOND included patients who were on csDMARD for three months or more and were considered as failed by one or more bDMARD; and EARLY included patients who

were either methotrexate naive or had at three weeks or less of methotrexate therapy but no exposure to any bDMARDs.

Baseline Characteristics

Patients included in the studies ranged in mean age from 53.4 years (standard deviation [SD] 12.73) in EARLY up to 57.1 years (SD 11.42) in BEYOND. The majority of patients in all studies were female and white. Within studies, imbalances between treatment groups can be noted in a numerically disproportional number of patients with positive rheumatoid factor in BEYOND, a numerically disproportional number of patients with history of treatment with methotrexate alone or with csDMARD other than methotrexate in NEXT, a numerically disproportional number of male patients randomized to treatment groups in NEXT, a numerically disproportional number of male patients randomized to treatment groups in NEXT, and a numerically disproportional number of patients with oral steroid use at baseline and anti-CCP positive antibodies in EARLY. Otherwise, there were no notable between-group differences in baseline characteristics. Across studies, patients in EARLY were younger and had experienced shorter duration of the disease; patients in BEYOND were oldest and had experienced the longest duration of the disease. Exposure to previous treatment varied across studies as dictated by the inclusion and exclusion criteria of each study, with COMPARE having the highest proportion of patients receiving oral steroids at baseline.

Table 5: Summary of Baseline Characteristics (NEXT, BEYOND)

Characteristic	SELECT-NEXT (M13-549)		SELECT-BEYOND (M13-542)	
	Placebo (N = 221)	UPA 15 mg (N = 221)	Placebo (N = 169)	UPA 15 mg (N = 164)
Age in years, mean (SD)	56.0 (12.22)	55.3 (11.47)	57.6 (11.39)	56.3 (11.34)
Male, n (%)	55 (24.9)	39 (17.6)	26 (15.4)	27 (16.5)
Race, n (%)				
White	187 (84.6)	188 (85.1)	143 (84.6)	142 (86.6)
Black or African-American	10 (4.5)	13 (5.9)	21 (12.4)	17 (10.4)
American Indian or Alaska Native	1 (0.5)	0	0	3 (1.8)
Asian	19 (8.6)	19 (8.6)	5 (3.0)	2 (1.2)
Multiple	4 (1.8)	1 (0.5)	0	0
Weight in kg, mean (SD)	81.9 (21.21)	80.8 (22.71)	80.2 (20.83)	83.1 (20.20)
Duration of RA diagnosis in years, mean (SD)	7.2 (7.45)	7.3 (7.89)	14.5 (9.22)	12.4 (9.38)
Joint counts, mean (SD)				
TJC68	24.7 (14.96)	25.2 (13.80)	28.5 (15.27)	27.8 (16.31)
SJC66	15.4 (9.24)	16.0 (10.04)	16.3 (9.58)	17.0 (10.75)
Biomarkers				
DAS28-CRP, mean (SD)	5.6 (0.84)	5.7 (0.97)	5.8 (1.00)	5.9 (0.95)
CRP in mg/L, mean (SD)	12.6 (13.96)	16.6 (19.17)	16.3 (21.10)	16.2 (18.62)
RF positive, n (%)	164 (74.2)	163 (73.8)	113 (66.9)	119 (73.0)
Anti-CCP positive, n (%)	167 (75.9)	174 (79.1)	117 (69.2)	119 (72.6)
Prior bDMARD use, n (%)	29 (13.1)	27 (12.2)	NA	NA

Characteristic	SELECT-NEXT (M13-549)		SELECT-BEYOND (M13-542)	
Prior failed bDMARDs				
1 MOA and ≤ 2 prior bDMARDs, n (%)	NA	NA	117 (69.2)	116 (70.7)
> 1 MOA and/or > 2 prior bDMARDs n (%)	NA	NA	52 (30.8)	48 (29.3)
Concomitant csDMARD at baseline				
MTX alone, n (%)	141 (64.1)	122 (55.5)	122 (72.6)	118 (73.3)
MTX + other csDMARD, n (%)	49 (22.3)	47 (21.4)	17 (10.1)	19 (11.8)
csDMARD other than MTX, n (%)	30 (13.6)	51 (23.2)	29 (17.3)	24 (14.9)
Missing, n (%)	1	1	-	-
Taking oral steroid at baseline, n (%)	106 (48.0)	96 (43.4)	74 (43.8)	83 (50.6)
Oral steroid dose in mg, mean (SD)	6.3 (2.55)	6.0 (2.36)	6.3 (2.42)	537 (2.37)
Taking MTX at baseline, n (%)	190 (86.0)	169 (76.5)	NA	NA
Mean MTX dose in mg, mean (SD)	16.3 (4.89)	17.0 (4.87)	NA	NA

bDMARD= biologic DMARD; CCP = cyclic citrullinated peptide; CRP = C-reactive protein; csDMARD = conventional synthetic DMARD; DAS28 = Disease Activity Score 28; DMARD = disease-modifying antirheumatic drug; MOA = mechanism of action; MTX = methotrexate; NA = not available; RA = rheumatoid arthritis; RF = rheumatoid factor; SD = standard deviation; SJC = swollen joint count; TJC = tender joint count; UPA = upadacitinib.

Source: SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report.[4]

Table 6: Baseline Characteristics (COMPARE, MONOTHERAPY, EARLY)

Characteristic	SELECT-COMPARE (M14-465)			SELECT-MONOTHERAPY (M15-555)		SELECT-EARLY (M13-545)	
	Placebo (N = 651)	ADA 40 mg (N = 327)	UPA 15 mg (N = 651)	MTX (N = 216)	UPA 15 mg (N = 217)	MTX (N = 314)	UPA 15 mg (N = 317)
Age in years, mean (SD)	53.6 (12.24)	53.7 (11.70)	54.2 (12.08)	55.3 (11.12)	54.5 (12.20)	53.3 (12.89)	51.9 (12.58)
Male, n (%)	139 (21.4)	68 (20.8)	130 (20.0)	37 (17.1)	43 (19.8)	74 (23.6)	76 (24.0)
Race, n (%)							
White	561 (86.2)	292 (89.3)	576 (88.5)	176 (81.5)	173 (79.7)	256 (81.5)	256 (80.8)
Black or African-American	38 (5.8)	17 (5.2)	33 (5.1)	11 (5.1)	15 (6.9)	12 (3.8)	8 (2.5)
American Indian or Alaska Native	2 (0.3)	1 (0.3)	1 (0.2)	3 (1.4)	4 (1.8)	2 (0.6)	8 (2.5)
Asian	39 (6.0)	15 (4.6)	31 (4.8)	24 (11.1)	24 (11.1)	37 (11.8)	35 (11.0)
Multiple	10 (1.5)	2 (0.6)	10 (1.5)	2 (0.9)	1 (0.5)	5 (1.6)	7 (2.2)
Weight in kg, mean (SD)	76.5 (18.46)	75.9 (18.83)	77.3 (19.83)	76.9 (20.43)	75.5 (20.30)	74.5 (19.13)	75.0 (18.72)
Duration of RA diagnosis in years, mean (SD)	8.3 (8.00)	8.3 (8.42)	8.1 (7.73)	5.8 (6.63)	7.5 (8.88)	2.6 (5.14)	2.9 (5.38)
Joint counts, mean (SD)							
TJC68	26.0 (14.30)	26.4 (15.16)	26.4 (15.15)	25.2 (15.99)	24.5 (15.10)	24.6 (16.15)	25.4 (14.42)
SJC66	16.2 (8.97)	16.3 (9.19)	16.6 (10.31)	16.9 (11.52)	16.4 (10.94)	16.9 (10.58)	16.9 (10.35)
Biomarkers							

Characteristic	SELECT-COMPARE (M14-465)			SELECT-MONOTHERAPY (M15-555)		SELECT-EARLY (M13-545)	
DAS28-CRP, mean (SD)	5.8 (0.94)	5.9 (0.96)	5.8 (0.97)	5.6 (1.04)	5.6 (0.92)	5.9 (0.97)	5.9 (0.97)
CRP in mg/L, mean (SD)	18.0 (21.52)	19.8 (21.51)	17.9 (22.49)	14.5 (17.33)	14.0 (16.49)	21.2 (22.05)	23.0 (27.37)
RF positive, n (%)	517 (79.4)	265 (81.0)	521 (80.0)	151 (69.9)	155 (71.4)	232 (73.9)	251 (79.4)
Anti-CCP positive, n (%)	529 (81.5)	264 (80.7)	525 (80.6)	153 (70.8)	159 (73.3)	236 (75.2)	258 (81.4)
Taking oral steroid at baseline, n (%)	391 (60.2)	202 (61.8)	388 (59.6)	115 (53.24)	112 (51.61)	162 (51.6)	146 (46.06)
Oral steroid dose in mg, mean (SD)	6.3 (2.41)	6.5 (2.44)	6.2 (2.27)	6.2 (2.56)	6.1 (2.52)	6.4 (2.41)	6.4 (3.10)
Taking MTX at baseline, n	650	326	650	215	215	NA	NA
Mean MTX dose in mg, mean (SD)	16.8 (3.82)	17.1 (3.76)	17.0 (4.17)	16.7 (4.41)	16.8 (4.21)	NA	NA

ADA = adalimumab; CCP = cyclic citrullinated peptide; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; MTX = methotrexate; NA = not available; RA = rheumatoid arthritis; RF = rheumatoid factor; SD = standard deviation; SJC = swollen joint count; TJC = tender joint count; UPA = upadacitinib.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[5] SELECT-EARLY Clinical Study Report.[5]

Interventions

While all the included studies randomized patients to a 15 mg once daily upadacitinib arm, they differed in comparison, background therapy, and availability of rescue therapy. Four of the five studies also randomized patients to a 30 mg upadacitinib arm; however, further description of the 30 mg upadacitinib arm will not be included in this systematic review as it is higher than the dose approved by Health Canada. All patients in the COMPARE study received matching placebo injection or pill (double-dummy design) and appropriate methotrexate background therapy and were able to receive rescue treatment (either upadacitinib for placebo or adalimumab treatment group patients or adalimumab for upadacitinib treatment group patients) in accordance with specific criteria in each treatment group. Patients enrolled in the BEYOND and NEXT studies were randomized to upadacitinib once daily or matching placebo and had csDMARD background therapy and no rescue therapy conditions. Patients enrolled in EARLY and MONOTHERAPY were randomized to upadacitinib or methotrexate with matching placebo and no background therapy, with rescue therapy (ability to initiate or change back to background RA medications, including corticosteroids, NSAIDs, acetaminophen or paracetamol, and csDMARDs) available after the primary outcome date.

Outcomes

The primary outcome in all the studies was the percentage of patients with ACR20 responses at 12 weeks, with the exception of MONOTHERAPY, in which the ACR20 primary outcome was assessed at 14 weeks. The ACR criteria provide a composite measure of improvement in both swollen and tender joint counts and at least three of five additional disease criteria: patient global assessment of disease activity; physician global assessment of disease activity; patient assessment of pain; HAQ; CRP; and ESR. The ACR joint count for RA assesses 68 joints for tenderness and 66 joints for swelling. Patient and physician assessments are conducted using visual analogue scale (VAS) or Likert scale

measurements. ACR20, 50, or 70 responses represent at least a 20%, 50%, or 70% improvement, respectively, in tender and swollen joint counts as well as in three of the five aforementioned core measures. In all the studies, the sponsor described a "joint evaluator" as assessing whether a particular joint was "tender or painful." Aspects related to the patient evaluation (e.g., pain, global assessment) were reported by the patient, while the physician global assessment was completed by the physician. No specific measures beyond the established allocation concealment described earlier were taken to maintain blinding in the ACR-related outcomes.

The change from the baseline DAS28-CRP at week 12 was a secondary outcome in all included studies, with the exception of MONOTHERAPY, which reported the change from baseline at 14 weeks. DAS28-CRP is based on a 28-joint count that includes hands, wrists, elbows, shoulders, and knees. The formula used to calculate the DAS28-CRP is as follows:

$$\text{DAS28-CRP} = 0.56 \times \sqrt{(t28)} + 0.28 \times \sqrt{(sw28)} + 0.014 \times GH + 0.36 \times \ln(CRP+1) + 0.96$$

Where DAS28 = Disease Activity Score 28; CRP = C-reactive protein; t28 = tender joint count of 28 joints; sw28 = swollen joint count of 28 joints; GH = general health measured by a patient's global assessment of disease activity on a VAS of 100 mm.

DAS28 indicates an absolute level of disease activity, with a score of 5.1 or greater being considered high disease activity, a score lower than 3.2 being considered LDA, and a score lower than 2.6 indicating remission.[20-22]

A minimal important difference for the DAS28 has not been determined. However, a clinical change based on the EULAR response criteria can be used to interpret a clinical response according to the DAS28, as described in Table 7.

Table 7: EULAR Improvement Response Criteria (DAS28)

Baseline DAS28	DAS28 improvement over time points		
	> 1.2	0.6 to 1.2	< 0.6
< 3.2	Good response	Moderate response	No response
3.2 to 5.1	Moderate response	Moderate response	No response
> 5.1	Moderate response	No response	No response

DAS28 = Disease Activity Score 28; EULAR = European League Against Rheumatism.

Source: Matsui et al. (2007).[23]

The change from baseline to week 12 in HAQ-DI scores was a secondary outcome of all included studies (except in MONOTHERAPY it was reported at week 14). The full HAQ collects data on five generic patient-centred health dimensions: 1) to avoid disability, 2) to be free of pain and discomfort, 3) to avoid adverse treatment effects, 4) to keep dollar costs of treatment low, and 5) to postpone death.[22] The HAQ-DI is the disability assessment component of the HAQ. There are 20 questions that assess a patient's physical functional status in eight categories: dressing, arising, eating, walking, hygiene, reach, grip, and common activities.[24,25] For each of these categories, patients report the amount of difficulty they have in performing specific activities, and their responses are made on a scale from 0 (no difficulty) to 3 (unable to do). The eight category scores are averaged into an overall HAQ-DI score on a scale from 0 (no disability) to 3 (completely disabled). A number of investigators have estimated the minimal important difference of the HAQ-DI to be 0.22; however, differences as small as 0.10 have been suggested as clinically important.[24]

A change from baseline in the mTSS was a secondary radiographic outcome in COMPARE and EARLY studies. The outcome was assessed centrally by two qualified physicians or radiologists, who were blinded to the site number, patient number, treatment allocation, time sequence, and clinical response. The score includes 16 joints from the hands and wrists (graded from 0 to 5) and six joints from the feet (graded from 0 to 10). The joint space narrowing score includes 15 areas from the hands and wrists (graded from 0 to 4) and six areas from the feet (also graded from 0 to 4). The maximum erosion score is 160 for hands and wrists and 120 for feet, while the maximum joint space narrowing score is 120 for hands and 48 for feet.[26] Maximum total scores for both erosion and joint space narrowing are calculated as follows:

Erosion = (32 joints in hands and wrists × 5) + (12 joints in feet × 10) = 280

Joint space narrowing = (30 joints in hands and wrists × 4) + (12 joints in feet × 4) = 168

A change from baseline in mTSS ≤ 0 was considered to define a patient with no radiographic progression.

The change from baseline to week 12 in the SF-36 was a secondary outcome in all the studies (although at week 14 in MONOTHERAPY). SF-36 is a generic health assessment questionnaire that has been used in clinical trials to study the impact of chronic disease on HRQoL. The SF-36 consists of eight subdomains: physical functioning, pain, vitality, social functioning, psychological functioning, general health perceptions, role limitations due to physical problems, and role limitations due to emotional problems.[12] The SF-36 also provides two component summaries: the physical component summary (PCS) and the mental component summary (MCS). The eight subdomains and component summaries are each measured on scales of 0 to 100, with an increase in score indicating improvement in health status. The MCID for either the PCS or MCS of the SF-36 is typically between 2.5 and 5 points.[27-29]

Other outcomes include the CDAI (a composite continuous index to assess disease activity without using the high-sensitivity CRP measurement). The CDAI can be calculated on the basis of the tender joint count of 28 joints, the swollen joint count of 28 joints, the patient's global assessment of disease activity (in centimetres; scale of 0 to 10), the physician's global assessment of disease activity (in centimetres; scale of 0 to 10), clinical remission (CDAI ≤ 2.8, or DAS28 < 2.6, or SDAI ≤ 3.3), LDA (DAS28 ≤ 3.2), and the Functional Assessment of Chronic Illness Therapy (FACIT)–Fatigue scale.

Harms were reported through a description of adverse events, serious adverse events, and withdrawal due to adverse events.

Statistical Analysis

Power analysis in COMPARE indicated that 1,500 patients on a 2:2:1 randomization to upadacitinib, placebo, and adalimumab would provide 90% power to detect a 22% difference in ACR20 response rate at week 12 versus placebo, assuming a 37% placebo response rate, a 10% dropout rate, and a two-sided significance level of 0.05. This sample size also provided at least 90% power for testing the noninferiority of upadacitinib versus adalimumab in LDA or ACR50 response rate at week 12, with a noninferiority margin of 10%, assuming 35% and 40% LDA or ACR50 response rates for adalimumab and upadacitinib, respectively. In MONOTHERAPY and NEXT studies, a sample size of 600 would provide 90% power to detect a 21% difference in ACR20 response rate, with an assumption of a 37% response rate in the methotrexate treatment or placebo groups, at

two-sided alpha = 0.025 and accounting for a 10% dropout rate. In BEYOND, a sample size of 450 would provide 90% power to detect a 20% difference in ACR20 rates at week 12 (assuming a placebo ACR20 response rate of 27%), at two-sided alpha = 0.025 and accounting for a 10% dropout rate. In EARLY, a sample size of 900 would provide 90% power to detect a 20% difference in ACR50 response rate, with an assumption of a 20% response rate in the methotrexate treatment group, at two-sided alpha = 0.025 and accounting for a 10% dropout rate. It was not clear what the basis was for the assumed ACR20 or ACR50 responses in either intervention, methotrexate, or placebo groups. No rationale was provided regarding the assumption for the sample size.

In all the trials, the point estimate, 95% CI, and P values were reported for the treatment comparisons between each upadacitinib dose group and the comparison group (placebo, methotrexate, or adalimumab). P values were constructed using the Cochran-Mantel-Haenszel test, adjusting for stratification factor. Analysis in COMPARE and NEXT was stratified by geographic region and prior exposure to bDMARD. Analysis in MONOTHERAPY and EARLY was stratified by geographic region. Analysis in BEYOND was stratified by the number of failed bDMARDs a patient had experienced. In all the trials except COMPARE, to account for comparison of multiple arms (two doses), the statistical significance level was set at 0.025. In COMPARE, statistical significance was set at 0.05 for upadacitinib versus placebo, and assessment of upadacitinib versus adalimumab was based on a noninferiority margin of 10% in the outcomes of ACR50 and LDA at week 12.

In all the trials, comparison to placebo in change from baseline in the DAS28-CRP and HAQ-DI outcomes (and the SF-36 outcome in the EARLY study) was calculated as a least squares mean, with a 95% CI and P value, using an analysis of covariance model, with treatment, baseline value, and stratification factors as covariates. The exception was the analysis method in EARLY, where the analysis of covariance model included treatment and geographic region as the fixed factors and the corresponding baseline values as the covariates. Other continuous outcomes (e.g., change in the duration of morning stiffness) in all the trials, with the exception of EARLY, were calculated using mixed models for repeated measures, with fixed effects of treatment, visit, and treatment-by-visit interaction, stratification factor, and baseline value as covariate.

In all the studies, subgroup analysis was planned for the primary efficacy outcome for gender, age, body mass index, weight, race, region, duration of RA, baseline rheumatoid factor status, baseline anti-CCP antibodies level, and baseline DAS28-CRP, unless the subgroup analysis was less than 10% of the planned study size, in which case the subgroup analysis would not be conducted. Subgroup analyses were to be presented without P value.

In all the studies, missing data in binary outcomes were handled through a nonresponder imputation approach. Further sensitivity analysis using observed cases was conducted for the primary outcome. Continuous data analyzed through analysis of variance used a multiple imputation approach, in which missing data would be imputed multiple times under random variation to generate multiple imputed “pseudo-complete” datasets. Missing data in the radiographic outcome of mTSS were imputed from linear extrapolation, in which the X-ray at the time point of interest was imputed, assuming a linear relationship between the baseline, the X-ray collected at rescue, and the time point of interest.

A hierarchical testing procedure was employed to adjust for multiple statistical comparisons in all the studies (i.e., to control the type I error rate). In studies with two upadacitinib doses,

the significance level was set at 0.025 for all statistical testing. Table 8 displays the hierarchical testing procedure of each of the studies.

Table 8: Multiple Testing Hierarchical Testing Procedure

COMPARE (M14-465)	MONOTHERAPY (M15-555)	NEXT (M13-549)	BEYOND (M13-542)	EARLY (M13-545)
ACR20	ACR20	ACR20	ACR20	ACR50
DAS28-CRP	DAS28-CRP	DAS28-CRP	DAS28-CRP	DAS28-CRP
mTSS	HAQ-DI	HAQ-DI	HAQ-DI	HAQ-DI
HAQ-DI	SF-36	SF-36	LDA	mTSS
ACR50 (noninferiority vs. adalimumab)	LDA	LDA	SF-36	LDA
SF-36	CR	CR	-	CR
LDA based on DAS28-CRP ≤ 3.2	Morning stiffness	LDA based on CDAI	-	SF-36
CR	-	Morning stiffness	-	-
LDA based on CDAI ≤ 10	-	-	-	-
Morning stiffness (duration)	-	-	-	-
FACIT-F	-	-	-	-
ACR50 (superiority vs. adalimumab)	-	-	-	-
Patient's global assessment of pain	-	-	-	-

ACR = American College of Rheumatology; CDAI = Clinical Disease Activity Index; CR = clinical remission; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue; HAQ-DI = Health Assessment Questionnaire–Disability Index; LDA = low disease activity; mTSS = modified total Sharp score; SF-36 = Short Form (36) Health Survey; vs. = versus.

Analysis Populations

In each study, the full analysis set includes all randomized patients who received at least one dose of the study drug. The full analysis set is used for all efficacy and baseline analyses. The per-protocol analysis set is a subset of the full analysis set and consists of all full analysis set patients who did not meet any major protocol deviations up to the reported date for the primary outcome. The safety analysis set included all patients who received at least one dose of the study drug.

Results

Patient Disposition

Overall, discontinuation rates in the period from randomization to the analysis of the primary outcome were less than 10%, with the exception of EARLY, in which the discontinuation rate was approximately 11.5%. The discontinuation rate was relatively balanced between treatment groups within the studies, with two notable exceptions: in the COMPARE study, the adalimumab discontinuation rate of 8.3% contrasted with the upadacitinib discontinuation rate of 4.8%, and in the EARLY study, the methotrexate discontinuation rate of 14.5% contrasted with the upadacitinib discontinuation rate of 8.5%. The most common reason for discontinuation in the control groups was withdrawal of consent, followed closely by adverse events.

Table 9: Patient Disposition up to Week 14 (COMPARE, MONOTHERAPY)

	SELECT-COMPARE (M14-465)			SELECT-MONOTHERAPY (M15-555)	
	UPA 15 mg q.d.	ADA 40 mg e.o.w.	Placebo	UPA 15 mg q.d.	MTX
Screened, N	1,629			648[a]	
Randomized, N	651	327	651	217	216
Completed period 1[b] study participation, N (%)	246 (37.79)	122 (37.31)	246 (37.79)	201 (92.6)	202 (93.5)
Discontinued study, N (%)	NA	NA	NA	16 (4.7)	14 (6.5)
Reason for discontinuation, N (%)					
Adverse events	NA	NA	NA	5 (2.3)	1 (0.5)
Lost to follow-up	NA	NA	NA	4 (1.8)	0
Patient withdrawal of consent	NA	NA	NA	6 (2.8)	11 (5.1)
Other	NA	NA	NA	1 (0.5)	2 (0.9)
Completed week 14 on study drug, N (%)	620 (95.2)	300 (91.7)	620 (95.2)	199 (91.7)	197 (91.2)
Discontinued study drug by week 14, N (%)	31 (4.8)	27 (8.3)	31 (4.8)	18 (8.3)	19 (8.8)
Reason for discontinuation, N (%)					
Adverse events	15 (2.3)	15 (4.6)	10 (1.5)	6 (2.8)	5 (2.3)
Lack of efficacy	0	0	1 (0.2)	1 (0.5)	4 (1.9)
Lost to follow-up	3 (0.5)	1 (0.3)	4 (0.6)	4 (1.8)	0
Patient withdrawal of consent	9 (1.4)	9 (2.8)	14 (2.2)	7 (3.2)	7 (3.2)
Other	4 (0.6)	2 (0.6)	2 (0.3)	0	3 (1.4)
FAS, N	651	327	651	217	216
PP, N	591	299	596	211	208
Safety, N	650[c]	327[c]	652[c]	217	216

ADA = adalimumab; e.o.w. = every other week; FAS = full analysis set; MTX = methotrexate; NA = not available; PP = per protocol; q.d. = once daily; UPA = upadacitinib.

[a] All patients were randomized; UPA 30 mg q.d. (n = 215) is not shown.

[b] Period 1 had a duration of 48 weeks for SELECT-COMPARE and 14 weeks for SELECT-MONOTHERAPY. The values reported for SELECT-COMPARE reflect those patients who had completed period 1 at the time of reporting, which was based on a week 26 interim analysis.

[c] Analysis up to week 14.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report.[2]

Table 10: Patient Disposition up to Week 12 (NEXT, BEYOND)

	SELECT-NEXT (M13-549)		SELECT-BEYOND (M13-542)	
	Placebo	UPA 15 mg	Placebo	UPA 15 mg
Screened, N	661[a]		499[b]	
Randomized, N	221	221	169	165
Completed period 1, N (%)	208 (94.1)	213 (96.4)	151 (89.3)	157 (95.7)
Discontinued study, N (%)	13 (5.9)	8 (3.6)	18 (10.7)	7 (4.3)
Reason for discontinuation, N (%)				
Adverse events	6 (2.7)	2 (0.9)	4 (2.4)	1 (0.6)
Lost to follow-up	1 (0.5)	0	3 (1.8)	0
Patient withdrawal of consent	3 (1.4)	5 (2.3)	3 (1.8)	4 (2.4)
Other	3 (1.4)	1 (0.5)	8 (4.7)	2 (1.2)

	SELECT-NEXT (M13-549)		SELECT-BEYOND (M13-542)	
Completed period 1 on study drug, N (%)	207 (93.7)	210 (95.0)	147 (87.0)	156 (95.1)
Discontinued study drug, N (%)	14 (6.3)	11 (5.0)	22 (13.0)	8 (4.9)
Reason for discontinuation, N (%)				
Adverse events	5 (2.3)	5 (2.3)	7 (4.1)	3 (1.8)
Lack of efficacy	4 (1.8)	0	10 (5.9)	1 (0.6)
Lost to follow-up	1 (0.5)	0	3 (1.8)	0
Patient withdrawal of consent	2 (0.9)	5 (2.3)	1 (0.6)	3 (1.8)
Other	2 (0.9)	1 (0.5)	1 (0.6)	1 (0.6)
FAS, N	221	221	169	164
PP, N	204	209	153	150
Safety, N	221	221	169	164

FAS = full analysis set; PP = per protocol; UPA = upadacitinib.

[a] All patients were randomized; UPA 30 mg q.d. (n = 219) is not shown.

[b] All patients were randomized; UPA 30 mg q.d. (n = 165) is not shown. Also, Of these patients, 498 patiens received study drug (1 patient who was a screen fail was randomized in error and did not receive study drug).

Source: SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND CSR.[4]

Table 11: Patient Disposition up to Week 24 (EARLY)

	SELECT-EARLY (M13-549)	
	UPA 15 mg	**MTX**
Screened, N	975[a]	
Randomized, N	317	315[b]
Completed period 1 (24 weeks) on study drug, N (%)	290 (91.5)	268 (85.1)
Discontinued study drug, N (%)	27 (8.5)	46 (14.6)
Reason for discontinuation of study drug, N (%)		
Adverse events	13 (4.1)	13 (4.1)
Lack of efficacy	0	10 (3.2)
Lost to follow-up	4 (1.3)	3 (1.0)
Patient withdrawal of consent	8 (2.5)	15 (4.8)
Other	2 (0.6)	5 (1.6)
FAS, N	317	314
PP, N	298	295
Safety, N	317	314

FAS = full analysis set; MTX = methotrexate; PP = per protocol; UPA = upadacitinib.

[a] All patients were randomized; UPA 7.5 mg (n = 75) and UPA 30 mg q.d. (n = 300) are not shown.

[b] One patient was randomized but not dosed.

Source: SELECT-EARLY Clinical Study Report.[5]

Exposure to Study Treatments

In reported period 1, in COMPARE, the mean (SD) duration of exposure was 129.5 (53.87) days with upadacitinib, 129.8 (54.48) with adalimumab, and 140.2 (44.07) with placebo. In MONOTHERAPY, the mean (SD) duration of exposure was 93.7 (19.49) days with

upadacitinib and 89.7 (23.84) days with methotrexate . In NEXT, the mean (SD) duration of exposure was 81.8 (12.87) days with upadacitinib and 81.7 (11.55) days with placebo. In BEYOND, the mean (SD) duration of exposure was 137.1 (42.86) days with upadacitinib and 77.9 (18.3) days with placebo. In EARLY, the mean (SD) duration of exposure was 160.6 (30.50) days with upadacitinib and 153.0 (41.43) days with methotrexate. Discrepancy was noted in the mean number of days of exposure between the two groups in the BEYOND study.

Table 12: Exposure to Study Treatments (Safety Analysis Set)

	N	Period of analysis	Number of days, mean (SD)	Number of days, median (min, max)
SELECT-COMPARE (M14-465)				
UPA 15 mg q.d.	1,031	Up to week 26	129.5 (53.87)	128.0 (2, 200)
ADA 40 mg e.o.w.	452		129.8 (54.48)	126.5 (14, 126)
Placebo	652		140.2 (44.07)	154.0 (1, 191)
SELECT-MONOTHERAPY (M15-555)				
UPA 15 mg q.d.	217	Up to week 14	93.7 (19.49)	98.0 (4, 140)
MTX	216		89.7 (23.84)	98.0 (7, 112)
SELECT-NEXT (M13-549)				
UPA 15 mg q.d.	221	Up to week 12	81.8 (12.87)	84.0 (6, 112)
Placebo	221		81.7 (11.55)	84.0 (7, 97)
SELECT-BEYOND (M13-542)				
UPA 15 mg q.d.	236	Up to week 24	137.1 (42.86)	167.0 (6, 176)
Placebo	169		77.9 (18.30)	84.0 (7, 92)
SELECT-EARLY (M13-545)				
UPA 15 mg q.d.	317	Up to week 24	160.6 (30.50)	168.0 (4, 211)
MTX	314		153.0 (41.43)	168.0 (7, 189)

ADA = adalimumab; e.o.w. = every other week; MTX = methotrexate; q.d. = once daily; SD = standard deviation; UPA = upadacitinib.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

Efficacy

Only those efficacy outcomes and analyses of subgroups identified in the review protocol are reported in the following sections. See Appendix 3 for other efficacy data (subgroups and FACIT-Fatigue).

Clinical Response (ACR20/50/70)

The primary outcome in SELECT-EARLY was the proportion of patients achieving ACR50 at week 12. All other pivotal studies used the proportion of patients achieving ACR20 at week 12, with the exception of SELECT-MONOTHERAPY, which used ACR20 at week 14 as the primary outcome. Sensitivity analyses using observed case imputations and per-protocol analysis set, as well as reported subgroups analyses, were consistent with base-case primary outcome analyses. Throughout the five studies, upadacitinib 15 mg once daily consistently achieved superiority versus comparators on the ACR20 primary outcome at week 12 (week 14 in MONOTHERAPY). When compared to placebo, the response rate was higher in upadacitinib-treated patients, at 34.1 (95% CI, 29.0 to 39.2) in COMPARE,

28.1 (95% CI, 19.1 to 37.0) in NEXT, and 36.2 (95% CI, 26.2 to 46.2) in BEYOND. When compared to methotrexate, the response rate was higher in upadacitinib-treated patients, at 26.5 (95% CI, 17.5 to 35.6) in MONOTHERAPY and 21.6 (95% CI, 14.3 to 28.8) in EARLY. When compared to adalimumab, the response rate was higher in upadacitinib-treated patients, at 7.5 (95% CI, 1.2 to 13.8). A similar magnitude of difference versus placebo and methotrexate was reported for ACR50 at week 12, with the exception of a numerically higher response rate difference in ACR50 than in ACR20 of upadacitinib versus adalimumab, at 16.1 (95% CI, 9.9 to 22.3). The ACR70 response rate of upadacitinib versus the control group also showed a statistically significant benefit of upadacitinib, with the exception of the ACR70 comparison in BEYOND; none of the ACR50 or ACR70 were included in the statistical hierarchy, with the exception of EARLY ACR50.

Table 13: Clinical Response (ACR20/50/70) at Week 12 or Week 14

	Total N	Responder n	Response rate (95% CI)	Response rate difference versus control	
				Point estimate (95% CI)	P value
ACR20 response at week 12 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	459	70.5 (67.0 to 74.0)	7.5 (1.2 to 13.8) [UPA vs. ADA] 34.1 (29.0 to 39.2) [UPA vs. PBO]	0.018[a] [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	206	63.0 (57.8 to 68.2)		
Placebo	651	237	36.4 (32.7 to 40.1)		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	147	67.7 (61.5 to 74.0)	26.5 (17.5 to 35.6)	< 0.001
MTX	216	89	41.2 (34.6 to 47.8)		
SELECT-NEXT					
UPA 15 mg q.d.	221	141	63.8 (57.5 to 70.1)	28.1 (19.1 to 37.0)	< 0.001
Placebo	221	79	35.7 (29.4 to 42.1)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	106	64.6 (57.3 to 72.0)	36.2 (26.2 to 46.2)	< 0.001
Placebo	169	48	28.4 (21.6 to 35.2)		
SELECT-EARLY					
UPA 15 mg q.d.	317	240	75.7 (71.0 to 80.4)	21.6 (14.3 to 28.8)	< 0.001[a]
MTX	314	170	54.1 (48.6 to 59.7)		
ACR50 response at week 12 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	294	45.2 (41.3 to 49.0)	16.1 (9.9 to 22.3) [UPA vs. ADA] 30.3 (25.6 to 35.0) [UPA vs. PBO]	NI met; Sup.: < 0.001 [UPA vs. ADA] < 0.001[a] [UPA vs. PBO]
ADA 40 mg e.o.w.	327	95	29.1 (24.1 to 34.0)		
Placebo	651	97	14.9 (12.2 to 17.6)		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	91	41.9 (35.4 to 48.5)	26.7 (18.5 to 34.8)	< 0.001[a]
MTX	216	33	15.3 (10.5 to 20.1)		
SELECT-NEXT					
UPA 15 mg q.d.	221	84	38.0 (31.6 to 44.4)	23.1 (15.1 to 31.0)	< 0.001[a]
Placebo	221	33	14.9 (10.2 to 19.6)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	56	34.1 (26.9 to 41.4)	22.3 (13.6 to 31.1)	< 0.001[a]
Placebo	169	20	11.8 (7.0 to 16.7)		

	Total N	Responder n	Response rate (95% CI)	Response rate difference versus control	
				Point estimate (95% CI)	P value
SELECT-EARLY					
UPA 15 mg q.d.	317	165	52.1 (46.6 to 57.5)	23.7 (16.3 to 31.1)	< 0.001
MTX	314	89	28.3 (23.4 to 33.3)		
ACR70 response at week 12 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	162	24.9 (21.6 to 28.2)	11.4 (6.5 to 16.4) [UPA vs. ADA] 20.0 (16.3 to 23.7) [UPA vs. PBO]	< 0.001[a] [UPA vs. ADA] < 0.001[a] [UPA vs. PBO]
ADA 40 mg e.o.w.	327	44	13.5 (9.8 to 17.2)		
Placebo	651	32	4.9 (3.3 to 6.6)		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	49	22.6 (17.0 to 28.1)	19.8 (13.8 to 25.8)	< 0.001[a]
MTX	216	6	2.8 (0.6 to 5.0)		
SELECT-NEXT					
UPA 15 mg q.d.	221	46	20.8 (15.5 to 26.2)	14.9 (8.7 to 21.1)	< 0.001[a]
Placebo	221	13	5.9 (2.8 to 9.0)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	19	11.6 (6.7 to 16.5)	5.1 (–1.1 to 11.2)	0.110[a]
Placebo	169	11	6.5 (2.8 to 10.2)		
SELECT-EARLY					
UPA 15 mg q.d.	317	103	32.5 (27.3 to 37.6)	18.5 (12.1 to 24.9)	< 0.001[a]
MTX	314	44	14.0 (10.2 to 17.9)		

ACR = American College of Rheumatology; ADA = adalimumab; CI = confidence interval; e.o.w. = every other week; FAS = full analysis set; MTX = methotrexate; NI = non-inferiority; NRI = nonresponder imputation; PBO = placebo; sup = superiority; q.d. = once daily; UPA = upadacitinib; vs. = versus.

[a] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

Radiographic Response

Radiographic response outcomes were not reported in SELECT-MONOTHERAPY, SELECT-NEXT, or SELECT-BEYOND. The mTSS and the proportion of patients to show no radiographic progression were both statistically significantly better in the upadacitinib arms than in the methotrexate and placebo arms. However, no statistically significant difference in these two outcomes was noted in the comparison of upadacitinib versus adalimumab. When compared to placebo, the radiographic response rate was higher in upadacitinib-treated patients, at 7.5 (95% CI, 3.0 to 12.1) in COMPARE, and when compared to methotrexate, the radiographic response rate was higher in upadacitinib-treated patients, at 9.8 (95% CI, 3.5 to 16.2) in EARLY.

Table 14: Radiographic Response: mTSS at Week 24 or 26 (COMPARE, EARLY)

	Total N	Baseline	Week 24/26		Treatment group difference versus control		
		Mean	Mean	LS mean change from baseline (95% CI)	N	LS mean difference (95% CI)	P value
Modified total Sharp score, change from baseline at week 24 or 26 (linear extrapolation, FAS)							
SELECT-COMPARE (at week 26)							
UPA 15 mg q.d.	651	34.73	34.93	0.24 (–0.04 to 0.53)	593	0.14 (–0.23 to 0.51) [UPA vs. ADA] –0.67 (–0.97 to –0.37) [UPA vs. PBO]	0.448 [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	35.09	35.15	0.10 (–0.25 to 0.46)	296		
Placebo	651	35.47	36.35	0.92 (0.64 to 1.20)	599		
SELECT-EARLY (at week 24)							
UPA 15 mg q.d.	317	17.03	17.16	0.14 (–0.09 to 0.37)	279	– 0.53 (–0.85 to –0.20)	0.001[a]
MTX	315	13.89	14.55	0.67 (0.43 to 0.90)	264		

ADA = adalimumab; CI = confidence interval; e.o.w. = every other week; FAS = full analysis set; LS = least squares; mTSS = modified total Sharp score; MTX = methotrexate; PBO = placebo; q.d. = once daily; UPA = upadacitinib; vs. = versus.

[a] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-EARLY Clinical Study Report.[5]

Table 15: Radiographic Response: No Radiographic Progression at Week 24 or 26 (COMPARE, EARLY)

	Total N	Responder	Response rate (95% CI)	Response rate difference versus control		
		n		N	Point estimate (95% CI)	P value
Proportion of patients with no radiographic progression (change from baseline mTSS ≤ 0), at week 24 or 26[a] (linear extrapolation, FAS)						
SELECT-COMPARE						
UPA 15 mg q.d.	651	495	83.5 (80.5 to 86.5)	593	–3.4 (–8.2 to 1.5) [UPA vs. ADA] 7.5 (3.0 to 12.1) [UPA vs. PBO]	0.187 [UPA vs. ADA] 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	257	86.8 (83.0 to 90.7)	296		
Placebo	651	455	76.0 (72.5 to 79.4)	599		
SELECT-EARLY						
UPA 15 mg q.d.	317	244	87.5 (83.6 to 91.3)	279	9.8 (3.5 to 16.2)	0.002[b]
MTX	315	205	77.7 (72.6 to 82.7)	264		

ADA = adalimumab; CI = confidence interval; e.o.w. = every other week; FAS = full analysis set; mTSS = modified total Sharp score; MTX = methotrexate; PBO = placebo; q.d. = once daily; UPA = upadacitinib; vs. versus.

[a] Nominal P value was constructed using the Cochran-Mantel-Haenszel test, adjusting for the stratification factor of prior biologic disease-modifying antirheumatic drug use.

[b] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-EARLY Clinical Study Report.[5]

Disease Activity

The DAS28-CRP at week 12 (week 14 in EARLY) was the first secondary outcome in the statistical testing hierarchy of all the studies. In all the reported studies, upadacitinib was statistically superior to all comparators, including versus adalimumab (comparison to adalimumab was not in the statistical testing hierarchy). This trend was also observed in the LDA and clinical response outcomes, where upadacitinib was statistically significantly superior to all comparators.

Table 16: Disease Activity: DAS28-CRP at Week 12 or 14

	Total N	Baseline	Week 12/14		Treatment group difference versus control		
		Mean	Mean	LS mean change from baseline (95% CI)	N	LS mean difference (95% CI)	P value
DAS28-CRP, change from baseline at week 12 (MI, FAS)							
SELECT-COMPARE							
UPA 15 mg q.d.	651	5.78	3.33	–2.48 (–2.61 to –2.35)	634	–0.47 (–0.638 to –0.295) [UPA vs. ADA] –1.33 (–1.47 to –1.19) [UPA vs. PBO]	< 0.001[a] [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	5.87	3.84	–2.01 (–2.18 to -1.85)	319		
Placebo	651	5.83	4.69	–1.15 (–1.28 to –1.02)	643		
SELECT-MONOTHERAPY (at week 14)							
UPA 15 mg q.d.	217	5.61	3.36	–2.29 (–2.476 to –2.098)	215	–1.08 (–1.319 to –0.848)	< 0.001
MTX	216	5.59	4.43	–1.20 (–1.393 to –1.013)	215		
SELECT-NEXT							
UPA 15 mg q.d.	221	5.65	3.39	–2.20 (–2.401 to –1.995)	217	–1.18 (–1.420 to –0.939)	< 0.001
Placebo	221	5.55	4.51	–1.02 (–1.220 to –0.817)	220		
SELECT-BEYOND							
UPA 15 mg q.d.	164	5.87	3.48	–2.31 (–2.517 to –2.096)	163	–1.29 (–1.574 to –1.008)	< 0.001
Placebo	169	5.83	4.75	–1.02 (–1.228 to –0.804)	165		
SELECT-EARLY							
UPA 15 mg q.d.	317	5.90	3.17	–2.73 (–2.87 to –2.58)	317	–0.88 (–1.09 to –0.67)	< 0.001
MTX	315	5.89	4.05	–1.85 (–2.00 to –1.69)	312		

ADA = adalimumab; CI = confidence interval; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; e.o.w. = every other week; FAS = full analysis set; LS = least squares; MI = multiple imputation; MTX = methotrexate; PBO = placebo; q.d. = once daily; UPA = upadacitinib; vs. = versus.

[a] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-EARLY Clinical Study Report.[5]

Table 17: Disease Activity Outcomes: LDA, CR

	Total N	Responder	Response rate (95% CI)	Response rate difference versus control	
		n		Point estimate (95% CI)	P value
Proportion of patients achieving LDA based on DAS28-CRP ≤ 3.2 at week 12 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	293	45.0 (41.2 to 48.8)	16.3 (10.0 to 22.5) [UPA vs. ADA] 31.2 (26.5 to 35.8) [UPA vs. PBO]	< 0.001[a] [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	94	28.7 (23.8 to 33.7)		
Placebo	651	90	13.8 (11.2 to 16.5)		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	97	44.7 (38.1 to 51.3)	25.3 (16.8 to 33.7)	< 0.001
MTX	216	42	19.4 (14.2 to 24.7)		
SELECT-NEXT					
UPA 15 mg q.d.	221	107	48.4 (41.8 to 55.0)	31.2 (23.0 to 39.5)	< 0.001
Placebo	221	38	17.2 (12.2 to 22.2)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	71	43.3 (35.7 to 50.9)	29.1 (19.9 to 38.3)	< 0.001
Placebo	169	24	14.2 (8.9 to 19.5)		
SELECT-EARLY					
UPA 15 mg q.d.	317	169	53.3 (47.8 to 58.8)	25.0 (17.6 to 32.4)	< 0.001
MTX	314	89	28.3 (23.4 to 33.3)		
Proportion of patients achieving CR based on DAS28-CRP at week 12 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	187	28.7 (25.2 to 32.2)	10.7 (5.3 to 16.1) [UPA vs. ADA] 22.6 (18.6 to 26.5) [UPA vs. PBO]	< 0.001[a] [UPA vs. ADA] < 0.001[UPA vs. PBO]
ADA 40 mg e.o.w.	327	59	18.0 (13.9 to 22.2)		
Placebo	651	40	6.1 (4.3 to 8.0)		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	61	28.1 (22.1 to 34.1)	19.8 (12.8 to 26.8)	< 0.001
MTX	216	18	8.3 (4.6 to 12.0)		
SELECT-NEXT					
UPA 15 mg q.d.	221	68	30.8 (24.7 to 36.9)	20.8 (13.6 to 28.1)	< 0.001
Placebo	221	22	10.0 (6.0 to 13.9)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	47	28.7 (21.7 to 35.6)	19.2 (11.0 to 27.4)	< 0.001[a]
Placebo	169	16	9.5 (5.1 to 13.9)		
SELECT-EARLY					
UPA 15 mg q.d.	317	113	35.6 (30.4 to 40.9)	22.0 (15.5 to 28.5)	< 0.001[a]
MTX	314	43	13.7 (9.9 to 17.5)		
Proportion of patients achieving LDA based on CDAI ≤ 10 at week 12 (NRI, FAS)					

	Total N	Responder	Response rate (95% CI)	Response rate difference versus control	
		n		Point estimate (95% CI)	P value
SELECT-COMPARE					
UPA 15 mg q.d.	651	263	40.4 (36.6 to 44.2)	10.4 (4.2 to 16.7) [UPA vs. ADA] 24.1 (19.4 to 28.8) [UPA vs. PBO]	0.001 [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	98	30.0 (25.0 to 34.9)		
Placebo	651	106	16.3 (13.4 to 19.1)		
SELECT-MONOTHERAPY (at week 14)					
UPA 15 mg q.d.	217	75	34.6 (28.2 to 40.9)	10.0 (1.5 to 18.6)	0.021[a]
MTX	216	53	24.5 (18.8 to 30.3)		
SELECT-NEXT					
UPA 15 mg q.d.	221	89	40.3 (33.8 to 46.7)	21.3 (13.0 to 29.5)	< 0.001
Placebo	221	42	19.0 (13.8 to 24.2)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	52	31.7 (24.6 to 38.8)	17.5 (8.7 to 26.4)	< 0.001[a]
Placebo	169	24	14.2 (8.9 to 19.5)		
SELECT-EARLY					
UPA 15 mg q.d.	317	147	46.4 (40.9 to 51.9)	16.8 (9.3 to 24.2)	< 0.001[a]
MTX	314	93	29.6 (24.6 to 34.7)		

ADA = adalimumab; CDAI = Clinical Disease Activity Index; CI = confidence interval; CR = clinical remission; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; e.o.w. = every other week; FAS = full analysis set; LDA = low disease activity; MTX = methotrexate; NRI = nonresponder imputation; PBO = placebo; q.d. = once daily; UPA = upadacitinib; vs. = versus.

[a] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

Quality of Life

The HAQ-DI was a secondary outcome of all the studies within the statistical testing hierarchy. In all the reported studies, upadacitinib was statistically superior to all comparators, including versus adalimumab (comparison to adalimumab was not in the statistical testing hierarchy). This trend was also observed in the SF-36 and the patient global assessment of pain.

Table 18: Quality of Life Outcomes at Week 12 or 14

	Total N	Baseline	Week 12/14		Treatment group difference versus control		
		Mean	Mean	LS mean change from baseline (95% CI)	N	LS mean difference (95% CI)	P value
HAQ-DI, change from baseline at week 12 (MI, FAS)							
SELECT-COMPARE							
UPA 15 mg q.d.	651	1.63	0.99	–0.60 (–0.653 to –0.538)	644	–0.11 (–0.184 to –0.036) [UPA vs. ADA] –0.31 (–0.372 to –0.253) [UPA vs. PBO]	0.004 [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	1.65	1.11	–0.49 (–0.556 to –0.415)	324		
Placebo	651	1.61	1.28	–0.28 (–0.339 to –0.227)	648		
SELECT-MONOTHERAPY (at week 14)							
UPA 15 mg q.d.	217	1.47	0.87	–0.65 (–0.734 to –0.565)	216	–0.33 (–0.431 to –0.220)	< 0.001
MTX	216	1.47	1.19	–0.32 (–0.413 to –0.235)	216		
SELECT-NEXT							
UPA 15 mg q.d.	221	1.48	0.85	–0.59 (–0.672 to –0.505)	216	–0.33 (–0.432 to –0.236)	< 0.001
Placebo	221	1.43	1.15	–0.25 (–0.340 to –0.170)	220		
SELECT-BEYOND							
UPA 15 mg q.d.	164	1.67	1.24	–0.39 (–0.475 to –0.304)	163	–0.22 (–0.343 to –0.100)	< 0.001
Placebo	169	1.57	1.38	–0.17 (–0.260 to –0.075)	165		
SELECT-EARLY							
UPA 15 mg q.d.	317	1.60	0.77	–0.83 (–0.90 to –0.76)	317	–0.34 (–0.44 to –0.25)	< 0.001
MTX	315	1.60	1.11	–0.49 (–0.55 to –0.42)	313		
SF-36 (physical component score), change from baseline at week 12 (MMRM, FAS)							
SELECT-COMPARE							
UPA 15 mg q.d.	651	32.46	40.92	7.89 (7.11 to 8.68)	616	1.62 (0.62 to 2.62) [UPA vs. ADA] 4.33 (3.52 to 5.15) [UPA vs. PBO]	0.002[a] [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	32.15	39.07	6.27 (5.31 to 7.23)	309		
Placebo	651	32.46	36.57	3.56 (2.79 to 4.33)	616		
SELECT-MONOTHERAPY (at week 14)							
UPA 15 mg q.d.	217	33.22	41.30	8.28 (7.17 to 9.40)	200	3.97 (2.52 to 5.42)	< 0.001
MTX	216	33.24	37.08	4.32 (3.19 to 5.44)	195		
SELECT-NEXT							
UPA 15 mg q.d.	221	33.26	41.34	7.58 (6.43 to 8.74)	209	4.55 (3.13 to 5.98)	< 0.001
Placebo	221	33.18	36.85	3.03 (1.88 to 4.18)	207		
SELECT-BEYOND							
UPA 15 mg q.d.	164	30.71	36.99	5.83 (4.60 to 7.05)	156	3.44 (1.72 to 5.15)	< 0.001
Placebo	169	31.84	34.60	2.39 (1.14 to 3.64)	145		

	Total N	Baseline	Week 12/14		Treatment group difference versus control		
		Mean	Mean	LS mean change from baseline (95% CI)	N	LS mean difference (95% CI)	P value
SELECT-EARLY (MI)							
UPA 15 mg q.d.	317	32.74	43.13	9.99 (9.11 to 10.88)	315	4.25 (3.00 to 5.50)	< 0.001
MTX	315	33.11	39.10	5.74 (4.84 to 6.64)	311		

ADA = adalimumab; CI = confidence interval; e.o.w. = every other week; FAS = full analysis set; HAQ-DI= Health Assessment Questionnaire–Disability Index; LS = least squares; MI = multiple imputation; MMRM = mixed models for repeated measures; MTX = methotrexate; PBO = placebo; q.d. = once daily; SF-36 = Short Form (36) Health Survey; UPA = upadacitinib; vs. = versus.

[a] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

Table 19: Patient Global Assessment of Pain at Week 12 or 14

	Total N	Baseline	Week 12/14		Treatment group difference versus control		
		Mean	Mean	LS mean change from baseline (95% CI)	N	LS mean difference (95% CI)	P value
PGA of pain (VAS, mm), change from baseline at week 12 (MMRM, FAS)							
SELECT-COMPARE							
UPA 15 mg q.d.	651	65.77	33.18	–31.76 (–33.96 to –29.56)	614	–6.45 (–9.63 to –3.27) [UPA vs. ADA] –16.30 (–18.89 to –13.71) [UPA vs. PBO]	< 0.001 [UPA vs. ADA] < 0.001[a] [UPA vs. PBO]
ADA 40 mg e.o.w.	327	66.36	39.90	–25.31 (–28.16 to –22.47)	307		
Placebo	651	64.82	49.24	–15.46 (–17.63 to –13.29)	616		
SELECT-MONOTHERAPY (at week 14)							
UPA 15 mg q.d.	217	62.06	35.31	–26.15 (–29.69 to –22.60)	198	–12.27 (–16.98 to –7.56)	< 0.001[a]
MTX	216	62.36	48.14	–13.88 (–17.44 to –10.31)	195		
SELECT-NEXT							
UPA 15 mg q.d.	221	64.51	32.87	–29.92 (–33.40 to –26.44)	207	–19.67 (–24.15 to –15.18)	< 0.001[a]
Placebo	221	62.08	50.83	–10.26 (–13.71 to –6.80)	206		
SELECT-BEYOND							
UPA 15 mg q.d.	164	67.96	40.40	–25.91 (–30.05 to –21.76)	156	–15.52 (–21.36 to –9.69)	< 0.001[a]
Placebo	169	69.00	55.09	–10.38 (–14.60 to –6.16)	145		
SELECT-EARLY (LOCF for rescue)							
UPA 15 mg q.d.	317	68.51	30.55	–36.28 (–39.08 to –33.49)	302	–10.92 (–14.89 to –6.96)	< 0.001[a]
MTX	315	65.15	40.72	–25.36 (–28.28 to –22.44)	278		

ADA = adalimumab; CI = confidence interval; e.o.w. = every other week; FAS = full analysis set; LOCF = last observation carried forward; LS = least squares; MMRM = mixed models for repeated measures; MTX = methotrexate; PBO = placebo; PGA = patient global assessment; q.d. = once daily; UPA = upadacitinib; VAS = visual analogue scale; vs. = versus.

[a] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

Functional and Disability Outcomes

Duration of morning joint stiffness was a secondary outcome of COMPARE, MONOTHERAPY, and NEXT; it was an exploratory outcome in BEYOND and EARLY. In all the reported studies, upadacitinib was statistically superior to all comparators, including versus adalimumab (comparison to adalimumab was not in the statistical testing hierarchy).

Table 20: Functional and Disability Outcomes at Week 12 or 14

	Total N	Baseline	Week 12/14		Treatment group difference versus control		
		Mean	Mean	LS mean change from baseline (95% CI)	N	LS mean difference (95% CI)	P value
Morning stiffness duration (minutes), change from baseline at week 12 (MMRM, FAS)							
SELECT-COMPARE							
UPA 15 mg q.d.	651	141.08	48.16	–92.63 (–103.03, –82.23)	618	–9.92 (–23.89, 4.05) [UPA vs. ADA] –44.04 (–55.39, –32.69) [UPA vs. PBO]	0.164[a] [UPA vs. ADA] < 0.001[UPA vs. PBO]
ADA 40 mg e.o.w.	327	149.06	59.75	–82.71 (–95.80, –69.62)	306		
Placebo	651	144.21	91.84	–48.59 (–58.84, –38.34)	619		
SELECT-MONOTHERAPY (at week 14)							
UPA 15 mg q.d.	217	147.65	55.79	–94.56 (–113.57, –75.54)	199	–41.53 (–66.56, –16.50)	0.001
MTX	216	155.70	102.26	–53.03 (–72.18, –33.88)	196		
SELECT-NEXT							
UPA 15 mg q.d.	221	147.52	54.27	–85.28 (–105.61, 64.95)	207	–51.01 (–78.14, –23.87)	< 0.001
Placebo	221	141.53	95.67	–34.27 (–54.63, –13.91)	202		
SELECT-BEYOND							
UPA 15 mg q.d.	164	141.71	67.82	–81.47 (–109.52, –53.42)	157	–66.40 (–105.50, –27.31)	< 0.001[a]
Placebo	169	140.74	135.33	–15.07 (–43.30, 13.16)	147		
SELECT-EARLY (LOCF for rescue)							
UPA 15 mg q.d.	317	172.78	43.20	–105.97 (–116.87, –95.08)	301	–33.61 (–49.00, –18.21)	< 0.001[a]
MTX	315	127.86	71.62	–72.37 (–83.67, –61.06)	278		

ADA = adalimumab; CI = confidence interval; e.o.w. = every other week; FAS = full analysis set; LOCF = last observation carried forward; LS = least squares; MMRM = mixed models for repeated measures; MTX = methotrexate; PBO = placebo; q.d. = once daily; SD = standard deviation; UPA = upadacitinib; vs. = versus.

[a] Outcome was outside the statistical testing hierarchy.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

Harms

Only those harms identified in the review protocol are reported in the following sections. See Table 21, Table 22, and Table 23 for detailed harms data.

Adverse Events

In COMPARE, 64.2% of upadacitinib patients, 60.2% of adalimumab patients, and 53.2% of placebo patients experienced an adverse event. In MONOTHERAPY, the percentages were 47.5% in the upadacitinib and 47.2% in the placebo groups. In NEXT, the percentages were 56.6% in the upadacitinib and 48.9% in the placebo groups. In BEYOND, the percentages were 55.5% in the upadacitinib and 56.2% in the placebo groups. In EARLY,

the percentages were 64.0% in the upadacitinib and 65.3% in the placebo groups. Respiratory tract infections were the most common adverse events in all the included studies.

Serious Adverse Events

In COMPARE, 3.7% of upadacitinib patients, 4.3% of adalimumab patients, and 2.9% of placebo patients experienced an adverse event. In MONOTHERAPY, the percentages were 5.1% in the upadacitinib and 2.8% in the methotrexate groups. In NEXT, the percentages were 4.1% in the upadacitinib and 2.3% in the placebo groups. In BEYOND, the percentages were 4.9% in the upadacitinib and 0% in the placebo groups. In EARLY, the percentages were 4.7% in the upadacitinib and 4.1% in the methotrexate groups. No single serious adverse event was most common across the five studies.

Withdrawals Due to Adverse Events

In COMPARE, 3.5% of upadacitinib patients, 6.1% of adalimumab patients, and 2.3% of placebo patients withdrew owing to adverse events. In MONOTHERAPY, the percentages were 3.7% in the upadacitinib and 2.8% in the methotrexate groups. In NEXT, the percentages were 3.2% in the upadacitinib and 3.2% in the placebo groups. In BEYOND, the percentages were 2.4% in the upadacitinib and 5.3% in the placebo groups. In EARLY, the percentages were 2.4% in the upadacitinib and 5.3% in the methotrexate groups.

Mortality

In COMPARE, three deaths were recorded in the adalimumab and two in the placebo groups. One death was recorded in the upadacitinib arm in the MONOTHERAPY study. In EARLY, two deaths were recorded in the upadacitinib arm and one in the methotrexate arm. Cardiovascular-related causes were responsible for almost one-third of the deaths.

Notable Harms

Notable harms identified for this review did not show explicit imbalance between groups, with the exception of a numerically higher proportion of neutropenia in COMPARE, BEYOND, and EARLY.

Table 21: Summary of Harms (COMPARE up to Week 26, Censored at Treatment Switching)

	COMPARE (M14-465)		
	UPA 15 mg q.d. N = 650	ADA 40 mg e.o.w. N = 327	PBO N = 652
Patients with ≥ 1 adverse event			
n (%)	417 (64.2)	197 (60.2)	347 (53.2)
Most common events,[a] n (%)			
Upper respiratory tract infection	37 (5.7)	7 (2.1)	24 (3.7)
Nasopharyngitis	36 (5.5)	9 (2.8)	19 (2.9)
Bronchitis	30 (4.6)	14 (4.3)	15 (2.3)
Urinary tract infection	29 (4.5)	16 (4.9)	23 (3.5)
Alanine aminotransferase increased	28 (4.3)	5 (1.5)	21 (3.2)
Hypertension	25 (3.8)	6 (1.8)	16 (2.5)
Aspartate aminotransferase increased	22 (3.4)	7 (2.1)	13 (2.0)

	COMPARE (M14-465)		
	UPA 15 mg q.d. N = 650	ADA 40 mg e.o.w. N = 327	PBO N = 652
Diarrhea	21 (3.2)	12 (3.7)	15 (2.3)
Pharyngitis	21 (3.2)	7 (2.1)	10 (1.5)
Blood creatine phosphokinase increased	20 (3.1)	2 (0.6)	10 (1.5)
Headache	19 (2.9)	5 (1.5)	16 (2.5)
Cough	17 (2.6)	4 (1.2)	8 (1.2)
Gastroenteritis	17 (2.6)	2 (0.6)	8 (1.2)
Back pain	16 (2.5)	4 (1.2)	13 (2.0)
Neutropenia	16 (2.5)	2 (0.6)	2 (0.3)
Nausea	14 (2.2)	9 (2.8)	14 (2.1)
Influenza	14 (2.2)	3 (0.9)	3 (0.5)
Leukopenia	13 (2.0)	2 (0.6)	4 (0.6)
Anemia	8 (1.2)	4 (1.2)	13 (2.0)
Sinusitis	7 (1.1)	9 (2.8)	4 (0.6)
Oral herpes	7 (1.1)	7 (2.1)	4 (0.6)
Rheumatoid arthritis (disease worsening)	6 (0.9)	6 (1.8)	27 (4.1)
Patients with ≥ 1 SAE			
n (%)	24 (3.7)	14 (4.3)	19 (2.9)
Most common events[b], n (%)			
Appendicitis	2 (0.3)	0	0
Cellulitis	0	2 (0.6)	0
Gastroenteritis	2 (0.3)	0	3 (0.5)
Pneumocystis jiroveci pneumonia	0	0	2 (0.3)
Abortion spontaneous	2 (0.3)	0	0
Pulmonary embolism	1 (0.2)	3 (0.9)	1 (0.2)
Patients who stopped treatment owing to adverse events			
n (%)	23 (3.5)	20 (6.1)	15 (2.3)
Most common events[b], n (%)			
Anemia	2 (0.3)	0	0
Pneumocystis jiroveci pneumonia	0	0	2 (0.3)
Alanine aminotransferase increased	2 (0.3)	2 (0.6)	1 (0.2)
Aspartate aminotransferase increased	2 (0.3)	2 (0.6)	1 (0.2)
Blood creatine increased	2 (0.3)	0	0
Pulmonary embolism	0	2 (0.6)	0
Deaths			
n (%)	0	3 (0.9)	2 (0.3)
Sudden death	0	0	1 (0.2)
Pneumocystis jiroveci pneumonia	0	0	1 (0.2)
Craniocerebral injury	0	1 (0.3)	0
Left ventricular failure	0	1 (0.3)	0

	COMPARE (M14-465)		
	UPA 15 mg q.d. N = 650	ADA 40 mg e.o.w. N = 327	PBO N = 652
Notable harms, n (%)			
Herpes zoster infection	5 (0.8)	1 (0.3)	3 (0.5)
Neutropenia	18 (2.8)	3 (0.9)	2 (0.3)
Lymphopenia	11 (1.7)	2 (0.6)	9 (1.4)
Thrombocytopenia	2 (0.3)	0	1 (0.2)
Malignancy (any)	0	1 (0.3)	2 (0.3)
Thrombosis (incl. increased platelets)[c]	2 (0.3)	3 (0.9)	1 (0.2)
MACE[d]	0	2 (0.6)	3 (0.5)
GI perforations	2 (0.3)	0	0
Hepatic disorder	43 (6.6)	12 (3.7)	32 (4.9)
Dyslipidemia	3 (0.5)	1 (0.3)	2 (0.3)

ADA = adalimumab; e.o.w. = every other week; GI = gastrointestinal; incl. = including; MACE = major adverse cardiovascular event; PBO = placebo; q.d. = once daily; SAE = serious adverse event; UPA = upadacitinib.

[a] Frequency ≥ 2% in any group.

[b] Frequency > 1 patient in any group.

[c] Deep vein thrombosis and fatal/nonfatal pulmonary embolism.

[d] Cardiovascular death, nonfatal myocardial infarction, and nonfatal stroke.

Source: SELECT-COMPARE Clinical Study Report.[1]

Table 22: Summary of Harms (MONOTHERAPY up to Week 14, EARLY up to Week 24)

	SELECT-MONOTHERAPY (M15-555)		SELECT-EARLY (M13-545)	
	UPA 15 mg q.d. N = 217	MTX N = 216	UPA 15 mg q.d. N = 317	MTX N = 314
Patients with ≥ 1 adverse event				
n (%)	103 (47.5)	102 (47.2)	203 (64.0)	205 (65.3)
Most common events[a]				
Urinary tract infection	9 (4.1)	5 (2.3)	17 (5.4)	20 (6.4)
Blood creatine phosphokinase increased	5 (2.3)	0	9 (2.8)	3 (1.0)
Upper respiratory tract infection	9 (4.1)	13 (6.0)	20 (6.3)	13 (4.1)
Herpes zoster	3 (1.4)	1 (0.5)	7 (2.2)	1 (0.3)
Rheumatoid arthritis (disease worsening)	2 (0.9)	10 (4.6)	6 (1.9)	14 (4.5)
Alanine aminotransferase increased	3 (1.4)	3 (1.4)	15 (4.7)	11 (3.5)
Leukopenia	1 (0.5)	1 (0.5)	6 (1.9)	4 (1.3)
Bronchitis	4 (1.8)	7 (3.2)	7 (2.2)	6 (1.9)
Nasopharyngitis	3 (1.4)	7 (3.2)	18 (5.7)	13 (4.1)
Headache	4 (1.8)	1 (0.5)	7 (2.2)	6 (1.9)
Nausea	2 (0.9)	4 (1.9)	17 (5.4)	15 (4.8)
Hypertension	4 (1.8)	4 (1.9)	12 (3.8)	8 (2.5)

	SELECT-MONOTHERAPY (M15-555)		SELECT-EARLY (M13-545)	
	UPA 15 mg q.d. N = 217	MTX N = 216	UPA 15 mg q.d. N = 317	MTX N = 314
Cough	1 (0.5)	1 (0.5)	10 (3.2)	4 (1.3)
Neutropenia	1 (0.5)	0	7 (2.2)	2 (0.6)
Hypertriglyceridemia	3 (1.4)	0	11 (3.5)	6 (1.9)
Diarrhea	1 (0.5)	3 (1.4)	10 (3.2)	9 (2.9)
Alopecia	1 (0.5)	0	6 (1.9)	7 (2.2)
Gastroenteritis	0	0	5 (1.6)	7 (2.2)
Dyspepsia	1 (0.5)	0	8 (2.5)	12 (3.8)
Respiratory tract infection	0	0	1 (0.3)	7 (2.2)
Fatigue	4 (1.8)	1 (0.5)	7 (2.2)	1 (0.3)
Patients with ≥ 1 SAE				
n (%)	11 (5.1)	6 (2.8)	15 (4.7)	13 (4.1)
Most common events[b]				
Acute myocardial infarction	0	0	0	2 (0.6)
Cholecystitis acute	0	2 (0.9)	NA	NA
Pneumonia	NA	NA	1 (0.3)	2 (0.6)
Patients who stopped treatment owing to adverse events				
n (%)	8 (3.7)	6 (2.8)	14 (4.4)	16 (5.1)
Most common events[b]				
Rheumatoid arthritis (disease worsening)	0	2 (0.9)	0	0
Gastroenteritis	0	0	0	2 (0.6)
Liver function test increased	0	0	0	2 (0.6)
Deaths				
n (%)	1 (0.5)	0	2 (0.6)	1 (0.3)
Hemorrhagic stroke	1 (0.5)	0	0	0
Acute myocardial infarction	0	0	0	1 (0.3)
Myocardial infarction	0	0	1 (0.3)	0
Hepatic vein thrombosis	0	0	1 (0.3)	0
Notable harms, n (%)				
Herpes zoster infection	3 (1.4)	1 (0.5)	7 (2.2)	1 (0.3)
Neutropenia	2 (0.9)	1 (0.5)	10 (3.2)	2 (0.6)
Lymphopenia	0	2 (0.9)	2 (0.6)	6 (1.9)
Thrombocytopenia	2 (0.9)	0	0	0
Malignancy (any)	2 (0.9)	1 (0.5)	3 (0.9)	1 (0.3)
Thrombosis (incl. increased platelets)[c]	1 (0.5)	0	0	1 (0.3)
MACE[d]	1 (0.5)	0	1 (0.3)	1 (0.3)

	SELECT-MONOTHERAPY (M15-555)		SELECT-EARLY (M13-545)	
	UPA 15 mg q.d. N = 217	MTX N = 216	UPA 15 mg q.d. N = 317	MTX N = 314
GI perforations	0	0	0	0
Hepatic disorder (any)	4 (1.8)	4 (1.9)	19 (6.0)	17 (5.4)
Dyslipidemia	2 (0.9)	0	3 (0.9)	0

GI = gastrointestinal; incl. = including; MACE = major adverse cardiovascular event; MTX = methotrexate; NA = not available; q.d. = once daily; SAE = serious adverse event; UPA = upadacitinib.

[a] Frequency ≥ 2% in any group.

[b] Frequency > 1 patient in any group.

[c] Deep vein thrombosis and fatal/nonfatal pulmonary embolism.

[d] Cardiovascular death, nonfatal myocardial infarction, and nonfatal stroke.

Source: SELECT-MONOTHERAPY Clinical Study Report;[2] and SELECT-EARLY Clinical Study Report.[5]

Table 23: Summary of Harms (NEXT up to Week 12, BEYOND up to Week 12)

	SELECT-NEXT (M13-549)		SELECT-BEYOND (M13-542)	
	UPA 15 mg q.d. N = 221	PBO N = 221	UPA 15 mg q.d. N = 164	PBO N = 169
Patients with ≥ 1 adverse event				
n (%)	125 (56.6)	108 (48.9)	91 (55.5)	95 (56.2)
Most common events[a]				
Nausea	16 (7.2)	7 (3.2)	6 (3.7)	4 (2.4)
Nasopharyngitis	12 (5.4)	9 (4.1)	7 (4.3)	11 (6.5)
Upper respiratory tract infection	12 (5.4)	9 (4.1)	13 (7.9)	13 (7.7)
Headache	9 (4.1)	12 (5.4)	7 (4.3)	8 (4.7)
Urinary tract infection	8 (3.6)	9 (4.1)	15 (9.1)	10 (5.9)
Blood creatine phosphokinase increased	5 (2.3)	0	NA	NA
Neutropenia	3 (1.4)	1 (0.5)	NA	NA
Bronchitis	4 (1.8)	5 (2.3)	7 (4.3)	4 (2.4)
Pyrexia	3 (1.4)	0	NA	NA
Sinusitis	6 (2.7)	1 (0.5)	4 (2.4)	2 (1.2)
Back pain	6 (2.7)	2 (0.9)	2 (1.2)	4 (2.4)
Diarrhea	5 (2.3)	9 (4.1)	4 (2.4)	6 (3.6)
Gastroenteritis	5 (2.3)	0	NA	NA
Rheumatoid arthritis (disease worsening)	4 (1.8)	10 (4.5)	4 (2.4)	10 (5.9)
Hypertension	3 (1.4)	5 (2.3)	3 (1.8)	4 (2.4)
Cough	8 (3.6)	2 (0.9)	4 (2.4)	2 (1.2)
Dizziness	6 (2.7)	1 (0.5)	2 (1.2)	5 (3.0)
Palpitations	NA	NA	0	4 (2.4)
Arthralgia	NA	NA	1 (0.6)	5 (3.0)

	SELECT-NEXT (M13-549)		SELECT-BEYOND (M13-542)	
	UPA 15 mg q.d. N = 221	PBO N = 221	UPA 15 mg q.d. N = 164	PBO N = 169
Patients with ≥ 1 SAE				
n (%)	9 (4.1)	5 (2.3)	8 (4.9)	0
Most common events[b]				
Wrist fracture	2 (0.9)	0	NA	NA
Patients who stopped treatment owing to adverse events				
n (%)	7 (3.2)	7 (3.2)	4 (2.4)	9 (5.3)
Most common events[b]				
Rheumatoid arthritis (disease worsening)	0	2 (0.9)	0	4 (2.4)
Deaths				
n (%)	0	0	0	0
Notable harms				
Herpes zoster infection	1 (0.5)	1 (0.5)	1 (0.6)	1 (0.6)
Neutropenia	4 (1.8)	1 (0.5)	5 (3.0)	0
Lymphopenia	1 (0.5)	1 (0.5)	2 (1.2)	2 (1.2)
Malignancy (any)	0	0	1 (0.6)	0
MACE[c]	0	0	1 (0.6)	0
Other adjudicated cardiovascular events	2 (0.9)	0	0	0
GI perforations	NA	NA	0	0
Creatine phosphokinase elevation	5 (2.3)	0	2 (1.2)	0
Renal dysfunction (any)	0	2 (0.9)	0	0
Hepatotoxicity	NA	NA	1 (0.6)	0
Dyslipidemia	1 (0.5)	0	NA	NA
Hepatic disorder (any)	4 (1.8)	5 (2.3)	2 (1.2)	2 (1.2)

GI = gastrointestinal; MACE = major adverse cardiovascular event; NA = not available; PBO = placebo; q.d. = once daily; SAE = serious adverse event; UPA = upadacitinib.

[a] Frequency ≥ 2% in any group.

[b] Frequency > 1 patient in any group.

[c] Cardiovascular death, nonfatal myocardial infarction, and nonfatal stroke.

Source: SELECT-NEXT Clinical Study Report;[3] and SELECT-BEYOND Clinical Study Report.[4]

Critical Appraisal

Internal Validity

The baseline and disease characteristics were comparable across treatment arms in all included trials, with the exception of imbalances within the studies, which can be noted in the percentage of patients with positive rheumatoid factor in BEYOND, patients with history of treatment with methotrexate alone or csDMARD other than methotrexate in NEXT, the number of male patients randomized to treatment groups in NEXT, and the number of patients with oral steroid use at baseline and anti-CCP positive antibodies in EARLY. Immune suppression and thus infection are known side effects of cytokine inhibitors for RA. However, infections are also very common events outside of clinical trials; therefore, it is

unlikely that an infection in a given patient would have been interpreted by that patient as a sign they were on upadacitinib, unless it were a serious, uncommon infection like herpes zoster, which occurred very infrequently in the included trials, with no obvious differences between groups.

The missing imputation method in binary outcomes employed a nonresponder imputation, supported by observed cases. Sensitivity analysis suggests that missing data had a small impact on the results. The impact of missing data on binary secondary outcomes (including the proportion of patients achieving ACR50, no radiographic progression, clinical remission, and LDA) was less clear owing to the lack of sensitivity analyses on these end points. Continuous outcome employed a multiple imputation approach to missing data, which requires a missing-at-random assumption.

There was a relatively large number of withdrawals in the placebo group in EARLY and a numerical difference between groups (14.6% withdrawals in methotrexate and 8.5% in upadacitinib). To a lesser numerical extent, there were more withdrawals in the adalimumab arm in COMPARE (8.3%) than in the upadacitinib arm (4.8%). It is not clear if these imbalances in discontinuation could have led to any sort of bias in the outcomes. Requirement of rescue medication was largely limited until after the primary outcome end point and thus unlikely to have affected the outcomes.

External Validity

The clinical expert consulted by CADTH on this review noted that while the inclusion criteria of the studies of upadacitinib were similar to other RA RCTs, in the expert's opinion, less than half of patients in the expert's practice would be eligible to participate in these trials. Most commonly, patients with negative rheumatoid factor are underrepresented in RA research. However, the clinical expert reported that there is no clear evidence that rheumatoid factor is a treatment effect modifier.

The included studies, collectively, cover a wide range of patients, who have been through various treatment options. However, Health Canada approval limits the indication to patients who have had an inadequate response or intolerance to methotrexate. As such, although EARLY is a pivotal study where patients were treatment naive, it does not match the Health Canada indication and may not inform the expected efficacy in Canadian practice.

One of the included studies compared upadacitinib with adalimumab directly. In addition, upadacitinib was compared to methotrexate as monotherapy. These two direct comparisons are valuable in informing the expected efficacy of upadacitinib in the Canadian practice setting. However, direct comparison against other existing JAK inhibitors would have been valuable in informing the clinical decision process as to which JAK inhibitor (e.g., baricitinib) could be useful.

The included studies covered a double-blind period that ranged from 12 weeks to 48 weeks. Although this is relatively a short time considering the chronic nature of the disease, all five studies have an extension phase that is ongoing.

Indirect Evidence

Objectives and Methods for the Summary of Indirect Evidence

Because of the lack of direct evidence comparing upadacitinib with other bDMARDs, other than adalimumab, the sponsor performed an NMA to estimate the efficacy of upadacitinib in patients with RA relative to other cDMARDs and bDMARDs. The objective of this section is to summarize and critically review the unpublished NMA performed by the sponsor and other available published indirect evidence that examines the relative efficacy and harms of upadacitinib compared with other treatments for RA.

Description of Indirect Comparison(s)

Two ITCs are discussed here: one submitted by the sponsor and one identified in the literature search as matching the inclusion criteria listed in this review. An overall description of the included studies is presented in Table 24.

Table 24: Study Selection Criteria and Methods for ITCs

	Sponsor-submitted ITC	Song et al. 2019[30]
Population	Pop. 1: Adult patients (≥ 18 years of age) meeting the ACR classification criteria for RA and an inadequate response to csDMARDs Pop. 2: Adult patients (≥ 18 years of age) meeting the ACR classification criteria for RA and an inadequate response to bDMARDs	csDMARD or bDMARD inadequate responders
Intervention	Upadacitinib 15 mg or 30 mg q.d. in monotherapy or in combination	Tofacitinib
Comparator	TNF-alpha inhibitors: • adalimumab • etanercept • infliximab • golimumab • certolizumab pegol JAK inhibitors: • tofacitinib • baricitinib • filgotinib • peficitinib Anti–B-cell therapy: • rituximab • co-stimulatory inhibitor molecules • abatacept • anti–IL-6 therapy • tocilizumab • sirukumab	Upadacitinib

	Sponsor-submitted ITC	Song et al. 2019[30]
	• sarilumab Anti–IL-1 therapy: • anakinra Additional interventions: • filgotinib and peficitinib Biosimilars to any of the previously interventions listed	
Outcome	Efficacy • ACR20/50/70 response rate to treatment • HAQ-DI • EULAR response • DAS28 score • DAS28 remission • CDAI score • patient assessment of functional ability • radiographic progression • patient's assessment of pain • patient or physician assessment of disease activity • morning stiffness • fatigue • disease activity (SDAI) • physical function • RA-related mortality • extra-articular manifestations of the disease • safety • incidence of AEs • thromboembolic events • incidence of SAEs, including MACEs • treatment withdrawal • health-related quality of life • EQ-5D-5LWPAI-RA	ACR20
Study design	RCTs, with no restriction on phase or study design (long-term extensions will be included if randomization is maintained) Observational studies, to include prospective cohort studies, case control studies, registries for upadacitinib only	RCTs
Publication characteristics	Not reported	Not reported
Exclusion criteria	Not matching the inclusion criteria	Not matching the inclusion criteria
Databases searched	Embase, MEDLINE, Cochrane library	MEDLINE, Embase, the Cochrane Controlled Trials Register, and the conference proceedings of the ACR and

	Sponsor-submitted ITC	Song et al. 2019[30]
		EULAR to identify available articles (up to November 2018)
Selection process	2 independent reviewers	2 independent reviewers
Data extraction process	1 reviewer and data check by another; no clear process for resolving discrepancies was reported	2 independent reviewers; no clear process for resolving discrepancies was reported
Quality assessment	NICE single technology appraisal criteria checklist	Jadad scoring

ACR = American College of Rheumatology; AE = adverse event; bDMARD = biologic DMARD; CDAI = Clinical Disease Activity Index; csDMARD = conventional synthetic DMARD; DAS28 = Disease Activity Score 28; DMARD = disease-modifying antirheumatic drug; EQ-5D-5L = EuroQol 5-Dimensions 5-Levels; EULAR = European League Against Rheumatism; HAQ-DI = Health Assessment Questionnaire–Disability Index; IL = interleukin; ITC = indirect treatment comparison; JAK = Janus kinase; MACE = major adverse cardiovascular event; NICE = National Institute for Health and Care Excellence; pop. = population; q.d. = once daily; RA = rheumatoid arthritis; RCT = randomized controlled trial; SAE = serious adverse event; SDAI = Simplified Disease Activity Index; TNF = tumour necrosis factor; WPAI = work productivity and activity impairment.

Source: CDR submission: TBD (upadacitinib), 15mg once daily oral tablet [CONFIDENTIAL manufacturer's submission]. In: Pointe-Claire (QC): AbbVie Corporation; 2019 Jul 04.[31]. Song GG, Choi SJ, Lee YH. Comparison of the efficacy and safety of tofacitinib and upadacitinib in patients with active rheumatoid arthritis: A Bayesian network meta-analysis of randomized controlled trials. *Int J Rheum Dis.* 2019.[30]

Methods of Sponsor-Submitted ITC

Objectives

The objective of the ITC was to assess the comparative efficacy and safety of upadacitinib with relevant intervention comparators, as listed in Table 24. Populations of interest were:

1. csDMARD inadequate responders
2. bDMARD inadequate responders.

Study Selection Methods

The authors developed a search strategy and conducted the search on three bibliographic databases (MEDLINE, Embase, and Cochrane library), with a manual screen of references of included studies. The inclusion and exclusion criteria focused on a population of csDMARD and bDMARD inadequate responders. The intervention was defined as upadacitinib 15 or 30 mg, and the comparator included any csDMARD, bDMARD, or JAK inhibitor. The criteria only included studies in English and did not have a publication date limit. Two reviewers independently screened the retrieved reports at two stages, with any disagreement adjudicated by a third reviewer. Data extraction was handled by one reviewer, and another reviewer conducted the data check. Quality assessment was carried out using a quality (risk of bias) assessment for all the publications identified in the clinical review and was performed using the seven-criteria checklist provided in section 4.6 of the National Institute for Health and Care Excellence single technology appraisal user guide.[10] The outcomes to be assessed are specified in Table 24.

ITC Analysis Methods

Two main network analyses were planned for csDMARD-experienced and bDMARD-experienced populations. The authors conducted a Bayesian NMA using an ordered multinomial likelihood with an ordered probit link to estimate the probabilities of achieving different levels of ACR response under a random-effects model at the 12 and 24 week time points. The authors reported that the choice of fixed or random-effects model would be based on the measurement of model fit. However, these results were not provided. The

NMA was conducted with noninformative priors, 10,000 burn-in iterations, and three chains each with 20,000 posterior iterations. The probabilities of achieving each level of ACR response were summarized using posterior medians and their associated 95% CrIs. The authors adjusted the response of the csDMARD reference arm as a trial-level covariate, under the assumption that the csDMARD reference arm response rate is an important proxy for both measured and unmeasured patient- and trial-level characteristics that can collectively influence a patient's response to treatment. [redacted] The authors provide further model diagnostics, including network diagram, diagnostic information criterion (DIC), and total residual deviance. However, the authors do not provide a description of methods of assessing inconsistency.

[redacted]

Table 25: ITC Analysis Methods

	csDMARD inadequate responders	bDMARD inadequate responders
ITC methods	Bayesian network meta-analysis	Bayesian network meta-analysis
Priors	Noninformative	Informative
Assessment of model fit	Diagnostic information criterion	Diagnostic information criterion
Assessment of consistency	Not reported	Not reported
Assessment of convergence	Not reported	Not reported
Outcomes	ACR20, 50, and 70 at weeks 12 and 24	ACR20, 50, and 70 at weeks 12 and 24
Follow-up time points	Not reported	Not reported
Construction of nodes	Not reported	Not reported
Sensitivity analyses	Not reported	Not reported
Subgroup analysis	Not reported	Not reported
Methods for pairwise meta-analysis	Not reported	Not reported

ACR = American College of Rheumatology; bDMARD = biologic disease-modifying antirheumatic drug; csDMARD = conventional synthetic disease-modifying antirheumatic drug; ITC = indirect treatment comparison.

Source: CDR submission: TBD (upadacitinib), 15mg once daily oral tablet [CONFIDENTIAL manufacturer's submission]. In: Pointe-Claire (QC): AbbVie Corporation; 2019 Jul 04.[31]

Results of Sponsor-Submitted ITC

Summary of Included Studies

For the patient population of csDMARD inadequate responders, the authors included 50 trials in the week 12 ACR outcome and 55 trials in the week 24 ACR outcome. Across the csDMARD inadequate responder study arms, mean ages ranged from 46 years to 58 years. The proportion of males across the studies ranged from 4% to 26%, and the mean DAS28-ESR baseline scores ranged from 5.0 to 6.9 (data are not shown here but were

provided by the sponsor). The clinical expert consulted on this review did not identify any of the observed differences in baseline characteristics between the included studies as potentially affecting the outcome. The authors reported a quality assessment for each of the included studies but did not provide a summary of the full inclusion criteria. The most common at-risk item was the unclear adequate randomization and allocation concealment.

For the patient population of bDMARD inadequate responders, the authors included 10 studies. Across the bDMARD inadequate responder study arms, the mean ages ranged from 50 years to 56 years. The proportion of males across the studies ranged from 15% to 23%, and the mean DAS28-ESR baseline scores ranged from 5.4 to 6.9 (data are not shown here but were provided by the sponsor). The clinical expert consulted on this review did not identify any of the observed differences in baseline characteristics between the included studies as potentially affecting the outcome. The authors reported quality assessment for each of the included studies but did not provide a summary of the full inclusion criteria. The most common at-risk item was the unclear adequate randomization and allocation concealment.

Results

[redacted]

[redacted]

Table 26: Probability of ACR Response at 12 Weeks: csDMARD Inadequate Responders

Treatment	ACR20		ACR50		ACR70	
	Median ACR20 response probability (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR50 response probability (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR70 response probability (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)
csDMARD	[redacted]					
Abatacept 10 mg/kg + csDMARD	[redacted]					
Abatacept 125 mg + csDMARD	[redacted]					
Adalimumab 40 mg	[redacted]					
Adalimumab 40 mg + csDMARD	[redacted]					

Treatment	ACR20		ACR50		ACR70	
	Median ACR20 response probability (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR50 response probability (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR70 response probability (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)
Baricitinib 2 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baricitinib 4 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Certolizumab 200 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Etanercept 50 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Etanercept 50 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Golimumab 50 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Infliximab 3 mg/kg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Intensive csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sarilumab 200 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tocilizumab 8 mg/kg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tocilizumab 8 mg/kg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tocilizumab 162 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib 5 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib 5 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib 10 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upadacitinib 15 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upadacitinib 30 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Treatment	ACR20		ACR50		ACR70	
	Median ACR20 response probability (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR50 response probability (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR70 response probability (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)
Upadacitinib 15 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upadacitinib 30 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ACR = American College of Rheumatology; CrI = credible interval; csDMARD = conventional synthetic disease-modifying antirheumatic drug; vs. = versus.

Source: CDR submission: TBD (upadacitinib), 15mg once daily oral tablet [CONFIDENTIAL manufacturer's submission]. In: Pointe-Claire (QC): AbbVie Corporation; 2019 Jul 04.[31]

Table 27: ACR Response at 24 Weeks: csDMARD Inadequate Responders

Treatment	ACR20 at 24 weeks		ACR50 at 24 weeks		ACR70 at 24 weeks	
	Median ACR20 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR50 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR70 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)
csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Abatacept 10 mg/kg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Abatacept 125 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adalimumab 40 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adalimumab 40 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baricitinib 2 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baricitinib 4 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Certolizumab 200 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Etanercept 50 mg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Etanercept 50 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Treatment	ACR20 at 24 weeks		ACR50 at 24 weeks		ACR70 at 24 weeks	
	Median ACR20 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR50 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR70 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)
Golimumab 50 mg + csDMARD	[redacted]					
Infliximab 3 mg/kg + csDMARD	[redacted]					
Intensive csDMARD	[redacted]					
Placebo	[redacted]					
Rituximab 2,000 mg + csDMARD	[redacted]					
Sarilumab 150 mg + csDMARD	[redacted]					
Sarilumab 200 mg	[redacted]					
Sarilumab 200 mg + csDMARD	[redacted]					
Sirukumab 50 mg	[redacted]					
Sirukumab 100 mg	[redacted]					
Tocilizumab 8 mg/kg	[redacted]					
Tocilizumab 8 mg/kg + csDMARD	[redacted]					
Tocilizumab 162 mg	[redacted]					
Tocilizumab 162 mg + csDMARD	[redacted]					
Tofacitinib 5 mg	[redacted]					
Tofacitinib 5 mg + csDMARD	[redacted]					
Tofacitinib 10 mg + csDMARD	[redacted]					

Treatment	ACR20 at 24 weeks		ACR50 at 24 weeks		ACR70 at 24 weeks	
	Median ACR20 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR50 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR70 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)
Upadacitinib 15 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ACR = American College of Rheumatology; CrI = credible interval; csDMARD = conventional synthetic disease-modifying antirheumatic drug; vs. = versus.

Source: CDR submission: TBD (upadacitinib), 15mg once daily oral tablet [CONFIDENTIAL manufacturer's submission]. In: Pointe-Claire (QC): AbbVie Corporation; 2019 Jul 04.[31]

Table 28: ACR Response at 12 Weeks: bDMARD Inadequate Responders

Treatment	ACR20 at 12 weeks		ACR50 at 12 weeks		ACR70 at 12 weeks	
	Median ACR20 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR50 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR70 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)
csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Abatacept 10 mg/kg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baricitinib 2 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Baricitinib 4 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Certolizumab 200 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Golimumab 50 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Rituximab 2,000 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sarilumab 150 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sarilumab 200 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sirukumab 50 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Treatment	ACR20 at 12 weeks		ACR50 at 12 weeks		ACR70 at 12 weeks	
	Median ACR20 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR50 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)	Median ACR70 response (95% CrI)	Median odds ratio vs. csDMARD (95% CrI)
Sirukumab 100 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tocilizumab 8 mg/kg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib 5 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tofacitinib 10 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upadacitinib 15 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Upadacitinib 30 mg + csDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ACR = American College of Rheumatology; bDMARD = biologic disease-modifying antirheumatic drug; CrI = credible interval; csDMARD = conventional synthetic disease-modifying antirheumatic drug; vs. = versus.

Source: CDR submission: TBD (upadacitinib), 15mg once daily oral tablet [CONFIDENTIAL manufacturer's submission]. In: Pointe-Claire (QC): AbbVie Corporation; 2019 Jul 04.[31]

Critical Appraisal of Sponsor-Submitted ITC

The authors provided a comprehensive and transparent approach to their systematic review, for which they have provided the search strategy; they conducted the search over several databases, used two independent reviewers for screening, and outlined a comprehensive list of inclusion and exclusion criteria.

The authors divided their population into csDMARD inadequate responders and bDMARD inadequate responders; for the csDMARD inadequate responders, the authors adjusted the response of the csDMARD arms to control for potential heterogeneity. It is not clear why the csDMARD was used as opposed to placebo. Also, the authors did not provide the results of the unadjusted model to allow assessment of fit between the two models. Furthermore, in the bDMARD population, the authors used informative priors with little evidence to justify the choice.

While the authors did provide the diagnostic measure regarding the model fit, the use of these diagnostic measures was limited since the authors only provided the random-effects model and did not provide an analysis using a fixed-effects model.

Lastly, the authors only provided comparison against the csDMARD as a reference arm, thus drastically decreasing the value of the NMA in providing indirect comparison with other active treatment, such as biologics or other JAK inhibitors.

Given the small evidence network, consisting of 10 studies, the values reported in the bDMARD population showed wide CrIs, reflecting high statistical uncertainty.

Methods of Song et al. (2019)

Objectives

The objective of the ITC was to assess the comparative efficacy and safety of upadacitinib with tofacitinib in patients who are csDMARD and/or bDMARD inadequate responders.

Study Selection Methods

The authors developed a search strategy and conducted the search on several bibliographic databases (MEDLINE, Embase, and Cochrane library). The inclusion and exclusion criteria focused on a population of csDMARD and bDMARD inadequate responders. The intervention was defined as tofacitinib and the comparator as upadacitinib 15 or 30 mg. Two reviewers independently screened the retrieved reports at two stages, with any disagreement adjudicated by a third reviewer. Data extraction was handled by two independent reviewers. Quality assessment was carried out using modified Jadad scoring.[32] The authors assessed the outcome of ACR20 at an unspecified time point.

ITC Analysis Methods

The authors conducted a Bayesian NMA. The NMA was conducted with noninformative priors, 10,000 burn-in iterations, and three chains each with 10,000 posterior iterations. The authors reported pairwise OR and 95% CrI and defined a statistically significant finding if the credible interval excluded the 1. The authors also conducted an inconsistency modelling and sensitivity testing by comparing the performance of the random-effects and fixed-effects models. The authors do not provide further model diagnostics, including network diagram, diafnostic information criterion (DIC), or total residual deviance.

It is not clear how the authors handled potential outcome definition differences. The authors also do not provide information about how they handled potential heterogeneity in the included studies doses, background therapy, outcome assessment point, or the definition of patients with RA, or the potential variations in the definition of methotrexate intolerance. Results were considered statistically significant if the 95% CrI did not include the value 1.

Results of Song et al. (2019)

Summary of Included Studies

Nine studies were included in the analysis. The authors do not provide a description of the characteristics of the included studies, the baseline characteristics of the enrolled patients in each study, or the study quality scores.

Results

Efficacy shows upadacitinib 15 mg + methotrexate to have the highest statistically significant OR versus placebo + methotrexate (OR 4.90 [95% CrI, 2.62 to 9.82]). While comparison versus upadacitinib 15 mg + methotrexate and tofacitinib 5 mg + methotrexate showed a higher OR of achieving ACR20, it was not statistically significant (OR 1.52 [95% CrI, 0.64 to 3.26]). See Table 29.

Table 29: League Tables Showing Odds Ratios and 95% Credible Intervals: (A) Efficacy and (B) Safety

A					
Upadacitinib 15 mg + MTX					
1.05 (0.48-2.30)	**Upadacitinib 30 mg + MTX**				
1.27 (0.56-2.85)	1.22 (0.47-3.04)	**Tofacitinib 10 mg + MTX**			
1.52 (0.64-3.26)	1.46 (0.54-3.54)	1.20 (0.70-1.88)	**Tofacitinib 5 mg + MTX**		
1.64 (0.71-3.90)	1.57 (0.57-4.43)	1.29 (0.60-2.92)	1.08 (0.52-2.58)	**Adalimumab + MTX**	
4.90 (2.62-9.82)	4.69 (2.17-10.79)	3.87 (2.44-6.68)	3.23 (2.09-5.91)	2.98 (1.49-6.30)	**Placebo + MTX**
B					
Placebo + MTX					
0.79 (0.25-2.39)	**Adalimumab + MTX**				
0.63 (0.25-1.49)	0.80 (0.22-2.87)	**Tofacitinib 10 mg + MTX**			
0.62 (0.25-1.51)	0.78 (0.22-2.88)	0.98 (0.42-2.37)	**Tofacitinib 5 mg + MTX**		
0.51 (0.15-1.23)	0.65 (0.15-2.06)	0.81 (0.18-2.67)	0.84 (0.18-2.67)	**Upadacitinib 15 mg + MTX**	
0.38 (0.08-1.19)	0.49 (0.08-2.05)	0.61 (0.11-2.52)	0.62 (0.11-2.57)	0.75 (0.22-2.47)	**Upadacitinib 30 mg + MTX**

MTX = methotrexate.

Reproduced with permission from Song GG, Choi SJ, Lee YH. Comparison of the efficacy and safety of tofacitinib and upadacitinib in patients with active rheumatoid arthritis: a Bayesian network meta-analysis of randomized controlled trials. Int J Rheum Dis. 2019; 22(8):1563-1571.© 2019 Asia Pacific League of Associations for Rheumatology and John Wiley & Sons Australia, Ltd.

Critical Appraisal of Song et al. (2019)

Although the authors outlined a systematic approach to the NMA and a planned analysis, the publication lacks essential information to allow the reader to determine potential clinical heterogeneity in the results. The authors do not include information regarding the baseline characteristics of the patients and are not clear about the characteristics of the studies included in the analysis. Thus, it is impossible to determine if methodological or clinical heterogeneity exist. This adds a high degree of uncertainty in the results.

Summary

Two ITCs were discussed in this section, both of which carry several limitations that increase uncertainty in the results. The two ITCs are different in scope and approach to synthesizing data; thus, contrasting the results between the two may not be feasible. However, this uncertainty may have been already reflected statistically through the presence of wide CrIs in most of the results presented through these two ITCs. Interpretation of the results indicates that upadacitinib 15 mg in the csDMARD-experienced population achieves a probability of ACR20 response that is similar to that of other

bDMARDs. The results of upadacitinib versus tofacitinib also do not conclude the superiority of either intervention over the other.

Other Relevant Studies

Long-Term Extension Studies

Each of the five pivotal studies consisted of two periods, with the first ranging from 12 weeks to 48 weeks in duration and the second ranging from 192 weeks to five years. At the time of this review, data up to 48 weeks were available for SELECT-COMPARE, SELECT-MONOTHERAPY, and SELECT-EARLY. SELECT-NEXT and SELECT-BEYOND included data up until week 60. The relevant efficacy and safety outcomes were not available in the interim clinical study reports and thus were derived from the integrated datasets provided by the sponsor.[31] The long-term data have been summarized in the following sections.

Methods

The five included studies were double-blind RCTs (period 1) followed by an open-label extension (SELECT-COMPARE and SELECT-EARLY) or blinded long-term extension (SELECT-MONOTHERAPY, SELECT-NEXT, and SELECT-BEYOND) (period 2). Patients were unblinded in SELECT-COMPARE and SELECT-EARLY after all patients had completed their last visit during period 1. In addition, an unblinded analysis was conducted for regulatory purposes partway through period 1, at week 26 for SELECT-COMPARE and week 24 for SELECT-EARLY. Unblinded interim analyses were also conducted at the end of period 1 of SELECT-MONOTHERAPY (week 14), SELECT-NEXT (week 12), and SELECT-BEYOND (week 24). Patients and sites remained blinded during the interim analyses in all studies.

The long-term efficacy data were presented by each study individually owing to the differences in study design and comparators. The sponsor provided a series of pooled or integrated datasets, which were used to assess long-term safety data. One of the datasets of interest for this review provided data up to week 26 (six months) for methotrexate-controlled studies (i.e., SELECT-EARLY and SELECT-MONOTHERAPY) (Table 40). A long-term dataset, including pooled data from each of the five pivotal studies concerning upadacitinib 15 mg once daily over one year (“any phase III” dataset), was also provided and summarized (Table 41). The dataset including patients who received upadacitinib 15 mg once daily in any of the phase III studies was also presented; patients who switched from adalimumab to upadacitinib in SELECT-COMPARE were removed.

Populations

As the long-term data were derived from the five pivotal studies, the populations at baseline were the same as those reported earlier in this report. A detailed summary of patient disposition through week 60 was not provided; however, it was reported that 78% and 74% of patients in SELECT-NEXT and SELECT-BEYOND, respectively, remained in the study through week 60.[33] For SELECT-COMPARE, SELECT-MONOTHERAPY, and SELECT-EARLY, 87%, 84%, and 80% of patients, respectively, remained in the study through week 48.[33]

Interventions

Patients who initially received placebo in SELECT-COMPARE were switched to upadacitinib 15 mg at week 26. The long-term efficacy outcomes presented here exclude this treatment arm and only apply to patients who received upadacitinib or adalimumab for all of period 1 (48 weeks).

All patients participating in SELECT-MONOTHERAPY received either upadacitinib (15 mg or 30 mg) or methotrexate during period 1 (14 weeks), after which patients receiving methotrexate were switched to upadacitinib 15 mg or 30 mg as well. Only data for patients who were initially randomized to upadacitinib 15 mg up until week 48 have been summarized here. Similarly, patients participating in SELECT-NEXT and SELECT-BEYOND were initially randomized to upadacitinib (15 mg or 30 mg) or placebo, then all patients on placebo switched to receive upadacitinib 15 mg or 30 mg after week 12. Thus, only data for patients who were initially randomized to upadacitinib 15 mg up until week 60 have been summarized for these two studies.

Lastly, patients in SELECT-EARLY were initially randomized to upadacitinib 15 mg or methotrexate, which were continued throughout the 48 weeks of period 1. There was an option for rescue therapy at week 26, which included adding on upadacitinib or methotrexate or optimizing background RA medications. The data presented in this summary are reflective of those patients who received upadacitinib 15 mg or methotrexate without rescue therapy for the duration of period 1.

Background therapy was permitted in all studies. More specifically, initiation of background RA medications (corticosteroids, NSAIDs, and acetaminophen/paracetamol) was permitted beginning at week 26 in SELECT-COMPARE and SELECT-MONOTHERAPY. This also applies to SELECT-NEXT and SELECT-BEYOND but beginning at week 24. Also, the addition or increase of up to two csDMARDs was also permitted. A dietary supplement of folic acid (or equivalent) was also permitted throughout the studies as per the instruction of the investigator.

Outcomes

The efficacy outcomes that have been presented herein correspond to the long-term data available for the outcomes reported in the clinical evidence section of this report.

Statistical Analysis

Type I error rate adjustments were not planned for any of the efficacy outcomes reported in the integrated efficacy analyses. The following outcomes in the SELECT-COMPARE study handled missing data using the nonresponder imputation approach, which categorized visits with missing data as nonresponders for that visit: ACR20/50/70, proportion of patients achieving LDA based on DAS28, proportion of patients achieving clinical remission based on DAS28, and proportion of patients achieving LDA based on CDAI. Patients with early discontinuation were also categorized as nonresponders. All other efficacy outcomes that have been reported for SELECT-COMPARE (except the change from baseline in mTSS) used the last observation carried forward imputation method for patients who received rescue therapy. The last observation carried forward for rescue method was also used for the continuous outcomes reported for SELECT-MONOTHERAPY and SF-36 for SELECT-NEXT. The remainder of the outcomes reported used the as-observed analysis set, which

did not involve imputation for missing data, but rather excluded patients from the as-observed analysis for visits where data were missing.

Exposure to Study Treatments

The mean (SD) number of days exposed to the study drug was not reported for the six-month methotrexate-controlled studies. For the “any phase III” dataset, the mean (SD) number of days exposed to the study drug was 368.7 (172.29) days for all patients and 372.4 (173.88) days for patients who did not cross over and receive adalimumab.[31]

Efficacy

ACR20/50/70

The proportion of patients meeting the ACR20, 50, and 70 response criteria at week 48 or week 60 has been summarized in Table 30, Table 32, Table 34, Table 36, and Table 38. The SELECT-COMPARE study used the nonresponder imputation method to account for missing data and reported 64.7% (95% CI, 61.0 to 68.3), 49.5% (95% CI, 45.6 to 53.3), and 36.1% (95% CI, 32.4 to 39.8) of patients met the criteria for ACR20/50/70, respectively, at week 48. Compared to patients receiving adalimumab, the response rate difference was 10.8 (95% CI, 4.3 to 17.4), 9.7 (95% CI, 3.2 to 16.3), and 12.9 (95% CI, 7.0 to 18.7) for ACR20, 50, and 70, respectively (Table 30). In the SELECT-EARLY study, 91.9% (95% CI, 88.5 to 95.3) of patients met the ACR20 criteria, 79.3% (95% CI, 74.2 to 84.3) met the ACR50 criteria, and 63.3% (95% CI, 57.2 to 69.3) met the ACR70 criteria at week 48; however, the as-observed dataset, which did not account for missing data, was used to describe these results. Also, the corresponding response rate for patients receiving methotrexate was 85.8 (95% CI, 81.2 to 90.5) who met the ACR20 criteria, 65.2 (95% CI, 58.8 to 71.7) who met the ACR50 criteria, and 43.7 (95% CI, 37.0 to 50.3) who met the ACR70 criteria (Table 38).

The three remaining studies used the as-observed dataset to describe their results and reported 87.2% (95% CI, 82.2 to 92.2) met the ACR20 criteria at week 48 (SELECT-MONOTHERAPY, Table 32) and 76.7% (95% CI, 69.5 to 83.9) and 85.0% (95% CI, 79.6 to 90.3) did at week 60 (SELECT-NEXT, Table 34; SELECT-BEYOND, Table 36). Further, 69.5% (95% CI, 62.7 to 76.4) of patients met the ACR50 criteria at week 48 in SELECT-MONOTHERAPY and 52.2% (95% CI, 43.8 to 60.7) and 72.6% (95% CI, 65.9 to 79.4) did at week 60 in SELECT-BEYOND and SELECT-NEXT, respectively. The ACR70 response rate at week 48 was 45.5% (95% CI, 38.1 to 52.8) in SELECT-MONOTHERAPY ,and at week 60, it was between 33.3% (95% CI, 25.4 to 41.3) and 51.5% (95% CI, 44.0 to 59.0) in SELECT-BEYOND and SELECT-NEXT.

LDA and Clinical Remission

The proportion of patients achieving LDA based on DAS28-CRP and CDAI criteria, and the proportion of patients achieving clinical remission based on DAS28-CRP criteria, were reported in all five of the included studies. To summarize, 49.9% (95% CI, 46.1 to 53.8) of patients receiving upadacitinib and 35.2% (95% CI, 30.0 to 40.3) receiving adalimumab achieved LDA (based on DAS28-CRP criteria) at week 48 in SELECT-COMPARE (Table 30). In SELECT-EARLY, 77.6% (95% CI, 72.4 to 82.8) of patients receiving upadacitinib and 61.0% (95% CI, 54.5 to 67.6) receiving methotrexate achieved LDA based on DAS28-CRP criteria (Table 38). The proportion of patients achieving LDA based on CDAI criteria was similar in both studies, with the exception of the group of patients receiving methotrexate in SELECT-EARLY, which reported 67.9% (95% CI, 61.6 to 74.3) with LDA.

Clinical remission based on the DAS28-CRP criteria was achieved at week 48 by 38.2% (95% CI, 34.5 to 42.0) and 27.5% (95% CI, 22.7 to 32.4) of patients treated with upadacitinib and adalimumab, respectively, in SELECT-COMPARE and by 63.4% (95% CI, 57.4 to 69.4) and 44.6% (95% CI, 37.9 to 51.3) of patients treated with upadacitinib and methotrexate in SELECT-EARLY.

At week 48 in SELECT-MONOTHERAPY and week 60 in SELECT-NEXT and SELECT-BEYOND, the proportion of patients receiving upadacitinib who achieved LDA ranged from 69.2% to 74.6% based on the DAS28-CRP criteria, with similar results based on the CDAI criteria. Clinical remission based on DAS28-CRP criteria was similar across the three studies as well, ranging from 52.6% to 59.0% of patients (Table 32, Table 34, and Table 36).

DAS28-CRP, HAQ-DI, SF-36, and Duration of Morning Stiffness

The change from baseline to week 48 or week 60 was reported for DAS28-CRP, HAQ-DI, SF-36, and duration of morning stiffness in all five studies. The results from SELECT-COMPARE reflected a reduction in disability or improvement in health status for each of the outcomes at week 48, in addition to an improvement in the mTSS for patients receiving upadacitinib and patients receiving adalimumab (Table 31). Similarly, the results for patients receiving upadacitinib or methotrexate in SELECT-EARLY reflected a reduction in disability or an improvement in health status based on the change from baseline in DAS28-CRP, HAQ-DI, SF-36, and duration of morning stiffness at week 48 (Table 39).

The change from baseline in DAS28-CRP, HAQ-DI, SF-36, and duration of morning stiffness at week 48 in SELECT-MONOTHERAPY and at week 60 in SELECT-NEXT and SELECT-BEYOND was reported as well and reflected a reduction in disability or an improvement in health status (Table 33, Table 35, and Table 37).

Table 30: Long-Term Efficacy Results (SELECT-COMPARE): Dichotomous Outcomes

	Total N	Responder	Response rate (95% CI)	Response rate difference versus control	
		n		Point estimate (95% CI)	P value
ACR20 response rate at week 48 (NRI, FAS)					
UPA 15 mg q.d.	651	421	64.7 (61.0 to 68.3)	10.8 (4.3 to 17.4)	0.001
ADA 40 mg e.o.w.	327	176	53.8 (48.4 to 59.2)		
ACR50 response rate at week 48 (NRI, FAS)					
UPA 15 mg q.d.	651	322	49.5 (45.6 to 53.3)	9.7 (3.2 to 16.3)	0.005
ADA 40 mg e.o.w.	327	130	39.8 (34.5 to 45.1)		
ACR70 response rate at week 48 (NRI, FAS)					
UPA 15 mg q.d.	651	235	36.1 (32.4 to 39.8)	12.9 (7.0 to 18.7)	< 0.001
ADA 40 mg e.o.w.	327	76	23.2 (18.7 to 27.8)		
Proportion of patients achieving LDA based on DAS28-CRP ≤ 3.2 at week 48 (NRI, FAS)					
UPA 15 mg q.d.	651	325	49.9 (46.1 to 53.8)	14.8 (8.3 to 21.2)	< 0.001
ADA 40 mg e.o.w.	327	115	35.2 (30.0 to 40.3)		
Proportion of patients achieving CR based on DAS28-CRP < 2.6 at week 48 (NRI, FAS)					
UPA 15 mg q.d.	651	249	38.2 (34.5 to 42.0)	10.7 (4.6 to 16.8)	0.001
ADA 40 mg e.o.w.	327	90	27.5 (22.7 to 32.4)		
Proportion of patients achieving LDA based on CDAI ≤ 10 at week 48 (NRI, FAS)					

	Total N	Responder n	Response rate (95% CI)	Response rate difference versus control	
				Point estimate (95% CI)	P value
UPA 15 mg q.d.	651	308	47.3 (43.5 to 51.1)	13.1 (6.6 to 19.5)	< 0.001
ADA 40 mg e.o.w.	327	112	34.3 (29.1 to 39.4)		
Proportion of patients with no radiographic progression, change from baseline at week 48 (LOCF for rescue, FAS)					
UPA 15 mg q.d.	651	522	86.4 (83.7 to 89.2)	–1.5 (–6.1 to 3.1)	0.517
ADA 40 mg e.o.w.	327	262	87.9 (84.2 to 91.6)		

ACR = American College of Rheumatology; ADA = adalimumab; CDAI = Clinical Disease Activity Index; CI = confidence interval; CR = clinical remission; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; e.o.w. = every other week; FAS = full analysis set; LDA = low disease activity; LOCF = last observation carried forward; NRI = nonresponder imputation; q.d. = once daily; UPA = upadacitinib.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Table 31: Long-Term Efficacy Results (SELECT-COMPARE): Continuous Outcomes

	Total N	Baseline	Week 48		Treatment group difference versus control		
		Mean	Mean	LS mean change from baseline (95% CI)	N	LS mean difference (95% CI)	P value
DAS28-CRP, change from baseline at week 48 (LOCF for rescue, FAS)							
UPA 15 mg q.d.	651	5.80	2.99	–2.80 (–2.98 to –2.62)	574	–0.44 (–0.64 to –0.24)	< 0.001
ADA 40 mg e.o.w.	327	5.85	3.46	–2.36 (–2.57 to –2.15)	271		
HAQ-DI, change from baseline at week 48 (LOCF for rescue, FAS)							
UPA 15 mg q.d.	651	1.65	0.86	–0.73 (–0.81 to –0.65)	573	–0.13 (–0.22 to –0.05)	0.002
ADA 40 mg e.o.w.	327	1.65	1.00	–0.60 (–0.69 to –0.51)	272		
SF-36 (physical component score), change from baseline at week 48 (LOCF for rescue, FAS)							
UPA 15 mg q.d.	651	32.33	42.72	9.75 (8.71 to 10.80)	573	1.70 (0.52 to 2.88)	0.005
ADA 40 mg e.o.w.	327	32.05	40.80	8.06 (6.82 to 9.29)	276		
Morning stiffness duration (minutes), change from baseline at week 48 (LOCF for rescue, FAS)							
UPA 15 mg q.d.	651	144.09	37.95	–101.65 (–113.12 to –90.19)	579	–6.15 (–19.00 to 6.69)	0.347
ADA 40 mg e.o.w.	327	142.39	44.06	–95.50 (–109.08 to –81.92)	270		
PGA of pain (VAS, mm), change from baseline at week 48 (LOCF for rescue, FAS)							
UPA 15 mg q.d.	651	66.10	28.37	–36.68 (–39.89 to –33.47)	574	–4.62 (–8.22 to –1.01)	0.012
ADA 40 mg e.o.w.	327	66.71	33.26	–32.07 (–35.86 to –28.28)	272		
mTSS, change from baseline at week 48 (linear extrapolation, FAS)							
UPA 15 mg q.d.	651	33.25	33.40	0.28 (–0.20 to 0.77)	569	–0.11 (–0.74 to 0.51)	0.730
ADA 40 mg e.o.w.	327	33.14	33.40	0.39 (–0.21 to 1.00)	276		

ADA = adalimumab; CI = confidence interval; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; e.o.w. = every other week; FAS = full analysis set; HAQ-DI = Health Assessment Questionnaire–Disability Index; LOCF = last observation carried forward; LS = least squares; mTSS = modified total Sharp score; PGA = patient global assessment; q.d. = once daily; SF-36 = Short Form (36) Health Survey; UPA = upadacitinib; VAS = visual analogue scale.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Table 32: Long-Term Efficacy Results (SELECT-MONOTHERAPY): Dichotomous Outcomes

	Total N[a]	Responder	Response rate (95% CI)
		n	
ACR20 response rate at week 48 (AO, FAS)			
UPA 15 mg q.d.	172	150	87.2 (82.2 to 92.2)
ACR50 response rate at week 48 (AO, FAS)			
UPA 15 mg q.d.	174	121	69.5 (62.7 to 76.4)
ACR70 response rate at week 48 (AO, FAS)			
UPA 15 mg q.d.	176	80	45.5 (38.1 to 52.8)
Proportion of patients achieving LDA based on DAS28-CRP ≤ 3.2 at week 48 (AO, FAS)			
UPA 15 mg q.d.	174	126	72.4 (65.8 to 79.1)
Proportion of patients achieving CR based on DAS28-CRP < 2.6 at week 48 (AO, FAS)			
UPA 15 mg q.d.	174	96	55.2 (47.8 to 62.6)
Proportion of patients achieving LDA based on CDAI ≤ 10 at week 48 (AO, FAS)			
UPA 15 mg q.d.	172	123	71.5 (64.8 to 78.3)

ACR = American College of Rheumatology; AO = as observed; CDAI = Clinical Disease Activity Index; CI = confidence interval; CR = clinical remission; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; FAS = full analysis set; LDA = low disease activity; q.d. = once daily; UPA = upadacitinib.

[a] Total N represents the number of patients included in the analysis of change from baseline to week 48. A total of 217 patients were randomized to UPA 15 mg q.d. at baseline.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Table 33: Long-Term Efficacy Results (SELECT-MONOTHERAPY): Continuous Outcomes

	Total N[a]	Baseline	Week 48	
		Mean	Mean	Mean change from baseline (95% CI)
DAS28-CRP, change from baseline at week 48 (LOCF for rescue, FAS)				
UPA 15 mg q.d.	173	5.61	2.64	–2.97 (–3.17 to –2.77)
HAQ-DI, change from baseline at week 48 (LOCF for rescue, FAS)				
UPA 15 mg q.d.	173	1.47	0.74	–0.73 (–0.83 to –0.62)
SF-36 (physical component score), change from baseline at week 48 (LOCF for rescue, FAS)				
UPA 15 mg q.d.	168	33.64	44.13	10.49 (9.12 to 11.87)
Morning stiffness duration (minutes), change from baseline at week 48 (LOCF for rescue, FAS)				
UPA 15 mg q.d.	175	151.3	36.6	–114.6 (–147.5 to –81.8)

CI = confidence interval; CRP = C-reactive protein; DAS28 = Disease Activity Score; FAS = full analysis set; HAQ-DI = Health Assessment Questionnaire–Disability Index; LOCF = last observation carried forward; q.d. = once daily; SF-36 = Short Form (36) Health Survey; UPA = upadacitinib.

[a] Total N represents the number of patients included in the analysis of change from baseline to week 48. A total of 217 patients were randomized to UPA 15 mg q.d. at baseline.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Table 34: Long-Term Efficacy Results (SELECT-NEXT): Dichotomous Outcomes

	Total N[a]	Responder	Response rate (95% CI)
		n	
ACR20 response rate at week 60 (AO, FAS)			
UPA 15 mg q.d.	173	147	85.0 (79.6 to 90.3)
ACR50 response rate at week 60 (AO, FAS)			
UPA 15 mg q.d.	168	122	72.6 (65.9 to 79.4)
ACR70 response rate at week 60 (AO, FAS)			
UPA 15 mg q.d.	171	88	51.5 (44.0 to 59.0)
Proportion of patients achieving LDA based on DAS28-CRP ≤ 3.2 at week 60 (AO, FAS)			
UPA 15 mg q.d.	173	129	74.6 (68.1 to 81.1)
Proportion of patients achieving CR based on DAS28-CRP < 2.6 at week 60 (AO, FAS)			
UPA 15 mg q.d.	173	102	59.0 (51.6 to 66.3)
Proportion of patients achieving LDA based on CDAI ≤ 10 at week 60 (AO, FAS)			
UPA 15 mg q.d.	171	129	75.4 (69.0 to 81.9)

ACR = American College of Rheumatology; AO = as observed; CDAI = Clinical Disease Activity Index; CI = confidence interval; CR = clinical remission; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; FAS = full analysis set; LDA = low disease activity; q.d. = once daily; UPA = upadacitinib.

[a] Total N represents the number of patients included in the analysis of change from baseline to week 48. A total of 221 patients were randomized to UPA 15 mg q.d. at baseline.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Table 35: Long-Term Efficacy Results (SELECT-NEXT): Continuous Outcomes

	Total N[a]	Baseline	Week 48/60	
		Mean	Mean	Mean change from baseline (95% CI)
DAS28-CRP, change from baseline at week 60 (AO, FAS)				
UPA 15 mg q.d.	169	5.68	2.53	–3.15 (–3.37 to –2.93)
HAQ-DI, change from baseline at week 60 (AO, FAS)				
UPA 15 mg q.d.	172	1.48	0.66	–0.83 (–0.93 to –0.72)
SF-36 (physical component score), change from baseline at week 48 (LOCF for rescue, FAS)				
UPA 15 mg q.d.	179	33.45	44.12	10.67 (9.45 to 11.89)
Morning stiffness duration (minutes), change from baseline at week 60 (AO, FAS)				
UPA 15 mg q.d.	175	136.6	36.1	–100.5 (–127.8 to –73.1)
Fatigue using the FACIT-F, change from baseline at week 48 (AO, FAS)				
UPA 15 mg q.d.	175	27.6	38.6	11.0 (9.5 to 12.5)

AO = as observed; CI = confidence interval; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue; FAS = full analysis set; HAQ-DI = Health Assessment Questionnaire–Disability Index; LOCF = last observation carried forward; q.d. = once daily; SF-36 = Short Form (36) Health Survey; UPA = upadacitinib.

[a] Total N represents the number of patients included in the analysis of change from baseline to week 48. A total of 221 patients were randomized to UPA 15 mg q.d. at baseline.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Table 36: Long-Term Efficacy Results (SELECT-BEYOND): Dichotomous Outcomes

	Total N[a]	Responder	Response rate (95% CI)
		n	
ACR20 response rate at week 60 (AO, FAS)			
UPA 15 mg q.d.	133	102	76.7 (69.5 to 83.9)
ACR50 response rate at week 60 (AO, FAS)			
UPA 15 mg q.d.	134	70	52.2 (43.8 to 60.7)
ACR70 response rate at week 60 (AO, FAS)			
UPA 15 mg q.d.	135	45	33.3 (25.4 to 41.3)
Proportion of patients achieving LDA based on DAS28-CRP ≤ 3.2 at week 60 (AO, FAS)			
UPA 15 mg q.d.	133	92	69.2 (61.3 to 77.0)
Proportion of patients achieving CR based on DAS28-CRP < 2.6 at week 60 (AO, FAS)			
UPA 15 mg q.d.	133	70	52.6 (44.1 to 61.1)
Proportion of patients achieving LDA based on CDAI ≤ 10 at week 60 (AO, FAS)			
UPA 15 mg q.d.	132	91	68.9 (61.0 to 76.8)

ACR = American College of Rheumatology; AO = as observed; CDAI = Clinical Disease Activity Index; CI = confidence interval; CR = clinical remission; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; FAS = full analysis set; LDA = low disease activity; q.d. = once daily; UPA = upadacitinib.

[a] Total N represents the number of patients included in the analysis of change from baseline to week 48. A total of 165 patients were randomized to UPA 15 mg q.d. at baseline.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Table 37: Long-Term Efficacy Results (SELECT-BEYOND): Continuous Outcomes

	Total N[a]	Baseline	Week 48/60	
		Mean	Mean	Mean change from baseline (95% CI)
DAS28-CRP, change from baseline at week 60 (AO, FAS)				
UPA 15 mg q.d.	132	5.88	2.78	–3.10 (–3.30 to –2.90)
HAQ-DI, change from baseline at week 60 (AO, FAS)				
UPA 15 mg q.d.	133	1.66	1.14	–0.52 (–0.62 to –0.42)
SF-36 (physical component score), change from baseline at week 48 (AO, FAS)				
UPA 15 mg q.d.	134	30.90	39.74	8.83 (7.30 to 10.37)
Morning stiffness duration (minutes), change from baseline at week 60 (AO, FAS)				
UPA 15 mg q.d.	111	131.2	59.9	–71.3 (–100.4 to –42.2)

AO = as observed; CI = confidence interval; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; FAS = full analysis set; HAQ-DI = Health Assessment Questionnaire–Disability Index; q.d. = once daily; SF-36 = Short Form (36) Health Survey; UPA = upadacitinib.

[a] Total N represents the number of patients included in the analysis of change from baseline to week 48. A total of 165 patients were randomized to UPA 15 mg q.d. at baseline.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Table 38: Long-Term Efficacy Results (SELECT-EARLY): Dichotomous Outcomes

	Total N[a]	Responder	Response rate (95% CI)
		n	
ACR20 response rate at week 48 (AO, FAS)			
UPA 15 mg q.d.	247	227	91.9 (88.5 to 95.3)
MTX	212	182	85.8 (81.2 to 90.5)
ACR50 response rate at week 48 (AO, FAS)			
UPA 15 mg q.d.	246	195	79.3 (74.2 to 84.3)
MTX	210	137	65.2 (58.8 to 71.7)
ACR70 response rate at week 48 (AO, FAS)			
UPA 15 mg q.d.	245	155	63.3 (57.2 to 69.3)
MTX	213	93	43.7 (37.0 to 50.3)
Proportion of patients achieving LDA based on DAS28-CRP ≤ 3.2 at week 48 (AO, FAS)			
UPA 15 mg q.d.	246	191	77.6 (72.4 to 82.8)
MTX	213	130	61.0 (54.5 to 67.6)
Proportion of patients achieving CR based on DAS28-CRP < 2.6 at week 48 (AO, FAS)			
UPA 15 mg q.d.	246	156	63.4 (57.4 to 69.4)
MTX	213	95	44.6 (37.9 to 51.3)
Proportion of patients achieving LDA based on CDAI ≤ 10 at week 48 (AO, FAS)			
UPA 15 mg q.d.	244	192	78.7 (73.6 to 83.8)
MTX	209	142	67.9 (61.6 to 74.3)

ACR = American College of Rheumatology; AO = as observed; CDAI = Clinical Disease Activity Index; CI = confidence interval; CR = clinical remission; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; FAS = full analysis set; LDA = low disease activity; MTX = methotrexate; q.d. = once daily; UPA = upadacitinib.

Note: The results presented in this table only include data for patients who did not switch treatments.

[a] Total N represents the number of patients included in the analysis of change from baseline to week 48. A total of 317 patients were randomized to UPA 15 mg q.d., and 315 to MTX, at baseline.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Table 39: Long-Term Efficacy Results (SELECT-EARLY): Continuous Outcomes

	Total N[a]	Baseline	Week 48	
		Mean	Mean	Mean change from baseline (95% CI)
DAS28-CRP, change from baseline at week 48 (AO, FAS)				
UPA 15 mg q.d.	246	5.91	2.35	–3.56 (–3.72 to –3.39)
MTX	213	5.92	2.92	–3.00 (–3.18 to –2.82)
HAQ-DI, change from baseline at week 48 (AO, FAS)				
UPA 15 mg q.d.	246	1.58	0.54	–1.03 (–1.13 to –0.94)
MTX	213	1.58	0.74	–0.84 (–0.94 to –0.74)
SF-36 (physical component score), change from baseline at week 48 (AO, FAS)				
UPA 15 mg q.d.	246	32.84	46.10	13.26 (12.07 to 14.46)
MTX	216	33.31	43.35	10.04 (8.84 to 11.25)
Morning stiffness duration (minutes), change from baseline at week 48 (AO, FAS)				
UPA 15 mg q.d.	247	160.9	19.8	–141.0 (–165.4 to –116.7)
MTX	213	131.2	44.6	–86.6 (–107.0 to –66.1)

	Total N[a]	Baseline	Week 48	
		Mean	Mean	Mean change from baseline (95% CI)
Fatigue using the FACIT-F, change from baseline at week 48 (AO, FAS)				
UPA 15 mg q.d.	246	27.2	38.8	11.6 (10.1 to 13.2)
MTX	216	27.6	37.5	9.9 (8.4 to 11.4)

AO = as observed; CI = confidence interval; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue; FAS = full analysis set; HAQ-DI = Health Assessment Questionnaire–Disability Index; MTX = methotrexate; q.d. = once daily; SF-36 = Short Form (36) Health Survey; UPA = upadacitinib.

Note: The results presented in this table only include data for patients who did not switch treatments.

[a] Total N represents the number of patients included in the analysis of change from baseline to week 48. A total of 317 patients were randomized to UPA 15 mg q.d., and 315 to MTX, at baseline.

Source: Integrated Summary of Efficacy Clinical Study Report.[31]

Harms

Methotrexate-Controlled Studies, Harms Data up to Six Months

The pooled harms data up to week 26 (six months) from the methotrexate-controlled studies, SELECT-EARLY and SELECT-MONOTHERAPY, have been summarized in Table 40. Of those patients receiving upadacitinib, 64.4% reported at least one adverse event, 5.8% reported a serious adverse event, 5.1% stopped receiving treatment owing to adverse events, and three patients died. For patients receiving methotrexate, 59.8% of patients reported an adverse event, 4.0% reported serious adverse events, 4.7% stopped treatment owing to adverse events, and one patient died. None of the individual notable harms were reported in more than 4.7% of patients receiving either treatment, and there were no major differences between groups, although the percentage of patients with herpes zoster infection and neutropenia was higher with upadacitinib than with methotrexate by a difference of 1.8% for both events.

Table 40: Long-Term Safety Outcomes: Methotrexate-Controlled Upadacitinib 15 mg, at Six Months (Data From SELECT-EARLY and SELECT-MONOTHERAPY)

	MTX-controlled upadacitinib 15 mg	
	UPA 15 mg q.d. N = 534	MTX N = 530
Patients with ≥ 1 adverse event, n (%)	[redacted]	[redacted]
Patients with ≥ 1 SAE, n (%)	[redacted]	[redacted]
Patients who stopped treatment owing to adverse events, n (%)	[redacted]	[redacted]
Deaths, n (%)	[redacted]	[redacted]
Notable harms, n (%)		
Herpes zoster infection	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]
Lymphopenia	[redacted]	[redacted]
Thrombocytopenia	[redacted]	[redacted]
Malignancy (any)	[redacted]	[redacted]
Thrombosis (incl. increased platelets)[a]	[redacted]	[redacted]
MACE[b]	[redacted]	[redacted]
GI perforations	[redacted]	[redacted]
Hepatic disorder	[redacted]	[redacted]

	MTX-controlled upadacitinib 15 mg	
	UPA 15 mg q.d. N = 534	MTX N = 530
Dyslipidemia	[redacted]	[redacted]

GI = gastrointestinal; incl. = including; MACE = major adverse cardiovascular event; MTX = methotrexate; q.d. = once daily; SAE = serious adverse event; UPA = upadacitinib.

[a] Venous thromboembolic events include deep vein thrombosis and pulmonary embolism.

[b] Cardiovascular death, nonfatal myocardial infarction, and nonfatal stroke.

Source: Integrated Summary of Safety Clinical Study Report.[31]

Pooled Upadacitinib 15 mg Dataset, up to One Year of Exposure

The pooled harms data, which included data from patients treated with upadacitinib 15 mg for up to one year, in any of the five studies, have been summarized in Table 41. In summary, 2,630 patients were included. Of the 2,630 patients, 73.8% reported experiencing an adverse event, 9.0% experienced a serious adverse event, 5.8% stopped treatment owing to adverse events, and 12 died. [redacted]

The results were similar when the 159 patients who were initially treated with adalimumab and then switched to upadacitinib were excluded (the "no adalimumab crossover" group).

Table 41: Long-Term Safety Outcomes: Any Phase III Upadacitinib 15 mg Analysis Set, One-Year Exposure

	UPA 15 mg q.d. N = 2,630	UPA 15 mg q.d., no ADA crossover[a] N = 2,471
Patients with ≥ 1 adverse event, n (%)	[redacted]	[redacted]
Patients with ≥ 1 SAE, n (%)	[redacted]	[redacted]
Patients who stopped treatment owing to adverse events, n (%)	[redacted]	[redacted]
Deaths, n (%)	[redacted]	[redacted]
Notable harms, n (%)		
Herpes zoster infection	[redacted]	[redacted]
Neutropenia	[redacted]	[redacted]
Lymphopenia	[redacted]	[redacted]
Thrombocytopenia	[redacted]	[redacted]
Malignancy (any)	[redacted]	[redacted]
Thrombosis (incl. increased platelets)[b]	[redacted]	[redacted]
MACE[c]	[redacted]	[redacted]
GI perforations	[redacted]	[redacted]
Hepatic disorder	[redacted]	[redacted]
Dyslipidemia	[redacted]	[redacted]

ADA = adalimumab; GI = gastrointestinal; incl. = including; MACE = major adverse cardiovascular event; q.d. = once daily; SAE = serious adverse event; UPA = upadacitinib.

[a] No ADA crossover: Patients who switched from adalimumab to upadacitinib 15 mg q.d. in Study M14-465 were excluded.

[b] Venous thromboembolic events include deep vein thrombosis and pulmonary embolism.

[c] Cardiovascular death, nonfatal myocardial infarction, and nonfatal stroke.

Source: Integrated Summary of Safety Clinical Study Report.[31]

Critical Appraisal

The long-term data associated with each of the five pivotal trials provided an overview of the efficacy of upadacitinib over a 48-week or 60-week period and the safety of upadacitinib over a period of up to one year. Overall, efficacy was maintained during this time period, and there were no major safety signals to report; however, the data are subject to certain limitations. All the studies were initially double blind, but patients in SELECT-EARLY and SELECT-COMPARE were unblinded after the last visit of period 1 (48 weeks) was complete. Moreover, and as noted in the methods section, an unblinded analysis (sites and patients remained blinded) was conducted partway through period 1 for SELECT-EARLY and SELECT-COMPARE and at the end of period 1 for SELECT-MONOTHERAPY, SELECT-NEXT, and SELECT-BEYOND. The unblinding of investigators may introduce bias into the analysis of long-term data, or patient evaluation, which is a major limitation of these studies. This is particularly true for the ACR criteria, which are largely composed of subjective measures, such as the patient and physician global assessment of disease activity.

The conclusions of the long-term efficacy and safety outcomes are also limited by a lack of comparator in SELECT-NEXT, SELECT-BEYOND, and SELECT-MONOTHERAPY. In addition, only descriptive statistics were provided, and any statistical testing that was performed was not included in the statistical hierarchy and is thus subject to the type I error rate. Missing data were only accounted for in SELECT-COMPARE, which used a nonresponder imputation method to label missing data as nonresponders. The amount of missing data in the remaining four studies that was unaccounted for ranged from 18.2% to 33.7%, which is another limitation. Consequently, the treatment effect of upadacitinib in patients in SELECT-COMPARE was smaller than in the other four studies, which used nonimputed, observed data. It is uncertain whether this difference in treatment effect is due to the conservative nature of the nonresponder imputation or an overestimation of the results by not imputing for missing data. Patient disposition was also not available.

The generalizability of the long-term safety and efficacy results is similar to what was outlined earlier in this report. The patients included in these studies represent a subset of the patients seen in Canadian clinical practice. Background medication and optimization of background RA treatments such as NSAIDs, corticosteroids, or low-potency analgesics were permitted, which is applicable to the Canadian context; however, these treatments should be taken into consideration when assessing the treatment effect of upadacitinib at week 48 in all studies.

Discussion

Summary of Available Evidence

Five pivotal multinational double-blind RCTs met the criteria for this systematic review: SELECT-COMPARE, SELECT-MONOTHERAPY, SELECT-NEXT, SELECT-BEYOND, and SELECT-EARLY. All five studies enrolled adults with adult-onset RA, and in all but the EARLY trial (where patients were methotrexate naive), the patients' symptoms had been inadequately controlled by a DMARD: In COMPARE (N = 1,629) patients had to be inadequately controlled on methotrexate, in NEXT (N = 661) patients had to be inadequately controlled on any csDMARD, and in BEYOND (N = 499) patients had to be inadequately controlled on bDMARD. The five included studies reflected patients with RA who had received various prior treatments and had switched to different regimens. In

EARLY, patients could not have been considered as methotrexate inadequate responders and could not have been on any DMARD beside methotrexate for longer than three weeks. Patients were randomized to receive upadacitinib 15 mg once daily, upadacitinib 30 mg once daily, or methotrexate (7.5 or 10 mg) once weekly. In MONOTHERAPY, patients had had a csDMARD experience and were randomized to receive upadacitinib 15 mg once daily, upadacitinib 30 mg once daily, or methotrexate (to continue on the prior-to-enrolment stable dose) with no other added therapy in the background. In COMPARE, patients had been considered as csDMARD inadequate responders but not bDMARD inadequate responders and moved into the trial with a background therapy of methotrexate; they were randomized (2:1:2) into upadacitinib 15 mg once daily, adalimumab 40 mg injection every other week, or placebo. In NEXT, patients had been considered csDMARD inadequate responders but not bDMARD inadequate responders and moved into the study with csDMARD background therapy; they were randomized into upadacitinib 15 mg once daily, upadacitinib 30 mg once daily, or placebo. In BEYOND, patients had been considered bDMARD inadequate responders and moved into the study with csDMARD background therapy; they were randomized into upadacitinib 15 mg once daily, upadacitinib 30 mg once daily, or placebo. The primary outcome in the COMPARE, NEXT, and BEYOND studies was the proportion of patients achieving an ACR20 response at 12 weeks; the primary outcome in MONOTHERAPY was ACR20 at 14 weeks, and the primary outcome in EARLY was ACR50 at 12 weeks (N = 947). Key secondary outcomes that were accounted for type I error rate included HRQoL on the HAQ-DI, the DAS28-CRP, and the mTSS. The primary outcome was reported at 12 weeks in all studies, except MONOTHERAPY, in which it was reported at week 14. Each of these studies has an ongoing long-term extension study.

The primary outcome in EARLY of ACR50 at week 12 showed a response rate of 52.1% (95% CI, 46.6 to 57.5) in the upadacitinib and 28.3% (95% CI, 23.4 to 33.3) in the methotrexate arms, with a rate difference of 23.7 (95% CI, 16.3 to 31.1). The primary outcome in MONOTHERAPY of ACR20 at week 14 showed a response rate of 67.7% (95% CI, 61.5 to 74.0) in the upadacitinib and 41.2% (95% CI, 34.6 to 47.8) in the methotrexate arms, with a rate difference of 26.5 (95% CI, 17.5 to 35.6). The primary outcome in COMPARE of ACR20 at week 12 showed a response rate of 70.5% (95% CI, 67.0 to 74.0) in the upadacitinib, 63.0% (95% CI, 57.8 to 68.2) in the adalimumab, and 36.4% (95% CI, 32.7 to 40.1) in the placebo arms, with a rate difference of 7.5 (95% CI, 1.2 to 13.8) versus adalimumab and 34.1 (95% CI, 29.0 to 39.2) versus placebo. The primary outcome in NEXT of ACR20 at week 12 showed a response rate of 63.8% (95% CI, 57.5 to 70.1) in the upadacitinib and 35.7% (95% CI, 29.4 to 42.1) in the placebo groups, with a rate difference of 28.1 (95% CI, 19.1 to 37.0). In BEYOND, the primary outcome of ACR20 at week 12 showed a response rate of 64.6% (95% CI, 57.3 to 72.0) in the upadacitinib and 28.4% (95% CI, 21.6 to 35.2) in the placebo groups, with a rate difference of 36.2 (95% CI, 26.2 to 46.2).

The HAQ-DI was a secondary outcome in all studies. Results of upadacitinib 15 mg versus the placebo and methotrexate groups achieved a statistically significant magnitude of difference greater than the estimated minimal important difference of 0.22 in all the studies at week 12 (week 14 in MONOTHERAPY). The DAS28-CRP was a secondary outcome in all studies. Results of upadacitinib 15 mg versus the placebo and methotrexate groups showed a statistically significant mean difference in the results in all the studies. The mTSS was reported in COMPARE (week 24) and EARLY (week 26) as a secondary outcome, with a responder analysis, in which a “responder” was defined as having no change in mTSS. In both studies, the mean difference was statistically significantly in favour of upadacitinib and, similarly, the response rate in upadacitinib-treated patients was statistically significantly

higher than in placebo-treated patients (COMPARE, rate difference 7.5 [95% CI, 3.0 to 12.1]) and in methotrexate-treated patients (EARLY, rate difference 9.8 [95% CI, 3.5 to 16.2]). Other secondary outcomes included the proportion of patients with LDA, SF-36, and morning stiffness. Overall, these results are consistent in showing the benefit of upadacitinib 15 mg once daily over placebo or methotrexate control arms. Each of these studies has an ongoing long-term extension phase. In addition, two ITCs were included in this review.

One of the comparisons provided in the studies was that of upadacitinib versus adalimumab in the COMPARE study. The sponsor's initial outcome for this comparison was to achieve a noninferiority on ACR50 at week 12; the result showed upadacitinib to be statistically superior to adalimumab (rate difference 16.1 [95% CI, 9.9 to 22.3]). In addition, upadacitinib was superior to adalimumab in the HAQ-DI measure, but the treatment difference did not exceed the identified minimal important difference of 0.22 points (least squares mean difference –0.11 [95% CI, –0.184 to –0.036]). Other outcomes beyond ACR50, HAQ-DI, and the patient's assessment of pain comparing upadacitinib 15 mg to adalimumab were outside the statistical testing hierarchy. However, upadacitinib showed better results than adalimumab in all examined outcomes except in mTSS.

The main limitations of the pivotal studies include an imbalance in the discontinuation rate between groups in the EARLY study, as well as a disproportional discontinuation rate between adalimumab and upadacitinib in the COMPARE study. As well, the EARLY study included treatment-naive patients, which is not aligned with the Health Canada indication for the treatment of adults with moderately to severely active RA who have had an inadequate response or intolerance to methotrexate. Additionally, although the trials' inclusion and exclusion criteria, as well as the baseline demographic characteristics of the enrolled patients, are generally in line with other RA clinical trials, according to the clinical expert the enrolled patients in these trials do not represent the majority of the patients present in clinical practice. Specifically, the high ratio of female, white, and rheumatoid factor–positive patients may not be representative of Canadian patients. However, there is no clear evidence that these factors will lead to variation in the response to the treatment. Limitations in the ITCs include wide CrIs where there is statistical uncertainty and lack of reporting key items to allow the reader to judge methodological and clinical heterogeneity.

Interpretation of Results

Efficacy

The five included studies reflected patients with RA who have received various prior treatments and have switched to different regimens. In EARLY, patients could not have been considered as methotrexate inadequate responders and could not have been on any DMARD beside methotrexate for longer than three weeks. The primary end point in all five studies demonstrated a statistically significant clinical benefit over placebo and methotrexate on the ACR20 responders' rate, which ranged between 24% and 36%. This benefit over the placebo and methotrexate control groups was consistent in the results of secondary outcomes. Also, a primary outcome of one study (COMPARE) versus adalimumab demonstrated a statistically significant clinical benefit of upadacitinib in the ACR50 response rate at a magnitude of 16.1%. These results can be interpreted as upadacitinib providing a meaningful improvement in patients with RA with various histories of treatment when compared to treatment with methotrexate or placebo with csDMARD background. The COMPARE pivotal study also suggests that upadacitinib provides greater

benefit to patients with RA who are csDMARD inadequate responders than does adalimumab 40 mg every other week, with uncertainties regarding the benefit of radiographic progression.

Indirect evidence from two ITCs indicates that upadacitinib 15 mg in a csDMARD-experienced population achieves a probability of ACR20 response that is more efficacious than treatment with csDMARD. However, because of the lack of reported comparative estimates of upadacitinib versus bDMARDs in the sponsor's ITC, we cannot estimate the magnitude of treatment difference between these agents, despite that, numerically, upadacitinib-treated patients show similar to better probability of achieving ACR20 than those treated with other bDMARDs. Another published ITC (Song et al. 2019) attempted to compare upadacitinib to tofacitinib; the report suggested that neither agent was superior to the other and that they appear to be numerically similar to other bDMARDs. Results of upadacitinib versus tofacitinib also do not show the superiority of either intervention over the other.

A limitation in the evidence is the lack of direct or indirect comparisons and estimated results versus baricitinib and the lack of availability of direct or high-quality indirect comparisons versus tofacitinib. The indirect comparison results available for upadacitinib versus tofacitinib are based on one ITC by Song et al. 2019, which carries high uncertainty due to missing key reporting items. The sponsor-submitted ITC did not provide the estimate results for an indirect comparison between upadacitinib and tofacitinib or baricitinib and only provided the probability for each to achieve ACR20/50/70 results; the comparative estimates were relative to csDMARD.

Additional follow-up and extension study data were provided. These results are from interim analyses, as the extension studies are still ongoing. The evidence of the long-term extension studies is challenging to interpret owing to the lack of control and descriptive nature of the results.

Subgroup analysis for the primary outcome did not provide results that are drastically different to the base-case analysis or suggest potential subgroups that would benefit from upadacitinib in a different manner than the general study population.

Harms

In COMPARE, 64.2% of upadacitinib patients, 60.2% of adalimumab patients, and 53.2% of placebo patients experienced an adverse event. In MONOTHERAPY, the percentages were 47.5% in upadacitinib and 47.2% in placebo groups. In NEXT, the percentages were 56.6% in upadacitinib and 48.9% in placebo groups. In BEYOND, the percentages were 55.5% in upadacitinib and 56.2% in placebo groups. In EARLY, the percentages were 64.0% in upadacitinib and 65.3% in placebo groups. Respiratory tract infections were the most common adverse events in all the included studies. Serious adverse events generally occurred in less than 5% of cases across the studies. In COMPARE, 3.7% of upadacitinib patients, 4.3% of adalimumab patients, and 2.9% of placebo patients experienced a serious adverse event. In MONOTHERAPY, the percentages were 5.1% in upadacitinib and 2.8% in methotrexate groups. In NEXT, the percentages were 4.1% in upadacitinib and 2.3% in placebo groups. In BEYOND, the percentages were 4.9% in upadacitinib and 0% in placebo groups. In EARLY, the percentages were 4.7% in upadacitinib and 4.1% in methotrexate groups. No single serious adverse event was most common across the five included studies.

According to the available data, there was a numerically higher incidence of herpes zoster infection in the upadacitinib treatment groups when contrasted to non-upadacitinib treatment groups. Over the course of the efficacy and extension phases of the studies, malignancies were reported in 1.1% of all patients who started and stayed on upadacitinib and 1.2% of patients who switched over to upadacitinib for the extension phase. Overall, notable harms identified for this review did not show explicit imbalance between groups, with the exception of a numerically higher proportion of neutropenia in COMPARE, BEYOND, and EARLY. Also, there was no explicit imbalance in the number of thromboembolic events between upadacitinib-treated patients and other groups.

Conclusions

The included studies showed that upadacitinib at 15 mg, orally once daily, after 12 weeks to 14 weeks of treatment improved clinical response in terms of the ACR20 outcome compared to placebo (two studies) and methotrexate (two studies) and in terms of the ACR50 outcome compared to adalimumab (one study) in a population of patients with RA who had either an inadequate response to csDMARDs (three studies), an inadequate response to bDMARDs (one study), or an undetermined response to either (one study). Three out of five studies had patients on either methotrexate or other csDMARD background therapy, alone or in combination. One study, BEYOND, specifically recruited patients who had failed bDMARD therapy. In all five studies, there was a statistically significant improvement in HRQoL and in disease activity with upadacitinib versus methotrexate or placebo groups, with a difference larger than the minimal important difference in the HAQ-DI outcome. Upadacitinib showed better treatment outcomes than adalimumab in disease activity measures and HRQoL measures but not in radiographic progression. The benefit of upadacitinib versus other JAK inhibitors remains uncertain owing to the lack of direct or indirect comparative estimates versus baricitinib and the indirect evidence versus tofacitinib, which carries a large degree of uncertainty. According to the indirect evidence, upadacitinib is likely at least as efficacious as other bDMARDs, but because of the lack of reported comparative estimates versus bDMARDs, no magnitude of treatment difference between upadacitinib and other bDMARDs could be reported. The risk of notable harms such as serious infections, malignancies, cardiovascular events, dyslipidemia, and elevated hepatic enzymes did not appear to differ between upadacitinib and placebo, although the included studies were not designed to assess outcomes such as these. Long-term extension studies are ongoing.

Appendix 1: Literature Search Strategy

Clinical Literature Search

OVERVIEW	
Interface:	Ovid
Databases:	MEDLINE All (1946–present) Embase (1974–present) **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	August 01, 2019
Alerts:	Biweekly search updates until project completion
Study Types:	All study types
Limits:	Conference abstracts: excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
MeSH	Medical Subject Heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
.ti	Title
.ab	Abstract
.hw	Heading word; usually includes subject headings and controlled vocabulary
.kf	Author keyword heading word (MEDLINE)
.kw	Author keyword (Embase)
.pt	Publication type
.ot	Original title (MEDLINE)
.rn	Registry number
.dq	Candidate term word (Embase)
medall	Ovid database code: MEDLINE All, 1946 to present, updated daily
oemezd	Ovid database code; Embase, 1974 to present, updated daily

MULTI-DATABASE STRATEGY	
Line #	**Search Strategy**
1	(upadacitinib* or ABT 494 or ABT494 or 4RA0KN46E0).ti,ab,kf,ot,hw,rn,nm
2	1 use medall
3	*upadacitinib/
4	(upadacitinib* or ABT 494 or ABT494).ti,ab,kw,dq
5	or/3-4
6	5 use oemezd
7	6 not conference abstract.pt.
8	2 or 7
9	remove duplicates from 8

CLINICAL TRIAL REGISTRIES	
ClinicalTrials.gov	Produced by the US National Library of Medicine. Targeted search used to capture registered clinical trials. [Search: Studies with results | rheumatoid arthritis AND (upadacitinib OR ABT-494)]
WHO ICTRP	International Clinical Trials Registry Platform, produced by the World Health Organization. Targeted search used to capture registered clinical trials. [Search terms: Rheumatoid Arthritis AND (upadacitinib OR ABT 494 OR ABT494)]
Health Canada Clinical Trials Database	Produced by Health Canada. Targeted search used to capture registered clinical trials. [Search terms: rheumatoid arthritis AND upadacitinib]

OTHER DATABASES	
PubMed	Searched to capture records not found in MEDLINE. Same MeSH, keywords, limits, and study types used as per the MEDLINE search, with appropriate syntax used.
Cochrane Central Register of Controlled Trials	Same MeSH, keywords, and limits used as per the MEDLINE search, excluding study types and human restrictions. Syntax adjusted for Wiley platform.

Grey Literature

Dates for Search:	July 24, 2019, to July 29, 2019
Keywords:	[(rheumatoid arthritis OR RA) AND (upadacitinib OR ABT-494)]
Limits:	Publication years: all years

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

- health technology assessment agencies
- health economics
- clinical practice guidelines
- drug and device regulatory approvals
- advisories and warnings
- drug class reviews
- clinical trial registries
- databases (free)
- health statistics
- internet search
- open access journals.

Appendix 2: Excluded Studies

Table 42: Excluded Studies

Reference	Reason for exclusion
Marker-Hermann et al.[34]	Commentary
BALANCE-I (M13-550)[35]	Intervention and study design (phase II)
BALANCE-II (M13-537)[36]	Intervention and study design (phase II)
BALANCE-EXTEND (M13-538)[37,38]	Intervention and study design (phase II)

Appendix 3: Detailed Outcome Data

Table 43: Clinical Response (ACR20/50/70) at Week 2

	Total N	Responder n	Response rate (95% CI)	Response rate difference versus control	
				Point estimate (95% CI)	P value
ACR20 response at week 2 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	219	33.6 (30.0 to 37.3)	0.3 (–6.0 to 6.6) [UPA vs. ADA] 19.7 (15.2 to 24.2) [UPA vs. PBO]	0.920[a] [UPA vs. ADA] < 0.001[a] [UPA vs. PBO]
ADA 40 mg e.o.w.	327	109	33.3 (28.2 to 38.4)		
Placebo	651	91	14.0 (11.3 to 16.6)		
SELECT-MONOTHERAPY					
UPA 15 mg q.d.	217	72	33.2 (26.9 to 39.4)	26.2 (19.1 to 33.4)	< 0.001[a]
MTX	216	15	6.9 (3.6 to 10.3)		
SELECT-NEXT					
UPA 15 mg q.d.	221	82	37.1 (30.7 to 43.5)	18.6 (10.4 to 26.7)	< 0.001[a]
Placebo	221	41	18.6 (13.4 to 23.7)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	66	40.2 (32.7 to 47.7)	19.5 (9.9 to 29.2)	< 0.001[a]
Placebo	169	35	20.7 (14.6 to 26.8)		
SELECT-EARLY					
UPA 15 mg q.d.	317	127	40.1 (34.7 to 45.5)	23.8 (17.1 to 30.6)	< 0.001[a]
MTX	314	51	16.2 (12.2 to 20.3)		
ACR50 response at week 2 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	75	11.5 (9.1 to 14.0)	3.0 (–0.9 to 6.9) 9.1 (6.3 to 11.8)	0.149[a] [UPA vs. ADA] < 0.001[a] [UPA vs. PBO]
ADA 40 mg e.o.w.	327	28	8.6 (5.5 to 11.6)		
Placebo	651	16	2.5 (1.3 to 3.6)		
SELECT-MONOTHERAPY					
UPA 15 mg q.d.	217	18	8.3 (4.6 to 12.0)	6.9 (2.9 to 10.9)	< 0.001[a]
MTX	216	3	1.4 (0.0 to 2.9)		
SELECT-NEXT					
UPA 15 mg q.d.	221	24	10.9 (6.8 to 15.0)	8.1 (3.5 to 12.8)	< 0.001[a]
Placebo	221	6	2.7 (0.6 to 4.9)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	24	14.6 (9.2 to 20.0)	9.3 (2.9 to 15.7)	0.005[a]
Placebo	169	9	5.3 (1.9 to 8.7)		
SELECT-EARLY					
UPA 15 mg q.d.	317	48	15.1 (11.2 to 19.1)	12.3 (7.9 to 16.6)	< 0.001[a]
MTX	314	9	2.9 (1.0 to 4.7)		
ACR70 response at week 2 (NRI, FAS)					
SELECT-COMPARE					
UPA 15 mg q.d.	651	18	2.8 (1.5 to 4.0)	0.9 (–1.0 to 2.9) [vs. ADA]	0.403

	Total N	Responder n	Response rate (95% CI)	Response rate difference versus control	
				Point estimate (95% CI)	P value
ADA 40 mg e.o.w.	327	6	1.8 (0.4 to 3.3)	2.2 (0.8 to 3.5) [vs. PBO]	0.003
Placebo	651	4	0.6 (0.0 to 1.2)		
SELECT-MONOTHERAPY					
UPA 15 mg q.d.	217	5	2.3 (0.3 to 4.3)	26.7 (18.5 to 34.8)	< 0.001[a]
MTX	216	0	0		
SELECT-NEXT					
UPA 15 mg q.d.	221	7	3.2 (0.9 to 5.5)	2.7 (0.2 to 5.2)	0.033[a]
Placebo	221	1	0.5 (0.0 to 1.3)		
SELECT-BEYOND					
UPA 15 mg q.d.	164	8	4.9 (1.6 to 8.2)	3.7 (0.0 to 7.4)	0.050[a]
Placebo	169	2	1.2 (0.0 to 2.8)		
SELECT-EARLY					
UPA 15 mg q.d.	317	16	5.0 (2.6 to 7.5)	3.8 (1.1 to 6.5)	0.007[a]
MTX	314	4	1.3 (0.0 to 2.5)		

ACR = American College of Rheumatology; ADA = adalimumab; CI = confidence interval; e.o.w. = every other week; FAS = full analysis set; MTX = methotrexate; NRI = nonresponder imputation; PBO = placebo; q.d. = once daily; UPA = upadacitinib; vs. = versus.

[a] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

Table 44: Functional and Disability Outcomes: Fatigue

	Total N	Baseline	Week 12/14		Treatment group difference versus control		
		Mean	Mean	LS mean change from baseline (95% CI)	N	LS mean difference (95% CI)	P value
Fatigue using the FACIT-F, change from baseline at week 12 (MMRM, FAS)							
SELECT-COMPARE							
UPA 15 mg q.d.	651	26.68	35.48	8.95 (7.98 to 9.93)	612	1.51 (0.27 to 2.76) [vs. ADA] 4.15 (3.13 to 5.16) [vs. PBO]	0.017[a] [UPA vs. ADA] < 0.001 [UPA vs. PBO]
ADA 40 mg e.o.w.	327	26.31	33.77	7.44 (6.25 to 8.64)	307		
Placebo	651	27.05	31.56	4.81 (3.85 to 5.77)	613		
SELECT-NEXT							
UPA 15 mg q.d.	221	27.84	36.16	7.91 (6.56 to 9.27)	207	4.95 (3.31 to 6.60)	< 0.001
Placebo	221	28.33	31.63	2.96 (1.62 to 4.30)	207		
SELECT-EARLY (LOCF for rescue)							
UPA 15 mg q.d.	317	26.37	37.07	10.01 (8.94 to 11.07)	301	3.20 (1.70 to 4.70)	< 0.001[b]
MTX	315	27.17	34.29	6.80 (5.70 to 7.91)	277		

ADA = adalimumab; CI = confidence interval; e.o.w. = every other week; FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue; FAS = full analysis set; LOCF = last observation carried forward; LS = least squares; MMRM = mixed models for repeated measures; MTX = methotrexate; PBO = placebo; q.d. = once daily; UPA = upadacitinib; vs. = versus.

[a] Outcome was not included in the ranked key end points and therefore not controlled for type I error rate.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-EARLY Clinical Study Report;[5] SELECT-NEXT Clinical Study Report.[3]

Subgroup Analyses – Clinical Response (ACR20)

Table 45: Subgroup Data: ACR20 at Week 12 (Primary Outcome) by Disease Activity

		Total N	Responder, n (%)	Response rate (95% CI)	Response rate difference versus control, point estimate (95% CI)
ACR20 at week 12 (NRI, FAS), by baseline DAS28-CRP or disease activity					
SELECT-COMPARE					
DAS28-CRP ≤ 5.1	UPA 15 mg q.d.	149	102 (68.5)	68.5 (61.0 to 75.9)	33.8 (22.8 to 44.9) [UPA vs. PBO]
	ADA 40 mg e.o.w.	71	46 (64.8)	64.8 (53.7 to 75.9)	
	Placebo	130	45 (34.6)	34.6 (26.4 to 42.8)	
DAS28-CRP > 5.1	UPA 15 mg q.d.	498	357 (71.7)	71.7 (67.7 to 75.6)	34.7 (29.0 to 40.4) [UPA vs. PBO]
	ADA 40 mg e.o.w.	253	160 (63.2)	63.2 (57.3 to 69.2)	
	Placebo	519	192 (37.0)	37.0 (32.8 to 41.1)	
SELECT-MONOTHERAPY (at week 14)					
DAS28-CRP ≤ 5.1	UPA 15 mg q.d.	72	43 (59.7)	59.7 (48.4 to 71.1)	22.7 (6.9 to 38.6) [UPA vs. MTX]
	MTX	73	27 (37.0)	37.0 (25.9 to 48.1)	
DAS28-CRP > 5.1	UPA 15 mg q.d.	144	104 (72.2)	72.2 (64.9 to 79.5)	28.9 (17.9 to 39.8) [UPA vs. MTX]
	MTX	143	62 (43.4)	43.4 (35.2 to 51.5)	
SELECT-BEYOND					
DAS28-CRP ≤ 5.1	UPA 15 mg q.d.	39	24 (61.5)	61.5 (46.3 to 76.8)	24.7 (3.1 to 46.3) [UPA vs. PBO]
	Placebo	38	14 (36.8)	36.8 (21.5 to 52.2)	
DAS28-CRP > 5.1	UPA 15 mg q.d.	124	82 (66.1)	66.1 (57.8 to 74.5)	39.6 (28.3 to 50.9) [UPA vs. PBO]
	Placebo	128	34 (26.6)	26.6 (18.9 to 34.2)	
SELECT-NEXT					
DAS28-CRP ≤ 5.1	UPA 15 mg q.d.	66	36 (54.5)	54.5 (42.5 to 66.6)	21.2 (4.8 to 37.6) [UPA vs. PBO]
	Placebo	69	23 (33.3)	33.3 (22.2 to 44.5)	
DAS28-CRP > 5.1	UPA 15 mg q.d.	151	105 (69.5)	69.5 (62.2 to 76.9)	32.7 (22.1 to 43.3) [UPA vs. PBO]
	Placebo	152	56 (36.8)	36.8 (29.2 to 44.5)	
SELECT-EARLY[a]					
DAS28-CRP ≤ 5.1	UPA 15 mg q.d.	72	49 (68.1)	68.1 (57.3 to 78.8)	19.7 (3.2 to 36.1) [UPA vs. MTX]
	MTX	62	30 (48.4)	48.4 (35.9 to 60.8)	
DAS28-CRP > 5.1	UPA 15 mg q.d.	245	191 (78.0)	78.0 (72.8 to 83.1)	22.4 (14.4 to 30.4) [UPA vs. MTX]
	MTX	252	140 (55.6)	55.6 (49.4 to 61.7)	

ACR = American College of Rheumatology; ADA = adalimumab; CI = confidence interval; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; e.o.w. = every other week; FAS = full analysis set; MTX = methotrexate; NRI = nonresponder imputation; PBO = placebo; q.d. = once daily; UPA = upadacitinib; vs. = versus.

[a] Subgroup analyses for ACR20 were performed for the Pharmaceuticals and Medical Devices Agency (Japan), not for the FDA.

Source: SELECT-COMPARE Clinical Study Report;[1] SELECT-MONOTHERAPY Clinical Study Report;[2] SELECT-NEXT Clinical Study Report;[3] SELECT-BEYOND Clinical Study Report;[4] SELECT-EARLY Clinical Study Report.[5]

A subgroup analysis of the primary outcome, ACR20 at week 12, by prior treatment failure was only conducted in SELECT-BEYOND. This subgroup analysis is not applicable to the remaining pivotal studies, as patients who were considered inadequate responders to bDMARD therapy were excluded from SELECT-COMPARE, SELECT-MONOTHERAPY, and SELECT-NEXT, and patients with exposure to any bDMARD were excluded from SELECT-EARLY.

Table 46: Subgroup Data: ACR20 at Week 12 (Primary Outcome) by Prior Treatment Failure

		Total N	Responder, n	Response rate (95% CI)	Response rate difference versus control, point estimate (95% CI)
SELECT-BEYOND					
ACR20 at week 12 (NRI, FAS), by prior failed bDMARD					
Failed ≤ 2 biologics with same MOA	UPA 15 mg q.d.	116	72 (62.1)	62.1 (53.2 to 70.9)	31.3 (19.1 to 43.5) [UPA vs. PBO]
	Placebo	117	36 (30.8)	30.8 (22.4 to 39.1)	
Failed > 2 biologics with same and/or multiple MOAs	UPA 15 mg q.d.	48	34 (70.8)	70.8 (58.0 to 83.7)	47.8 (30.5 to 65.0) [UPA vs. PBO]
	Placebo	52	12 (23.1)	23.1 (11.6 to 34.5)	
ACR20 at week 12 (NRI, FAS), by failure of ≥ 1 bDMARD due to lack of efficacy					
Yes	UPA 15 mg q.d.	146	95 (65.1)	65.1 (57.3 to 72.8)	35.5 (25.0 to 46.0) [UPA vs. PBO]
	Placebo	159	47 (29.6)	29.6 (22.5 to 36.7)	
No	UPA 15 mg q.d.	18	11 (61.1)	61.1 (38.6 to 83.6)	51.1 (21.9 to 80.3) [UPA vs. PBO]
	Placebo	10	1 (10.0)	10.0 (0.0 to 28.6)	
ACR20 at week 12 (NRI, FAS), by failure of anti–IL-6 due to lack of efficacy					
Yes	UPA 15 mg q.d.	27	15 (55.6)	55.6 (36.8 to 74.3)	35.6 (12.0 to 59.1) [UPA vs. PBO]
	Placebo	30	6 (20.0)	20.0 (5.7 to 34.3)	
No	UPA 15 mg q.d.	137	91 (66.4)	66.4 (58.5 to 74.3)	36.2 (25.2 to 47.2) [UPA vs. PBO]
	Placebo	139	42 (30.2)	30.2 (22.6 to 37.8)	

ACR = American College of Rheumatology; bDMARD = biologic disease-modifying antirheumatic drug; CI = confidence interval; FAS = full analysis set; IL = interleukin; MOA = mechanism of action; NRI = nonresponder imputation; PBO = placebo; q.d. = once daily; UPA = upadacitinib; vs. = versus.

Source: SELECT-BEYOND Clinical Study Report.[4]

Table 47: Summary of Harms (COMPARE up to Week 14, Prior to Availability of Rescue Therapy)

	COMPARE (M14-465)		
	UPA 15 mg q.d. N = 650	ADA 40 mg e.o.w. N = 327	PBO N = 652
Patients with ≥ 1 adverse event			
n (%)	348 (53.5)	158 (48.3)	303 (46.5)
Most common events[a], n (%)			
Nasopharyngitis	27 (4.2)	8 (2.4)	15 (2.3)
Upper respiratory tract infection	27 (4.2)	6 (1.8)	16 (2.5)
Alanine aminotransferase increased	23 (3.5)	5 (1.5)	18 (2.8)
Bronchitis	22 (3.4)	8 (2.4)	11 (1.7)

	COMPARE (M14-465)		
	UPA 15 mg q.d. N = 650	ADA 40 mg e.o.w. N = 327	PBO N = 652
Diarrhea	20 (3.1)	10 (3.1)	13 (2.0)
Urinary tract infection	18 (2.8)	13 (4.0)	16 (2.5)
Hypertension	18 (2.8)	4 (1.2)	11 (1.7)
Headache	17 (2.6)	4 (1.2)	16 (2.5)
Blood creatine phosphokinase increased	17 (2.6)	1 (0.3)	8 (1.2)
Aspartate aminotransferase increased	16 (2.5)	6 (1.8)	10 (1.5)
Nausea	14 (2.2)	8 (2.4)	13 (2.0)
Pharyngitis	13 (2.0)	7 (2.1)	7 (1.1)
Rheumatoid arthritis (disease worsening)	3 (0.5)	5 (1.5)	22 (3.4)
Patients with ≥ 1 SAE			
n (%)	18 (2.8)	8 (2.4)	14 (2.1)
Most common events[b], n (%)			
Appendicitis	2 (0.3)	0	0
Cellulitis	0	2 (0.6)	0
Gastroenteritis	2 (0.3)	0	3 (0.5)
Pneumocystis jiroveci pneumonia	0	0	2 (0.3)
Pulmonary embolism	0	3 (0.9)	1 (0.2)
Patients who stopped treatment owing to adverse events			
n (%)	18 (2.8)	16 (4.9)	12 (1.8)
Most common events[b], n (%)			
Anemia	2 (0.3)	0	0
Pneumocystis jiroveci pneumonia	0	0	2 (0.3)
Alanine aminotransferase increased	1 (0.2)	2 (0.6)	0
Aspartate aminotransferase increased	1 (0.2)	2 (0.6)	0
Blood creatine increased	2 (0.3)	0	0
Pulmonary embolism	0	2 (0.6)	0
Deaths			
n (%)	0	1 (0.3)	2 (0.3)
Sudden death	0	0	1 (0.2)
Pneumocystis jiroveci pneumonia	0	0	1 (0.2)
Craniocerebral injury	0	1 (0.3)	0
Notable harms			
n (%)			
Herpes zoster infection	5 (0.8)	1 (0.3)	1 (0.2)
Neutropenia	9 (1.4)	1 (0.3)	2 (0.3)
Lymphopenia	11 (1.7)	2 (0.6)	8 (1.2)
Thrombocytopenia	2 (0.3)	0	1 (0.2)
Malignancy (any)	0	1 (0.3)	2 (0.3)
Thrombosis (incl. increased platelets)[c]	1 (0.2)	3 (0.9)	1 (0.2)
MACE[d]	0	1 (0.3)	3 (0.5)

	COMPARE (M14-465)		
	UPA 15 mg q.d. N = 650	ADA 40 mg e.o.w. N = 327	PBO N = 652
GI perforations	2 (0.3)	0	0
Hepatic disorder	38 (5.8)	11 (3.4)	22 (3.4)
Dyslipidemia	2 (0.3)	0	1 (0.2)

ADA = adalimumab; e.o.w. = every other week; GI = gastrointestinal; incl. = including; MACE = major adverse cardiovascular event; PBO = placebo; q.d. = once daily; SAE = serious adverse event; UPA = upadacitinib.

[a] Frequency ≥ 2% in any group.

[b] Frequency > 1 patient in any group.

[c] Deep vein thrombosis and fatal/nonfatal pulmonary embolism.

[d] Cardiovascular death, nonfatal myocardial infarction, and nonfatal stroke.

Source: SELECT-COMPARE Clinical Study Report.[1]

Table 48: Summary of Harms (COMPARE up to Week 26, After Treatment Switching)

	COMPARE (M14-465)		
	UPAvs.ADA 40 mg e.o.w. N = 125 PYS = 23.1	ADA–UPA 15 mg q.d. N = 77 PYS = 15.5	PBO–UPA 15 mg q.d. N = 304 PYS = 60.4
Number of AEs			
E/100PYS (%)	96 (415.6)	40 (258.1)	225 (372.5)
Number of SAEs			
E/100PYS (%)	2 (8.7)	0	8 (13.2)
Number of AEs leading to discontinuation of treatment			
E/100PYS (%)	4 (17.3)	0	10 (16.6)
Number of AEs leading to death			
E/100PYS (%)	0	0	0

ADA = adalimumab; AE = adverse event; e.o.w. = every other week; PBO = placebo; PYS = patient-years; q.d. = once daily; SAE = serious adverse event; UPA = upadacitinib.

Source: SELECT-COMPARE Clinical Study Report.[1]

Appendix 4: Description and Appraisal of Outcome Measures

Aim

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

- ACR response criteria: ACR20, ACR50, and ACR70
- DAS28
- HAQ-DI
- mTSS
- SF-36
- CDAI.

A summary of the level of outcomes in each of the included studies is provided in Table 49, and a summary of the corresponding measurement properties is provided in Table 50.

Table 49: Outcome Measures Included in Each Study

Outcome measure	COMPARE	MONOTHERAPY	NEXT	BEYOND	EARLY
ACR20 ACR50 ACR70	Primary Secondary/exploratory Exploratory	Primary Exploratory Exploratory	Primary Exploratory Exploratory	Primary Exploratory Exploratory	Exploratory Primary Exploratory
DAS28	Secondary	Secondary	Secondary	Secondary	Secondary
HAQ-DI	Secondary	Secondary	Secondary	Secondary	Secondary
mTSS	Secondary	Not reported	Not reported	Not reported	Secondary
SF-36	Secondary	Secondary	Secondary	Exploratory	Secondary
CDAI	Exploratory	Not reported[a]	Secondary	Exploratory	Exploratory

ACR = American College of Rheumatology; CDAI = Clinical Disease Activity Index; DAS28 = Disease Activity Score 28; HAQ-DI = Health Assessment Questionnaire–Disability Index; mTSS = modified total Sharp score; SF-36 = Short Form (36) Health Survey.

[a] Included in assessment of low disease activity, not reported separately as a change from baseline.

Table 50: Summary of Outcome Measures and Their Measurement Properties

Outcome measure	Type	Conclusions about measurement properties	MID
ACR20 ACR50 ACR70	ACR20, ACR50, and ACR70 responses represent at least a 20%, 50%, and 70% improvement, respectively, in tender and swollen joint counts and in three of the five additional criteria: • patient global assessment of disease activity • physician global assessment of disease activity • patient assessment of pain • HAQ	Individual criteria were selected on the basis of their construct validity, face validity, content validity, criterion validity, and discriminant validity.[39] Validity, reliability, and responsiveness of ACR20/50/70 as a composite measure was not identified.	ACR50 represents a more robust clinical response when comparing active and control therapies;[40] however, ACR20 is widely accepted and sufficient for FDA regulatory purposes.[41]

Outcome measure	Type	Conclusions about measurement properties	MID
	• CRP or ESR		
DAS28	DAS28 is an abbreviated version of the DAS based on a 28-joint count that omits the feet and ankle joints.	DAS28-CRP: Test-retest reliability (r = 0.92 for patients and r = 0.87 for physicians).[42] Indirect assessment of concurrent validity for DAS28 by correlation with DAS (r > 0.94);[43] construct validity through correlation with HAQ (r = 0.49) and SF-36 Physical Functioning Scale (r = 0.46).[43] Responsiveness was also demonstrated using clinical trial data, which showed a statistically significant change from baseline for the DAS.[44]	Not identified.
HAQ-DI	The HAQ-DI is the disability assessment component of the HAQ, a self-reported assessment of functional status. The overall HAQ-DI score ranges from 0 (no disability) to 3 (completely disabled).	Validity (known-groups and convergent) and test-retest reliability have been demonstrated.[45]	0.22 points.
mTSS	The mTSS is a composite measure of joint erosion and joint space narrowing based on radiograph assessment.	Inter-rater reliability was demonstrated on the basis of a 95% level of agreement between two readers for radiographic images from patients with RA. Evidence regarding the responsiveness and validity were not identified.	3.0 to 4.6 units.
SF-36 v2	The SF-36 consists of eight subdomains. The SF-36 provides two component summaries: PCS and MCS. The eight subdomains are each measured on a scale of 0 to 100, with an increase in score indicating an improvement in health status.	Evidence of responsiveness for the physical functioning and bodily pain scales and PCS score is based on a SRM ≥ 0.50.[46]	Improvement in PCS: 7.2 (95% CI, 4.6 to 8.0).[46]
SDAI and CDAI	The SDAI integrates measures of physical examination, acute phase response, patient self-assessment, and evaluator assessment to simplify the assessment of disease activity in clinical practice. The CDAI is similar to the SDAI, but it allows for immediate scoring because it does not include a laboratory result.	Reliability: Intra-rater reliability was demonstrated for the SDAI and CDAI on the basis of ICC ranging from 0.85 to 0.89.[47] Also, internal consistency was demonstrated on the basis of fair agreement between the SDAI (κ = 0.382) and CDAI (κ = 0.354) and the DAS28.[48] Validity: Concurrent validity demonstrated through a strong correlation between the SDAI or	Not identified.

Outcome measure	Type	Conclusions about measurement properties	MID
		CDAI and DAS28 (Pearson r = 0.87 to 0.90).[49] Responsiveness: the SDAI and CDAI were able to distinguish between responders and nonresponders (data not provided).[47]	

ACR = American College of Rheumatology; CDAI = Clinical Disease Activity Index; CI = confidence interval; CRP = C-reactive protein; DAS28 = Disease Activity Score 28; ESR = erythrocyte sedimentation rate; HAQ = Health Assessment Questionnaire; HAQ-DI = Health Assessment Questionnaire–Disability Index; MCS = mental component summary; MID = minimal important difference; mTSS = modified total Sharp score; PCS = physical component summary; RA = rheumatoid arthritis; SDAI = Simplified Disease Activity Index; SF-36 = Short Form (36) Health Survey; SRM = standardized response mean.

Findings

ACR Criteria

The ACR criteria for assessing joint status were initially developed for patients with RA.[50] ACR criteria provide a composite measure of improvement in both swollen and tender joint counts and at least three of five additional disease criteria:

- patient global assessment of disease activity
- physician global assessment of disease activity
- patient assessment of pain
- HAQ
- levels of either CRP or ESR.

The ACR joint count for RA assesses 68 joints for tenderness and 66 joints for swelling. Patient and physician assessments are conducted using VAS or Likert scale measurements. ACR20, 50, or 70 responses represent at least a 20%, 50%, or 70% improvement, respectively, in tender and swollen joint counts as well as in three of the five additional core measures listed. This core set of measures included in the ACR response criteria was established through a consensus process of clinical experts. Individual criteria were selected on the basis of their construct validity, face validity, content validity, criterion validity, and discriminant validity.[39] In the assessment of criterion validity, standards for comparison included death, physical disability, and radiologic evidence of joint damage. It was considered that physical functioning capacity was a strong predictor of mortality, as measured by the HAQ, and that many other risk factors for premature mortality were insignificant after adjusting for functional capacity. Predictors of radiographic progression included swollen joint counts and levels of acute phase reactants indicated by, for example, ESR and CRP.[39] Pain assessments, global assessments, tender joint counts, and HAQ scores all had strong discriminant validity in their ability to detect change.

The ACR20 is most commonly used as the primary end point in RCTs evaluating biologics in the treatment of RA. The FDA considers ACR20 a well-validated composite end point for assessing the signs and symptoms of RA, as noted in guidance provided to industry on the conduct of trials in RA patients.[41] ACR50 and ACR70 are often reported in clinical trials and are considered more stringent outcome measures.

Chung et al.[40] conducted a meta-analysis of 21 RCTs of RA therapies published between 1997 and 2004 to compare the discriminant capabilities of the ACR50 and ACR20

responses and to determine whether ACR50 is as informative as ACR20 in distinguishing between active therapies and control groups. While both measures have the ability to distinguish an active therapy from a control therapy, the levels of improvement captured by the ACR20 response do not generally represent an optimal clinical improvement. Furthermore, since the development of the ACR20 response criteria, much more aggressive therapies have been introduced in the treatment of RA, and larger clinical responses can be expected. This meta-analysis concluded that ACR20 and ACR50 are similar in distinguishing between active and control therapies but that ACR50 represents a more robust clinical response and may be a preferred end point in clinical trials.[40]

ACR70 is considered even more rigorous than ACR50. It is a component of the definitions established by the FDA to satisfy labelling requirements for RA drugs. Specifically, a "major clinical response" as defined by the FDA refers to a statistically significant increase in the proportion of patients achieving an ACR70 response, maintained over six months, with active therapy compared with the control group.[41]

Through widespread use of the ACR criteria over the past 20 years, limitations associated with them have been identified. For example, while ACR response indicates the change from baseline, it does not indicate the final level of disease severity that the patient attains. This limitation also means that patients who are classified as ACR responders could have very different levels of disease.[51] Other criticisms of the ACR criteria include that most of their component measures are subjective, that dichotomous measures such as ACR lack sensitivity to change compared with continuous measures of response, and that the ACR20 response threshold is too low relative to the treatment goals applied in clinical practice.[52] In response to these criticisms, attempts have been made to develop improved outcome measures for RA, although none have widespread acceptance or are consistently used in clinical trials.[52,53]

Disease Activity Score 28

The DAS is a measure of RA disease activity and includes the Ritchie Articular Index (0 to 78), which is performed on 53 joints; a 44-joint swollen joint count (0 to 44); ESR or CRP; and a general health item using a VAS (0 to 100).[54] DAS28 is an abbreviated version of the DAS, based on a 28-joint count that omits the feet and ankle joints. Thus one obvious criticism of this scale is that a patient who only had inflammation at the feet and ankles would be counted as in remission.[55] The DAS components correlate well with one another and with the ACR criteria.[56-59] The DAS28 is a composite score derived using the following formula:

$$DAS28 = 0.56 \times \sqrt{(t28)} + 0.28 \times \sqrt{(sw28)} + 0.70 \times \ln(ESR) + 0.014 \times GH$$

Where DAS28 = Disease Activity Score 28; t28 = tender joint count of 28 joints; sw28 = swollen joint count of 28 joints; ESR = erythrocyte sedimentation rate; GH = general health measured by patient's global assessment of disease activity on a VAS of 100 mm.

The formula was developed by comparing serial assessments of tender and swollen joint counts, ESR, and patient global assessment (global health) for a panel of patients with RA both at times of poorly controlled RA and when well-controlled.[60] DAS28 indicates an absolute level of disease activity, with a score of 5.1 or greater being considered high disease activity, a score lower than 3.2 indicating LDA, and a score lower than 2.6 indicating remission.[20-22]

In recent years, CRP has been used in place of ESR to calculate the DAS28. The trend of using CRP levels as opposed to ESR is mainly driven by greater availability, reduced cost, and increased sensitivity of CRP to short-term changes in disease activity.[20,61] The formula used to calculate the DAS28-CRP is as follows:

$$DAS28\text{-}CRP = 0.56 \times \sqrt{(t28)} + 0.28 \times \sqrt{(sw28)} + 0.014 \times GH + 0.36 \times \ln(CRP+1) + 0.96$$

Where DAS28 = Disease Activity Score 28; CRP = C-reactive protein; t28 = tender joint count of 28 joints; sw28 = swollen joint count of 28 joints; GH = general health measured by patient's global assessment of disease activity on a VAS of 100 mm.

The test-retest reliability of the DAS28-CRP was examined in a study that included 30 out-clinic patients with RA with stable disease. Reliability was evaluated by comparing patient and physician scores of a joint count performed twice, one week apart. Both groups demonstrated strong test-retest reliability based on a correlation of r = 0.92 for patients and 0.87 for physicians.[42] Concurrent validity of the DAS was shown through correlation with 12 other common estimators of disease activity (mean r = 0.61), and by extension, the DAS28 was well correlated (r > 0.94) with the DAS.[43] Construct validity was demonstrated through correlation with the HAQ (r = 0.49) and the SF-36 Physical Functioning Scale (r = 0.46).[43] Lastly, responsiveness was examined using data from a study including 155 patients with early RA who demonstrated "excellent clinical response."[44] Responsiveness was evaluated using the standardized response mean for the change from baseline at week 16. The DAS showed a mean change of –2.1 (standard error = 0.1), which exceeds the 0.35 difference that the author considered to be significant at a two-sided 0.05 level.[44]

Overall, the DAS28-CRP correlates well with DAS28-ESR, and both are validated measures for assessing disease activity in RA.[20,21,23,62,63] However, studies have shown that the DAS28-CRP value is usually lower than the DAS28-ESR value.[21,23,62-67] The difference (DAS28-CRP minus DAS28-ESR) ranges from −0.2[62] to −0.8.[64] Because the definitions of remission (score lower than 2.6) are the same for both DAS28-CRP and DAS28-ESR, it was concluded that DAS28-CRP underestimates disease activity and overestimates the improvement in disease activity and the remission rate compared with DAS28-ESR. It was also suggested that DAS28-CRP should be evaluated using different criteria from those used for DAS28-ESR.[23] Furthermore, EULAR recommended that the clinical implications of the DAS28 score (e.g., good response, moderate response, no response) should be determined on the basis of the baseline DAS28 (see Table 51).[68] Finally, an minimal important difference for change in DAS28 was not identified; however, a clinical change based on the EULAR response criteria can be used to interpret a clinical response according to the DAS28, as described.

Table 51: EULAR Improvement Response Criteria (DAS28)

Baseline DAS28 score	DAS28 improvement over time points		
	> 1.2	0.6 to 1.2	< 0.6
< 3.2	Good response	Moderate response	No response
3.2 to 5.1	Moderate response	Moderate response	No response
> 5.1	Moderate response	No response	No response

DAS28 = Disease Activity Score 28; EULAR = European League Against Rheumatism.

Source: Matsui et al. (2007).[23]

HAQ and Disability Index (HAQ-DI)

The HAQ was originally developed in 1978 at Stanford University.[69] It was one of the first self-reported functional status (disability) measures and has become the dominant instrument in many disease areas, including arthritis.[70] The HAQ has been widely validated in patients with RA.[70] The full HAQ collects data on five generic patient-centred health dimensions: 1) to avoid disability, 2) to be free of pain and discomfort, 3) to avoid adverse treatment effects, 4) to keep dollar costs of treatment low, and 5) to postpone death.[22]

The HAQ-DI is the disability assessment component of the HAQ. It assesses a patient's level of functional ability. There are 20 questions to assess a patient's physical functional status in eight categories: dressing, arising, eating, walking, hygiene, reach, grip, and common activities.[24,25] For each of these categories, patients report the amount of difficulty they have in performing specific activities, and their responses are made on a scale from 0 (no difficulty) to 3 (unable to do). The eight category scores are averaged into an overall HAQ-DI score on a scale from 0 (no disability) to 3 (completely disabled).

Observational studies and RCTs have demonstrated that the HAQ-DI possesses face validity, content validity, construct validity, predictive validity, and discriminant validity. For example, a study conducted by Linde et al. (2008) aimed to validate the HAQ among other outcomes for RA. Two samples of patients with RA (n = 200 and n = 150), recruited from outpatient clinics in Denmark, were included. Known-groups validity was demonstrated by the ability to distinguish between groups according to the DAS28, VAS for arthritis activity, and receivers versus nonreceivers of a disability pension.[45] Convergent validity was also demonstrated through a correlation greater than 0.70 with measures of physical function, pain, global RA, and overall health.[45] Lastly, reliability was demonstrated through agreement of patient-reported HAQ scores after two weeks, measured by an interclass correlation of 0.97 (95% CI, 0.96 to 0.98).[45]

There is evidence suggesting that baseline HAQ scores are predictive of radiographic damage, work disability, and quality of life.[45,71] A number of investigators have suggested that the MID is 0.22; however, differences as small as 0.10 have been suggested as clinically important.[24]

Modified Total Sharp Score

The Sharp scoring system, first developed in 1971, has undergone modifications over time and is now referred to as the modified Sharp. This method allows for the assessment of two aspects of joint damage: articular erosions (representing direct invasion of cartilage and bone by the proliferating synovial pannus) and joint space narrowing (representing destruction of surface cartilage). Data on the progression of joint structural damage are obtained by taking X-rays of specific joints (typically in the hands and feet) before treatment and at various points after treatment has been initiated. The most recent modification of the Sharp scoring system was performed by van der Heijde.[72] Van der Heijde scores erosions as listed in Table 52.

Table 52: Scoring for the Modified Total Sharp Score

Sharp/van der Heijde[73]	
Erosions	
Score	Description
0	Normal
1	Discrete erosions
2 to 3	Larger erosions according to surface area involved
4	Erosion extending over the middle of the bone
5	Complete collapse
Joint space narrowing	
Score	Description
0	Intact bony outlines and normal joint space
1	Erosion < 1 mm in diameter or joint space narrowing
2	One or several small erosions (diameter > 1 mm)
3	Marked erosions
4	Severe erosions (usually no joint space left, and the original bony outlines are only partly preserved)
5	Mutilating changes (original bony outlines have been destroyed)

The van der Heijde erosion score includes 16 joints from the hands and wrists (graded from 0 to 5) and six joints from the feet (graded from 0 to 10). The joint space narrowing score includes 15 areas from the hands and wrists (graded from 0 to 4) and six areas from the feet (also graded from 0 to 4). The maximum erosion score is 160 for hands and wrists and 120 for feet, while the maximum joint space narrowing score is 120 for hands and 48 for feet.[26] Maximum total scores for both erosion and joint space narrowing are calculated as follows:

Erosion = (32 joints in hands and wrists × 5) + (12 joints in feet × 10) = 280

Joint space narrowing = (30 joints in hands and wrists × 4) + (12 joints in feet × 4) = 168

The van der Heijde modification has become the most commonly used for a few reasons: 1) it includes both hands and feet; 2) it measures erosions and joint space narrowing; and 3) it covers a broad spectrum of joints, providing sensitivity to change.[74]

In the early stages of RA, inflammation appears to be the main contributor to increased disability, rather than actual damage to joints.[75,76] The relationship between radiological and

functional changes has been studied. A reanalysis of published data performed by Welsing et al. found that patients must reach a certain amount of radiological damage before an increase in damage will impact disability. The authors also found that changes in Sharp scores had a greater impact on disability with advancing age. A study by Sabin et al. found that radiologic damage assessed by the van der Heijde method was highly correlated with HAQ scores in a population with a mean disease duration of seven years. They also cited findings from another study, which found that Sharp scores became correlated with HAQ after six years' disease duration. At the other end of the spectrum, a study by Clarke et al. found that radiological scores assessed using the Genant method were positively correlated with HAQ in patients with 20 years' disease duration.[77] Therefore, radiological changes assessed by Sharp scores, and functional changes assessed by the HAQ do not correlate with each other early in RA, but after several years of disease.

Several limitations exist with using radiographs to assess clinical status in RA. Radiographs tend to change slowly in RA, requiring at least six months to a year to detect changes in a single patient. Inter-rater and intra-rater reliability is also a concern owing to the subtle nature of changes and subjective interpretation. The images themselves can also vary between samples, owing to positioning and quality. Radiographs should be read in random order to reduce the potential bias of interpretation at different time points.[78] Given these limitations, beginning in the early 1990s, the use of MRI was being examined as an alternative for assessing disease progression.[79] However, the use of MRI for assessing the clinical status of RA is limited owing to cost and accessibility.

In a study by Bruynesteyn, authors determined an MCID of 4.6 units for the Sharp/van der Heijde method, using a panel of experts.[80] They defined the MCID as a progression in radiologic joint damage that makes a rheumatologist change therapy. This MCID was equal to, or slightly lower than, the smallest detectable difference for this scoring system. The smallest detectable difference represents the smallest change score that can be reliably discriminated from the measurement error of the scoring method.[81] The smallest detectable change score is another method of measuring reliability. Similar to the MCID, the smallest detectable change score can provide guidance for interpreting if there has been a real change in patient outcomes over time. A study by Navarro-Compan assessed the level of agreement between two readers on radiographic images from patients with RA.[82] The authors found a smallest detectable change of 3.1 (range: 2.3 to 4.3) using the 95% level of agreement method and suggest that a score of 3.0 units is a reasonable cut-off for interpreting radiographic progression as clinically meaningful.[82]

SF-36 Version 2

The SF-36 is a generic health assessment questionnaire that has been used in clinical trials to study the impact of chronic disease on HRQoL. The SF-36 consists of eight subdomains: physical functioning, pain, vitality, social functioning, psychological functioning, general health perceptions, role limitations due to physical problems and role limitations due to emotional problems.[83] The SF-36 also provides two component summaries: the PCS and the MCS. The eight subdomains are each measured on a scale of 0 to 100, with an increase in score indicating an improvement in health status. The MCID for either the PCS or MCS of the SF-36 is typically between 2.5 and 5 points.[27-29]

A systematic review was conducted to evaluate the responsiveness of the SF-36 in drug trials by assessing the concordance of primary clinical outcomes and the minimal important difference of the SF-36.[84] Fifteen studies for RA were included, with 93% achieving a net

improvement in the physical component score or mental component score greater than or equal to 3, thereby demonstrating responsiveness.

Another study by Ward et al. (2014) assessed the responsiveness of the SF-36 in 243 patients with active RA, who completed the survey before and after treatment escalation.[46] A standardized response mean greater than or equal to 0.50 was used to assess the responsiveness of the SF-36 scales, and only the physical functioning and bodily pain scales and the PCS scores were deemed adequate on the basis of a standardized response mean of 0.55 (95% CI, 0.46 to 0.65), 0.65 (95% CI, 0.57 to 0.73), and 0.63 (95% CI, 0.55 to 0.71), respectively. A minimum clinically important improvement was estimated for the scales that were considered responsive, using a receiver operating characteristic curve analysis to identify a change associated with specificity of 0.80. The minimum clinically important improvements determined for the physical functioning scale, bodily pain scale, and PCS score were 7.1 (95% CI, 4.2 to 8.9), 4.9 (95% CI, 3.9 to 12.9), and 7.2 (95% CI, 4.6 to 8.0), respectively.[46]

SDAI and CDAI

The SDAI is a tool for measuring disease activity that integrates measures of physical examination, acute phase response, patient self-assessment, and evaluator assessment.[49] It was originally developed to simplify the assessment of disease activity in clinical practice.[49]

SDAI is calculated by a simple numerical addition of the scores from the five following assessments:

- number of tender joints (0 to 28)
- number of swollen joints (0 to 28)
- CRP in mg/dL (0.1 to 10.0)
- patient global assessment of disease activity, VAS (0 to 10.0 cm)
- physician global assessment of disease activity, VAS (0 to 10.0 cm).

The CDAI is similar to the SDAI, but it allows for immediate scoring because it does not include a laboratory result.[49] Therefore, the CDAI is calculated by adding the scores from the four following assessments:

- number of tender joints (0 to 28)
- number of swollen joints (0 to 28)
- patient global assessment of disease activity, VAS (0 to 10.0 cm)
- physician global assessment of disease activity, VAS (0 to 10.0 cm).

Both the SDAI and CDAI have been validated and show correlation with each other as well as with the DAS28.[47-49] This was analyzed in the original dataset used to develop the instruments, as well as a series of additional datasets. According to one review, the SDAI or the CDAI and the DAS28 were generally well correlated, with a Pearson coefficient between 0.87 and 0.90 at baseline and after six to 12 months. Further, the review describes the correlation between the CDAI and the SDAI as “almost perfect,” with a correlation coefficient ranging from 0.94 to 1.00.[49] Disease remission is defined as an SDAI score less than or equal to 3.3 and as a CDAI score less than or equal to 2.8.[49,85] The internal consistency of the SDAI and the CDAI were measured in a study that assessed the disease activities of 250 patients living with RA who were recruited from an outpatient rheumatology

clinic.[48] As shown by the kappa statistic, there was fair agreement between the DAS28-CRP and both the SDAI (κ = 0.382) and the CDAI (κ = 0.354).[48,86] Lastly, the responsiveness of the SDAI and the CDAI was assessed through sensitivity to change and using the ACR criteria as an anchor.[47] In summary, both measures were able to distinguish between responders and nonresponders; however, the emasures did not reflect a significant difference between ACR50 and ACR70.

FACIT-Fatigue

The FACIT-Fatigue scale was originally developed for use in patients with cancer. It is one of a series of symptom subscales in the FACIT measurement system and has since been validated for use in patients with RA.[87]

FACIT-Fatigue is a patient self-report measure consisting of 13 statements. Patients are asked to indicate to what extent the statement applies to them over the course of the previous seven days. Each statement has five possible levels of response, scored on a scale of 0 to 4 (0 representing “not at all” and 4 representing “very much”), resulting in scores ranging from 0 to 52. Lower scores indicate higher levels of fatigue. A suggested MCID for FACIT-Fatigue in patients with RA is between 3 and 4 points.[87] This MCID was found in a sample of 271 patients (77% female, 81% white), median age of 56 years (range: 28 to 84 years), a median tender joint count of 26 (range: nine to 68) and a median swollen joint count of 15 (range: two to 43).[87]

Conclusion

ACR response, HAQ-DI, SF-36, DAS28, SDAI, CDAI, and the mTSS were used as efficacy measures in the upadacitinib trials. The ACR20, 50, and 70 indicate a percentage improvement from baseline (but not a final level of disease activity). ACR20 is most commonly reported in clinical trials; however, ACR50 or ACR70 are often cited as evidence of a more robust treatment effect. The HAQ is a comprehensive measure of the patient’s perception of functional status and has been widely validated in patients with RA. A suggested MCID in patients with RA is 0.22; however, differences as small as 0.10 have also been suggested. SF-36 is a generic health assessment questionnaire that consists of eight subdomains[83] but also provides two component summaries: the PCS and the MCS. The MCID is typically between 2.5 and 5 points, but approximately 7.2 (95% CI, 4.6 to 8.0) for the PCS in patients living with RA.[27-29,46] The DAS28 measures an absolute rather than relative level of disease activity, and its components correlate well with one another and with the ACR components. However, it was reported that DAS28-CRP overestimates the improvement in disease activity and the remission rate compared with DAS28-ESR. The MCID for a change in DAS28 values has not been specified; however, a meaningful change may be interpreted using the EULAR criteria.[88] The SDAI and the CDAI have been well validated in previous studies of patients living with RA; however, an MCID was not identified at this time. The mTSS allows for the assessment of two aspects of joint damage in the hands, wrists, and feet: articular erosions (representing direct invasion of cartilage and bone by the proliferating synovial pannus) and joint space narrowing (representing the destruction of surface cartilage). Some limitations of the mTSS include the time it takes for changes to appear on the radiographic image, inter-rater and intra-rater reliability, and the variability in images between samples due to positioning and quality. An MCID of 3.0 to 4.6 units on the mTSS has been suggested.

References

1. Clinical study report: M14-465. A phase 3, randomized, double-blind study comparing upadacitinib (ABT-494) to placebo and to adalimumab in subjects with moderately to severely active rheumatoid arthritis who are on a stable background of methotrexate (MTX) and who have an inadequate response to MTX (MTX-IR) [CONFIDENTIAL internal manufacturer's report]. In: North Chicago (IL): AbbVie; 2018 Aug 28.
2. Clinical study report: M15-555. A phase 3, randomized, double-blind study comparing upadacitinib (ABT-494) monotherapy to methotrexate (MTX) in subjects with moderately to severely active rheumatoid arthritis with inadequate response to MTX [CONFIDENTIAL internal manufacturer's report]. In: North Chicago (IL): AbbVie; 2018 Jun 07.
3. Clinical study report: M13-549. A phase 3, randomized, double-blind study comparing upadacitinib (ABT-494) to placebo in subjects with moderately to severely active rheumatoid arthritis who are on a stable dose of conventional synthetic disease-modifying anti-rheumatic drugs (csDMARDs) and have an inadequate response to csDMARDs [CONFIDENTIAL internal manufacturer's report]. In: North Chicago (IL): AbbVie; 2018 Mar 01.
4. Clinical study report: M13-542. A phase 3, randomized, double-blind study comparing upadacitinib (ABT-494) to placebo on stable conventional synthetic disease-modifying anti-rheumatic drugs (csDMARDs) in subjects with moderately to severely active rheumatoid arthritis with inadequate response or intolerance to biologic DMARDs (bDMARDs) [CONFIDENTIAL internal manufacturer's report]. In: North Chicago (IL): AbbVie; 2018 Jul 23.
5. Clinical study report: M13-545. A phase 3, randomized, double-blind study comparing upadacitinib (ABT-494) once daily monotherapy to methotrexate (MTX) monotherapy in MTX-naïve subjects with moderately to severely active rheumatoid arthritis [CONFIDENTIAL internal manufacturer's report]. In: North Chicago (IL): AbbVie; 2018 Aug 27.
6. e-CPS. In: Ottawa (ON): Canadian Pharmacists Association; 2017: https://www.e-therapueutics.ca. Accessed 2019 Oct 04.
7. Arthritis Society. Arthritis risk factors. 2019; https://arthritis.ca/about-arthritis/arthritis-risk-factors. Accessed 2019 Oct 04.
8. The impact of arthritis in Canada: today and over the next 30 years. In: Toronto (ON): Arthritis Alliance of Canada; 2011: http://www.arthritisalliance.ca/en/impactquick. Accessed 2019 May 28.
9. Venables PJW. Diagnosis and differential diagnosis of rheumatoid arthritis. In: Post TW, ed. *UpToDate.* Waltham (MA): UpToDate; 2018. Accessed 2019 Jul 24.
10. National Institute for Health Care Excellence. Rheumatoid arthritis in adults: management. *(NICE guideline NG100)* 2018; https://www.nice.org.uk/guidance/ng100. Accessed 2019 Oct 04.
11. Singh JA, Saag KG, Bridges SL, Jr., et al. 2015 American College of Rheumatology Guideline for the treatment of rheumatoid arthritis. *Arthritis rheumatol.* 2016;68(1):1-26.
12. Guidelines & Protocols Advisory Committee. Rheumatoid arthritis - diagnosis, management and monitoring. In: Victoria (BC): BCGuidelines.ca; 2012: https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/bc-guidelines/rheumatoid_arthritis.pdf. Accessed 2019 Oct 04.
13. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. *J Clin Epidemiol.* 2016;75:40-46.
14. Grey matters: a practical tool for searching health-related grey literature. In: Ottawa (ON): CADTH; 2019: https://www.cadth.ca/grey-matters. Accessed 2019 Jul 24.
15. Fleischmann R, Pangan AL, Song IH, et al. Upadacitinib versus placebo or adalimumab in patients with rheumatoid arthritis and an inadequate response to methotrexate: results of a phase 3, double-blind, randomized controlled trial. *Arthritis rheumatol.* 2019.
16. Fleischmann RM, Genovese MC, Enejosa JV, et al. Safety and effectiveness of upadacitinib or adalimumab plus methotrexate in patients with rheumatoid arthritis over 48 weeks with switch to alternate therapy in patients with insufficient response. *Ann Rheum Dis.* 2019.
17. Smolen JS, Pangan AL, Emery P, et al. Upadacitinib as monotherapy in patients with active rheumatoid arthritis and inadequate response to methotrexate (SELECT-MONOTHERAPY): a randomised, placebo-controlled, double-blind phase 3 study. *Lancet.* 2019;393(10188):2303-2311.
18. Burmester GR, Kremer JM, Van den Bosch F, et al. Safety and efficacy of upadacitinib in patients with rheumatoid arthritis and inadequate response to conventional synthetic disease-modifying anti-rheumatic drugs (SELECT-NEXT): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet.* 2018;391(10139):2503-2512.
19. Genovese MC, Fleischmann R, Combe B, et al. Safety and efficacy of upadacitinib in patients with active rheumatoid arthritis refractory to biologic disease-modifying anti-rheumatic drugs (SELECT-BEYOND): a double-blind, randomised controlled phase 3 trial. *Lancet.* 2018;391(10139):2513-2524.
20. Wells G, Becker JC, Teng J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. *Ann Rheum Dis.* 2009;68(6):954-960.
21. Inoue E, Yamanaka H, Hara M, Tomatsu T, Kamatani N. Comparison of Disease Activity Score (DAS)28- erythrocyte sedimentation rate and DAS28- C-reactive protein threshold values. *Ann Rheum Dis.* 2007;66(3):407-409.
22. Anderson J, Caplan L, Yazdany J, et al. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. *Arthritis Care Res (Hoboken).* 2012;64(5):640-647.

23. Matsui T, Kuga Y, Kaneko A, et al. Disease Activity Score 28 (DAS28) using C-reactive protein underestimates disease activity and overestimates EULAR response criteria compared with DAS28 using erythrocyte sedimentation rate in a large observational cohort of rheumatoid arthritis patients in Japan. *Ann Rheum Dis.* 2007;66(9):1221-1226.
24. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: dimensions and practical applications. *Health Qual Life Outcomes.* 2003;1:20.
25. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: a review of its history, issues, progress, and documentation. *J Rheumatol.* 2003;30(1):167-178.
26. Sokka T. Radiographic scoring in rheumatoid arthritis: a short introduction to the methods. *Bull NYU Hosp Jt Dis.* 2008;66(2):166-168.
27. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. *Ann Med.* 2001;33(5):350-357.
28. 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;15(2):141-155.
29. 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;14(4):239-253.
30. Song GG, Choi SJ, Lee YH. Comparison of the efficacy and safety of tofacitinib and upadacitinib in patients with active rheumatoid arthritis: A Bayesian network meta-analysis of randomized controlled trials. *Int J Rheum Dis.* 2019.
31. CDR submission: Rinvoq (upadacitinib), 15mg once daily oral tablet [CONFIDENTIAL manufacturer's submission]. In: Pointe-Claire (QC): AbbVie Corporation; 2019 Jul 04.
32. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17(1):1-12.
33. Common technical document: upadacitinib (ABT-494) summary of clinical efficacy 2.7.3. In: CDR submission: Rinvoq (upadacitinib), 15mg once daily oral tablet [CONFIDENTIAL manufacturer's submission]. In: Pointe-Claire (QC): AbbVie Corporation; 2019 Jul 04.
34. Marker-Hermann E, Nitschmann S. Upadacitinib in patients with rheumatoid arthritis: SELECT-NEXT and SELECT-BEYOND. [German]. *Internist.* 2019;60(5):549-552.
35. Kremer JM, Emery P, Camp HS, et al. A Phase IIb study of ABT-494, a selective JAK-1 Inhibitor, in patients with rheumatoid arthritis and an inadequate response to anti-tumor necrosis factor therapy. *Arthritis Rheumatol.* 2016;68(12):2867-2877.
36. Genovese MC, Smolen JS, Weinblatt ME, et al. Efficacy and safety of ABT-494, a selective JAK-1 inhibitor, in a phase IIb study in patients with rheumatoid arthritis and an inadequate response to methotrexate. *Arthritis Rheumatol.* 2016;68(12):2857-2866.
37. Genovese MC, Kremer J, Zhong S, Friedman A. Long-term safety and efficacy of upadacitinib (ABT-494), an oral JAK-1 inhibitor in patients with rheumatoid arthritis in an open label extension study [abstract]. *Arthritis Rheumatol.* 2017;69(suppl 10).
38. Strand V, Tundia N, Friedman A. Patient-reported outcomes of long-term upadacitinib use in patients with rheumatoid arthritis: interim analysis results of a phase 2, open-label extension study [abstract]. *Arthritis Rheumatol.* 2017;69(suppl 10).
39. Cohen S, Hurd E, Cush J, et al. Treatment of rheumatoid arthritis with anakinra, a recombinant human interleukin-1 receptor antagonist, in combination with methotrexate: results of a twenty-four-week, multicenter, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum.* 2002;46(3):614-624.
40. Chung CP, Thompson JL, Koch GG, Amara I, Strand V, Pincus T. Are American College of Rheumatology 50% response criteria superior to 20% criteria in distinguishing active aggressive treatment in rheumatoid arthritis clinical trials reported since 1997? A meta-analysis of discriminant capacities. *Ann Rheum Dis.* 2006;65(12):1602-1607.
41. Center for Drug Evaluation and Research, U.S. Food and Drug Administration. Guidance for industry: clinical development programs for drugs, devices, and biological products for the treatment of rheumatoid arthritis (RA). In: Rockville (MD): Food and Drug Administration; 1999: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071579.pdf. Accessed 2019 May 28.
42. Heegaard C, Dreyer L, Egsmose C, Madsen OR. Test-retest reliability of the disease activity score 28 CRP (DAS28-CRP), the simplified disease activity index (SDAI) and the clinical disease activity index (CDAI) in rheumatoid arthritis when based on patient self-assessment of tender and swollen joints. *Clin Rheumatol.* 2013;32(10):1493-1500.
43. van Riel PL. The development of the disease activity score (DAS) and the disease activity score using 28 joint counts (DAS28). *Clin Exp Rheumatol.* 2014;32(5 Suppl 85):S-65-74.
44. Verhoeven AC, Boers M, van Der Linden S. Responsiveness of the core set, response criteria, and utilities in early rheumatoid arthritis. *Ann Rheum Dis.* 2000;59(12):966-974.
45. Linde L, Sorensen J, Ostergaard M, Horslev-Petersen K, Hetland ML. Health-related quality of life: validity, reliability, and responsiveness of SF-36, 15D, EQ-5D RAQoL, and HAQ in patients with rheumatoid arthritis. *J Rheumatol.* 2008;35(8):1528-1537.
46. Ward MM, Guthrie LC, Alba MI. Clinically important changes in short form 36 health survey scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. *Arthritis Care Res (Hoboken).* 2014;66(12):1783-1789.
47. Gaujoux-Viala C, Mouterde G, Baillet A, et al. Evaluating disease activity in rheumatoid arthritis: which composite index is best? A systematic literature analysis of studies comparing the psychometric properties of the DAS, DAS28, SDAI and CDAI. *Joint Bone Spine.* 2012;79(2):149-155.

48. Fujiwara M, Kita Y. Reexamination of the assessment criteria for rheumatoid arthritis disease activity based on comparison of the Disease Activity Score 28 with other simpler assessment methods. *Mod Rheumatol.* 2013;23(2):260-268.

49. Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol.* 2005;23(5 Suppl 39):S100-108.

50. van Riel PL, van Gestel AM. Clinical outcome measures in rheumatoid arthritis. *Ann Rheum Dis.* 2000;59 Suppl 1:i28-i31.

51. Bansback N, Ara R, Karnon J, Anis A. Economic evaluations in rheumatoid arthritis: a critical review of measures used to define health states. *Pharmacoeconomics.* 2008;26(5):395-408.

52. American College of Rheumatology Committee to Reevaluate Improvement Criteria. A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. *Arthritis Rheum.* 2007;57(2):193-202.

53. Felson DT, Furst DE, Boers M. Rationale and strategies for reevaluating the ACR20. *J Rheumatol.* 2007;34(5):1184-1187.

54. Fransen J, Antoni C, Mease PJ, et al. Performance of response criteria for assessing peripheral arthritis in patients with psoriatic arthritis: analysis of data from randomized, controlled trials of two TNF inhibitors. *Ann Rheum Dis.* 2006;65(10):1373-1378.

55. Landewe R, van der Heijde D. The validity of a rheumatoid arthritis medical records-based index of severity compared with the DAS28. *Arthritis Res Ther.* 2006;8(3):107.

56. van der Heijde DM, van 't Hof MA, van Riel PL, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis.* 1990;49(11):916-920.

57. van der Heijde DM, van 't Hof MA, van Riel PL, van Leeuwen MA, van Rijswijk MH, van de Putte LB. Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. *Ann Rheum Dis.* 1992;51(2):177-181.

58. Prevoo ML, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum.* 1995;38(1):44-48.

59. van Gestel AM, Anderson JJ, van Riel PL, et al. ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. *J Rheumatol.* 1999;26(3):705-711.

60. Ranganath VK, Khanna D, Paulus HE. ACR remission criteria and response criteria. *Clin Exp Rheumatol.* 2006;24(6 Suppl 43):S14-S21.

61. Crowson CS, Rahman MU, Matteson EL. Which measure of inflammation to use? A comparison of erythrocyte sedimentation rate and C-reactive protein measurements from randomized clinical trials of golimumab in rheumatoid arthritis. *J Rheumatol.* 2009;36(8):1606-1610.

62. Siemons L, Vonkeman HE, Ten Klooster PM, van Riel PL, van de Laar MA. Interchangeability of 28-joint disease activity scores using the erythrocyte sedimentation rate or the C-reactive protein as inflammatory marker. *Clin Rheumatol.* 2014;33(6):783-789.

63. Fernandez-Montero JV, Barreiro P, Soriano V. HIV protease inhibitors: recent clinical trials and recommendations on use. *Expert Opin Pharmacother.* 2009;10(10):1615-1629.

64. Tamhane A, Redden DT, McGwin G, Jr., et al. Comparison of the disease activity score using erythrocyte sedimentation rate and C-reactive protein in African Americans with rheumatoid arthritis. *J Rheumatol.* 2013;40(11):1812-1822.

65. Clinical and economic overview: biological response modifier agents for adults with rheumatoid arthritis. In: *(CADTH Therapeutic review).* Ottawa (ON): CADTH; 2010: http://www.cadth.ca/media/pdf/TR_RA_Clinical_and_Economic_Overview_e.pdf. Accessed 2019 May 28.

66. McKesson Pharmaclick. 2014; https://www.mckesson.ca/. Accessed 2014.

67. Ontario Ministry of Health and Long-Term Care. Exceptional Access Program (EAP). In: Toronto (ON)2014.

68. Disease activity score in 28 joints (DAS28). In: Philadelphia (PA): The Institute for Continuing Healthcare Education; 2009.

69. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum.* 1980;23(2):137-145.

70. Ramey DR, Fries JF, Singh G. The Health Assessment Questionnaire 1995 -- status and review. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials.* 2nd ed ed. Philadelphia: Lippincott-Raven Publishers; 1996:227-237.

71. Cohen JD, Dougados M, Goupille P, et al. Health assessment questionnaire score is the best predictor of 5-year quality of life in early rheumatoid arthritis. *J Rheumatol.* 2006;33(10):1936-1941.

72. Boini S, Guillemin F. Radiographic scoring methods as outcome measures in rheumatoid arthritis: properties and advantages. *Ann Rheum Dis.* 2001;60(9):817-827.

73. Guillemin F, Billot L, Boini S, Gerard N, Odegaard S, Kvien TK. Reproducibility and sensitivity to change of 5 methods for scoring hand radiographic damage in patients with rheumatoid arthritis. *J Rheumatol.* 2005;32(5):778-786.

74. Landewe R, van der Heijde D. Radiographic progression in rheumatoid arthritis. *Clin Exp Rheumatol.* 2005;23(5 Suppl 39):S63-S68.

75. Voskuyl AE, Dijkmans BA. Remission and radiographic progression in rheumatoid arthritis. *Clin Exp Rheumatol.* 2006;24(6 Suppl 43):S37-S40.

76. Drossaers-Bakker KW, de Buck M, van Zeben D, Zwinderman AH, Breedveld FC, Hazes JM. Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. *Arthritis Rheum.* 1999;42(9):1854-1860.

77. Sahin F, Kotevoglu N, Taspinar S, Yilmaz F, Kuran B. Comparison of functional disability scales and their relevance to radiological progression in patients with rheumatoid arthritis in remission. *Clin Exp Rheumatol.* 2006;24(5):540-545.

78. Swinkels HL, Laan RF, van 't Hof MA, van der Heijde DM, de Vries N, van Riel PL. Modified sharp method: factors influencing reproducibility and variability. *Semin Arthritis Rheum.* 2001;31(3):176-190.

79. Foley-Nolan D, Stack JP, Ryan M, et al. Magnetic resonance imaging in the assessment of rheumatoid arthritis--a comparison with plain film radiographs. *Br J Rheumatol.* 1991;30(2):101-106.

80. Bruynesteyn K, van der Heijde D, Boers M, et al. Determination of the minimal clinically important difference in rheumatoid arthritis joint damage of the Sharp/van der Heijde and Larsen/Scott scoring methods by clinical experts and comparison with the smallest detectable difference. *Arthritis Rheum.* 2002;46(4):913-920.

81. Landewe R, van der Heijde D. Is radiographic progression a realistic outcome measure in clinical trials with early inflammatory arthritis? *Clin Exp Rheumatol.* 2003;21(5 Suppl 31):S37-S41.

82. Navarro-Compan V, van der HD, Ahmad HA, Miller CG, Wolterbeek R, Landewe R. Measurement error in the assessment of radiographic progression in rheumatoid arthritis (RA) clinical trials: the smallest detectable change (SDC) revisited. *Ann Rheum Dis.* 2014;73(6):1067-1070.

83. Gallagher EJ, Liebman M, Bijur PE. Prospective validation of clinically important changes in pain severity measured on a visual analog scale. *Ann Emerg Med.* 2001;38(6):633-638.

84. Frendl DM, Ware JE, Jr. Patient-reported functional health and well-being outcomes with drug therapy: a systematic review of randomized trials using the SF-36 health survey. *Med Care.* 2014;52(5):439-445.

85. Smolen JS, van der Heijde D, Machold KP, Aletaha D, Landewe R. Proposal for a new nomenclature of disease-modifying antirheumatic drugs. *Ann Rheum Dis.* 2014;73(1):3-5.

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

87. Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J. Validation of the functional assessment of chronic illness therapy fatigue scale relative to other instrumentation in patients with rheumatoid arthritis. *J Rheumatol.* 2005;32(5):811-819.

88. Maxwell LJ, Zochling J, Boonen A, et al. TNF-alpha inhibitors for ankylosing spondylitis. *Cochrane Database Syst Rev.* 2015(4):CD005468.