Common Drug Review
Clinical Review Report

January 2018

Drug	tofacitinib (Xeljanz) tablets
Indication	Tofacitinib, in combination with methotrexate (MTX), is indicated for reducing the signs and symptoms of rheumatoid arthritis (RA), in adult patients with moderately to severely active RA who have had an inadequate response to MTX. In cases of intolerance to MTX, physicians may consider the use of Xeljanz (tofacitinib) as monotherapy.
Listing request	Patients with moderately to severely active RA in a similar manner to the tumour necrosis factor (TNF) alpha inhibitors
Manufacturer	Pfizer Canada Inc.

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

Redactions: Confidential information in this document has been redacted at the request of the manufacturer in accordance with the *CADTH Common Drug Review Confidentiality Guidelines.*

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.

TABLE OF CONTENTS

Tables

Figures

ABBREVIATIONS

ACR	American College of Rheumatology
ALT	alanine transaminase
ANC	absolute neutrophil count
AST	aspartate transaminase
BOCF	baseline observation carried forward
BRM	biologic response modifier
CCP	cyclic citrullinated peptide
CDR	CADTH Common Drug Review
CK	creatine kinase
CRA	Canadian Rheumatology Association
CRP	C-reactive protein
DAS28-3 (CRP/ESR)	Disease Activity Score (including C-reactive protein or erythrocyte sedimentation rate)
DMARD	disease-modifying antirheumatic drug
EMEA	European Medicines Agency
EQ-5D	EuroQol Five-Dimension Health-Related Quality of Life Questionnaire
ESR	erythrocyte sedimentation rate
FACIT	Functional Assessment of Chronic Illness Therapy
FAS	full analysis set
FDA	Food and Drug Administration
HAQ-DI	Health Assessment Questionnaire–Disability Index
HDL	high-density lipoprotein
IL	interleukin
ITT	intention-to-treat
JAK	Janus kinase
JSN	joint space narrowing
LDL	low-density lipoprotein
LOCF	last observation carried forward
MDC	mean difference of change
MOS-SS	Medical Outcome Study Sleep Scale
mTSS	modified Total Sharp Score
MTX	methotrexate
NA	not applicable
NRI	non-responder imputation
PGA	physician global assessment
PP	per-protocol
PtGA	patient global assessment
RA	rheumatoid arthritis
SAE	serious adverse event
SF-36	Short-Form (36) Health Survey
TNF	tumour necrosis factor
TyK	tyrosine kinase
VAS	visual analogue scale
WLQ	Work Limitations Questionnaire

EXECUTIVE SUMMARY

Introduction

The objective of this report is to perform a systematic review of the beneficial and harmful effects of tofacitinib 5 mg twice daily for the treatment of rheumatoid arthritis (RA) in adults, alone or in combination with methotrexate (MTX). Tofacitinib (Xeljanz) is a Janus kinase (JAK) inhibitor that primarily targets JAK1, JAK2, and JAK3 (there is less inhibition of tyrosine kinase 2 [TyK2]).[1] These kinases are involved in the signal transduction process that leads to the production of cytokines, including interleukin (IL)-2, IL-4, IL-6, IL-7, IL-15, and IL-21, which are key components of the inflammatory process.[1] It is administered orally.

Rheumatoid arthritis is a chronic inflammatory disease characterized by joint swelling, joint tenderness, and destruction of synovial joints, leading to severe disability and premature mortality.[2] According to a report by the Arthritis Alliance of Canada, RA is the most common inflammatory joint disease, with a prevalence of 0.9% in 2010 (272,299 patients).

Indication under review
Xeljanz (tofacitinib), in combination with MTX, is indicated for reducing the signs and symptoms of RA, in adult patients with moderately to severely active RA who have had an inadequate response to MTX. In cases of intolerance to MTX, physicians may consider the use of Xeljanz (tofacitinib) as monotherapy.
Listing criteria requested by sponsor
Patients with moderately to severely active RA in a similar manner to the tumour necrosis factor (TNF) alpha inhibitors.

Results and Interpretation

Included Studies

Five manufacturer-sponsored, published, double-blind randomized controlled trials evaluating the efficacy and harms of tofacitinib 5 mg or 10 mg twice daily and adalimumab (only in Study 1064) versus placebo were included in the systematic review. Data from the 10 mg twice daily treatment groups are not presented in this report. The total sample size across the five trials was 3,322 patients, including 1,216 patients who were randomized to the tofacitinib approved dose of 5 mg twice daily. One study was performed in patients with previous TNF inhibitor inadequate response (Study 1032) whereas the other trials were performed in patients with previous disease-modifying antirheumatic drug (DMARD) or MTX inadequate response (Studies 1044, 1045, 1046, and 1064). Tofacitinib was given as monotherapy in one study (Study 1045) and with background DMARDs in the other studies (studies 1032, 1044, 1046, and 1064). The double-blind periods of the studies were six months to 24 months in duration, but the placebo-controlled periods were not longer than three months prior to early escape.

All five trials used the same three measures for the co-primary end points. The co-primary end point of American College of Rheumatology (ACR) 20 response was evaluated at month 3 (studies 1032 and 1045) or month 6 (studies 1044, 1046, and 1064). The co-primary end point of change from baseline for the Health Assessment Questionnaire–Disability Index (HAQ-DI) was evaluated at month 3 in all trials. The co-primary end point of Disease Activity Score (DAS) 28-4 (erythrocyte sedimentation rate [ESR]) was evaluated at month 3 (studies 1032 and 1045) or month 6 (studies 1044, 1046, and 1064). Study

1044 used a fourth co-primary end point of modified Total Sharp Score (mTSS). No other studies measured radiographic outcomes.

The studies allowed early escape or crossover to tofacitinib at month 3, and all patients taking placebo were given tofacitinib by month 6 at the latest. This design has numerous limitations, including the fact that patients who meet early escape are not randomized to dose escalation or another type of strategy. Early escape, while common in RA trials based on ethical considerations, limits the interpretation and clinical relevance of the trial data.

Efficacy

In all five studies, there was a statistically significant greater proportion of ACR 20, ACR 50, and ACR 70 responders at month 3 or month 6 in the tofacitinib 5 mg twice daily group compared with placebo (see Table 1). Response rates in the tofacitinib treatment groups ranged from 42% to 60% (ACR 20), 27% to 37% (ACR 50), and 13% to 20% (ACR 70). Response rates for placebo ranged from 25% to 31% (ACR 20), 8% to 13% (ACR 50), and 1% to 6% (ACR 70). There were also statistically significant improvements in the mean HAQ-DI score for tofacitinib 5 mg twice daily compared with placebo in four of five studies, with mean difference of change (MDC) between 0.2 and 0.3 points. DAS28-4(ESR) response rates were modest (6% to 9% in tofacitinib 5 mg twice daily treatment groups), and three of five studies showed statistically significant improvements for tofacitinib 5 mg twice daily versus placebo. There was no statistically significant difference in mTSS in the single study that measured this outcome (Study 1044). Statistical step-down testing of the co-primary end points stopped prematurely in this study because of this non-statistically significant finding in the mTSS (see Figure 3, APPENDIX 4: Detailed Outcome Data).

Secondary outcomes also showed statistically significant differences favouring tofacitinib 5 mg twice daily versus placebo in the five trials, including the Short-Form (36) Health Survey (SF-36) mental component summary (MDC 1.87 to 4.62) and physical component summary (MDC 1.73 to 4.16) and the Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue scale (MDC approximately 3 to 5 points). There were numerical improvements in almost all individual components of the ACR, including pain scores, for tofacitinib compared with placebo.

Harms

The only placebo-controlled data uncontaminated by patient early escape are from month 3 for all five studies. There were no obvious trends observed at this time point, but many safety concerns have emerged during the longer-term usage of tofacitinib. These were well documented in the product monograph.[1] The highlighted concerns include serious infections (e.g., tuberculosis, cryptococcus, esophageal candidiasis, pneumocystosis, multidermatomal herpes zoster, cytomegalovirus, BK virus infections, cellulitis, and urinary tract infections), lymphoma and other malignancies, heart rate decrease and PR interval prolongation, gastrointestinal perforation, liver enzyme elevations and drug-induced liver injury, interstitial lung disease, lymphopenia, neutropenia, and lipid elevations. The European Medicines Agency (EMEA) cited risk of harm and the difficulty in managing this risk in clinical practice as reasons for rejecting tofacitinib for market approval in Europe in April 2013.[3] Tofacitinib is approved for use in the US, Japan, and other countries.

While the manufacturer did report incidence of adverse events in long-term extension studies, and these data have some value in understanding the risks associated with tofacitinib compared with adalimumab (Study 1064), there was no placebo control group (see APPENDIX 6: Summary of Extension Studies). The randomized trials excluded patients who were at increased risk of developing specific adverse events associated with the use of tofacitinib (e.g., serious infections) and thus may not reflect

the incidence in clinical practice. The harms profile of tofacitinib in patients with RA deserves further study in long-term active-controlled and observational studies.

Conclusions

In five double-blind randomized controlled trials in patients with active RA, tofacitinib 5 mg twice daily was associated with higher rates of ACR 20, ACR 50, and ACR 70 response compared with placebo (with or without background DMARDs). Other outcomes, such as the HAQ-DI, DAS28 response, SF-36, and the FACIT-Fatigue scale, also showed statistically significant improvements favouring tofacitinib versus placebo at month 3 or month 6. Radiologic progression did not reach statistical significance in the single study that measured this outcome. Analyses needed to take into account the fact that many patients on placebo had early escape and this may have weakened the internal validity of the results.

There is risk of serious harm such as malignancies and infections with tofacitinib, similar to the risk for anti-TNF alpha drugs used to treat RA. Further research will be needed to ascertain the risks relative to other commonly used biologic drugs and non-biologic DMARDs.

TABLE 1: SUMMARY OF RESULTS

	1032 Month 3		1045 Month 3		1044 Month 6		1046 Month 6		1064 Month 6		
Unique Study Feature	**TNFi Failure Population**		**TOF Monotherapy**		**Radiographic Outcomes**				**ADA Treatment Group**		
	TOF 5 mg	**PL**	**TOF 5 mg**	**PL**	**TOF 5 mg**	**PL**	**TOF 5 mg**	**PL**	**TOF 5 mg**	**PL**	**ADA**
Randomized	133	132	243	122	321	160	315	159	204	108	204
Completed, n (%)	107 (81)	101 (75)	232 (95)	105 (86)	212 (66)	107 (67)	261 (82)	138 (86)	150 (74)	86 (80)	162 (79)
Total discontinued, n (%)[a]	26 (20)	31 (23)	11 (5)	17 (14)	109 (34)	53 (33)	54 (17)	21 (13)	54 (26)	22 (20)	21 (42)
WDAE, n (%)	12 (9)	6 (8)	3 (1)	5 (4)	41 (13)	12 (7)	22 (7)	5 (3)	125 (12)	7 (7)	25 (12)
ACR 20, n/N (%)	55/132 (42) *P* = 0.0025	32/131 (25)	144/241 (60) *P* < 0.0001	32/120 (27)	159/309 (52) *P* < 0.0001	39/154 (26)	164/311 (53) *P* < 0.0001	49/157 (31)	101/196 (52) *P* < 0.0001	30/106 (28)	94/199 (47) *P* = 0.0008
ACR 50, n/N (%)	35/132 (27) *P* < 0.0001	11/131 (8)	75/241 (31) *P* < 0.0001	15/120 (13)	100/309 (32) *P* < 0.0001	13/154 (8)	105/311 (34) *P* < 0.0001	20/157 (13)	72/196 (37) *P* < 0.0001	13/106 (12)	55/199 (28) *P* = 0.0006
ACR 70, n/N (%)	18/132 (13) *P* < 0.0001	2/131 (2)	37/241 (15) *P* = 0.0026	7/120 (6)	45/309 (15) *P* < 0.0001	2/154 (1)	41/311 (13) *P* < 0.0001	5/157 (3)	39/196 (20) *P* < 0.0001	2/106 (2)	18/199 (9) *P* = 0.0031
HAQ-DI MCFB	–0.5 *P* = 0.0002	–0.2	–0.5 *P* < 0.0001	–0.2	–0.4 NSS[b]	–0.2	–0.5 *P* < 0.0001	–0.2	–0.6 *P* < 0.0001	–0.3	–0.5
DAS28-4(ESR) < 2.6, n/N (%)	8/119 (7) *P* = 0.0497	2/120 (2)	13/232 (6) *P* = 0.618	5/114 (4)	19/265 (7) NSS[b]	2/129 (1)	24/263 (9) *P* = 0.004	4/148 (3)	11/177 (6) *P* = 0.015	1/92 (1)	12/178 (7) *P* = 0.009
mTSS, MCFB	NA	NA	NA	NA	0.12 *P* = 0.079	0.47	NA	NA	NA	NA	NA

ACR = American College of Rheumatology; ADA = adalimumab; DAS28-4(ESR) = Disease Activity Score (erythrocyte sedimentation rate); HAQ-DI = Health Assessment Questionnaire–Disability Index; MCFB = mean change from baseline; mTSS = modified Total Sharp Score; NA = not applicable; NSS = not statistically significant; PL = placebo; TNFi = tumour necrosis factor inhibitor; TOF = tofacitinib; WDAE = withdrawals due to adverse events.

[a] At end of study.

[b] NSS here means that the trial could not claim statistical significance because of the step-down statistical testing procedure.

Source: Manufacturer's Clinical Study Reports and publications for studies 1032,[4-6] 1045,[7,8] 1044,[9-11] 1046,[12,13] 1064,[14,15] and Food and Drug Administration Medical Review.[16]

1. INTRODUCTION

1.1 Disease Prevalence and Incidence

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterized by joint swelling, joint tenderness, and destruction of synovial joints, leading to severe disability and premature mortality.[2] According to a report by the Arthritis Alliance of Canada, RA is the most common inflammatory joint disease, with a prevalence of 0.9% in 2010 (272,299 patients), which is expected to increase to an estimated 1.3% (549,218 patients) of the Canadian population by 2040. More than one-half of all new RA cases occur between the ages of 40 and 70 years, although all age groups are affected, and the prevalence is approximately two times higher among women than among men.[17]

1.2 Standards of Therapy

1.2.1 Non-pharmacological Management

Guidelines for the management of RA emphasize the use of non-drug interventions in addition to pharmacological therapy.[18,19] Some modalities included in non-drug care are exercise therapy, electro-physical modalities, orthoses and assistive devices, and self-management interventions. There is evidence to support the utility of non-drug care to achieve symptomatic relief, including pain control and muscle stimulation, relief of strain or load on a joint, improved patterns of motion and function, and prevention of deformity, without detrimental effects on disease activity.[19] Education on self-management strategies, such as joint protection and energy conservation, exercises, or the use of assistive devices, equips RA patients with tools to cope with the disease.[19]

1.2.2 Pharmacological Management

The goal of RA treatment is to achieve remission and, when that is not possible, to minimize disease activity while controlling symptoms, halting damage, preventing disability, and improving quality of life.[18] Beginning treatment early and aggressively with non-biologic, synthetic, disease-modifying antirheumatic drugs (DMARDs) has been shown to alter the clinical course of RA and slow or halt radiographic progression.[18]

Methotrexate (MTX) is the preferred DMARD with respect to efficacy and safety and is usually the first-line DMARD in patients with RA unless contraindicated. Therapy with MTX is individualized, with doses rapidly titrated to a usual maximum dose of 25 mg per week for intramuscular or intravenous use, and 20 mg per week for oral use.[18] The Canadian Rheumatology Association (CRA) recommends parenteral administration of MTX in patients with an inadequate response or intolerance to oral MTX.[18] The initial treatment strategy with DMARDs can include nonsteroidal anti-inflammatory drugs or glucocorticoids (in the lowest effective dose possible), or both, as bridge therapy while waiting for DMARDs to take effect, to manage flares, or for symptom control if no other options exist.[18]

Currently, all Canadian provincial formularies require failure of at least two DMARDs prior to accessing a biologic response modifier (BRM), and many also require failure of an adequate trial of combination DMARD therapy.[18] Methotrexate is the preferred anchor drug in combination therapy with conventional DMARDs, unless contraindicated.[18] The CRA defines inadequate response to DMARD as moderate to high disease activity despite treatment with at least two DMARDs (including MTX unless contraindicated) in monotherapy or combination therapy after three months at target doses. The CRA provides a reference guide for the definition of moderate to high disease activity according to disease activity measure selected (e.g., DAS moderate > 2.4 to 3.6, high > 3.6; DAS28 moderate > 3.2 to 5.1, high > 5.1) with the outcome measure selection at the discretion of the rheumatologist.[18]

Most BRMs currently approved for use in Canada belong to the tumour necrosis factor (TNF) inhibitors class and include etanercept, infliximab, adalimumab, golimumab, and certolizumab pegol. Other approved BRMs are abatacept (T cell co-stimulatory inhibitor), rituximab (B lymphocyte-depleting drug), tocilizumab (interleukin [IL]-6 antagonist), and anakinra (IL-1 antagonist).[18] Although co-administration of MTX with BRMs (i.e., adalimumab, certolizumab, etanercept, abatacept, tocilizumab, and tofacitinib) is recommended for improved efficacy, each has an indication for use as monotherapy.[1,18,20-23] This is an important distinction, as not all patients tolerate MTX. In recently diagnosed patients who have not been previously treated with MTX, abatacept is to be used in combination with MTX.[22]

Based on the CRA guidelines,[18] patients who have failed treatment with one or two TNF inhibitors as a result of lack of efficacy or toxicity could be switched to another TNF inhibitor or another BRM with a different mechanism of action (Table 2). Both abatacept and tocilizumab are indicated for the treatment of patients with RA who have had an inadequate response to one or more DMARDs or TNF inhibitors or to both.[20,22] Rituximab, in combination with MTX, is indicated in RA patients who have had an inadequate response or intolerance to one or more TNF inhibitors.[23] In situations of inadequate response to a TNF inhibitor used as monotherapy, adding MTX or other DMARDs is recommended.[18]

According to the CRA recommendations, patients with active RA should be monitored every one to three months, and non-biologic and biologic DMARD therapy should be adjusted every three to six months if treatment targets have not been achieved.[18]

1.3 Drug

Xeljanz (tofacitinib) is a Janus kinase (JAK) inhibitor that primarily targets JAK1, JAK2, and JAK3 (there is less inhibition of tyrosine kinase 2 [TyK2]).[1] These kinases are involved in the signal transduction process that leads to the production of cytokines, including interleukin (IL)-2, IL-4, IL-6, IL-7, IL-15, and IL-21, which are key components of the inflammatory process.[1] The active ingredient is tofacitinib citrate. Tofacitinib has been approved by Health Canada for reducing the signs and symptoms of RA in adult patients with moderately to severely active RA who have had an inadequate response to MTX. It is approved for use as a 5 mg oral tablet taken twice daily either alone (in cases of intolerance to MTX) or in combination with MTX.[1]

Indication under review
Xeljanz (tofacitinib), in combination with MTX, is indicated for reducing the signs and symptoms of RA in adult patients with moderately to severely active RA who have had an inadequate response to MTX. In cases of intolerance to MTX, physicians may consider the use of Xeljanz (tofacitinib) as monotherapy.
Listing criteria requested by sponsor
Patients with moderately to severely active RA in a similar manner to the TNF alpha inhibitors.

TABLE 2: KEY CHARACTERISTICS OF BIOLOGIC DRUGS FOR RHEUMATOID ARTHRITIS AND TOFACITINIB

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

DMARD = disease-modifying antirheumatic drug; IL = interleukin; IV = intravenous; JAK = Janus kinase; MTX = methotrexate; SC = subcutaneous injection; 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).
[b] If patient is intolerant to MTX.
[c] If first-line treatment, give with MTX.
[d] The DMARD used is usually MTX.
[e] Other DMARDs may also be used.
Source: Health Canada product monographs.

2. OBJECTIVES AND METHODS

2.1 Objectives

To perform a systematic review of the beneficial and harmful effects of tofacitinib (Xeljanz) 5 mg orally twice daily for the treatment of moderately to severely active RA, with or without MTX, in adults who have had an inadequate response to MTX.

2.2 Methods

Studies selected for inclusion in the systematic review included the pivotal studies provided in the manufacturer's submission to the CADTH Common Drug Review (CDR) as well as those meeting the selection criteria presented in Table 3.

TABLE 3: INCLUSION CRITERIA FOR THE SYSTEMATIC REVIEW

Patient Population	Adults with moderately to severely active RA Subgroups of interest based on: • Taking concomitant MTX (versus not taking MTX) • Previously used BRMs (versus no prior BRM use) • Dose of concomitant MTX • Age • Body weight at baseline • Disease severity at baseline • Rheumatoid factor status • Cyclic citrullinated peptide antibody status
Intervention	Tofacitinib alone or in combination with MTX administered orally at recommended doses
Comparators	Individual or combination therapy with: • BRMs (e.g., infliximab, abatacept, rituximab, etanercept, adalimumab, certolizumab, golimumab, tocilizumab) or • Other DMARDs including MTX
Outcomes	**Key efficacy outcomes** Radiographic changes ACR response QoL, functional and disability outcomes Disease activity Health care resource utilization **Harms outcomes** Mortality, SAEs, AEs (including bacterial and viral infections and relation to neutrophil counts, malignancies, lipids, gastrointestinal perforation, cardiovascular morbidity), WDAEs
Study Design	Published and unpublished phase 3 RCTs

ACR = American College of Rheumatology; AE = adverse events; BRM = biologic response modifiers; DMARD = disease-modifying antirheumatic drug; MTX = methotrexate; QoL = quality of life; RA = rheumatoid arthritis; RCT = randomized controlled trial; SAE = serious adverse events; WDAE = withdrawal due to adverse events.

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

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

(Medical Subject Headings), and keywords. The main search concepts were Xeljanz (tofacitinib) and Rheumatoid Arthritis.

Methodological filters were not applied. Where possible, retrieval was limited to the human population. Retrieval was not limited by publication year or by language. Conference abstracts were excluded from the search results.

The initial search was completed on October 14, 2014. Regular alerts were established to update the search until the meeting of the Canadian Drug Expert Committee on February 18, 2015. Regular search updates were performed on databases that do not provide alert services.

Grey literature (literature that is not commercially published) was identified by searching relevant websites from the following sections of the Grey Matters checklist (http://www.cadth.ca/en/resources/finding-evidence-is/grey-matters): Health Technology Assessment Agencies, Health Economics, Clinical Practice Guidelines, Drug and Device Regulatory Approvals, Advisories and Warnings, Drug Class Reviews, Databases (free), Internet Search. Google and other Internet search engines were used to search for additional web-based materials. These searches were supplemented by reviewing the bibliographies of key papers and through contacts with appropriate experts. In addition, the manufacturer of the drug was contacted for information regarding unpublished studies.

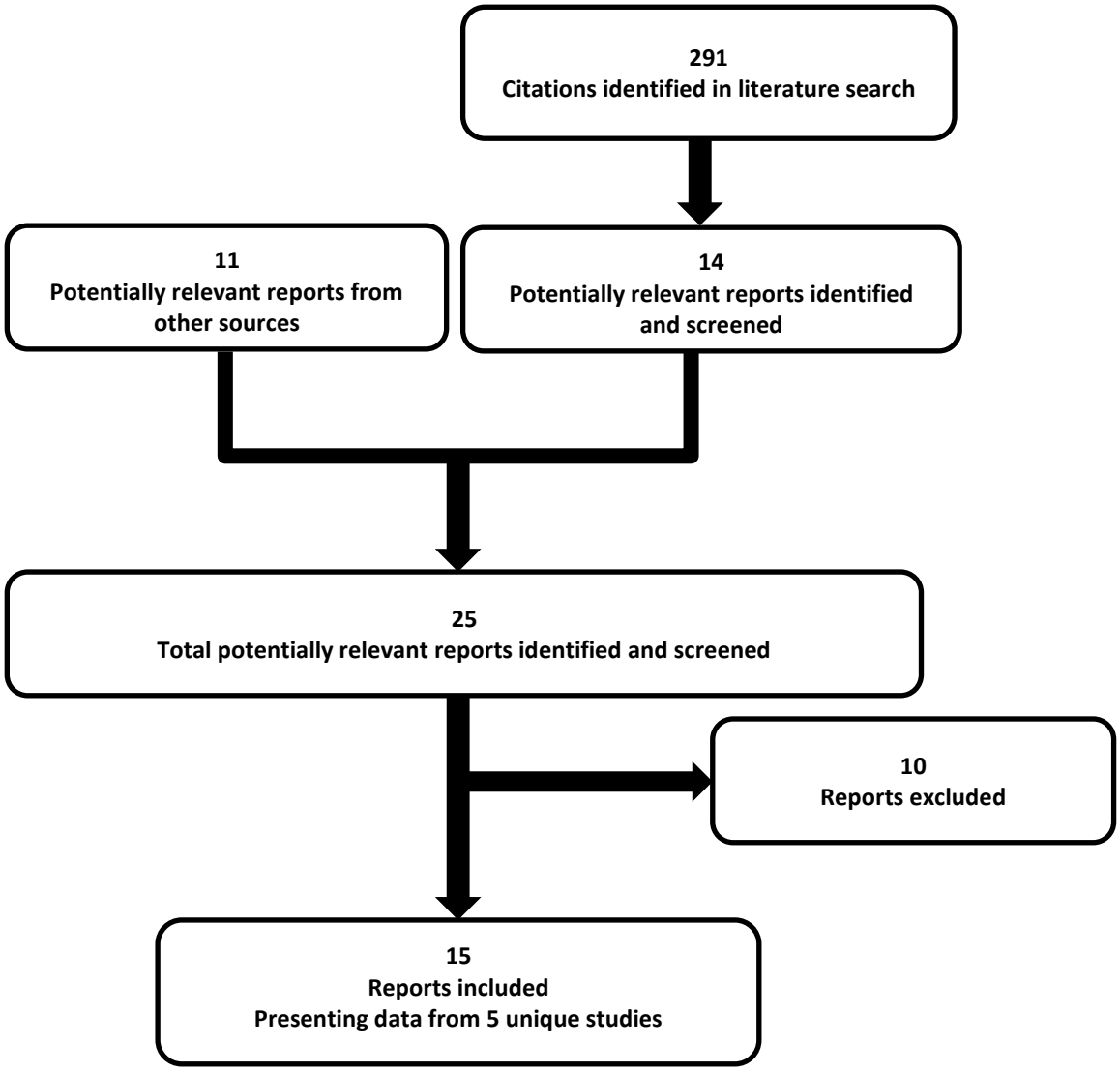
Two CDR clinical reviewers independently selected studies for inclusion in the review based on titles and abstracts, according to the predetermined protocol. Full-text articles of all citations considered potentially relevant by at least one reviewer were acquired. Reviewers independently made the final selection of studies to be included in the review, and differences were resolved through discussion. Included studies are presented in Table 4; excluded studies (with reasons) are presented in APPENDIX 3: Excluded Studies).

3. RESULTS

3.1 Findings From the Literature

A total of five studies were identified from the literature for inclusion in the systematic review (Figure 1). The included studies are summarized in Table 2 and described in Section 3.2. A list of excluded studies is presented in APPENDIX 3: Excluded Studies).

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



QUOROM = Quality of Reporting of Meta-analyses.

TABLE 4: DETAILS OF INCLUDED STUDIES

		6 Month Study Duration		≥ 12 Month Study Duration		
		1032	1045	1044	1046	1064
DESIGNS & POPULATIONS	**Study design**	DB RCT, parallel group; randomized 2:2:1:1[a]	DB RCT, parallel group; randomized 4:4:1:1	DB RCT, parallel group; randomized 4:4:1:1	DB RCT, parallel group; randomized 4:4:1:1	DB RCT, parallel group; randomized 4:4:1:1:4
	Background DMARD	MTX (mean dose 15.4 mg/week)	None	MTX (mean dose 15.1 mg/week)	DMARDs (mean MTX 14.0 mg/week)	MTX (mean dose 15.4 mg/week)
	Unique feature	TNFi failure population	TOF monotherapy	Radiographic outcomes		ADA treatment group
	Population	TNFi failure	DMARD failure	MTX failure	DMARD failure	MTX failure
	Locations	82 sites (Australia, Americas, Europe, Asia)	94 sites (Americas, Europe, India, Asia)	111 sites (Americas, Europe, India, Asia, Australia)	114 sites (Americas, Europe, Asia, Australia)	115 sites (Americas, Europe, Asia, Australia)
	Randomized (N)	399	610	800	795	717
	Inclusion criteria	Active RA (based on TJC and SJC) and ongoing treatment with an adequate and stable dose of MTX; must have: ESR > 28 mm/hour and/or CRP > 7 mg/L.	Active RA (based on TJC and SJC) and ongoing treatment with an adequate and stable dose of MTX; must have: ESR > 28 mm/hour and/or CRP > 7 mg/L.	Active RA (based on TJC and SJC) and ongoing treatment with an adequate and stable dose of MTX; must have: ESR > 28 mm/hour and/or CRP > 7 mg/L.	Active RA (based on TJC and SJC) and ongoing treatment with an adequate and stable dose of a traditional DMARD; must have: ESR > 28 mm/hour and/or CRP > 7 mg/L.	Active RA (based on TJC and SJC) and ongoing treatment with an adequate and stable dose of MTX; must have: ESR > 28 mm/hour and/or CRP > 7 mg/L.
	Exclusion criteria	Active or latent TB; blood dyscrasias, or other severe acute or chronic medical or psychiatric condition or lab abnormality.	Active or latent TB; blood dyscrasias, or other severe acute or chronic medical or psychiatric condition or lab abnormality.	Active or latent TB; blood dyscrasias, or other severe acute or chronic medical or psychiatric condition or lab abnormality.	Active or latent TB; blood dyscrasias, or other severe acute or chronic medical or psychiatric condition or lab abnormality.	Active or latent TB; blood dyscrasias, or other severe acute or chronic medical or psychiatric condition or lab abnormality; failed treatment with any TNFi.

		6 Month Study Duration		≥ 12 Month Study Duration		
		1032	1045	1044	1046	1064
DRUGS	**Intervention**	TOF 5 mg b.i.d. TOF 10 mg b.i.d.	TOF 5 mg b.i.d. TOF 10 mg b.i.d.	TOF 5 mg b.i.d. TOF 10 mg b.i.d.	TOF 5 mg b.i.d. TOF 10 mg b.i.d.	TOF 5 mg b.i.d. TOF 10 mg b.i.d.
	Comparator(s)	PL → TOF 5 mg b.i.d. PL → TOF 10 mg b.i.d. (all PL patients switched to TOF at month 3, remained blinded)	PL → TOF 5 mg b.i.d. PL → TOF 10 mg b.i.d. (all PL patients switched to TOF at month 3, remained blinded)	PL → TOF 5 mg b.i.d. PL → TOF 10 mg b.i.d. (all PL patients switched to TOF at month 6 but early escape to TOF at month 3 was allowed if < 20% improvement in TJC and SJC, remained blinded)	PL → TOF 5 mg b.i.d. PL → TOF 10 mg b.i.d. (all PL patients switched to TOF at month 6 but early escape to TOF at month 3 was allowed if < 20% improvement in TJC and SJC, remained blinded)	PL → TOF 5 mg b.i.d. PL → TOF 10 mg b.i.d. Adalimumab (all PL patients switched to TOF at month 6 but early escape to TOF at month 3 was allowed if < 20% improvement in TJC and SJC, remained blinded)
DURATION	**Phase**					
	PL-controlled phase	0 to 3 months	0 to 3 months	0 to 6 months	0 to 6 months	0 to 6 months
	Double-blind	0 to 6 months	0 to 6 months	0 to 24 months	0 to 12 months	0 to 12 months
OUTCOMES	**Primary end point(s)**	(1) ACR 20 at month 3; (2) HAQ-DI at month 3; (3) proportion of patients with a DAS28-4(ESR) < 2.6 at month 3	(1) ACR 20 at month 3; (2) HAQ-DI at month 3; (3) proportion of patients with a DAS28-4(ESR) < 2.6 at month 3	(1) ACR 20 at month 6; (2) HAQ-DI at month 3; (3) proportion of patients with a DAS28-4(ESR) < 2.6 at month 6 and (4) mTSS at month 6	(1) ACR 20 at month 6; (2) HAQ-DI at month 3; (3) proportion of patients with a DAS28-4(ESR) < 2.6 at month 6	(1) ACR 20 at month 6; (2) HAQ-DI at month 3; (3) proportion of patients with a DAS28-4(ESR) < 2.6 at month 6
	Other end points	Other time points for ACR 20/50/70, DAS28-3(CRP), DAS28-4(ESR), HAQ-DI; Pain; PtGA; PGA; SF-36; MOS-SS; FACIT-Fatigue; EQ-5D; HCRU Questionnaire;	Other time points for ACR 20/50/70, DAS28-3(CRP), DAS28-4(ESR), HAQ-DI; Pain; PtGA; PGA; SF-36; MOS-SS; FACIT-Fatigue; EQ-5D; HCRU Questionnaire; WLQ; AE (including serious infections)	Other time points for ACR 20/50/70 and components of ACR; mTSS score at other time points, PtGA; PGA; SF-36; MOS-SS; FACIT-Fatigue; EQ-5D; HCRU Questionnaire; WLQ; AE (including serious infections)	Other time points for ACR 20/50/70 and components of ACR; mTSS score at other time points, PtGA; PGA; SF-36; MOS-SS; FACIT-Fatigue; EQ-5D; HCRU Questionnaire; WLQ; AE (including serious infections)	Other time points for ACR 20/50/70 and components of ACR; mTSS score at other time points, PtGA; PGA; SF-36; MOS-SS; FACIT-Fatigue; EQ-5D; HCRU Questionnaire; WLQ; AE (including serious infections)

		6 Month Study Duration		≥ 12 Month Study Duration		
		1032	1045	1044	1046	1064
		WLQ; AE (including serious infections)				
NOTES	**Publications**	Burmester et al.[5] Strand et al.[6]	Fleischmann et al.[8]	van der Heidje et al.[11]	Kremer et al.[13]	van Vollenhoven et al.[15]

ACR = American College of Rheumatology; ADA = adalimumab; AE = adverse events; b.i.d. = twice daily; CRP = C-reactive protein; DAS28-3(CRP/ESR) = Disease Activity Score (including C-reactive protein or erythrocyte sedimentation rate); DB = double-blind; DMARD = disease-modifying antirheumatic drug; EQ-5D = EuroQol Five-Dimensions Health-Related Quality of Life Questionnaire; ESR = erythrocyte sedimentation rate; FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy-Fatigue Scale; HAQ-DI = Health Assessment Questionnaire–Disability Index; HCRU = health care resource utilization; MOS-SS = Medical Outcome Study Sleep Scale; mTSS = modified Total Sharp Score; MTX = methotrexate; PGA = physician global assessment; PL = placebo; PtGA = patient global assessment; RA = rheumatoid arthritis; RCT = randomized controlled trial; SF-36 = Short-Form (36) Health Survey; SJC = swollen joint count; TB = tuberculosis; TJC = tender joint count; TNFi = tumour necrosis factor inhibitor; TOF = tofacitinib; WLQ = Work Limitations Questionnaire.

[a] Randomization ratios are TOF 5 mg:TOF 10 mg : PL → TOF 5 mg: PL → TOF 10 mg: adalimumab (for Study 1064).

Note: Three additional reports were included: Food and Drug Administration (FDA) Medical Review,[16] FDA Statistical Review Report,[24] manufacturer's submission binder.[25]

Source: Manufacturer's Clinical Study Reports and publications for studies 1032,[4-6] 1045,[7,8] 1044,[9-11] 1046,[12,13] and 1064.[14,15]

3.2 Included Studies

3.2.1 Description of Studies

Five double-blind, placebo-controlled randomized controlled trials met the inclusion criteria for this systematic review. No information regarding stratification at randomization was provided. Adalimumab was used as an active comparator only in Study 1064. Randomization was unequal in the trials, with fewer patients randomized initially to placebo than to tofacitinib. Early escape was permitted at month 3 for patients initially randomized to placebo. At month 3 (studies 1032 and 1045) and month 6 (studies 1044, 1046, and 1064), all patients initially randomized to placebo were switched to tofacitinib.

The duration of the placebo-controlled phases of the trials was either three months (studies 1032 and 1045) or six months (studies 1044, 1046, and 1064). The double-blind phases of the trials ranged from six to 24 months.

3.2.2 Populations

a) Inclusion and Exclusion Criteria

Patients were required to meet the following criteria for all studies:

1. ACR classification criteria for the diagnosis of RA by satisfying at least four of the seven criteria
2. Active disease at both screening and baseline, as defined by having both:
   a. six or more tender/painful joints on motion; *and*
   b. six or more swollen joints
3. Active disease, as defined by one of the following criteria at screening:
   a. erythrocyte sedimentation rate (ESR) greater than 28 mm per hour; *or*
   b. C-reactive protein (CRP) greater than 7 mg/L
4. Class I, II, or III of the ACR 1991 Revised Criteria for Global Functional Status in RA

While the inclusion/exclusion criteria would exclude some patients with active disease (for example, those with ESR < 28 mm per hour), the clinical expert for this review believed that the criteria used were reasonable.

Beyond these common inclusion criteria, each study had unique inclusion/exclusion criteria and rules regarding concomitant therapy. These are summarized in Table 5.

TABLE 5: SUMMARY OF TRIAL INCLUSION/EXCLUSION CRITERIA: PAST AND CONCOMITANT TREATMENTS FOR RHEUMATOID ARTHRITIS

Study	Non-biologic DMARD Criteria	BRM Criteria	Notes Regarding Concomitant Therapy
6 Month Study Duration			
1032		In the opinion of the investigator, must have used at least one approved TNFi that was ineffective and/or not tolerated	Patients must have taken MTX for ≥ 4 months before study and be on a stable dose of 7.5 mg to 25 mg per week for at least 6 weeks before study start. No other DMARDs allowed except antimalarials.
1045	Inadequate response to at least 1 DMARD (traditional or biologic) due to lack of efficacy or toxicity.		No concomitant DMARDs allowed except antimalarials.

Study	Non-biologic DMARD Criteria	BRM Criteria	Notes Regarding Concomitant Therapy
≥ 12 Month Study Duration			
1044	All patients must have had inadequate clinical response to MTX defined as the presence of sufficient residual disease activity to meet the entry criteria.		Patients must have taken MTX for ≥ 4 months before study and be on a stable dose of 7.5 mg to 25 mg per week for at least 6 weeks before study start. No other DMARDs allowed.
1046	Patient must have had an inadequate response to at least 1 DMARD (traditional or biologic) due to lack of efficacy or toxicity.		Patients must have taken ≥ 1 traditional DMARD and stay on it throughout study.
1064	All patients must have had inadequate clinical response to MTX defined as the presence of sufficient residual disease activity to meet the entry criteria.	Patients were *excluded* who previously used adalimumab, or failed any TNFi for lack of efficacy or AE.	Patients must have taken MTX for ≥ 4 months before study and be on a stable dose of 7.5 mg to 25 mg per week for at least 6 weeks before study start. No other DMARDs allowed.

AE = adverse event; BRM = biologic response modifier; DMARD = disease-modifying antirheumatic drug; MTX = methotrexate; TNFi = tumour necrosis factor inhibitor.
Source: Manufacturer's Clinical Study Reports and publications for Studies 1032,[4-6] 1045,[7,8] 1044,[9-11] 1046,[12,13] and 1064.[14,15]

In this CDR report, the studies are categorized by duration of study. Alternatively, the studies could also be grouped by:

- Previous TNFi inadequate response (Study 1032) and previous DMARD/MTX inadequate response (studies 1044, 1045,1046, and 1064)
- Tofacitinib given as monotherapy (Study 1045) and tofacitinib given with background DMARD (studies 1032, 1044, 1046, and 1064).

b) Baseline Characteristics (See Appendix 4 for Detailed Data Presentation)

The population in the tofacitinib trials were adults with long-standing (approximately 7 to 13 years) moderate to severe RA. Patients' mean age was approximately 50 to 55 years. The majority of patients were female (approximately 77% to 87%). Moderate to high disease activity at baseline was reflected in the mean Disease Activity Score (including erythrocyte sedimentation rate) (DAS28-4[ESR]) of approximately 6.5. The majority of patients were rheumatoid factor positive (approximately 52% to 78%) and anti-cyclic citrullinated peptide (CCP) antibody positive (approximately 64% to 86%). In addition, the majority of patients were of Caucasian ethnicity in all studies except in Study 1044, which also recruited Asian patients. The average patient weight in most studies was approximately 70 kg but was slightly higher in Study 1032 (approximately 80 kg).

All patients (100%) had previously used MTX in studies 1032, 1044, and 1064; however, in studies 1045 and 1046, approximately 85% of patients had prior use of MTX. All patients had inadequate response to one or more DMARDs, except in Study 1032, in which they had inadequate response to one or more TNF inhibitors. In studies other than Study 1032, only a small percentage of patients had previously used a TNF inhibitor (approximately 6% to 19%). The number of DMARDs to which patients had a previous inadequate response was reported only in Study 1046 (mean of approximately 1.4 per patient). More than half of the patients used corticosteroids at baseline and during the studies.

The baseline characteristics were well balanced among study groups within the individual studies.

3.2.3 Interventions

Tofacitinib, placebo, and adalimumab were used in the five studies. The dose of tofacitinib that patients received after placebo was randomly assigned at the beginning of the study. The tofacitinib doses were either 5 mg twice daily (the Health Canada approved dose) or 10 mg twice daily. The sponsor, investigator, and patients were blinded to treatment assignment. Matching placebo tablets and injections (Study 1064) were used to maintain blinding.

Patients could continue on oral corticosteroids during the studies if they were already taking them at screening at a stable dose of ≤ 10 mg per day of oral prednisone or equivalent for four weeks before the first dose of study drug. Background therapy could also include stable doses of opioids, acetaminophen, nonsteroidal anti-inflammatory drugs, and cyclo-oxygenase-2 inhibitors.

3.2.4 Outcomes

a) Primary End Points

All five trials used the same three measures for the co-primary end points. The co-primary end point of ACR 20 response was evaluated at month 3 (studies 1032 and 1045) or month 6 (studies 1044, 1046, and 1064). The co-primary end point of change from baseline for the Health Assessment Questionnaire–Disability Index (HAQ-DI) was evaluated at month 3 in all trials. The co-primary end point of DAS28-4(ESR) was evaluated at month 3 (studies 1032 and 1045) or month 6 (studies 1044, 1046, and 1064). Study 1044 used a fourth co-primary end point of modified Total Sharp Score (mTSS). No other studies measured radiographic outcomes.

American College of Rheumatology Assessments[9]

A responder using ACR 20 criteria was defined as having ≥ 20% improvement in tender and swollen joint counts and ≥ 20% improvement from baseline in three of the five remaining ACR core set measures: patient global assessment, physician global assessment, pain, disability, and an acute-phase reactant. Similarly, ACR 50 and ACR 70 were calculated as the respective percentage improvements from baseline. The ACR 20/50/70 efficacy determinations were made at every study visit.

The specific components of the ACR assessments were:

- tender/painful joint count (68 joints)
- swollen joint count (66 joints)
- patient assessment of arthritis pain
- patient global assessment of arthritis (PtGA)
- physician global assessment of arthritis (PGA)
- CRP or ESR
- HAQ-DI.

Tender/Painful Joint Count

Sixty-eight joints were assessed by a joint assessor who was blinded to study treatment to determine the number of joints that were considered tender or painful. The response to pressure and motion on each joint was assessed using the following scale: present/absent/not done/not applicable (where not applicable was to be used for artificial joints).

Swollen Joint Count

The joint assessor who was blinded to study treatment also assessed joints for swelling using the following scale: present/absent/not done/not applicable (where not applicable was to be used for artificial joints).

The 66 joints assessed for swelling were the same as those assessed for tenderness/pain, except that the right and left hip joints were not included in the swollen joint count.

Patient Assessment of Arthritis Pain
Patients assessed the severity of their arthritis pain using a 100 mm visual analogue scale (VAS) by placing a mark on the scale between 0 (no pain) and 100 (most severe pain) that corresponded to the magnitude of their pain.

Patient Global Assessment of Arthritis
Patients answered the following question, "Considering all the ways your arthritis affects you, how are you feeling today?" The patient's response was recorded using a 100 mm VAS (0 = very well, 100 = very poorly).

Physician Global Assessment of Arthritis
The investigator assessed how the patient's overall arthritis appeared at the time of the visit. This was an evaluation based on the patient's disease signs, functional capacity, and physical examination, and was independent of the PtGA. The investigator's response was recorded using a 100 mm VAS (0 = very good, 100 = very poor).

Health Assessment Questionnaire–Disability Index
The HAQ-DI assessed the degree of difficulty a patient had experienced during the past week in eight domains of daily living activities: dressing and grooming, arising, eating, walking, hygiene, reach, grip, and other activities. Each activity category consisted of two to three items. For each question in the questionnaire, the level of difficulty was scored from 0 to 3, with 0 representing "no difficulty," 1 "some difficulty," 2 "much difficulty," and 3 "unable to do." Any activity that required assistance from another individual or required the use of an assistive device adjusted to a minimum score of 2 to represent a more limited functional status. The minimal clinically important difference (MCID) is estimated to be 0.22.

Disease Activity Score Assessments
The DAS assessment is a derived measurement, with differential weighting given to each component. The term "DAS28" refers to both DAS28-3(CRP) and DAS28-4(ESR). The DAS28-3(CRP) was calculated at each visit. The DAS28-4(ESR) was calculated at each visit only if ESR data were available. Higher DAS28 scores indicate greater disease activity. The components of the DAS28 arthritis assessments were as follows:

- tender/painful joint count (28 joints)
- swollen joint count (28 joints)
- CRP or ESR
- PtGA (for DAS28-4[ESR]).

Improvement of RA as measured by DAS28 (both DAS28-3[CRP] and DAS28-4[ESR]) is categorized according to the information in Table 6.

TABLE 6: THE EULAR IMPROVEMENT RESPONSE CRITERIA (DAS28)

Current DAS 28	Improvement in DAS 28 from Baseline (Decline in DAS)		
	≤ 0.6	> 0.6 to ≤ 1.2	> 1.2
≤ 3.2	None	Moderate	Good
> 3.2 to ≤ 5.1	None	Moderate	Moderate
> 5.1	None	None	Moderate

DAS 28 = Disease Activity Score; EULAR = European League Against Rheumatism

Modified Total Sharp Scores (used in Study 1044)[9]

Also known as the Sharp/van der Heijde scoring method, the mTSS is commonly used in rheumatology studies. The method of computing mTSS is to grade the presence of erosions in the hands and feet and the presence of joint space narrowing (JSN) in the hands, wrists, and feet. The scores for each feature for the individual joints are summed. For erosion scores, 16 locations in each hand and wrist and 12 locations in each foot were scored using a six-point scale from 0 to 5. For JSN, 15 locations in each hand and wrist and six locations in each foot were scored using a five-point scale from 0 to 4. The maximum mTSS = maximum erosion score (280) + maximum JSN score (168) and is therefore a total score of 448, with higher scores indicating greater disease severity.

Radiographs for each patient were graded by two independent readers, with the patient randomization sequence and visit or time of acquisition blinded. The two readers' grades for each patient were averaged, and this composite score was then compared by time point (study visit) to determine radiographic progression.

b) Secondary End Points

The outcomes used for the primary end points were also evaluated at other time points for the secondary end point analyses. Efficacy evaluations commonly occurred every six to 12 weeks during the double-blind phase of the studies. Other secondary outcomes included PGA and the patient-reported outcomes: Medical Outcome Study Sleep Scale (MOS-SS), Short-Form (36) Health Survey (SF-36), Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-Fatigue), EuroQol Five-Dimension Health-Related Quality of Life Questionnaire (EQ-5D), health care resource utilization, Work Limitations Questionnaire (WLQ), and PtGA.

Description of Selected Patient-Reported Outcomes[9]

Short-Form (36) Health Survey (Version 2, Acute): The SF-36 Version 2 (Acute) is a 36-item generic health status measure. It measures eight general health domains: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health. Higher scores indicate better health-related quality of life. The eight subdomains are each measured on a scale of 0 to 100, with an increase in score indicating improvement in health status. The MCID is estimated at 2.5 to 5 points.

Functional Assessment of Chronic Illness Therapy-Fatigue Scale: The FACIT-Fatigue scale is a patient-completed questionnaire consisting of 13 items that assess fatigue. Instrument scoring yields a range from 0 to 52, with higher scores representing better patient status (i.e., less fatigue). A suggested MCID for the FACIT-Fatigue in RA patients is between 3 and 4 points.

c) Harms

Serious adverse event and adverse event data were collected, with a particular interest in infections, including tuberculosis, treated infections, malignancies, potential cases of drug-induced liver injury (based on elevated aspartate transaminase [AST] and/or alanine transaminase levels [ALT]), and electrocardiographic abnormalities.

3.2.5 Statistical Analysis

a) Sample Size Estimation Methods

The sample size calculations for most trials (studies 1032, 1045, 1046, and 1064) were made separately for each of the three primary end points. Power estimates were provided for each of the three end points for a given delta (effect difference between placebo and drug).

For studies 1032, 1045, 1046, and 1064, sample size estimates were based on similar assumptions. Actual enrolment in these trials ranged from 399 to 792 and was often higher than the sample size estimates. Therefore, there appears to have been extra enrolment permitted beyond the original estimate in some studies.[7] For ACR 20 analysis, the sample size in each trial was planned to yield greater than 90% power, assuming a difference in response rates of at least 20% (with the placebo response at 30%). For analysis of the HAQ-DI, the planned sample size resulted in greater than 90% power for differences of 0.3 or greater, assuming a standard deviation of 0.75. For the analysis of DAS28-4(ESR) scores less than 2.6, the planned sample size resulted in more than 90% power for differences in response rates of at least 15% (with placebo response at 10%).

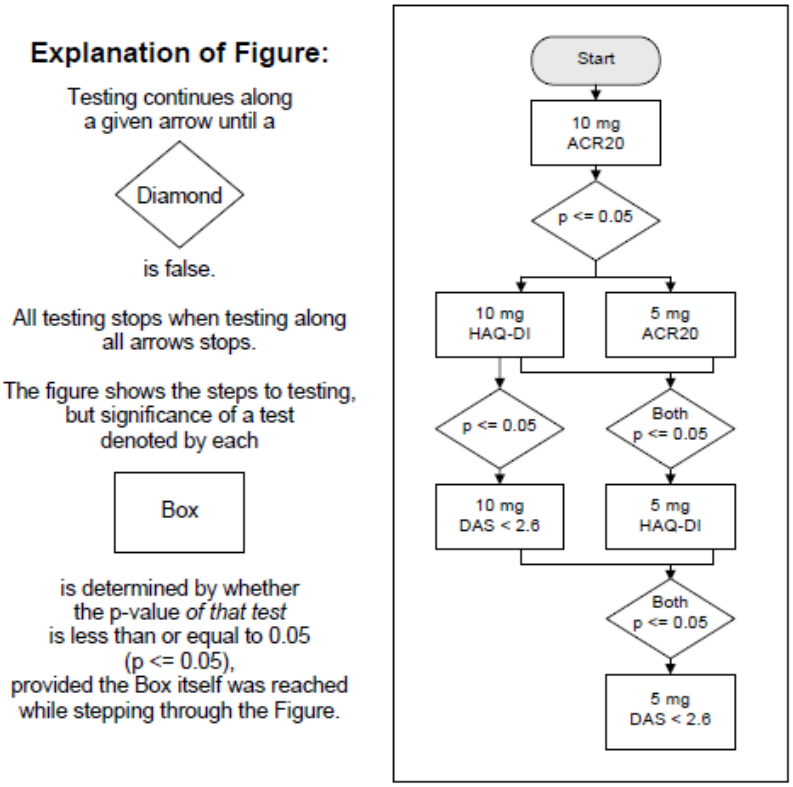
Sample size for Study 1044 was calculated differently from the other studies. The sample size was estimated using the mTSS.[9] Investigators estimated that 750 patients would provide approximately 90% power to detect a difference of 0.8 points between placebo and tofacitinib 10 mg twice daily at month 6.

b) Methods for Analysis of Co-Primary End Points

Each objective was assessed sequentially, using a step-down approach, where statistical significance could be claimed for the end point only if the previous end point in the sequence met the requirements for significance. Additionally, as there were two doses within each end point, the step-down approach was applied; i.e., the highest dose (10 mg twice daily) at a given end point could achieve significance only if the high dose at the prior end point was significant; the low dose (5 mg twice daily) at a given end point could achieve significance only if both the high dose at the same end point and the low dose at the prior end point were statistically significant.

The sequence of primary end points is depicted in Figure 2. The sequence for Study 1044 was different and is presented APPENDIX 4: Detailed Outcome Data. In this approach, treatment difference in ACR 20 between 10 mg and placebo was tested first (Family 1). If it was significant at $P \leq 0.05$, then treatment differences in 10 mg HAQ-DI and 5 mg ACR 20 were tested at $P \leq 0.05$, continuing as depicted in Figure 2.

FIGURE 2: STEP-DOWN APPROACH FOR ANALYSIS OF CO-PRIMARY END POINTS (STUDIES 1032, 1045, 1046, AND 1064)[4,7,12,14]



ACR = American College of Rheumatology; DAS = Disease Activity Score; HAQ-DI = Health Assessment Questionnaire–Disability Index.

For ACR 20 and incidence of DAS28-4(ESR) less than 2.6, the normal approximation for the difference in binomial proportions was used to test the superiority of each dose of tofacitinib to placebo. The HAQ-DI analysis was done using a mixed effects, repeated measures model that included the fixed effects of treatment, visit, treatment by visit interaction, geographic region, and the baseline value as a covariate.

In Study 1044, for mTSS, patients who were advanced at month 3 had their month 6 mTSS calculated using a linear extrapolation from the radiographs taken at baseline and month 3. For month 12, comparisons with placebo were done by linearly extrapolating a month 12 value based on baseline and month 6. All mTSS-related variables were imputed using this method. Binary variables (rates of patients with no progression in mTSS and rates of patients with no progression in mean erosion score) were analyzed using normal approximation to the binomial.[9] No interim analyses were performed in any of the studies except Study 1044 at 12 months.

Imputation of no response was used to account for missing data in the calculation of ACR 20 response rates and DAS28-4(ESR) and was applied to patients who discontinued the study drug for any reason (including patients who were lost to follow-up before month 6). In the trial publications and manufacturer's reports, this method was called baseline observation carried forward (BOCF) or non-responder imputation (NRI). The imputation of no response was also applied to patients who did not have a 20% reduction in the number of tender and swollen joints at month 3 ("non-responders" at month 3 in studies 1044, 1046, and 1064), regardless of treatment assignment; patients in the active-

treatment groups who did not meet the criteria for this response continued with the same treatment, whereas patients who were receiving placebo were switched to tofacitinib in a double-blind fashion. The imputation-of-no-response analysis assumes that patients who did not have a response to treatment by month 3 will not have a response during the remainder of the trial — even if they subsequently meet the criteria for an ACR 20 response.[4,15]

c) Secondary Outcome Analyses

The following is an example of how the secondary outcomes were analyzed (from the Clinical Study Report for Study 1044).[9]

All secondary analyses were based on the full analysis set (FAS).

Secondary analyses included the normal approximation for the difference in binomial proportions for the ACR variables (ACR 20, ACR 50, and ACR 70) obtained in separate analyses. The binomial variables — incidence of DAS28 ≤ 3.2, incidence of DAS28 < 2.6, DAS28 responses (good or moderate), and clinically meaningful decrease in HAQ-DI — were analyzed by considering the proportion of patients responding to each end point and using the same normal approximation to the binomial for the analyses.

The other six components of the ACR criteria — DAS28, the eight domains and two scores of SF-36, MOS-SS, EQ-5D, the four domain scores and the work loss index of WLQ, and FACIT-Fatigue scale — were each analyzed in the same way as HAQ-DI. Each end point's baseline values were used as covariates. The data from the RA health care resources utilization (HCRU) were listed, and descriptive statistics were generated.

In addition, change in HAQ-DI from baseline after month 3 was also analyzed for descriptive purposes; and the mixed effects model with repeated measures was applied as well to evaluate the effect after month 3. Further analyses described in the Statistical Analysis Plan using the sequence effects model were conducted for the HAQ-DI actual values at each visit. Other continuous variables, e.g., DAS28-4(ESR) and SF-36, also followed this sequence effects model.

d) Subgroup Analyses

The manufacturer performed many exploratory subgroup analyses for various subpopulations (see Figure 4, APPENDIX 4: Detailed Outcome Data), some of which did not appear to be pre-specified.

3.2.6 Analysis Populations[4,7,9,10,12,14]

The FAS was defined by the manufacturer as all patients who were randomized to the study and received at least one dose of tofacitinib or placebo. The primary efficacy statistical analysis was conducted on the FAS population. However, the Food and Drug Administration (FDA) statistical reviewers noted that, in their actual analyses, the manufacturer excluded subjects in the FAS who did not have baseline or at least one post-baseline measurement, and therefore it was not a true intention-to-treat (ITT) population.[24]

FAS patients who had a protocol deviation thought to affect the efficacy analysis were excluded from the per-protocol (PP) efficacy analysis. The safety analysis set was defined as patients who received at least one dose of tofacitinib or placebo.

There were some discrepancies between the true ITT population and the manufacturer-defined FAS population.[24] ITT data would include all randomized patients who took at least one dose of study

medication. FAS data used by the manufacturer excluded patients if they did not have one or more baseline or post-baseline measurement. The FAS populations used by the manufacturer for specific outcomes, such as ACR 20, HAQ-DI, or DAS28, were different from the overall FAS populations. The sizes of the outcome-specific FAS populations were as much as 10% smaller for specific end points.

3.3 Patient Disposition

The proportion of patients who discontinued from each study before the end of the double-blind phase is presented in Table 7: The overall discontinuation rate among studies ranged from 9% to 22% in the six-month studies, and from 18% to 32% in the trials of more than 12 months' duration.

At the end of each study, the percentage of patients who discontinued treatment was slightly higher for patients taking tofacitinib or adalimumab compared with placebo within each study (approximately 1% to 10% higher, data not shown). The exception to this was the monotherapy study (Study 1045), in which the study discontinuation rates were slightly lower between tofacitinib and placebo (approximately 3% to 9% lower, data not shown). Overall discontinuation rates tended to be slightly lower in treatment groups that were randomized to tofacitinib compared with those randomized to placebo followed by tofacitinib. These differences between treatment groups were more pronounced for the longer studies (≥ 12 months' duration) than the shorter studies (six months' duration).

TABLE 7: PATIENT DISPOSITION

	6 Month Study Duration			≥ 12 Month Study Duration	
Study	**1032**	**1045**	**1044**	**1046**	**1064**
Screened, N	589	954	1,291	1,281	1,042
Randomized, N	399	611	800	795	717
Discontinued, N (%)[a]	88 (22)	55 (9)	257 (32)	141 (18)	161 (22)
Died	1 (< 1)	1 (< 1)	6 (1)	2 (< 1)	1 (< 1)
Related to study drug	36 (9)	2 (3)	116 (15)	84 (11)	79 (11)
AE	18 (5)	12 (2)	91 (11)	34 (4)	54 (8)
Lack of efficacy	18 (5)	9 (1)	21 (3)	34 (4)	25 (3)
Not related to study drug	51 (13)	33 (5)	135 (17)	67 (8)	81 (11)
AE	14 (4)	5 (1)	33 (4)	20 (3)	23 (3)
Other	3 (1)	4 (1)	36 (4)	33 (4)	50 (7)
Protocol violation	17 (4)	12 (2)	22 (3)	0	0
Patient withdrew	17 (4)	12 (2)	44 (6)	14 (2)	8 (11)
Early escape at month 3 (non-responders)[b]	NA	NA	219 (27)	[redacted]	[redacted]

AE = adverse event; NA = not applicable.
[a] Data are N (% of randomized)
[b] See Table 8 for a detailed breakdown by treatment group.
Source: Manufacturer's Clinical Study Reports and publications for Studies 1032,[4-6] 1045,[7,8] 1044,[9-11] 1046,[12,13] and 1064.[14,15]

3.3.1 Early Escape Patients

Non-responder analysis was performed at month 3 in studies 1044, 1046, and 1064. If there was not at least a 20% improvement in both the tender/painful and swollen joint counts, the patient was considered a non-responder. If patients taking tofacitinib 5 mg twice daily or 10 mg twice daily were classified as non-responders at month 3, they continued on the same dose. If patients taking placebo were classified as non-responders at month 3, they were advanced to tofacitinib 5 mg or 10 mg, but the treatment remained blinded until the end of the study. These data are summarized in Table 8.

Approximately twice as many patients were non-responders at month 3 in the placebo groups as in the active-drug-treatment groups.

TABLE 8: NON-RESPONDERS AT MONTH 3 FOR PLACEBO EARLY ESCAPE

Study	1044[9]		1046[13]		1064[14]	
Randomized Treatment Sequence	**Randomized, N**	**Non-responders (PL Group Advanced to Early Escape) at month 3, n (%)**	**Randomized, N**	**Non-responders (PL Group Advanced to Escape) at month 3, n (%)**	**Randomized, N**	**Non-responders (PL Group Advanced to Early Escape) at month 3, n (%)**
TOF 5 mg b.i.d.	321	84 (26)	318	[redacted]	204	[redacted]
TOF 10 mg b.i.d.	316	56 (18)	318	[redacted]	201	[redacted]
PL→TOF 5 mg b.i.d.	81	42 (52)	79	[redacted]	56	[redacted]
PL→TOF 10 mg b.i.d.	79	37 (47)	80	[redacted]	52	[redacted]
ADA 40 mg SC q.2w.	–	–	–	–	204	[redacted]

ADA = adalimumab; b.i.d. = twice daily; PL = placebo; q.2w. = every two weeks; SC = subcutaneous injection; TOF = tofacitinib.
Source: Manufacturer's Clinical Study Reports and publications for Studies 1044,[9-11] 1046,[12,13] and 1064.[14,15]

3.4 Exposure to Study Treatments

If duration of treatment is defined as the total number of dosing days from the first to and including the last day of each study treatment, the mean duration is approximately 280 days for patients in the five studies (N = 3,332). This included time on placebo but excluded adalimumab patients (Summary of Clinical Safety).[25] This did not include the second treatment year for Study 1044.

3.5 Critical Appraisal

3.5.1 Internal Validity

- The trials applied appropriate allocation concealment, randomization, and blinding methods.
- A relevant comparator that is widely used in Canadian practice was selected for one of the studies (adalimumab, Study 1064), and some statistical comparisons were made between tofacitinib and adalimumab. However, the study was not powered to detect any differences between the two drugs. Four studies used placebo as a control, which is not as clinically relevant as a comparator group.
- The baseline characteristics were well balanced within the trials and across the treatment groups.
- A substantial number of patients switched from placebo to tofacitinib at month 3 in studies 1044, 1046, and 1064 after meeting criteria for early escape. This limits the ability to make assertions about the results beyond the three month time point. The proportion of early escape patients across the treatment groups differed. The manufacturer used the BOCF method to impute data for the primary analyses. The direction of bias associated with this analytical approach is unknown.
- The manufacturer did not use a true ITT population for the primary end point analyses. The sizes of the outcome-specific FAS populations were as much as 10% smaller than a true ITT population for many of the primary end point analyses. This had the effect of reducing the denominators, and the direction of bias that this may have introduced is unknown.[24]
- After month 3 in studies 1032 and 1045 and after month 6 in studies 1044, 1046, and 1064, all patients knew that they were taking active treatment, but they did not know the dose they were

taking. The only comparisons that could be made after the switch to active treatment are between the 5 mg and 10 mg doses (and between tofacitinib and adalimumab in Study 1064). The comparison of tofacitinib 5 mg to 10 mg is not relevant for the purposes of this review, since the 10 mg dose is not approved by Health Canada. The data beyond these time points may provide some information regarding the relative effects of tofacitinib 5 mg twice daily and adalimumab (open-label data from Study 1064), but, other than this, the data have little value for evaluating the efficacy of tofacitinib.

- Inherent in the use of the step-down procedure for testing the co-primary end points is that the order of testing can be arranged a priori to maximize the power of the entire testing procedure.
- The co-primary end points for the studies relied on results from different time points: three months and six months. While statistically possible, it is conceptually complex to interpret a study that has a trinity of co-primary end points at different time points.
- One weakness of the manufacturer's analyses is that, as patients discontinued the trial for various reasons, the denominator was reduced in some of the analyses; hence, the analyses were not based on the ITT principle. This may have resulted in overestimating the response rates at study end.
- For the patient-reported outcomes such as SF-36 and FACIT-Fatigue, the manufacturer stated: "For all continuous variables, it should be noted that patients who were non-responders (in the joint counts) advanced to tofacitinib at month 3. The placebo group after month 3 is comprised of responders. A bias may exist in the results due to the depletion of non-responders from the placebo group. The longitudinal-linear model may not properly penalize the values for patients who withdrew due to lack of efficacy, or were advanced before month 6."[12]

3.5.2 External Validity

- Between the trials, the baseline characteristics were similar. A few exceptions to this included a higher mean patient weight at baseline and longer duration of disease in Study 1032 (TNF inhibitor inadequate response) and a lower representation of Caucasian patients due to enrolment of Asian patients in Study 1044, which was a positive design feature of this trial.
- The patient populations reflect a spectrum of patients with RA, including tofacitinib monotherapy, patients with inadequate response to TNF inhibitors, and patients with inadequate response to DMARDs.
- Given the aging population phenomena in western societies, it would have been helpful to have greater representation of patients older than 70 years.
- RA is a chronic disease; it is expected that patients will be on treatment for many years. Although longer-term harms data were reported after all patients had switched to active drug (see APPENDIX 6: Summary of Extension Studies), most of the controlled data that exist for tofacitinib (prior to early escape) are from month 3, with some data available from month 6, after early escape.
- High-risk patients for specific harms (e.g., history of hospitalization for infection within the previous six months) were excluded. While this is a prudent approach, it limits generalizability of the harms results to clinical practice, in which patients at higher risk of harms may be prescribed the drug.

3.6 Efficacy

Only those efficacy outcomes identified in the review protocol are reported (Section 2.2, Table 3). See APPENDIX 4: Detailed Outcome Data for detailed efficacy data.

3.6.1 ACR Response

The proportion of patients achieving ACR 20 response was larger in the tofacitinib 5 mg twice daily group (42% to 60%) than in the placebo group (25% to 30%) at month 3 or month 6. Statistically significant results at the same time points were observed for tofacitinib 5 mg twice daily versus placebo for the proportion of patients achieving ACR 50 and ACR 70 response. Statistically significant results were also observed for adalimumab versus placebo for ACR 20, ACR 50, and ACR 70 response.

Patients on placebo could escape at three months to tofacitinib 5 mg twice daily or 10 mg twice daily. The patients who escaped early or withdrew were assigned "no response" and, therefore, they are not included in the numerators in the data in Table 9.

TABLE 9: ACR RESPONSE BASED ON MANUFACTURER'S FAS POPULATION (USING MANUFACTURER'S NON-RESPONDER IMPUTATION METHOD)

Study Number and Time of Evaluation	ACR 20, n/N (%) *P* Value vs. PL	ACR 50, n/N (%) *P* Value vs. PL	ACR 70, n/N (%) *P* Value vs. PL
1032 month 3			
TOF 5 mg b.i.d.	55/132 (42); *P* = 0.0025[a]	35/132 (27); *P* < 0.0001	18/132 (13); *P* < 0.0001
PL	32/131 (25)	11/131 (8)	2/131 (2)
1045 month 3			
TOF 5 mg b.i.d.	144/241 (60) ;*P* < 0.0001[a]	75/241 (31); *P* < 0.0001	37/241 (15); *P* = 0.0026
PL	32/120 (27)	15/120 (13)	7/120 (6)
1044 month 6			
TOF 5 mg b.i.d.	159/309 (52); *P* < 0.0001[a]	100/309 (32); *P* < 0.0001	45/309 (15); *P* < 0.0001
PL	39/154 (26)	13/154 (8)	2/154 (1)
1046 month 6			
TOF 5 mg b.i.d.	164/311 (53); *P* < 0.0001[a]	105/311 (34); *P* < 0.0001	41/311 (13); *P* < 0.0001
PL	49/157 (31)	20/157 (13)	5/157 (3)
1064 month 6			
TOF 5 mg b.i.d.	101/196 (52); *P* < 0.0001[a]	72/196 (37); *P* < 0.0001	39/196 (20); *P* < 0.0001
PL	30/106 (28)	13/106 (12)	2/106 (2)
ADA 40 mg q.2w.	94/199 (47); *P* = 0.0008	55/199 (28); *P* = 0.0006	18/199 (9); *P* = 0.0031
1064 month 12			
TOF 5 mg b.i.d.	97/196 (49); *P* < 0.0001	72/196 (37); *P* < 0.0001	45/196 (23); *P* < 0.0001
PL→TOF 5 mg b.i.d.	19/56 (34); *P* < 0.0001	12/56 (21); *P* < 0.0001	6/56 (11); *P* = 0.0095
ADA 40 mg q.2w.	98/199 (49); *P* < 0.0001	67/199 (34); *P* < 0.0001	33/199 (17); *P* < 0.0001

ACR = American College of Rheumatology; ADA = adalimumab; b.i.d. = twice daily; FAS = full analysis set; PL = placebo; q.2w. = every two weeks; TOF = tofacitinib; vs. = versus.

[a] Co-primary end points are in bold.

Note: all *P* values are versus placebo.

Source: Manufacturer's Clinical Study Reports and publications for studies 1032,[4-6] 1045,[7,8] 1044,[9-11] 1046,[12,13] 1064[14,15] and FDA Clinical Review.[16]

Data from the individual components of the ACR are presented in APPENDIX 4: Detailed Outcome Data. For each of the individual components of the ACR response criteria, tofacitinib 5 mg twice daily was numerically improved compared with placebo (no statistical testing was provided by the manufacturer).

Data for Month 12 (which was not a placebo-controlled time point) are also presented for Study 1064 in Table 9. These data showed ACR 20 response rates of 49% for both tofacitinib 5 mg twice daily and

adalimumab 40 mg every two weeks. Data beyond 12 months are presented for Study 1044 in APPENDIX 6: Summary of Extension Studies. ACR 20 rates appeared to decrease over time, and at month 24 ACR 20 response rates were 41% for the tofacitinib 5 mg twice daily group. Rates also tended to converge between treatment groups after 24 months of treatment.

3.6.2 Health Assessment Questionnaire–Disability Index

Baseline mean HAQ-DI scores ranged from 1.3 to 1.6 across treatment groups. At the primary end point time point (month 3 for all studies), the mean change in scores decreased (improved) from baseline for all treatment groups, including placebo. The difference in improvement in mean change score with tofacitinib 5 mg twice daily was statistically significant relative to placebo for all studies.

TABLE 10: HAQ-DI RESPONSE AT MONTH 3, BASED ON MANUFACTURER'S FAS POPULATION AND USING MANUFACTURER'S NON-RESPONDER IMPUTATION METHOD

Study and Time Point	N	Mean BL Score; MCFB	MDC (95% CI) vs. PL
1032 month 3			
TOF 5 mg b.i.d.	117	1.6; –0.5	**–0.3 (–0.4 to –0.1)[a]** ***P* = 0.0002**
PL	118	1.6; –0.2	
1045 month 3			
TOF 5 mg b.i.d.	237	1.5; –0.5	**–0.3 (–0.4 to –0.2)[a]** ***P* < 0.0001**
PL	109	1.5; –0.2	
1044 month 3			
TOF 5 mg b.i.d.	294	1.4; –0.4	**–0.3 (0.4 to –0.2)[a]** **NSS[b]**
PL	146	1.4; –0.2	
1046 month 3			
TOF 5 mg b.i.d.	292	1.4; –0.5	**–0.3 (–0.4 to –0.2)[a]** ***P* < 0.0001**
PL	147	1.3; –0.2	
1064 month 3			
TOF 5 mg b.i.d.	188	1.5; –0.6	**–0.3 (–0.4 to –0.2)[a]** ***P* < 0.0001**
PL	98	1.4; –0.3	
ADA 40 mg q.2w.	190	1.5; –0.5	–0.2 (–0.4 to –0.1) *P* < 0.0001

ADA = adalimumab; b.i.d. = twice daily; BL = baseline; CI = confidence interval; FAS = full analysis set; HAQ-DI = Health Assessment Questionnaire–Disability Index; MDC = mean difference of change; MCFB = mean change from baseline; PL = placebo; q.2w. = every two weeks; TOF = tofacitinib.

[a] Co-primary end points are in bold.

[b] NSS here means that the study could not claim statistical significance because of the step-down statistical testing procedure.

Source: Manufacturer's Clinical Study Reports and publications for studies 1032,[4-6] 1045,[7,8] 1044,[9-11] 1046,[12,13] 1064[14,15] and FDA Clinical Review.[16]

3.6.3 Disease Activity Score 28-4(ESR)

Baseline DAS28-4(ESR) mean values ranged from approximately 6.3 to 6.6, and baseline DAS28-4(CRP) mean values ranged from 5.2 to 5.7 across the studies. The proportion of patients achieving DAS28-4(ESR) less than 2.6 was low (fewer than 10%) in all tofacitinib 5 mg twice daily or placebo treatment groups. In the manufacturer's analyses, there were statistically significant differences in studies 1032, 1046, and 1064 for

the co-primary DAS28-4 end point. The FDA statistical reviewer noted that the manufacturer did not use a true ITT population in its analyses. The FDA statistical report re-analysis concluded that differences between tofacitinib 5 mg twice daily and placebo for DAS28-4(ESR) were only statistically significant in Study 1046.

Statistical significance for the month 6 DAS28-4(ESR) could not be claimed in Study 1044 because the step-down procedure was halted after the tofacitinib 5 mg twice daily dose was not statistically significant from placebo for mTSS (see Figure 3, APPENDIX 4: Detailed Outcome Data*).*

TABLE 11: DAS28-4(ESR) RESPONSE, BASED ON MANUFACTURER'S FAS POPULATION AND USING MANUFACTURER'S NON-RESPONDER IMPUTATION METHOD

Study and Time Point	DAS28-4(ESR) < 2.6 n/N (%)	DAS28-4(CRP) < 2.6 n/N (%)
1032 month 3		
TOF 5 mg b.i.d.	**8/119 (7); *P* = 0.0497**[a]	27/132 (20); *P* < 0.0001
PL	**2/120 (2)**	6/131 (5)
1045 month 3		
TOF 5 mg b.i.d.	**13/232 (6); *P* = 0.618**[a]	45/241 (19); *P* < 0.0001
PL	**5/114 (4)**	6/120 (5)
1044 month 6		
TOF 5 mg b.i.d.	**19/265 (7); NSS**[a,b]	71/309 (23); *P* < 0.0001
PL	**2/129 (1)**	8/154 (5)
1046 month 6		
TOF 5 mg b.i.d.	**24/263 (9); *P* = 0.004**[a]	69/311 (22); *P* < 0.0001
PL	**4/148 (3)**	14/157 (9)
1064 month 6		
TOF 5 mg b.i.d.	**11/177 (6); *P* = 0.015**[a]	36/196 (18); *P* = 0.0016
PL	**1/92 (1)**	7/104 (7)
ADA 40 mg q.2w.	12/178 (7); *P* = 0.009	33/199 (17); *P* = 0.006

ADA = adalimumab; b.i.d. = twice daily; CRP = C-reactive protein; DAS28-4(ESR) = Disease Activity Score (erythrocyte sedimentation rate); ESR = erythrocyte sedimentation rate; FAS = full analysis set; PL = placebo; q.2w. = every two weeks; TOF = tofacitinib.

[a] Co-primary end points are in bold.

[b] NSS here means that the trial could not claim statistical significance because of the step-down statistical testing procedure.

Source: Manufacturer's Clinical Study Reports and publications for studies 1032,[4-6] 1045,[7,8] 1044,[9-11] 1046,[12,13] and 1064.[14,15]

3.6.4 American College of Rheumatology and Disease Activity Score 28-4 (ESR) Response at Month 6 by Early Escape Status[26]

The data in Table 12 provide a more detailed breakdown of the response data in Table 9 and Table 11 for the three studies that were ≥ 12 months duration. These three tables provide information regarding the response status at month 6 (ACR 20/50/70 and DAS28-4[ESR]), grouped by response status at month 3 (based on joint counts). In the manufacturer's primary analysis of these three trials, early escape to tofacitinib at month 3 was allowed if there was less than 20% improvement in tender joint count and swollen joint count.

In Table 12, the ACR response rates at month 6 are provided for patients who did not respond at month 3. For example, in Study 1044, [redacted]% of the patients taking tofacitinib 5 mg twice daily who did not respond at month 3 achieved ACR 20 response at month 6. The corresponding rates in studies 1046 and 1064 were [redacted]% and [redacted]%, respectively. The rates of ACR 50 and ACR 70 response in this same subset of patients was [redacted] ([redacted]%). The rates of DAS28-4(ESR) response were [redacted]% in patients taking tofacitinib 5 mg twice daily at month 6 who had been non-responders at month 3.

TABLE 12: ACR 20/50/70 AND DAS28-4(ESR) RESPONSE AT MONTH 6 GIVEN RESPONSE STATUS AT MONTH 3

	Study 1044 (MTX Failure Patients)				Study 1046 (DMARD Failure Patients)				Study 1064 (MTX Failure Patients)			
	PL→TOF 5 mg b.i.d.		TOF 5 mg b.i.d.		PL→TOF 5 mg b.i.d.		TOF 5 mg b.i.d.		PL→TOF 5 mg b.i.d.		TOF 5 mg b.i.d.	
Number of patients randomized	81		321		79		318		56		204	
	Responder[b] at month 3 (stayed on placebo to month 6)	Non-responder[c] at month 3 (switched to TOF 5 mg b.i.d.)	Responder[b] at month 3	Non-responder[c] at month 3	Responder[b] at month 3 (stayed on placebo to month 6)	Non-responder[c] at month 3 (switched to TOF 5 mg b.i.d.)	Responder[b] at month 3	Non-responder[c] at month 3	Responder[b] at month 3 (stayed on placebo to month 6)	Non-responder[c] at month 3 (switched to TOF 5 mg b.i.d.)	Responder[b] at month 3	Non-responder[c] at month 3
Responders or non-responders at month 3[a]	[redacted]											
Response at month 6, n/N (%)												
ACR 20	[redacted]											
ACR 50	[redacted]											
ACR 70	[redacted]											
Responders or non-responders at month 3[d]	[redacted]											
DAS28-4(ESR) < 2.6	[redacted]											

ACR = American College of Rheumatology; b.i.d. = twice daily; DAS28-4(ESR) = Disease Activity Score (erythrocyte sedimentation rate); DMARD = disease-modifying antirheumatic drug; MTX = methotrexate; PL = placebo; TOF = tofacitinib.

[a] N = number of responders or non-responders at month 3 with response available at month 6; n = number of responders at month 6.

[b] At month 3, a responder is a patient who improved by at least 20% from baseline in number of swollen and tender/painful joints according to the site.

[c] At month 3, a non-responder is a patient who failed to improve by at least 20% from baseline in number of swollen and tender/painful joints according to the site.

[d] Some sites were not able to perform blinded ESR required for the DAS28-4(ESR) calculation, contributing to fewer patients with DAS response data available month 6.

Source: manufacturer's unpublished data.[26]

3.6.5 Modified Total Sharp Score

Study 1044 was the only study that measured radiographic progression. Mean change from baseline at six months in the mTSS was one of the four co-primary end points in Study 1044. The baseline total score in the randomized treatment groups ranged from 30 to 38 points (total range of scale: 0 to 448).[11] At six months, both tofacitinib groups and the placebo group showed worsening (increase in mTSS). The difference of –0.34 points between tofacitinib and placebo was not statistically significant (*P* = 0.079) at month 6. Therefore, the primary end point for Study 1044 was not achieved, and testing stopped in the step-down procedure (Figure 3, APPENDIX 4: Detailed Outcome Data).

TABLE 13: STUDY 1044 MODIFIED TOTAL SHARP SCORE CHANGE FROM BASELINE (FAS, IMPUTATION USING LINEAR EXTRAPOLATION)[9]

	Month 6 (Primary Outcome)			Month 12		
	N	MCFB	MDC (95% CI) vs. PL	N	MCFB	MDC (95% CI) vs. PL
TOF 5 mg b.i.d.	**278**	**0.12**	**–0.34 (–0.73 to 0.04)[a] *P* = 0.079**	286	0.29	–0.63 (–1.27 to 0.02) *P* = 0.056
PL	**140**	**0.47**		139	0.92	

b.i.d. = twice daily; CI = confidence interval; MDC = mean difference of change; MCFB = mean change from baseline; PL = placebo; TOF = tofacitinib; vs. = versus.
[a] Co-primary end points are in bold.
Source: Manufacturer's Clinical Study Reports and publications for Study 1044.[9-11]

3.6.6 Quality of Life: Short-Form (36) Health Survey

The SF-36 scores range from 0 to 100, with higher scores indicating better health status. Statistically significant improvements in the physical component summary scores were observed for tofacitinib versus placebo at the same time point as the primary outcome (month 3 or month 6) for four of five trials as per Table 14. The score difference was approximately three to five points versus placebo across studies. The MCID is approximately 2.5 to 5 points (see APPENDIX 5: Validity of Outcome Measures).

TABLE 14: SF-36 RESPONSE AT MONTH 3 AND MONTH 6

		MCS		PCS	
Study and Time of Evaluation	N	MCFB	MDC (95% CI)	MCFB	MDC (95% CI)
1032 month 3					
TOF 5 mg b.i.d.	118	3.52	3.15 (0.87 to 5.43); *P* = 0.0068	5.65	3.63 (1.94 to 5.31); *P* < 0.0001
PL	116	0.37		2.03	
1045 month 3					
TOF 5 mg b.i.d.	231	4.11	3.02 (0.93,5.12); *P* = 0.0048	6.79	4.16 (2.33 to 5.99); *P* < 0.0001
PL	107	1.09		2.63	
1044 month 6					
TOF 5 mg b.i.d.	201	4.75	3.30 (0.95 to 5.66); *P* = 0.0059	6.72	3.15 (1.48 to 4.82); *P* = 0.0002
PL	62	1.45		3.58	
1046 month 6					
TOF 5 mg b.i.d.	205	4.14	1.87 (–0.30 to 4.04); *P* = 0.091	7.44	1.73 (0.05 to 3.42); *P* = 0.044

		MCS		PCS	
Study and Time of Evaluation	**N**	**MCFB**	**MDC (95% CI)**	**MCFB**	**MDC (95% CI)**
PL	68	2.27		5.71	
1064 month 6					
TOF 5 mg b.i.d.	127	5.51	4.62 (2.02 to 7.22); *P* = 0.0005	8.52	3.80 (1.74 to 5.85); *P* = 0.0003
PL	46	0.89		4.72	
ADA 40 mg q.2w.	125	3.93	3.04 (0.44 to 5.65); *P* = 0.022	7.25	2.52 (0.46 to 4.58); *P* = 0.016

ADA = adalimumab; b.i.d. = twice daily; CI = confidence interval; MCFB = mean change from baseline; MCS = mental component summary; MDC = mean difference of change; PL = placebo; PCS = physical component summary; q.2w. = every two weeks; TOF = tofacitinib.
Note: No mean baseline values for SF-36 were provided by the manufacturer.
Source: Manufacturer's Clinical Study Reports and publications for studies 1032,[4-6] 1045,[7,8] 1044,[9-11] 1046,[12,13] and 1064.[14,15]

3.6.7 Functional Assessment of Chronic Illness Therapy-Fatigue Scale

Statistically significant improvements in the FACIT-Fatigue scale were observed for tofacitinib 5 mg twice daily versus placebo at the same time point as the primary outcome (month 3 or month 6) for all trials (Table 15). The score difference was approximately 3 to 5 points versus placebo across studies. The MCID for FACIT-Fatigue is estimated at 3 to 4 points (see APPENDIX 5: Validity of Outcome Measures).

TABLE 15: FACIT-FATIGUE SCORE CHANGES AT MONTH 3 AND MONTH 6

Study and Time of Evaluation	N	MCFB	MDC (95% CI) vs. PL
1032 month 3			
TOF 5 mg b.i.d.	117	6.27	5.15 (2.77 to 7.54); *P* < 0.0001
PL	114	1.11	
1045 month 3			
TOF 5 mg b.i.d.	236	6.70	3.86 (1.93 to 5.78); *P* < 0.0001
PL	109	2.84	
1044 month 6			
TOF 5 mg b.i.d.	201	5.64	3.50 (1.47 to 5.52); *P* = 0.0007
PL	60	2.14	
1046 month 6			
TOF 5 mg b.i.d.	203	6.56	2.85 (0.95 to 4.74); *P* = 0.003
PL	68	3.71	
1064 month 6			
TOF 5 mg b.i.d.	127	6.99	5.07 (2.76 to 7.38); *P* < 0.0001
PL	46	1.92	
ADA 40 mg q.2w.	126	6.47	4.55 (2.24 to 6.86); *P* = 0.0001

ADA = adalimumab; b.i.d. = twice daily; CI = confidence interval; FACIT = Functional Assessment of Chronic Illness Therapy; MCFB = mean change from baseline; MDC = mean difference of change; PL = placebo; q.2w. = every two weeks; TOF = tofacitinib; vs. = versus.
Note: No mean baseline values for FACIT were provided by the manufacturer.
Source: Manufacturer's Clinical Study Reports and publications for studies 1032,[4-6] 1045,[7,8] 1044,[9-11] 1046,[12,13] and 1064.[14,15]

3.6.8 American College of Rheumatology Components

Mean change from baseline of individual components of the ACR was tabulated (tender/painful joint count, swollen joint count, patient assessment of arthritis pain, patient global assessment of arthritis, physician global assessment of arthritis, CRP, and HAQ-DI; see APPENDIX 4: Detailed Outcome Data). No statistical analysis was performed by the manufacturer. Tofacitinib 5 mg twice daily showed numerical improvement relative to placebo for all comparisons.[16]

3.6.9 Subgroups

Several subgroups of interest were specified in the protocol. There were no strong signals from the data that tofacitinib is more effective than placebo in specific populations with a particular treatment history or demographic feature (see Figure 4, APPENDIX 4: Detailed Outcome Data).

3.7 Harms

Only those harms identified in the review protocol are reported (see Section 2.2.1, Protocol). See APPENDIX 4: Detailed Outcome Data for detailed harms data from individual studies.

3.7.1 Adverse Events

Adverse events of primary interest were those from the placebo-controlled period for the tofacitinib 5 mg dose, prior to the possibility of early escape. Most data were summarized in the product monograph and represent pooled data from the five studies.[1]

TABLE 16: SUMMARY OF ADVERSE EVENTS REPORTED BY 1% OR MORE OF PATIENTS TREATED WITH TOFACITINIB (ALL CAUSES) — STUDIES 1032, 1045, 1044, 1046, AND 1064 (UP TO 3 MONTHS)

Adverse Event, n (%)	TOF 5 mg b.i.d. N = 1,216	PL N = 681	ADA 40 mg q.2w. N = 204
Infections and infestations			
Upper respiratory tract infection	53 (4.4)	23 (3.4)	7 (3.4)
Nasopharyngitis	48 (3.9)	19 (2.8)	7 (3.4)
Urinary tract infection	25 (2.1)	12 (1.8)	7 (3.4)
Bronchitis	14 (1.2)	10 (1.5)	4 (2.0)
Herpes zoster	5 (0.4)	2 (0.3)	0
Blood and lymphatic system disorders			
Anemia	15 (1.2)	8 (1.2)	0
Metabolism and nutrition disorders			
Hypercholesterolemia	12 (1.0)	3 (0.4)	1 (0.5)
Nervous system disorders			
Headache	54 (4.4)	15 (2.2)	5 (2.5)
Dizziness	13 (1.1)	8 (1.2)	3 (1.5)
Vascular disorders			
Hypertension	20 (1.6)	7 (1.0)	0
Respiratory, thoracic and mediastinal disorders			
Cough	11 (0.9)	11 (1.6)	4 (2.0)
Gastrointestinal disorders			
Diarrhea	45 (3.7)	16 (2.3)	2 (1.0)
Nausea	32 (2.6)	18 (2.6)	3 (1.5)
Dyspepsia	19 (1.6)	11 (1.6)	3 (1.5)
Abdominal pain, upper	23 (1.9)	5 (0.7)	3 (1.5)
Vomiting	21 (1.7)	10 (1.5)	0
Constipation	16 (1.3)	6 (0.9)	2 (1.0)

Adverse Event, n (%)	TOF 5 mg b.i.d. N = 1,216	PL N = 681	ADA 40 mg q.2w. N = 204
Gastritis	12 (1.0)	7 (1.0)	0
Abdominal pain	10 (0.8)	7 (1.0)	2 (1.0)
Gastroenteritis	12 (1.0)	5 (0.7)	0
Hepatobiliary disorders			
Alanine transaminase increased	14 (1.2)	7 (1.0)	1 (0.5)
Musculoskeletal and connective tissue disorders			
Rheumatoid arthritis	17 (1.4)	17 (2.5)	1 (0.5)
Back pain	18 (1.5)	5 (0.7)	1 (0.5)
Arthralgia	13 (1.1)	16 (2.3)	4 (2.0)
General disorders and administration site conditions			
Edema, peripheral	17 (1.4)	16 (2.3)	3 (1.5)
Influenza	9 (0.7)	5 (0.7)	2 (1.0)
Pyrexia	13 (1.1)	5 (0.7)	1 (0.5)
Investigations			
Blood creatine phosphokinase increased	9 (0.7)	3 (0.4)	1 (0.5)
Weight increased	11 (0.9)	4 (0.6)	2 (1.0)
Injury, poisoning and procedural complications			
Fall	7 (0.6)	4 (0.6)	1 (0.5)

ADA = adalimumab; b.i.d. = twice daily; PL = placebo; q.2w. = every two weeks; TOF = tofacitinib.
Source: Product Monograph, Xeljanz.

a) Overall Infections

In the five controlled trials, from baseline to three months' exposure, the overall frequency of infections was 20% in the 5 mg twice daily tofacitinib group and 18% in the placebo group. The most commonly reported infections were upper respiratory tract infections and nasopharyngitis, and urinary tract infections.

b) Serious Infections

In the five controlled trials, from baseline to three months' exposure, serious infections were reported in one patient (0.6 events per 100 patient-years) who received placebo and in eight patients (2.8 events per 100 patient-years) who received tofacitinib 5 mg twice daily.

In the long-term safety all-exposure population, the overall frequency of serious infections was 2.4% (3.2 events per 100 patient-years) for the tofacitinib 5 mg twice daily group.

The most common serious infections reported with tofacitinib included pneumonia, urinary tract infection, and herpes zoster.

c) Tuberculosis

In the five controlled trials, from baseline to three months' exposure, no cases of tuberculosis were reported in patients who received placebo or tofacitinib 5 mg twice daily. [1]

During zero to 12 months' exposure, tuberculosis was reported in zero patients who received 5 mg twice daily of tofacitinib. Cases of disseminated tuberculosis were also reported. The median tofacitinib exposure prior to diagnosis of tuberculosis was 10 months (range from 152 to 960 days).

d) Opportunistic Infections (Excluding Tuberculosis)

In the five controlled trials, after three months' exposure, no opportunistic infections were reported in patients who received placebo whereas opportunistic infections were reported in 2 (0.2%) patients (0.7 events per 100 patient-years) who received tofacitinib 5 mg twice daily. From baseline to 12 months' exposure, opportunistic infections were reported in three (0.3%) patients (0.3 events per 100 patient-years) who received tofacitinib 5 mg twice daily. The median tofacitinib exposure prior to diagnosis of an opportunistic infection was eight months (range from 41 to 698 days).

e) Malignancy (Excluding Nonmelanoma Skin Cancer)

In the five controlled trials, from baseline to three months' exposure, no malignancies (excluding nonmelanoma skin cancer) were reported in patients who received placebo; however, malignancies were reported in two (0.2%) patients (0.7 events per 100 patient-years) who received tofacitinib 5 mg twice daily. From baseline to 12 months' exposure, malignancies (excluding nonmelanoma skin cancer) were reported in five (0.4%) patients (0.6 events per 100 patient-years) who received tofacitinib 5 mg twice daily.

The most common types of malignancy (excluding nonmelanoma skin cancer), including malignancies observed during the long-term extension, were lung and breast cancer, followed by gastric, colorectal, renal cell, prostate cancer, lymphoma, and malignant melanoma.

In the five controlled trials, from baseline to three months' exposure, nonmelanoma skin cancer was reported in one (0.2%) patient (0.6 events per 100 patient-years) who received placebo and two (0.2%) patients (0.7 events per 100 patient-years) who received tofacitinib 5 mg twice daily. After 12 months' exposure, nonmelanoma skin cancer was reported in three (0.3%) patients (0.3 events per 100 patient-years) who received tofacitinib 5 mg twice daily.

f) Creatine Kinase

Treatment with tofacitinib was associated with increases in creatine kinase (CK). Maximum effects were generally observed within six months. Rhabdomyolysis was reported in one patient in the tofacitinib RA clinical trials. CK levels should be checked in patients with symptoms of muscle weakness and/or muscle pain to evaluate for evidence of rhabdomyolysis.

g) Electrocardiographic Findings

In placebo-controlled phase 2 clinical trials in patients with RA, steady-state treatment with tofacitinib 5 mg or 10 mg twice daily was associated with statistically significant decreases in heart rate of four to seven beats per minute and increases in the PR interval of 4 ms to 10 ms compared with placebo.

h) Lipids

Elevations in lipid parameters (total cholesterol, low-density-lipoprotein [LDL] cholesterol, high-density-lipoprotein [HDL] cholesterol, and triglycerides) generally reached maximal effects at six weeks following initiation of tofacitinib in the controlled double-blind clinical trials. Changes in lipid parameters from baseline through the end of the study (six to 12 months) in the controlled clinical studies were the following:

- Mean LDL cholesterol increased by 14% in the tofacitinib 5 mg twice daily group.
- Mean HDL cholesterol increased by 16% in the tofacitinib 5 mg twice daily group.
- Mean LDL/HDL ratios were essentially unchanged in tofacitinib -treated patients.

In all five controlled studies, 4.4% of patients treated with tofacitinib 5 mg twice daily initiated lipid-lowering medication over the duration of the studies.

i) Liver Enzyme Tests

Confirmed increases in liver enzymes of more than three times the upper limit of normal (ULN) were uncommonly observed. In patients experiencing liver enzyme elevation, modification of treatment regimen, such as reduction in the dose of concomitant DMARD, interruption of tofacitinib, or reduction in tofacitinib dose, resulted in decrease or normalization of liver enzymes. In the controlled portion of Study 1045 (after three months' exposure), ALT elevations of more than three times the ULN were observed in 1.65% and 0.41% of patients receiving placebo and tofacitinib 5 mg twice daily, respectively. In this study, AST elevations of more than three times the ULN were observed in 1.65% and 0.41% of patients receiving placebo and tofacitinib 5 mg twice daily, respectively.

In the controlled portion of the phase 3 studies that included background DMARDs, after three months' exposure, ALT elevations of more than three times the ULN were observed in 0.9% and 1.24% of patients receiving placebo and tofacitinib 5 mg twice daily, respectively. In these studies, AST elevations of more than three times the ULN were observed in 0.72% and 0.52% of patients receiving placebo and tofacitinib 5 mg twice daily, respectively.

j) Lymphocytes

In all five controlled clinical trials, confirmed decreases in absolute lymphocyte counts to less than 500 cells/mm^3 occurred in 0.2% of patients in the tofacitinib 5 mg twice daily group during 12 months of exposure.

Confirmed lymphocyte counts of less than 500 cells/mm^3 were associated with an increased incidence of treated and serious infections.

k) Neutrophils

In the controlled clinical studies, confirmed decreases in absolute neutrophil count (ANC) below 1,000/mm^3 occurred in 0.08% of patients in the tofacitinib 5 mg twice daily group during 12 months of exposure. There were no confirmed decreases in ANC to less than 500/mm^3 observed in any treatment group. There was no clear relationship between neutropenia and the occurrence of serious infections. In the long-term safety population, the pattern and incidence of confirmed decreases in ANC remained consistent with those seen in the controlled clinical studies.

l) Serum Creatinine

In the controlled clinical trials, dose-related elevations in serum creatinine were observed with tofacitinib treatment. The mean increase in serum creatinine was less than 0.1 mg/dL in the 12-month pooled safety analysis; however, with increasing duration of exposure in the long-term extensions, up to 2% of patients were discontinued from tofacitinib treatment owing to the protocol-specified discontinuation criterion of an increase in creatinine of more than 50% of baseline. The clinical significance of the observed serum creatinine elevations is unknown.

3.7.2 Serious Adverse Events

Similar percentages of treatment-emergent serious adverse events (SAEs) were reported (approximately 3%) for patients in the tofacitinib 5 mg or 10 mg twice daily groups as in the placebo groups, during the first three months in all five controlled studies and from months 3 to 6. Most SAEs were reported for one patient in the tofacitinib all-dose groups. During the first three months, three SAEs each were

reported for cellulitis and pneumonia. During months 3 to 6, four SAEs of cellulitis and three SAEs of congestive cardiac failure were reported. The percentages of patients with SAEs in the tofacitinib 5 mg (3.1%) and 10 mg (2.7%) groups were similar to the percentage in the placebo group (3.5%) during the first three months of the studies.

For time points later than six months, SAEs were reported for 3.7% of patients in the tofacitinib all-dose groups; most SAEs were reported for one patient. Eight SAEs of pneumonia, five of chest pain, four of pulmonary tuberculosis, and three of cholelithiasis were reported.

3.7.3 Withdrawals Due to Adverse Events

The rate of withdrawals due to adverse events was approximately 2% to 6% of patients during the placebo-controlled portions of the studies.[25] The rate was slightly higher in patients assigned to tofacitinib at the outset of the trial compared with placebo patients. See APPENDIX 4: Detailed Outcome Data for detailed data.

3.7.4 Mortality

There were 10 deaths in patients taking either placebo or tofacitinib from baseline to 12 months during the five controlled studies.[25] This time period included the 30 days after the last dose of study drug. There was one death in the adalimumab group.[25] The all-cause mortality rate for patients receiving any dose of tofacitinib in the five studies, within 30 days after the last dose of study treatment, was 0.477 deaths per 100 patient-years.

4. DISCUSSION

4.1 Summary of Available Evidence

Five manufacturer-sponsored, published, double-blind randomized controlled trials (N = 3,322) evaluating the efficacy and harms of tofacitinib 5 mg twice daily and 10 mg twice daily, and adalimumab (Study 1064) versus placebo were included in the systematic review. One study was performed in patients with previous inadequate response to a TNF inhibitor (Study 1032) and the others were performed in patients with previous inadequate response to a DMARD or MTX (studies 1044, 1045, 1046, and 1064). Tofacitinib was given as monotherapy in one study (Study 1045) and with background DMARDs in the other studies (studies 1032, 1044, 1046, and 1064). The studies were six to 24 months in duration, but the placebo-controlled periods were not longer than three months prior to early escape.

All five trials used the same three measures for the co-primary end points. The co-primary end point of ACR 20 response was evaluated at month 3 (studies 1032 and 1045) or month 6 (studies 1044, 1046 and 1064). The co-primary end point of change from baseline for HAQ-DI was evaluated at month 3 in all trials. The co-primary end point of DAS28-4(ESR) was evaluated at month 3 (studies 1032 and 1045) or month 6 (studies 1044, 1046, and 1064). Study 1044 used a fourth co-primary end point of mTSS. No other studies measured radiographic outcomes. The step-down procedure proceeded as planned for all primary outcome analyses, except in Study 1044, where one of the co-primary end points did not reach statistical significance.

The studies allowed early escape or crossover to tofacitinib at month 3, and all patients taking placebo were given tofacitinib by month 6 at the latest. This design has numerous limitations, including the fact that patients who meet early escape are not randomized to dose escalation or another type of strategy.

Early escape, while common in RA trials based on ethical considerations, limits the interpretation and clinical relevance of the trial data.

4.2 Interpretation of Results

4.2.1 Efficacy

The trial populations were generally reflective of patients with RA treated in Canadian clinical practice and had reasonable representation of ethnicities and treatment history. The inclusion criteria for the trials were less specific than the criteria applied to using TNF inhibitor drugs for RA by some of the public drug plans in Canada. The Ontario Exceptional Access Program, for example, requires severe active disease (five or more swollen joints, rheumatoid factor positive status, and radiographic evidence of RA) and a trial of two to three DMARDs, including MTX.

Across all five studies and the various study populations, there were consistent statistically significant results favouring tofacitinib versus placebo (with DMARD background therapy or as monotherapy). There were no strong signals from the data that tofacitinib is more effective in specific populations with a particular treatment history or demographic feature (see Figure 4, APPENDIX 4: Detailed Outcome Data). The response rates for tofacitinib appeared to be slightly lower in patients who had previously used TNF inhibitor drugs, but the studies were not designed to test this hypothesis. Not all guidelines have commented on the place in therapy for tofacitinib; however, the European League Against Rheumatism (EULAR) 2013 update suggested that tofacitinib should be used after treatment with multiple biologic DMARD drugs has failed.[27] The efficacy of tofacitinib in this population recommended by EULAR is unknown.

The response rates in the adalimumab treatment group in Study 1064 for the co-primary end points were similar to the rates observed with tofacitinib 5 mg twice daily. The manufacturer submitted three indirect comparison analyses in populations of patients with inadequate response to (1) TNF inhibitors and MTX, (2) DMARDs, and (3) TNF inhibitors. There were no clear signals of tofacitinib being either statistically superior or inferior to other biologic DMARDs.

Tofacitinib is indicated for use with or without MTX. Most patients in the studies (except the monotherapy Study 1045) were taking concomitant MTX. Tofacitinib is indicated for RA in patients who have already used MTX; however, the manufacturer performed an additional study in patients who had not previously used MTX at therapeutic doses (. Relative to the MTX group, tofacitinib 5 mg twice daily showed statistically significant improvements in mTSS, ACR 20/50/70 response rates, HAQ-DI, and DAS-4(ESR).

An important treatment objective in RA is inhibition of progression of structural damage. No statistically significant improvements in radiographic scores were observed in Study 1044 for tofacitinib 5 mg twice daily versus placebo.

Outcomes showed statistically significant improvements in health-related quality of life for the SF-36 physical component and mental component summaries and the FACIT-Fatigue scale. Some of the differences reached the threshold for the MCID. Tofacitinib 5 mg twice daily appeared to meet the threshold for the MCID versus placebo for some scales (FACIT-Fatigue, HAQ-DI), but not for others (mTSS, see APPENDIX 5: Validity of Outcome Measures).

In studies 1044, 1046, and 1064, patients taking tofacitinib 5 mg twice daily who were classified as non-responsive at month 3, [redacted] rates of response at month 6, compared with patients

who were classified as responders at month 3. [redacted].

4.2.2 Harms

The only placebo-controlled data uncontaminated by patient early escape are from month 3 for all five studies. There were no obvious trends observed at this time point, but many safety concerns emerged during the longer-term use of tofacitinib. These were well documented in the product monograph.[1] The highlighted concerns include serious infections (e.g., tuberculosis, cryptococcus, esophageal candidiasis, pneumocystosis, multidermatomal herpes zoster, cytomegalovirus, BK virus infections, cellulitis, and urinary tract infections), lymphoma and malignancies, heart rate decrease and PR interval prolongation, gastrointestinal perforation, liver enzyme elevations and drug-induced liver injury, interstitial lung disease, lymphopenia, neutropenia, and lipid elevations. The European Medicines Agency (EMEA) cited risk of harm and the difficulty in managing this risk in clinical practice, as reasons for rejecting tofacitinib for market approval in Europe in April 2013.[3] Tofacitinib is approved for use in the US, Japan, and other countries.

While the manufacturer did report incidence of adverse events in long-term extension studies, and these data have some value in understanding the risks associated with tofacitinib compared with adalimumab (Study 1064), there was no placebo control group (see APPENDIX 6: Summary of Extension Studies). The randomized trials excluded patients who were at increased risk of developing specific adverse events associated with the use of tofacitinib (e.g., serious infections) and thus may not reflect the incidence in clinical practice. The harms profile of tofacitinib in patients with RA deserves further study in long-term active-controlled and observational studies.

The FDA medical reviewers provided a risk–benefit analysis that compared some of the known benefits and risks of tofacitinib.[16] The figures in Table 17 were derived from the five studies included in this CDR report. Data on number needed to treat were derived from randomized patient data at the time of the primary end point for the studies. Data on number needed to harm were derived from results from baseline to month 12.

TABLE 17: RISK–BENEFIT OVERVIEW OF TOFACITINIB

	Tofacitinib 5 mg b.i.d.
NNT (95% CI)	
ACR 20	4 (3 to 5)
ACR 50	5 (4 to 6)
ACR 70	8 (7 to 10)
NNH (95% CI)	
Malignancy, excluding NMSC	181 (75 to ∞)
Serious infections	58 (29 to ∞)
Tuberculosis	Not calculable
Opportunistic infections	226 (88 to ∞)
Major adverse cardiac event	1,995 (133 to ∞)
GI perforation	Not calculable

ACR = American College of Rheumatology; b.i.d. = twice daily; CI = confidence interval; GI = gastrointestinal; NMSC = nonmelanoma skin cancer; NNT = number needed to treat; NNH = number needed to harm.
Source: FDA Medical Review.[16]

5. CONCLUSIONS

In five double-blind randomized controlled trials in patients with active RA, tofacitinib 5 mg twice daily was associated with higher rates of ACR 20, ACR 50, and ACR 70 response compared with placebo (with or without background DMARDs). Other outcomes, such as the HAQ-DI, DAS28 response, SF-36, and the FACIT-Fatigue scale, also showed statistically significant improvements favouring tofacitinib versus placebo at month 3 or month 6. Radiologic progression did not reach statistical significance in the single study that measured this outcome. Analyses needed to take into account the fact that many patients on placebo had early escape and this may have weakened the internal validity of the results.

There is risk of serious harm such as malignancies and infections with tofacitinib, similar to the risk for anti-TNF alpha drugs used to treat RA. Further research will be needed to ascertain the risks relative to other commonly used biologic drugs and non-biologic DMARDs.

APPENDIX 1: PATIENT INPUT SUMMARY

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

1. Brief Description of Patient Group(s) Supplying Input

Three patient groups representing people with rheumatoid arthritis (RA) provided input.

The Canadian Arthritis Patient Alliance (CAPA) is a national education and advocacy organization that brings together individuals with arthritis from across Canada. CAPA's aim is to improve the quality of life of individuals with arthritis and to educate and encourage individuals to become more effective advocates. CAPA receives funding from private and public sources as well as from individual donors. Over the past 12 months, grants and support have been received from AbbVie, Amgen Canada, Arthritis Alliance of Canada, The Arthritis Society, Canadian Rheumatology Association, Janssen, Novartis, Ontario Rheumatology Association, and UCB Pharma. Previous support has also been received from the Canadian Institutes of Health Research, Hoffmann-La Roche, Pfizer Canada, Rx&D, Schering Canada, Scleroderma Society, and STA HealthCare Communications. CAPA declared no conflicts of interest in the preparation of the submission.

The Arthritis Society is a national organization that provides education, programs, and support to individuals with arthritis, as well as funding for arthritis research. The Arthritis Society receives funding from individual donors and various pharmaceutical companies. Over the past 12 months, these organizations have included AbbVie, Amgen, Bayer, Bristol-Myers Squibb, Celgene, Eli Lilly, GlaxoSmithKline, Janssen, Novartis, Pfizer, Roche, and UCB Pharma. The Arthritis Society declared no conflicts of interest in the preparation of the submission.

Arthritis Consumer Experts (ACE) is a national organization working to educate and empower individuals with arthritis to take control of their disease and improve their quality of life and self-esteem; to make evidence-based information more accessible to and interpretable by the general public, governments and media; and to train individuals with arthritis to be able to contribute meaningfully to research initiatives and governmental decision-making. ACE provides programs in both official languages. ACE receives unrestricted grants-in-aid from public and private sector organizations as well as unsolicited funding from individual donors, including AbbVie, Amgen Canada, Arthritis Research Centre of Canada, BIOTECanada, Bristol-Myers Squibb Canada, Canadian Institutes of Health Research, Canadian Rheumatology Research Consortium, Celgene, GlaxoSmithKline, Hoffmann-La Roche Canada, Janssen, Pfizer Canada, Purdue Pharma, and the University of British Columbia. ACE declared no conflicts of interest in the preparation of the submission.

2. Condition and Current Therapy-Related Information

This information was collected through patient interviews, one-to-one conversations, and other correspondence with patients and caregivers, as well as from survey responses and a review of the literature.

RA is a systemic, autoimmune, chronic condition for which there is no cure, and it results in severe limitations in the day-to-day and productive activities of a patient's life. When diagnosed in childhood or adolescent years, RA has an impact on patients' ability to stay in school and seek post-secondary education, which leads to challenges when seeking employment in adulthood. RA is an inflammatory condition involving swelling, pain, and joint destruction, which may lead to major surgery on affected

joints. The severity of the disease is variable in terms of the intensity, duration, and predictability of periods of inflammation. Cardiovascular disease, osteoporosis, lung disease, and other diseases often accompany RA. Patients commonly report pain as the most concerning symptom, causing difficulty with activities of daily living (e.g., lifting pots and pans, opening containers, bending over, climbing stairs, tying hair back), relationships, employment, and leisure activities (e.g., sports, hiking, riding a motorcycle, skiing). These activities become impossible if RA is not managed. Modified tools are required to do certain tasks. Fatigue, stiffness, lack of concentration, frustration with inability to do certain things, frustration with the progression of the disease, financial concerns, and limitations in mobility are also of importance to patients.

Caregivers of those living with RA also face considerable demands. For example, caregivers must perform all the household tasks that their loved ones are unable to help with in addition to their existing responsibilities. Family members may be required to take time off work to assist the patient, or, if they are unable to do so, the family may face financial hardship from having to hire home care assistance. Patients with RA can experience strain on their relationships, and it can be difficult for family members to see their loved one change in their level of activity and health. The frustration of not having a caregiver to help out is also recognized by patients.

Current treatments for RA include disease-modifying antirheumatic drugs (DMARDs, including biologics and methotrexate), nonsteroidal anti-inflammatory drugs (NSAIDs), corticosteroids, and analgesics. Patients often require multiple drugs in combination to manage their RA. When patients respond to treatment it can be very effective, yet for others, current therapies are partially or completely ineffective. Even when current treatment is effective, patients often fear that at some point it will stop working for them and they may not be able to find a suitable replacement. This is especially a concern for young patients who will require treatment for the rest of their lives.

Currently available RA medications have several adverse effects, including fever, night sweats, nausea, vomiting, fatigue, easy bruising or bleeding, dizziness, itching, weight loss, feeling full after eating a small amount, stomach pain, pale skin, shortness of breath, rapid heart rate, loss of appetite, dark urine, clay-coloured stools, jaundice, dry skin, hair loss and suppression of the immune system. Because of the route of administration (intravenous or infusion), adverse effects also include injection-site reactions, vein scarring and scar tissue. Administering injections also creates logistical challenges. Patients often have to take time off work to receive injections, travel to a clinic and pay parking fees, and face waiting lists to see a rheumatologist. Frequent injections also make travel difficult because the medication needs to be kept refrigerated, needles need to be brought through security checkpoints, and coverage of a larger treatment supply may be restricted by drug plans. RA medications are very expensive, and thus patients need to have private insurance or take on extra work to cover this cost. There is also a significant paperwork burden with provincial drug plans to approve requests for drug coverage. The patient groups emphasized that having a range of treatment options increases the likelihood that patients will have better access to affordable and effective medication with fewer side effects.

3. Related Information About the Drug Being Reviewed

Some members of the patient groups have had experience with tofacitinib. Tofacitinib was found to reduce swelling, pain and use of pain medication, and increase mobility and the ability to do creative work. One patient still experienced fatigue and had a mildly irritated stomach, while another had occasional headaches and muscle spasms. The pill form is expected to be a more convenient method of administration, especially when travelling (as the patient does not have to search for intravenous administration sites). According to the patients interviewed, the pill form can be more easily

administered than the injection form and is more portable. As a result, it was speculated that the pill form would improve adherence.

The availability of a new drug is always a good thing for patients with RA. Patient groups expected that tofacitinib would result in less pain and less suppression of the immune system; since it has a new mechanism of action, it would offer patients another alternative in terms of having a potentially helpful medication available to them. Furthermore, it was believed that it may be effective as a monotherapy.

4. Additional Information

It was expressed that new drugs are particularly valuable for RA because of the variable and unpredictable nature of the disease.

One patient group was concerned about the costs associated with tofacitinib affecting the decision-making process. There is a worry that patients will have to switch from an effective medication to a less costly medication, although they may not respond as well. The patient group wanted to reinforce that it often takes a lot of time for a patient to find an effective treatment, and if required to switch to a less costly one, the patient may spend unnecessary time with unstable RA, resulting in greater health care dollars spent.

APPENDIX 2: LITERATURE SEARCH STRATEGY

OVERVIEW	
Interface:	Ovid
Databases:	Embase 1974 to present MEDLINE Daily and MEDLINE 1946 to present MEDLINE In-Process & Other Non-Indexed Citations **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	September 18, 2014
Alerts:	Weekly search updates until (date of CDEC meeting)
Study Types:	No search filters were applied
Limits:	No date or language limits were used Conference abstracts were excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
.sh	At the end of a phrase, searches the phrase as a subject heading
MeSH	Medical Subject Heading
exp	Explode a subject heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
adj	Requires words are adjacent to each other (in any order)
adj#	Adjacency within # number of words (in any order)
.ti	Title
.ab	Abstract
.ot	Original title
.hw	Heading word; usually includes subject headings and controlled vocabulary
.pt	Publication type
.rn	CAS registry number
.nm	Name of substance word
pmez	Ovid database code; MEDLINE In-Process & Other Non-Indexed Citations, MEDLINE Daily and Ovid MEDLINE 1946 to Present
oemezd	Ovid database code; Embase 1974 to present, updated daily

MULTI-DATABASE STRATEGY	
1	(Xeljanz or tofacitinib or Jakvinus or CP-690550 or tasocitinib).ti,ot,ab,sh,rn,hw,nm.
2	(janus adj kinase adj inhibitor*).ti,ab.
3	540737-29-9.rn,nm.
4	or/1-3
5	exp arthritis rheumatoid/ use pmez
6	exp *rheumatoid arthritis/ use oemezd
7	((rheumatoid or inflammatory or rheumatic) adj2 arthritis).ti,ab.

MULTI-DATABASE STRATEGY	
8	((chronic or rheumatic) adj2 (polyarthritis or poly-arthritis)).ti,ab.
9	(arthritis deformans or arthrosis deformans or Beauvais disease or rheumarthritic or rheumatism or rheumatic or RA).ti,ab.
10	((still* or felty* or caplan* or sicca* or sjogren* or chauffard*) adj2 (syndrome* or disease*)).ti,ab.
11	or/5-10
12	4 and 11
13	exp animals/
14	exp animal experimentation/ or exp animal experiment/
15	exp models animal/
16	nonhuman/
17	exp vertebrate/ or exp vertebrates/
18	animal.po.
19	or/13-18
20	exp humans/
21	exp human experimentation/ or exp human experiment/
22	human.po.
23	or/20-22
24	19 not 23
25	12 not 24
26	25 not conference abstract.pt.
27	remove duplicates from 26

Other Databases	
PubMed	Same MeSH, keywords, limits, and study types used as per MEDLINE search, with appropriate syntax used.
Trial registries (Clinicaltrials.gov and others)	Same keywords, limits used as per MEDLINE search.

Grey Literature

Dates for Search:	September, 2014
Keywords:	Xeljanz (tofacitnib), rheumatoid arthritis, subcutaneous
Limits:	No date or language limits used

Relevant websites from the following sections of the CADTH grey literature checklist, "Grey matters: a practical tool for evidence-based searching" (http://www.cadth.ca/en/resources/finding-evidence-is/grey-matters) were searched:

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

APPENDIX 3: EXCLUDED STUDIES

Phase 2 Study; Focus is Not on Approved Dose
Coombs JH, Bloom BJ, Breedveld FC, Fletcher MP, Gruben D, Kremer JM, et al. Improved pain, physical functioning and health status in patients with rheumatoid arthritis treated with CP-690,550, an orally active Janus kinase (JAK) inhibitor: results from a randomised, double-blind, placebo-controlled trial. Ann Rheum Dis. 2010 Feb;69(2):413-6.

Fleischmann R, Cutolo M, Genovese MC, Lee EB, Kanik KS, Sadis S, et al. Phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) or adalimumab monotherapy versus placebo in patients with active rheumatoid arthritis with an inadequate response to disease-modifying antirheumatic drugs. Arthritis Rheum. 2012 Mar;64(3):617-29.

Kremer JM, Bloom BJ, Breedveld FC, Coombs JH, Fletcher MP, Gruben D, et al. The safety and efficacy of a JAK inhibitor in patients with active rheumatoid arthritis: Results of a double-blind, placebo-controlled phase IIa trial of three dosage levels of CP-690,550 versus placebo. Arthritis Rheum. 2009 Jul;60(7):1895-905.

Kremer JM, Cohen S, Wilkinson BE, Connell CA, French JL, Gomez-Reino J, et al. A phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) versus placebo in combination with background methotrexate in patients with active rheumatoid arthritis and an inadequate response to methotrexate alone. Arthritis Rheum. 2012 Apr;64(4):970-81.

Tanaka Y, Suzuki M, Nakamura H, Toyoizumi S, Zwillich SH, Tofacitinib Study Investigators. Phase II study of tofacitinib (CP-690,550) combined with methotrexate in patients with rheumatoid arthritis and an inadequate response to methotrexate. Arthritis Care Res (Hoboken). 2011 Aug;63(8):1150-8.

Included Patients Who Did Not Have Prior Use of Methotrexate
Lee EB, Fleischmann R, Hall S, Wilkinson B, Bradley JD, Gruben D, et al. Tofacitinib versus methotrexate in rheumatoid arthritis. N Engl J Med. 2014 Jun 19;370(25):2377-86.

Retrospective Cohort Study
Kim JW, Choi IA, Lee EY, Song YW, Lee EB. Tofacitinib prevents radiographic progression in rheumatoid arthritis. J Korean Med Sci. 2013 Aug [cited 2014 Oct 22];28(8):1134-8. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3744699/pdf/jkms-28-1134.pdf

Tofacitinib Patients Not Randomized
McInnes IB, Kim HY, Lee SH, Mandel D, Song YW, Connell CA, et al. Open-label tofacitinib and double-blind atorvastatin in rheumatoid arthritis patients: a randomised study. Ann Rheum Dis. 2014 Jan;73(1):124-31.

Methotrexate-Naive (Include in Supplemental Issue)
Clinical Study Report: A3921069 Phase 3 randomized, double-blind study of the efficacy and safety of 2 doses of CP-690,550 compared to methotrexate in methotrexate-naïve patients with rheumatoid arthritis (1-Year analysis) [**CONFIDENTIAL** internal manufacturer's report]. Kirkland (QC): Pfizer, Inc; 2012 Oct 18.

Long-term Extension Studies (Include in Supplemental Issue)

Clinical Study Report: A3921024 and A3921041 Amended. Summary of long-term extension studies study A3921024: a long-term, open-label follow-up study of tasocitinib (CP-690,550) for treatment of rheumatoid arthritis [and] study A3921041: a long-term, open-label study of CP-690,550 to confirm the safety following long-term administration of CP-690,550 in the treatment of rheumatoid arthritis [**CONFIDENTIAL** internal manufacturer's report]. Kirkland (QC): Pfizer, Inc; 2012 Jan 31.

APPENDIX 4: DETAILED OUTCOME DATA

FIGURE 3: STEP-DOWN APPROACH FOR ANALYSIS OF CO-PRIMARY END POINTS FOR STUDY 1044[9]

Explanation of Figure:

Testing continues along a given arrow until a

Diamond

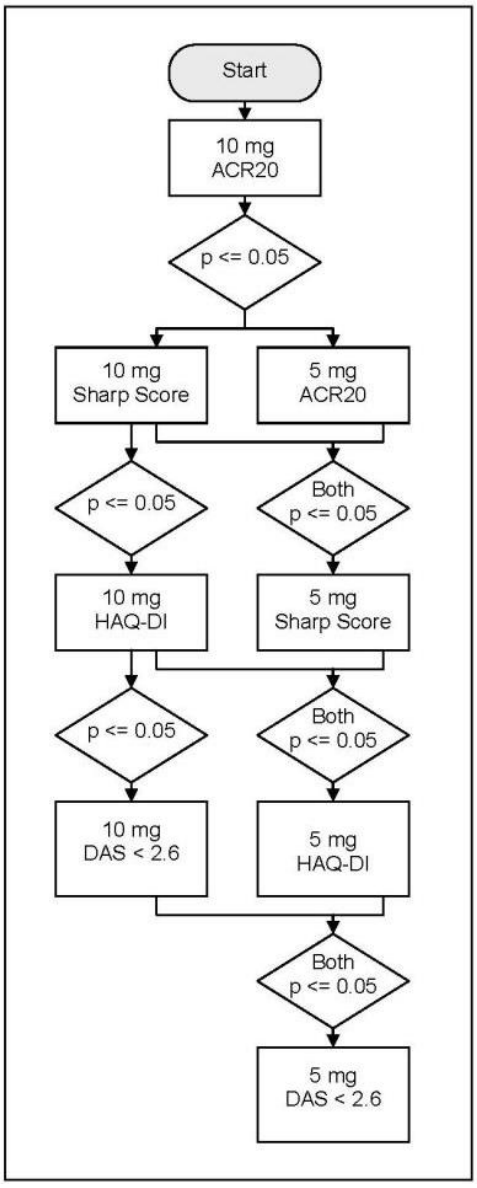
is false.

All testing stops when testing along all arrows stops.

The figure shows the steps to testing, but significance of a test denoted by each

Box

is determined by whether the p-value *of that test* is less than or equal to 0.05 (p <= 0.05), provided the Box itself was reached while stepping through the Figure.



ACR = American College of Rheumatology; DAS = Disease Activity Score; HAQ-DI = Health Assessment Questionnaire–Disability Index.

TABLE 18: SUMMARY OF BASELINE CHARACTERISTICS

	STUDY 1032				STUDY 1045			
	Tofacitinib 5 mg b.i.d. N = 133	Tofacitinib 10 mg b.i.d. N = 134	Placebo → Tofacitinib 5 mg b.i.d. N = 66	Placebo → Tofacitinib 10 mg b.i.d. N = 66	Tofacitinib 5 mg b.i.d. N = 243	Tofacitinib 10 mg b.i.d. N = 245	Placebo → Tofacitinib 5 mg b.i.d. N = 61	Placebo → Tofacitinib 10 mg b.i.d. N = 61
Age, years, mean (SD)	55.4 (11.5)	55.1 (11.3)	54.3 (11.7)	54.5 (11.0)	52.2 (11.5)	52.4 (11.7)	50.7 (12.8)	48.8 (11.9)
Female sex, N (%)	113 (85)	116 (87)	53 (80)	53 (80)	207 (85)	216 (88)	54 (89)	51 (84)
Baseline weight, kg, mean (SD)	77.6 (21.3)	78.8 (20.1)	80.4 (27.7)	81.7 (20.6)	72.2 (20.2)	71.6 (19.6)	69.8 (15.1)	75.3 (21.6)
Disease duration, years, mean (SD)	13.0 (NR)	12.6 (NR)	11.3 (NR)	11.2 (NR)	8.0 (NR)	8.6 (NR)	7.3 (NR)	8.1 (NR)
Disease severity								
Tender joints, mean number (SD)	28.4 (18.3)	27.6 (15.7)	26.7 (16.4)	29.7 (17.1)	29.4 (15.0) n = 240	29.1 (15.6) n = 243	28.4 (15.3)	29.4 (16.6)
Swollen joints, mean number (SD)	16.2 (10.1)	16.6 (9.9)	15.1 (9.0)	19.3 (11.8)	16.3 (8.6) n = 240	17.0 (10.4) n = 243	16.8 (10.0)	17.7 (11.5)
HAQ-DI score, mean (SD)	1.6 (0.7)	1.5 (0.6)	1.6 (0.6)	1.7 (0.7)	1.5 (0.7) n = 240	1.5 (0.6) n = 241	1.5 (0.6)	1.6 (0.7)
Assessment of disease activity								
Pain (patient), 0–100 mm VAS, mean (SD)	65.7 (22.8) n = 127	60.1 (23.2) n = 128	61.6 (22.9) n = 66	59.9 (24.3) n = 65	61.4 (22.3) n = 241	62.0 (23.6) n = 243	60.6 (20.7)	63.0 (21.9)
Disease activity (patient), 0–100 mm VAS, mean (SD)	64.7 (23.2) n = 127	58.8 (23.6) n = 128	63.4 (21.9) n = 66	60.3 (24.0) n = 65	61.7 (22.0) n = 240	63.5 (23.2) n = 240	63.3 (21.6)	62.0 (22.3)
Disease activity (physician), 0–100 mm VAS, mean (SD)	65.4 (18.2) n = 126	58.7 (19.4) n = 128	65.0 (14.2) n = 65	63.7 (18.9) n = 66	61.5 (16.9) n = 241	60.8 (16.8) n = 243	62.3 (14.3)	62.1 (19.2)
DAS28(CRP), mean (SD)	5.4 (1.0)	5.3 (0.9)	5.2 (0.9)	5.6 (1.0)	5.7 (0.9) n = 239	5.6 (0.9) n = 242	5.6 (0.8)	5.6 (0.9) n = 60
DAS28(ESR), mean (SD)	6.5 (1.1)	6.4 (0.9)	6.3 (1.0)	6.6 (1.1)	6.7 (0.9) n = 236	6.7 (0.9) n = 234	6.6 (0.9) n = 58	6.7 (1.0) n = 57
CRP, mg/L, mean (SD)	19.3 (27.5)	15.7 (21.6)	15.4 (15.9)	18.0 (22.8)	22.8 (27.0) n = 242	19.0 (19.9) n = 244	14.1 (13.0)	21.5 (32.7) n = 60
Positive RF status, N (%)	80 (61)	83 (62)	40 (61)	46 (71)	171 (71) N = 240	157 (65) N = 241	35 (57)	29 (48)
Anti-CCP positive, N (%)	89 (68)	90 (70)	48 (74)	49 (78)	172 (71) N = 242	169 (69) N = 244	44 (72)	33 (55) N = 60

	STUDY 1032				STUDY 1045			
	Tofacitinib 5 mg b.i.d. N = 133	Tofacitinib 10 mg b.i.d. N = 134	Placebo → Tofacitinib 5 mg b.i.d. N = 66	Placebo → Tofacitinib 10 mg b.i.d. N = 66	Tofacitinib 5 mg b.i.d. N = 243	Tofacitinib 10 mg b.i.d. N = 245	Placebo → Tofacitinib 5 mg b.i.d. N = 61	Placebo → Tofacitinib 10 mg b.i.d. N = 61
Therapy prior to enrolment, N (%)								
MTX	133 (100)	134 (100)	66 (100)	66 (100)	209 (86)	207 (84)	51 (84)	51 (84)
Non-biologic DMARDs other than MTX	53 (40)	37 (28)	16 (24)	17 (26)	158 (65)	164 (67)	44 (72)	39 (64)
Failed DMARDS (inadequate response, mean number/patient)	NR	NR	NR	NR	NR	NR	NR	NR
TNF inhibitor	132 (99)	132 (99)	66 (100)	66 (100)	34 (14)	41 (17)	12 (20)	12 (20)
Adalimumab	65 (49)	74 (55)	36 (55)	42 (64)	16 (7)	21 (9)	5 (8)	5 (8)
Certolizumab	2 (2)	2 (1)	3 (5)	2 (3)	0	0	0	0
Etanercept	65 (49)	57 (43)	29 (44)	28 (42)	21 (9)	20 (8)	7 (12)	4 (7)
Golimumab	5 (4)	8 (6)	2 (3)	5 (8)	0	2 (< 1)	0	0
Infliximab	56 (42)	42 (31)	27 (41)	16 (24)	13 (5)	22 (9)	6 (10)	6 (10)
Number of previous TNF inhibitors								
1	83	89	85		NR	NR	NR	NR
2	37	30	37		NR	NR	NR	NR
≥ 3	11	12	9		NR	NR	NR	NR
Previous non-TNF inhibitor biologics	21 (16)	11 (8)	4 (6)	10 (15)	12 (5)	19 (8)	6 (10)	4 (7)
Concomitant therapy, N (%)								
Concomitant DMARDs other than MTX	1 (< 1)	1 (< 1)	1 (2)	0	1 (< 1)	0	0	1(2)
Concomitant corticosteroids	85 (64)	81 (60)	43 (65)	40 (61)	140 (57)	148 (60)	35 (57)	42 (69)
Mean concomitant prednisone dose, mg, mean (SD)	NR	NR	NR	NR	NR	NR	NR	NR
Concomitant antimalarials	12 (9)	7 (5)	2 (3)	3 (5)	45 (18)	41 (17)	8 (13)	7 (12)
Lipid-lowering therapy	2 (2)	3 (2)	2 (3)	2 (3)	28 (12)	36 (15)	6 (10)	2 (3)

b.i.d. = twice daily; CCP = cyclic citrullinated peptide; CRP = C-reactive protein; DAS = Disease Activity Score; DMARD = disease-modifying antirheumatic drug; ESR = erythrocyte sedimentation rate; HAQ-DI = Health Assessment Questionnaire–Disability Index; MTX = methotrexate; NR = not reported; RF = rheumatoid factor; SD = standard deviation; TNF = tumour necrosis factor; VAS = visual analogue scale.

TABLE 19: SUMMARY OF BASELINE CHARACTERISTICS

	STUDY 1044				STUDY 1046			
	Tofacitinib 5 mg b.i.d. N = 321	Tofacitinib 10 mg b.i.d. N = 316	Placebo → Tofacitinib 5 mg b.i.d. N = 81	Placebo → Tofacitinib 10 mg b.i.d. N = 79	Tofacitinib 5 mg b.i.d. N = 315	Tofacitinib 10 mg b.i.d. N = 318	Placebo → Tofacitinib 5 mg b.i.d. N = 79	Placebo → Tofacitinib 10 mg b.i.d. N = 80
Age, years, mean (SD)	53.7 (11.6)	52.0 (11.4)	53.2 (11.5)	52.1 (11.8)	52.7 (11.7)	51.9 (11.8)	50.8 (11.2)	53.3 (10.8)
Female sex, N (%)	269 (84)	273 (86)	65 (80)	72 (91)	264 (84)	258 (81)	63 (80)	60 (75)
Baseline weight, kg, mean (SD)	68.7 (19.8)	66.8 (18.4)	65.6 (18.7)	70.3 (21.9)	69.6 (18.8)	71.0 (18.7)	71.9 (19.1)	70.9 (20.8)
Disease duration, years, mean (SD)	8.9 (NR)	9.0 (NR)	8.8 (NR)	9.5 (NR)	8.1 (NR)	9.2 (NR)	9.5 (NR)	10.2 (NR)
Disease severity								
Tender joints, number (SD)	24.1 (14.0)	23.0 (14.5)	23.3 (13.5)	22.6 (12.9)	25.0 (15.3) n = 312	26.6 (16.1) n = 315	27.2 (16.8)	21.9 (13.0) n = 79
Swollen joints, number (SD)	14.1 (8.2)	14.4 (7.7)	14.0 (7.9)	14.5 (8.9)	14.5 (10.3) n = 312	14.4 (9.7) n = 315	14.6 (9.7)	13.9 (8.6) n = 79
HAQ-DI score, number (SD)	1.4 (0.7)	1.4 (0.7)	1.4 (0.6)	1.2 (0.7)	1.4 (0.7) n = 311	1.4 (0.7) n = 315	1.5 (0.6)	1.2 (0.7) n = 78
Assessment of disease activity								
Pain (patient), 0 to 100 mm VAS, mean (SD)	58.4 (23.1) n = 319	57.6 (24.1) n = 309	57.9 (21.3) n = 79	52.0 (26.1) n = 77	57.1 (23.8) n = 311	58.6 (22.2) n = 315	58.3 (22.6)	55.9 (23.1) n = 79
Disease activity (patient), 0 to 100 mm VAS, mean (SD)	58.1 (23.6) n = 316	56.5 (23.0) n = 308	54.6 (20.1) n = 79	53.6 (25.5) n = 77	59.0 (22.9) n = 311	60.2 (22.5) n = 315	59.1 (23.2)	56.8 (23.6) n = 79
Disease activity (physician), 0 to 100 mm VAS, mean (SD)	59.4 (15.9) n = 316	58.4 (17.1) n = 309	55.4 (18.4) n = 79	56.5 (16.4) n = 77	60.5 (17.8) n = 310	59.7 (17.0) n = 313	59.0 (15.7)	58.8 (17.4) n = 79
DAS28-CRP, mean (SD)	5.2 (0.9) n = 316	5.2 (0.9) n = 309	5.1 (0.9) n = 79	5.2 (0.9) n = 77	5.2 (0.9) n = 312	5.3 (1.0) n = 315	5.3 (0.9)	5.1 (1.0) n = 79
DAS28-ESR,mean (SD)	6.3 (1.0) n = 319	6.3 (1.0) n = 306	6.3 (1.0) n = 78	6.3 (1.1) n = 76	6.3 (1.0) n = 310	6.4 (1.1) n = 315	6.4 (1.0)	6.1 (1.0) n = 79
CRP, mg/L,mean (SD)	15.5 (19.1) n = 319	17.0 (26.9) n = 309	12.2 (14.5) n = 79	15.3 (15.1) n = 77	17.7 (21.4) n = 312	17.7 (21.9) n = 313	16.9 (16.5)	16.5 (18.2) n = 79
Positive RF status, N (%)	237 (75) N = 315	239 (78) N = 308	63 (80) N = 79	58 (75) N = 77	227 (74) n = 307	228 (73) n = 313	57 (73) n = 78	57 (72) n = 79
Anti-CCP positive, N (%)	275 (86) N = 320	266 (84) N = 315	68 (84) N = 81	65 (82) N = 79	237 (77) n = 307	241 (76) n = 316	59 (76) n = 78	61 (76) n = 80

	STUDY 1044				STUDY 1046			
	Tofacitinib 5 mg b.i.d. N = 321	Tofacitinib 10 mg b.i.d. N = 316	Placebo → Tofacitinib 5 mg b.i.d. N = 81	Placebo → Tofacitinib 10 mg b.i.d. N = 79	Tofacitinib 5 mg b.i.d. N = 315	Tofacitinib 10 mg b.i.d. N = 318	Placebo → Tofacitinib 5 mg b.i.d. N = 79	Placebo → Tofacitinib 10 mg b.i.d. N = 80
Therapy prior to enrolment, N (%)								
MTX	321 (100)	315 (100)	81 (100)	79 (100)	273 (87)	263 (83)	66 (84)	66 (83)
Non-biologic DMARDs other than MTX	193 (60)	192 (61)	62 (77)	46 (58)	232 (74)	242 (76)	55 (70)	62 (78)
Failed DMARDS (inadequate response, mean number/patient)	NR	NR	NR	NR	1.4	1.4	1.3	1.4
TNF inhibitor	62 (19)	50 (16)	8 (10)	7 (9)	23 (7)	19 (6)	5 (6)	5 (6)
Adalimumab	21 (7)	20 (6)	2 (2)	1 (1)	8 (3)	10 (3)	2 (3)	2 (3)
Certolizumab	3 (1)	1 (< 1)	0	0	1 (< 1)	0	0	0
Etanercept	25 (8)	29 (9)	2 (2)	3 (4)	11 (3)	7 (2)	1 (1)	3 (4)
Golimumab	4 (1)	1 (< 1)	2 (2)	0	2 (1)	1 (< 1)	1 (1)	0
Infliximab	25 (8)	18 (6)	5 (6)	4 (5)	6 (2)	6 (2)	1 (1)	2 (3)
Number of previous TNF inhibitors								
1	NR	NR	NR	NR	NR	NR	NR	NR
2	NR	NR	NR	NR	NR	NR	NR	NR
≥ 3	NR	NR	NR	NR	NR	NR	NR	NR
Previous non-TNF inhibitors biologics	17 (5)	15 (5)	3 (4)	2 (3)	7 (2)	10 (3)	6 (8)	0
Concomitant therapy, N (%)								
Concomitant DMARDs other than MTX	4 (1)	1 (< 1)	0	0	NR	NR	NR	NR
Concomitant corticosteroids	200 (62)	190 (60)	43 (53)	42 (53)	195 (62)	182 (57)	47 (60)	47 (59)
Mean concomitant prednisone dose, mg (SD)	NR	NR	NR	NR	6.4 (2.8)	6.6 (3.1)	7.4 (2.8)	7.1 (3.0)
Concomitant antimalarials	NR	NR	NR	NR	NR	NR	NR	NR
Lipid-lowering therapy	20 (6)	28 (9)	7 (9)	2 (3)	15 (5)	22 (7)	2 (3)	1 (1)

b.i.d. = twice daily; CCP = cyclic citrullinated peptide; CRP = C-reactive protein; DAS = Disease Activity Score; DMARD = disease-modifying antirheumatic drug; ESR = erythrocyte sedimentation rate; HAQ-DI = Health Assessment Questionnaire–Disability Index; MTX = methotrexate; NR = not reported; RF = rheumatoid factor; SD = standard deviation; TNF = tumour necrosis factor; VAS = visual analogue scale.

TABLE 20: SUMMARY OF BASELINE CHARACTERISTICS FOR STUDY 1064

	Tofacitinib 5 mg b.i.d. N = 204	Tofacitinib 10 mg b.i.d. N = 201	Placebo → Tofacitinib 5 mg b.i.d. N = 56	Placebo → Tofacitinib 10 mg b.i.d. N = 52	Adalimumab 40 mg SC q.2w. N = 204
Age, years, mean (SD)	53.0 (11.9)	52.9 (11.8)	55.5 (13.7)	51.9 (13.7)	52.5 (11.7)
Female sex, N (%)	174 (85)	168 (84)	43 (77)	39 (75)	162 (79)
Baseline weight, kg, mean (SD)	71.8 (18.9)	73.1 (17.8)	67.7 (14.2)	71.4 (24.3)	72.4 (16.1)
Disease duration, years, mean (SD)	7.6 (NR)	7.4 (NR)	6.9 (NR)	9.0 (NR)	8.1 (NR)
Disease severity					
Tender joints, mean (SD)	28.5 (15.0) n = 201	26.1 (14.1) n = 199	26.6 (14.4) n = 55	28.1 (14.4) n = 51	26.7 (15.3) n = 201
Swollen joints, mean (SD)	16.7 (8.8) n = 201	15.8 (7.8) n = 199	16.9 (10.0) n = 55	16.4 (7.5) n = 51	16.4 (8.7) n = 201
HAQ-DI score, mean (SD)	1.5 (0.6) n = 201	1.5 (0.6) n = 199	1.5 (0.7) n = 55	1.4 (0.7) n = 51	1.5 (0.6) n = 201
Assessment of disease activity					
Pain (patient), 0 to 100 mm VAS, mean (SD)	59.3 (21.0) n = 201	59.0 (22.2) n = 199	57.6 (20.6) n = 55	52.6 (21.8) n = 51	56.5 (21.9) n = 201
Disease activity (patient), 0 to 100 mm VAS, mean (SD)	59.9 (21.4) n = 201	56.6 (23.8) n = 199	59.1 (20.7) n = 55	49.4 (20.9) n = 51	57.2 (22.2) n = 201
Disease activity (physician), 0 to 100 mm VAS, mean (SD)	59.9 (16.8) n = 199	59.6 (16.7) n = 199	62.4 (18.7) n = 55	58.0 (13.6) n = 51	58.6 (16.0) n = 201
DAS28-CRP, mean (SD)	5.4 (0.9) n = 200	5.4 (0.8) n = 199	5.6 (1.0) n = 55	5.3 (0.8) n = 51	5.3 (0.9) n = 201
DAS28-ESR, mean (SD)	6.6 (0.9) n = 195	6.5 (0.9) n = 194	6.6 (1.0) n = 54	6.3 (0.8) n = 49	6.4 (0.9) n = 194
CRP, mg/L, mean (SD)	14.9 (18.6) n = 200	17.3 (19.5) n = 199	20.3 (20.1) n = 56	11.6 (16.3) n = 51	17.5 (22.5) n = 201
Positive RF status, N (%)	133 (67) n = 199	131 (66) n = 198	40 (71) n = 56	31 (61) n = 51	137 (68) n = 201
Anti-CCP positive, N (%)	144 (71) n = 202	126 (64) n = 197	42 (76) n = 55	31 (62) n = 50	151 (75) n = 202
Therapy prior to enrolment, N (%)					
MTX	200 (98)	201 (100)	56 (100)	51 (98)	202 (99)
Non-biologic DMARDs other than MTX	109 (53)	115 (57)	30 (54)	29 (56)	114 (56)
Failed DMARDS (inadequate response, mean number/patient)	NR	NR	NR	NR	NR

	Tofacitinib 5 mg b.i.d. N = 204	Tofacitinib 10 mg b.i.d. N = 201	Placebo → Tofacitinib 5 mg b.i.d. N = 56	Placebo → Tofacitinib 10 mg b.i.d. N = 52	Adalimumab 40 mg SC q.2w. N = 204
TNF inhibitor	12 (6)	14 (7)	4 (7)	5 (10)	16 (8)
Adalimumab	NR	NR	NR	NR	NR
Certolizumab	1 (< 1)	1 (< 1)	1 (2)	1 (2)	3 (1)
Etanercept	7 (3)	7 (3)	3 (5)	1 (2)	5 (2)
Golimumab	2 (1)	3 (1)	0	1 (2)	5 (2)
Infliximab	4 (2)	6 (3)	0	2 (4)	4 (2)
Number of previous TNF inhibitors					
1	NR	NR	NR	NR	NR
2	NR	NR	NR	NR	NR
≥ 3	NR	NR	NR	NR	NR
Previous non-TNF inhibitor biologics	2 (1)	4 (2)	4 (7)	2 (4)	3 (2)
Concomitant Therapy, N (%)					
Concomitant DMARDs other than MTX	1 (1)	1 (1)	0	1 (2)	2 (1)
Concomitant corticosteroids	117 (57)	121 (60)	40 (71)	29 (56)	116 (57)
Mean concomitant prednisone dose, mg (SD)	NR	NR	NR	NR	NR
Concomitant antimalarials	NR	NR	NR	NR	NR
Lipid-lowering therapy	8 (4)	10 (5)	1 (2)	3 (6)	10 (5)

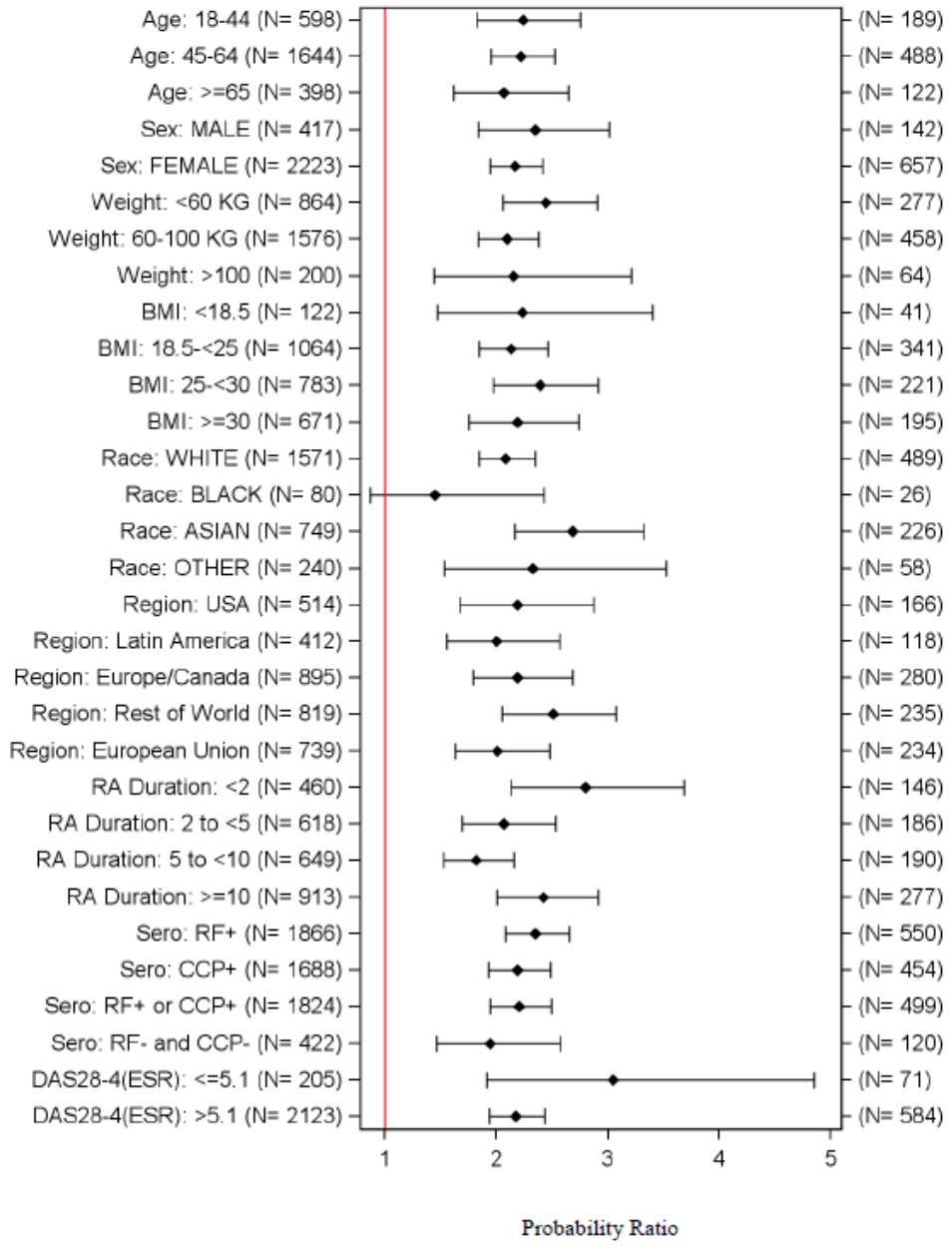
b.i.d. = twice daily; CCP = cyclic citrullinated peptide; CRP = C-reactive protein; DAS = Disease Activity Score; DMARD = disease-modifying antirheumatic drug; ESR = erythrocyte sedimentation rate; HAQ-DI = Health Assessment Questionnaire–Disability Index; MTX = methotrexate; NR = not reported; q.2w. = every two weeks; RF = rheumatoid factor; SC = subcutaneous injection; SD = standard deviation; TNF = tumour necrosis factor; VAS = visual analogue scale.

TABLE 21: SUMMARY OF PATIENTS WITH ACR 20 RESPONSE (FAS POPULATION USING NON-RESPONDER IMPUTATION) AT MONTH 3, BY PREVIOUS TREATMENT WITH A TNF INHIBITOR

	N	ACR20 Responders n (%)	N	ACR20 Responders n (%)	N	ACR20 Responders n (%)
Monotherapy						
	Placebo		**CP-690,550 5 mg BID**		**CP-690,550 10 mg BID**	
No previous TNFi	200	47 (23.5)	298	187 (62.8)	306	214 (69.9)
One TNFi	22	5 (22.7)	28	17 (60.7)	32	21 (65.6)
Adalimumab	9	3 (33.3)	7	1 (14.3)	10	7 (70.0)
Certolizumab pergol	0		0		0	
Etanercept	4	0	13	10 (76.9)	7	3 (42.9)
Golimumab	0		0		0	
Infliximab	9	2 (22.2)	8	6 (75.0)	15	11 (73.3)
D/C due to LOE	14	3 (21.4)	14	7 (50.0)	20	14 (70.0)
D/C due to AE	2	0	0		4	3 (75.0)
D/C due LOE and AE	0		4	3 (75.0)	1	0
Two TNFi	6	1 (16.7)	15	5 (33.3)	11	7 (63.6)
Three or more TNFi	3	0	1	1 (100)	7	4 (57.1)
Background DMARD						
	Placebo + DMARD		**CP-690,550 5 mg BID + DMARD**		**CP-690,550 10 mg BID +DMARD**	
No previous TNFi	479	134 (28.0)	817	479 (58.6)	836	540 (64.6)
One TNFi	110	30 (27.3)	155	69 (44.5)	146	73 (50.0)
Adalimumab	46	14 (30.4)	47	24 (51.1)	49	19 (38.8)
Certolizumab pergol	5	1 (20.0)	5	2 (40.0)	6	3 (50.0)
Etanercept	25	5 (20.0)	47	20 (42.6)	48	30 (62.5)
Golimumab	6	1 (16.7)	8	1 (12.5)	9	2 (22.2)
Infliximab	28	9 (32.1)	48	22 (45.8)	34	19 (55.9))
D/C due to LOE	70	20 (28.6)	90	38 (42.2)	82	40 (48.8)
D/C due to AE	10	2 (20.0)	12	6 (50.0)	12	7 (58.3)
D/C due LOE and AE	2	1 (50.0)	3	1 (33.3)	8	2 (25.0)
Two TNFi	46	6 (13.0)	58	30 (51.7)	46	25 (54.3)
Three or more TNFi	10	3 (30.0)	16	5 (31.3)	19	9 (47.4)

ACR = American College of Rheumatology; AE = adverse events; BID = twice daily; CP-690,550 = tofacitinib; D/C = discontinued; DMARD = disease-modifying antirheumatic drug; FAS = full analysis set; LOE = lack of effectiveness; TNF = tumour necrosis factor; TNFi = tumour necrosis factor inhibitor.
Source: Manufacturer pooled data from phase 2 and phase 3 studies, Summary of Clinical Efficacy.[25]

FIGURE 4: FOREST PLOT OF ACR 20 PROBABILITY RATIO (TOFACITINIB ANY DOSE VERSUS PLACEBO)



ACR = American College of Rheumatology; AE = adverse events; BID = twice daily; BMI = body mass index; CP-690,550 = tofacitinib; D/C = discontinued; DMARD = disease-modifying antirheumatic drug; ESR = erythrocyte sedimentation rate; FAS = full analysis set; LOE = lack of effectiveness; RA = rheumatoid arthritis; TNF = tumour necrosis factor; TNFi = tumour necrosis factor inhibitor.

Source: Manufacturer pooled data from Phase 2 and Phase 3 studies, Summary of Clinical Efficacy.[25]

TABLE 22: MEAN CHANGE FROM BASELINE IN ACR COMPONENTS AT THREE MONTHS

Mean Change from Baseline in ACR Components at 3 Months, by Phase 3 Trial												
	Placebo			CP 5 mg BID			CP 10 mg BID			Adalimumab		
Study	N	LS Mean	SE	N	LS Mean	SE	N	LS Mean	SE	N	LS Mean	SE
Tender painful joint count												
1032	118	-6.74	1.27	119	-12.49	1.25	125	-14.12	1.28			
1044	146	-5.66	0.83	294	-11.33	0.59	300	-12.99	0.59			
1046	148	-8.57	0.85	294	-13.62	0.62	292	-15.6	0.63			
1064	98	-6.83	1.09	188	-14.21	0.81	185	-14.77	0.81	190	-13.52	0.81
1045*	109	-9.14	1.14	237	-15.01	0.79	229	-16.41	0.8			
Swollen joint count												
1032	118	-4.78	0.8	119	-9.28	0.79	125	-10	0.8			
1044	146	-4.06	0.54	294	-8.35	0.38	300	-8.72	0.38			
1046	148	-5.09	0.6	294	-7.76	0.43	292	-8.71	0.44			
1064	98	-5.98	0.63	188	-10.15	0.47	185	-9.89	0.47	190	-9.28	0.46
1045*	109	-5.14	0.74	237	-9.62	0.51	229	-10.52	0.52			
Patient's assessment of arthritis pain												
1032	115	-8.26	2.41	114	-27.16	2.43	119	-24.95	2.48			
1044	146	-8.88	1.75	295	-23.32	1.24	300	-28.05	1.24			
1046	148	-11.38	1.72	293	-24.18	1.24	292	-26.78	1.25			
1064	98	-9.57	2.17	188	-26.85	1.61	185	-28.11	1.62	190	-22.26	1.6
1045*	108	-10.71	2.14	237	-26.94	1.47	228	-31.06	1.5			
Patient's global assessment of arthritis												
1032	117	-9.19	2.42	114	-23.39	2.43	120	-25.05	2.48			
1044	146	-8.59	1.73	295	-22.31	1.22	299	-27.67	1.22			
1046	148	-12.54	1.72	293	-24.82	1.24	292	-28.19	1.25			
1064	98	-7.37	2.22	188	-23.89	1.65	185	-26.6	1.66	190	-21.37	1.64
1045*	108	-11.19	2.1	237	-26.99	1.45	229	-30.94	1.47			
Physician's global assessment of arthritis												
1032	115	-19.26	2.18	111	-30.26	2.2	119	-30.14	2.22			
1044	146	-13.5	1.45	294	-28.93	1.03	300	-33.83	1.02			
1046	148	-21.22	1.42	291	-30.37	1.03	290	-33.25	1.04			
1064	98	-16.76	1.91	184	-31.61	1.43	185	-31.47	1.43	189	-28.85	1.42
1045*	109	-18.44	1.81	238	-32.39	1.24	229	-36.8	1.27			
C-reactive protein												
1032	118	3.12	1.4	118	-11.77	1.39	124	-12.17	1.41			
1044	145	-0.81	1.1	296	-8.91	0.77	299	-11.53	0.77			
1046	147	-1.99	0.99	281	-11.79	0.72	273	-12.22	0.73			
1064	99	1.42	1.61	186	-9.66	1.2	184	-8.97	1.21	190	-8.24	1.19
1045*	108	-3.26	1.25	237	-12.6	0.85	228	-15.02	0.87			
HAQ-DI												
1032	118	-0.18	0.04	117	-0.43	0.04	125	-0.46	0.04			
1044	146	-0.15	0.04	294	-0.4	0.03	300	-0.54	0.03			
1046	147	-0.21	0.04	292	-0.46	0.03	292	-0.56	0.03			
1064	98	-0.24	0.05	188	-0.55	0.04	185	-0.61	0.04	190	-0.49	0.04
1045*	109	-0.19	0.05	237	-0.5	0.03	227	-0.57	0.03			

ACR = American College of Rheumatology; b.i.d. = twice daily; CP = tofacitinib; HAQ-DI = Health Assessment Questionnaire–Disability Index; LS = least squares; SE = standard error.
Source: FDA Medical Review Report.[16]

TABLE 23: HARMS FOR STUDY 1032 FROM BASELINE TO THREE MONTHS

	Tofacitinib 5 mg b.i.d. N = 133	Tofacitinib 10 mg b.i.d. N = 134	Placebo N = 132
AEs			
Patients with > 0 AEs, N (%)	71 (53)	76 (57)	75 (57)
Most common AEs[a]			
Infections and Infestations, N (%)	24 (18)	29 (22)	26 (20)
Gastrointestinal sisorders, N (%)	24 (18)	25 (19)	21 (16)
Diarrhea	8 (6)	5 (4)	5 (4)

	Tofacitinib 5 mg b.i.d. N = 133	Tofacitinib 10 mg b.i.d. N = 134	Placebo N = 132
Nausea	4 (3)	2 (2)	9 (7)
Musculoskeletal and connective tissue disorders, N (%)	12 (9)	13 (10)	23 (17)
Nervous system disorders, N (%)	6 (5)	13 (10)	9 (7)
Headache	3 (2)	8 (6)	1 (< 1)
General disorders and administration site conditions, N (%)	6 (5)	8 (6)	11 (8)
Injury, poisoning, and procedural complications, N (%)	7 (5)	9 (7)	11 (8)
Investigations, N (%)	4 (3)	8 (6)	3 (2)
Metabolism and nutrition disorders, N (%)	1 (< 1)	5 (4)	9 (7)
Respiratory, thoracic, and mediastinal disorders, N (%)	7 (5)	5 (4)	5 (4)
Skin and subcutaneous tissue disorders	7 (5)	9 (7)	8 (6)
SAEs			
Patients with > 0 SAEs, N (%)	2 (2)	2 (2)	6 (5)
Most common SAEs[a]			
Infections and infestations	0	0	0
WDAEs			
WDAEs, N (%)	8 (6)	6 (5)	7 (5)
Most common reasons			
Infections and infestations			
Deaths			
Number of deaths, N (%)	0	0	0
Other notable harms at month 3			
Malignancies, N (%)	0	0	0
Gastrointestinal perforations, N (%)	NR	NR	NR
Congestive heart failure, N (%)	4 (3)	1 (< 1)	5 (4)
Ischemic heart disease, N (%)	0	3 (2)	0
Myocardial infarction, N (%)	0	3 (2)	0
Cerebrovascular accident, N (%)	1 (< 1)	0	0
Dyslipidemia, N (%)	2 (2)	1 (< 1)	2 (2)
Neutrophil counts, 10^9/L LSM change (SE) from baseline	−0.9 (0.2)	−0.8 (0.2)	0.1 (0.2)
95% CI versus placebo	−1.5 to 0.6	−1.4 to −0.5	
P value versus placebo	< 0.0001	< 0.0001	
Hemoglobin, g/L mean change (SD) from baseline	1.1 (7.0)	0.1 (8.2)	−1.0 (7.8)
95% CI versus placebo	−0.2 to 4.0	−0.9 to 3.1	
P value versus placebo	0.03	0.29	
LDL, LSM % change (SE) from baseline at month 3	0.3 (0.1)	0.3 (0.1)	−0.01 (0.1)
95% CI versus placebo	0.2 to 0.4	0.2 to 0.4	
P value versus placebo	< 0.0001	< 0.0001	
HDL, LSM % change (SE) from baseline	0.4 (0.1)	0.4 (0.1)	0.0 (0.1)
95% CI versus placebo	0.2 to 0.5	0.3 to 0.5	
P value versus placebo	< 0.0001	< 0.0001	
Serum creatinine, µmol/L, LSM change (SE) from baseline	3.1 (1.5)	3.8 (1.5)	3.8 (1.5)
95% CI versus placebo	–4.6 to 3.1	–3.8 to 3.8	
P value versus placebo	0.71	0.90	
Mild (1.5 to 1.999 × 10^9/L) neutropenia incidence, N (%)	3 (3), n = 116	2 (2), n = 124	2 (2), n = 118
Moderate (1.0 to 1.499 × 10^9/L) to severe (0.5 to 0.999 × 10^9/L) neutropenia incidence, N (%)	1 (1), n = 116	0, n = 124	0, n = 118
Decreased hemoglobin, mild to moderate (≥ 10 g/L to	9 (8)	16 (13)	12 (10)

	Tofacitinib 5 mg b.i.d. N = 133	Tofacitinib 10 mg b.i.d. N = 134	Placebo N = 132
≤ 20 g/L), N (%)			
Decreased hemoglobin, potentially life-threatening (decrease of ≥ 30 g/L or hemoglobin ≤ 70 g/L)	0	0	1 (1)
AST > 3 × ULN incidence, N (%)	0	0	0
ALT > 3 × ULN incidence, N (%)	0	2 (2), n = 133	0

AE = adverse event; ALT = alanine aminotransferase; AST = aspartate transaminase; b.i.d. = twice daily; CI = confidence interval; HDL = high-density lipoprotein; LDL = low-density lipoprotein; LSM = least squares mean; NR = not reported; SAE = serious adverse events; SD = standard deviation; SE = standard error; ULN = upper limit of normal; WDAE = withdrawal due to adverse event.

[a] Frequency > 5%.

[b] Frequency > 1%.

Note: Harms refer to outcomes identified as important to the review (see Table 2 for review protocol).

TABLE 24: HARMS FOR STUDY 1045 FROM BASELINE TO THREE MONTHS

	Tofacitinib 5 mg b.i.d. N = 243	Tofacitinib 10 mg b.i.d. N = 245	Placebo N = 122
AEs			
Patients with > 0 AEs, N (%)	124 (51)	139 (57)	67 (55)
Most common AEs[a]			
Infections and Infestations, N (%)	40 (17)	47 (19)	23 (19)
Upper respiratory tract infection	11 (5)	8 (3)	6 (5)
Gastrointestinal disorders, N (%)	48 (20)	47 (19)	26 (21)
Diarrhea			
Musculoskeletal and connective tissue disorders, N (%)	21 (9)	19 (8)	13 (11)
Nervous System Disorders, N (%)	30 (12)	20 (8)	14 (12)
Headache	13 (5)	11 (5)	3 (3)
General disorders and administration site conditions, N (%)	14 (6)	9 (4)	8 (7)
Injury, poisoning, and procedural complications, N (%)	11 (5)	10 (4)	1 (1)
Investigations, N (%)	12 (5)	26 (11)	6 (5)
Skin and subcutaneous tissue disorders	11 (5)	9 (4)	4 (3)
SAEs			
Patients with > 0 SAEs, N (%)	1 (< 1)	5 (2)	6 (5)
Most common SAEs[b]			
Infections and infestations	0	1 (< 1)	0
WDAEs			
WDAEs, N (%)	2 (1)	6 (2)[c]	5 (4)
Deaths			
Number of deaths, N (%)	0	0[d]	0
Other notable harms at month 3			
Malignancies, N (%)	0	0	0
Gastrointestinal perforations, N (%)	NR	NR	NR
Congestive heart failure, N (%)	8 (3)	5 (2)	3 (3)
Ischemic heart disease, N (%)	3 (1)	10 (4)	1 (1)

	Tofacitinib 5 mg b.i.d. N = 243	Tofacitinib 10 mg b.i.d. N = 245	Placebo N = 122
Myocardial infarction, N (%)	3 (1)	10 (4)	1 (1)
Cerebrovascular accident, N (%)	NR	NR	NR
Dyslipidemia, N (%)	8 (3)	10 (4)	1 (1)
Neutrophil counts, $10^{-3}/mm^3$ LSM change (SE) from baseline	–0.8 (0.1)[e]	–1.4 (0.1)[e]	–0.1 (0.2)
Hemoglobin, g/dL mean change (SD) from baseline	0.3 (0.9)	0.03 (1.0)	−0.1 (0.8)
LDL, mean change (SE) from baseline, %	13.6 (1.6)[e]	19.1 (1.6)[e]	3.5 (2.3)
HDL, mean change (SE) from baseline, %	12.2 (1.3)[e]	15.0 (1.3)[e]	−0.8 (1.9)
Serum creatinine, mg/dL, mean change (SE) from baseline	0.04 (0.01)	0.05 (0.01)	0 (0.02)
Mild neutropenia (1,500 to 1,999 cells/mm^3), %	3.1	3.7	< 1.0
Moderate neutropenia (1,000 to 1,499 cells/mm^3), %	1.3	0	
Severe neutropenia (500 to 999 cells/mm^3), %	0	< 1.0	0
Decreased hemoglobin (−1.0 to −3.0 g/dL), %	5.8	14.4	14.6
AST > 3 × ULN, %[f]	< 1.0	0	< 1.0
ALT > 3 × ULN, %[f]	< 1.0	0	< 1.0

AE = adverse event; ALT = alanine transaminase; AST = aspartate transaminase; b.i.d. = twice daily; HDL = high-density lipoprotein; LDL = low-density lipoprotein; NR = not reported; SAE = serious adverse event; SD = standard deviation; SE = standard error; ULN = upper limit of normal; WDAE = withdrawal due to adverse event.
[a] Frequency > 5%
[b] Frequency > 1%
[c] One patient who died (on Study Day 107) was included here in the Clinical Study Report for Study 1045.
[d] See footnote c.
[e] *P* value is < 0.001 for comparison with placebo.
[f] Incidence is shown for patients who had normal values at baseline.
Note: Outcomes identified as important to the review (see Table 2 for review protocol).

TABLE 25: HARMS FOR STUDY 1044 FROM BASELINE TO THREE MONTHS

	Tofacitinib 5 mg b.i.d. N = 321	Tofacitinib 10 mg b.i.d. N = 316	Placebo N = 160
AEs			
Patients with > 0 AEs, N (%)	157 (49)	171 (54)	73 (46)
Most common AEs[a]			
Infections and infestations, N (%)	65 (20)	70 (22)	19 (12)
Gastrointestinal disorders, N (%)	49 (15)	42 (13)	15 (9)
Musculoskeletal and connective tissue disorders, N (%)	17 (5)	13 (4)	14 (9)
Nervous system disorders, N (%)	26 (8)	17 (5)	9 (6)
Headache	18 (6)	4 (1)	3 (2)
General disorders and administration site conditions, N (%)	16 (5)	11 (4)	6 (4)
Investigations, N (%)	19 (6)	26 (8)	11 (7)
Skin and subcutaneous tissue disorders	11 (3)	20 (6)	8 (5)
Respiratory, thoracic, and mediastinal disorders	8 (3)	20 (6)	4 (3)
SAEs			
Patients with > 0 SAEs, N (%)	12 (4)	10 (3)	5 (3)

	Tofacitinib 5 mg b.i.d. N = 321	Tofacitinib 10 mg b.i.d. N = 316	Placebo N = 160
Most common SAEs[b]			
Infections and infestations	2 (1)	2 (1)	0
WDAEs			
WDAEs, N (%)	15 (5)	14 (4)	5 (3)
Most common reasons			
Headache	4 (1)	0	0
Rheumatoid arthritis	0	0	3 (2)
Arthralgia	1 (< 1)	2 (1)	0
Deaths			
Number of deaths, N (%)	2 (1)[c]	0	0
Other notable harms at month 3			
Malignancies, N (%)	2 (1)	2 (1)	0
Gastrointestinal perforations, N (%)	NR	NR	NR
Congestive heart failure, N (%)	4 (1)	5 (2)	3 (2)
Ischemic heart disease, N (%)	1 (< 1)	3 (1)	0
Myocardial infarction, N (%)	0	3 (1)	0
Cerebrovascular accident, N (%)	0	0	0
Dyslipidemia, N (%)	9 (3)	16 (5)	3 (2)
Neutrophil counts, $10^3/mm^3$ change (SD) from baseline	–1.1 (2.1) n = 291	–1.2 (2.1) n = 293	–0.2 (1.9) n = 144
Hemoglobin, g/dL mean change (SD) from baseline	0.2 (0.8) n = 291	0.1 (0.9) n = 294	0.0 (0.8) n = 144
LDL, LSM change (SE) from baseline, %	16.2 (1.4)[d] n = 289	19.3 (1.4)[d] n = 290	2.0 (1.9) n = 142
HDL, LSM change (SE) from baseline, %	9.8 (1.3)[d] n = 291	15.1 (1.3)[d] n = 295	–2.0 (1.8) n = 143
Serum creatinine, mg/dL, LSM change (SE) from baseline	0.05 (0.01) n = 295	0.06 (0.01) n = 299	0.01 (0.01) n = 145
Mild neutropenia (1.5 to < $2 \times 10^3/\mu L$), N (%)	5 (2) n = 293	11 (4) n = 296	2 (1) n = 144
Moderate to severe neutropenia (≥ 0.5 to < $1.5 \times 10^3/\mu L$), N (%)	3 (1) n = 293	2 (< 1) n = 296	0
Potentially life-threatening neutropenia (< $0.5 \times 10^3/\mu L$), N (%)	0	0	0
Decreased hemoglobin, mild to moderate (decrease of ≥ 1 g/dL to ≤ 2 g/dL), N (%)	19 (7) n = 293	35 (12) n = 297	15 (10) n = 144
Decreased hemoglobin, severe (decrease of > 2 g/dL to < 3 g/dL or hemoglobin > 7 g/dL and < 8 g/dL), N (%)	1 (< 1) n = 293	1 (< 1) n = 297	2 (1) n = 144
Decreased hemoglobin, potentially life-threatening (decrease of ≥ 3 g/dL or hemoglobin ≤ 7 g/dL), N (%)	0 n = 293	1 (< 1) n = 297	0 n = 144

	Tofacitinib 5 mg b.i.d. N = 321	Tofacitinib 10 mg b.i.d. N = 316	Placebo N = 160
AST > 3 × ULN, N (%)	1 (< 1) n = 318	2 (< 1) n = 315	2 (1) n = 159
ALT > 3 × ULN, N (%)	3 (< 1) n = 318	6 (2) n = 315	3 (2) n = 159

AE = adverse event; ALT = alanine transaminase; AST = aspartate transaminase; b.i.d. = twice daily; CI = confidence interval; HDL = high-density lipoprotein; LDL = low-density lipoprotein; LSM = least squares mean; NR = not reported; SAE = serious adverse events; SD = standard deviation; SE = standard error; ULN = upper limit of normal; WDAE = withdrawal due to adverse event.

[a] Frequency > 5%

[b] Frequency > 1%

[c] Deaths occurred following discontinuation of study drug.

[d] *P* value is < 0.0001 for comparison with placebo.

Note: Outcomes identified as important to the review (see Table 2 for review protocol).

TABLE 26: HARMS FOR STUDY 1044 FROM THREE TO SIX MONTHS

	Tofacitinib 5 mg b.i.d. N = 321	Tofacitinib 10 mg b.i.d. N = 316	Placebo → Tofacitinib 5 mg b.i.d. n = 42	Placebo → Tofacitinib 10 mg b.i.d. n = 37	Placebo N = 81
AEs					
Patients with > 0 AEs, N (%)	145 (45)	111 (35)	18 (43)	15 (41)	21 (26)
Most common AEs[a]					
Infections and infestations, N (%)	77 (24)	48 (15)	5 (12)	7 (19)	10 (12)
Upper respiratory tract infection, N (%)	15 (5)	6 (2)	2 (5)	3 (8)	0
Nasopharyngitis, N (%)	12 (4)	7 (2)	1 (2)	2 (5)	4 (5)
Gastrointestinal DISORDERS, n (%)	25 (8)	17 (5)	2 (5)	5 (14)	3 (4)
Stomatitis, N (%)	0	1 (< 1)	0	3 (8)	1 (1)
Musculoskeletal and connective tissue disorders, N (%)	17 (5)	11 (4)	2 (5)	4 (11)	2 (3)
Nervous system disorders, n (%)	5 (2)	7 (2)	1 (2)	2 (5)	2 (3)
Injury, poisoning, and procedural complications, N (%)	10 (3)	8 (3)	3 (7)	1 (3)	1 (1)
Skin laceration, N (%)	2 (1)	0	2 (5)	0	0
Investigations, N (%)	18 (6)	15 (5)	3 (7)	0	2 (3)
Metabolism and nutrition disorders, n (%)	4 (1)	8 (3)	2 (5)	0	2 (3)
Skin and subcutaneous tissue disorders	17 (5)	8 (3)	0	1 (3)	1 (1)
SAEs					
Patients with > 0 SAEs, N (%)	17 (5)	7 (2)	1 (2)	1 (3)	5 (6)
Most common SAEs[b]					
Infections and infestations	8 (3)	2 (1)	1 (1)		1 (1)
WDAEs					
WDAEs, N (%)	16 (5)	8 (3)	2 (5)	2 (5)	3 (4)
Most common reasons					
Infections and infestations	9 (3)	2 (1)	1 (2)	0	2 (3)
Deaths					
Number of deaths, N (%)	0	1 (< 1)[c]	0	0	NR

	Tofacitinib 5 mg b.i.d. N = 321	Tofacitinib 10 mg b.i.d. N = 316	Placebo → Tofacitinib 5 mg b.i.d. n = 42	Placebo → Tofacitinib 10 mg b.i.d. n = 37	Placebo N = 81
Other notable harms at month 6					
Malignancies, N (%)	2 (1)	1 (< 1)	0	0	2 (2)
Gastrointestinal perforations, N (%)	NR	NR	NR	NR	NR
Patients with congestive heart failure AE, N (%)	0	5 (2)	0	0	0
Patients with ischemic heart disease AE, N (%)	2 (1)	2 (1)	2 (5)	0	0
Patients with myocardial infarction AE, N (%)	1 (< 1)	2 (1)	1 (2)	0	0
Cerebrovascular accident, N (%)	NR	NR	NR	NR	NR
Dyslipidemia, N (%)	3 (1)	4 (1)	2 (5)	0	1 (1)
Neutrophil counts, $10^3/mm^3$ change (SD) from baseline	–0.9 (2.2) n = 278	–1.2 (2.1) n = 280	–0.8 (1.9) n = 40	–1.2 (1.7) n = 37	–0.2 (1.9) n = 62
Hemoglobin, g/dL mean change (SD) from baseline	0.2 (0.9) n = 278	0.2 (1.0) n = 280	–0.1 (0.7) n = 40	0.0 (1.0) n = 37	0.1 (0.9) n = 62
LDL, LSM change (SE) from baseline, %	17.7 (1.6)[d] n = 195	21.4 (1.5)[d] n = 224	NR	NR	–1.6 (2.6) n = 60
HDL, LSM change (SE) from baseline, %	9.6 (1.4)[b] n = 199	16.0 (1.4)[d] n = 230	NR	NR	–2.7 (2.4) n = 61
Serum creatinine, mg/dL, LSM change (SE) from baseline	0.05 (0.01) n = 202 *P* = 0.0039[e]	0.07 (0.01)[d] n = 232	NR	NR	0.02 (0.01) n = 62
Mild neutropenia (1.5 to < 2 × $10^3/\mu L$), N (%)	6 (2) n = 281	11 (4) n = 283	2 (5) n = 40	2 (5) n = 37	0 n = 62
Moderate to severe neutropenia (≥ 0.5 to < 1.5 × $10^3/\mu L$), N (%)	3 (1) n = 281	4 (1) n = 283	0	1 (3) n = 37	0
Potentially life-threatening neutropenia (< 0.5×$10^3/\mu L$), n (%)	0	0	0	0	0
Decreased hemoglobin, mild to moderate (decrease of ≥ 1 g/dL to ≤ 2 g/dL), N (%)	28 (10) n = 281	27 (10) n = 283	5 (12) n = 40	4 (11) n = 37	4 (7) n = 62
Decreased hemoglobin, severe (decrease of > 2 g/dL to < 3 g/dL or hemoglobin > 7 g/dL and < 8 g/dL), N (%)	1 (< 1)	3 (1)	0	0	1 (2)
Decreased hemoglobin, potentially life-threatening (decrease of ≥ 3 g/dL or hemoglobin ≤ 7 g/dL), N (%)	1 (< 1)	0	0	0	0
AST > 3 × ULN, N (%)	2 (< 1)	2 (< 1)	1 (2)	0	0
ALT > 3 × ULN, N (%)	2 (< 1)	4 (1)	2 (5)	1 (3)	0

AE = adverse event; ALT = alanine transaminase; AST = aspartate transaminase; b.i.d. = twice daily; CI = confidence interval; HDL = high-density lipoprotein; LDL = low-density lipoprotein; LSM = least squares mean; NR = not reported; SAE = serious adverse events; SD = standard deviation; SE = standard error; ULN = upper limit of normal; WDAE = withdrawal due to adverse event.

[a] Frequency > 5%

[b] Frequency > 1%

[c] Death occurred following discontinuation of study drug.

[d] *P* value is < 0.0001 for comparison with placebo.

[e] *P* value is for comparison with placebo.

Note: Outcomes identified as important to the review (see Table 2 for review protocol).

TABLE 27: HARMS FOR STUDY 1046 FROM BASELINE TO THREE MONTHS

	Tofacitinib 5 mg b.i.d. N = 315	Tofacitinib 10 mg b.i.d. N = 318	Placebo N = 159
AEs			
Patients with > 0 AEs, N (%)	166 (53)	173 (54)	97 (61)
Most common AEs[a]			
Infections and infestations, N (%)	77 (24)	76 (24)	47 (30)
Nasopharyngitis	16 (5)	7 (2)	12 (8)
Upper respiratory tract infection	19 (6)	23 (7)	7 (4)
Gastrointestinal disorders, N (%)	59 (19)	54 (17)	23 (15)
Blood and lymphatic system disorders, N (%)	15 (5)	7 (2)	7 (4)
Musculoskeletal and connective tissue disorders, N (%)	24 (8)	20 (6)	13 (8)
Nervous system disorders, N (%)	18 (6)	16 (5)	10 (6)
General disorders and administration site conditions, N (%)	8 (3)	15 (5)	6 (4)
Investigations, N (%)	23 (7)	31 (10)	11 (7)
Skin and subcutaneous tissue disorders	16 (5)	14 (4)	7 (4)
Respiratory, thoracic, and mediastinal disorders	17 (5)	12 (4)	6 (4)
SAEs			
Patients with > 0 SAEs, N (%)	9 (3)	8 (3)	6 (4)
Most common SAEs[b]			
Infections and infestations, %	1	1	0
WDAEs			
WDAEs, N (%)	13 (4)	13 (4)	2 (1)
Deaths			
Number of deaths, N (%)	0	0	0
Notable harms at month 3			
Malignancies, N (%)	0	2 (1)	0
Gastrointestinal perforations, N (%)	NR	NR	NR
Patients with congestive heart failure AE, N (%)	0	8 (3)	3 (2)
Patients with ischemic heart disease AE, N (%)	6 (2)	6 (2)	1 (1)
Patients with myocardial infarction AE, N (%)	5 (2)	6 (2)	1 (1)
Cerebrovascular accident, N (%)	1 (< 1)	0	0
Dyslipidemia, N (%)	9 (3)	12 (4)	1 (1)
Neutrophil counts, 10^9 cells/L, LSM change from baseline (95% CI)[c]	–0.8 (–1.0 to –0.6)	–1.0 (–1.2 to –0.8)	–0.02 (–0.3 to 0.3)
Hemoglobin level, g/L, LSM change from baseline (95% CI)[c]	0.3 (0.2 to 0.4)	0.1 (–0.01 to 0.2)	–0.04 (–0.2 to 0.1)
LDL, LSM change from baseline (95% CI)[c], %	15.7 (13.0 to 18.3)	18.3 (15.6 to 20.9)	–0.6 (–4.2 to 3.0)
HDL, LSM change from baseline (95% CI)[c], %	12.2 (9.9 to 14.5)	14.7 (12.4 to 17.1)	–0.7 (–4.0 to 2.5)
Serum creatinine, mg/dL, LSM change from baseline (95% CI)[c]	0.03 (0.02 to 0.04)	0.06 (0.05 to 0.07)	0.02 (0.00 to 0.03)
Mild neutropenia (1.5 to < 2 × 10^3/µL), N (%)	4 (1) n = 291	10 (4) n = 288	3 (2) n = 147
Moderate to severe neutropenia (≥ 0.5 to < 1.5 × 10^3/µL), N (%)	1 (< 1) n = 291	8 (3) n = 288	1 (1) n = 147
Potentially life-threatening neutropenia (< 0.5 × 10^3/µL), N (%)	0	0	0

	Tofacitinib 5 mg b.i.d. N = 315	Tofacitinib 10 mg b.i.d. N = 318	Placebo N = 159
Decreased hemoglobin, mild to moderate (decrease of ≥ 1 g/dL to ≤ 2 g/dL), N (%)	16 (6) n = 291	23 (8) n = 288	12 (8) n = 147
Decreased hemoglobin, severe (decrease of > 2 g/dL to < 3 g/dL or hemoglobin > 7 g/dL and < 8 g/dL), N (%)	1 (< 1) n = 291	1 (< 1) n = 288	3 (2) n = 147
Decreased hemoglobin, potentially life-threatening (decrease of ≥ 3 g/dL or hemoglobin ≤ 7 g/dL), N (%)	0	0	0
AST > 3 × ULN, N (%)	3 (< 1)	1 (< 1)	1 (< 1)
ALT > 3 × ULN, N (%)	6 (2)	3 (< 1)	1 (< 1)

AE = adverse event; ALT = alanine transaminase; AST = aspartate transaminase; b.i.d. = twice daily; CI = confidence interval; HDL = high-density lipoprotein; LDL = low-density lipoprotein; LSM = least squares mean; NR = not reported; SAE = serious adverse events; SD = standard deviation; SE = standard error; ULN = upper limit of normal; WDAE = withdrawal due to adverse event.

[a] Frequency > 5%

[b] Frequency > 1%

[c] The CIs from baseline to three months represent effect of treatment versus placebo.

Note: Outcomes identified as important to the review (see Table 2 for review protocol).

TABLE 28: HARMS FOR STUDY 1046 FROM THREE TO SIX MONTHS

	Tofacitinib 5 mg b.i.d. N = 315	Tofacitinib 10 mg b.i.d. N = 318	Placebo → Tofacitinib 5 mg b.i.d. n = 38	Placebo → Tofacitinib 10 mg b.i.d. n = 40	Placebo N = 81
AEs					
Patients with > 0 AEs, N (%)	121 (38)	124 (39)	16 (42)	18 (45)	21 (26)
Most common AEs[a]					
Infections and infestations, N (%)	53 (17)	52 (16)	3 (8)	9 (23)	6 (7)
Upper respiratory tract infection, N (%)	13 (4)	9 (3)	1 (3)	3 (8)	0
Urinary tract infection, N (%)	5 (2)	5 (2)	1 (3)	2 (5)	1 (1)
Gastrointestinal disorders, N (%)	25 (8)	12 (4)	4 (11)	5 (13)	2 (3)
General disorders and administration site conditions, N (%)	8 (3)	14 (4)	6 (16)	1 (3)	3 (4)
Peripheral edema, N (%)	4 (1)	8 (3)	3 (8)	0	1 (1)
Pyrexia, N (%)	1 (< 1)	1 (< 1)	3 (8)	1 (3)	0
Musculoskeletal and connective tissue disorders, N (%)	16 (5)	19 (6)	2 (5)	3 (8)	5 (6)
Nervous system disorders, N (%)	12 (4)	8 (3)	2 (5)	0	3 (4)
Injury, poisoning, and procedural complications, N (%)	6 (2)	16 (5)	2 (5)	2 (5)	1 (1)
Investigations, N (%)	12 (4)	21 (7)	2 (5)	4 (10)	3 (4)
Psychiatric disorders, N (%)	2 (1)	1 (< 1)	0	2 (5)	0
Respiratory, thoracic, and mediastinal disorders, N (%)	8 (3)	6 (2)	5 (13)	3 (8)	0
Cough, N (%)	1 (< 1)	2 (1)	2 (5)	2 (5)	0
Skin and subcutaneous tissue disorders, N (%)	15 (5)	6 (2)	0	4 (10)	2 (3)
Rash, N (%)	3 (1)	4 (1)	1 (3)	0	1 (1)

	Tofacitinib 5 mg b.i.d. N = 315	Tofacitinib 10 mg b.i.d. N = 318	Placebo → Tofacitinib 5 mg b.i.d. n = 38	Placebo → Tofacitinib 10 mg b.i.d. n = 40	Placebo N = 81
SAEs					
Patients with > 0 SAEs, N (%)	5 (2)	7 (2)	0	0	0
Most common SAEs[b]					
Infections and infestations, %	0	0	< 1	< 1	0
WDAEs					
WDAEs, N (%)	6 (2)	8 (3)	0	1 (3)	1 (1)
Deaths					
Number of deaths, N (%)	0	1	0	0	0
Other notable harms at month 6					
Malignancies, N (%)	0	0	0	0	0
Gastrointestinal perforations, N (%)	NR	NR	NR	NR	NR
Patients with congestive heart failure AE, N (%)	5 (2)	10 (3)	3 (8)	0	1 (1)
Patients with ischemic heart disease AE, N (%)	1 (1)	5 (2)	1 (3)	1 (3)	0
Patients with myocardial infarction AE, N (%)	3 (1)	5 (2)	1 (3)	1 (3)	0
Cerebrovascular accident, N (%)	NR	NR	NR	NR	NR
Dyslipidemia, N (%)	6 (2)	5 (2)	0	1 (3)	0
Neutrophil counts, 10^9 cells/L, LSM change from baseline (95% CI)[c]	–1.0 (–1.2 to –0.8)	–1.1 (–1.3 to –0.9)	–0.5 (–0.9 to –0.2)	–1.0 (–1.3 to –0.6)	–0.4 (–0.7 to –0.03)
Hemoglobin level, g/L, LSM change from baseline (95% CI)[c]	0.3 (0.2 to 0.4)	0.1 (–0.02 to 0.2)	0.01 (–0.2 to 0.2)	–0.1 (–0.3 to 0.1)	–0.2 (–0.3 to 0.01)
LDL, LSM change from baseline (95% CI)[c], %	15.2 (12.4 to 18.0)	17.9 (15.0 to 20.7)	4.5 (–0.9 to 9.9)	5.7 (0.2 to 11.2)	–2.6 (–7.5 to 2.3)
HDL, LSM change from baseline (95% CI)[c], %	12.3 (9.8 to 14.8)	14.0 (11.5 to 16.5)	3.9 (–0.9 to 8.7)	10.8 (6.0 to 15.7)	–1.7 (–5.9 to 2.5)
Serum creatinine, mg/dL, LSM change from baseline (95% CI)[c]	0.05 (0.04 to 0.06)	0.06 (0.04 to 0.07)	0.03 (0.01 to 0.05)	0.06 (0.03 to 0.08)	0.02 (0.00 to 0.04)
Mild neutropenia (1.5 to < 2 × 10^3/µL), N (%)	7 (3) n = 276	12 (4) n = 280	2 (3) n = 71	5 (7) n = 72	NR
Moderate to severe neutropenia (≥ 0.5 to < 1.5 × 10^3/µL), N (%)	1 (< 1) n = 276	8 (3) n = 280	1 (1) n = 71	1 (1) n = 72	NR
Potentially life-threatening neutropenia (< 0.5×10^3/µL), N (%)	0	0	0	0	NR
Decreased hemoglobin, mild to moderate (decrease of ≥ 1 g/dL to ≤ 2 g/dL), N (%)	24 (9) n = 276	27 (10) n = 280	7 (10) n = 71	9 (13) n = 72	NR
Decreased hemoglobin, severe (decrease of > 2 g/dL to < 3 g/dL or hemoglobin > 7 g/dL and < 8 g/dL), N (%)	1 (< 1) n = 276	2 (< 1) n = 280	0 n = 71	2 (3) n = 72	NR

	Tofacitinib 5 mg b.i.d. N = 315	Tofacitinib 10 mg b.i.d. N = 318	Placebo → Tofacitinib 5 mg b.i.d. n = 38	Placebo → Tofacitinib 10 mg b.i.d. n = 40	Placebo N = 81
Decreased hemoglobin, potentially life-threatening (decrease of ≥ 3 g/dL or hemoglobin ≤ 7 g/dL), N (%)	0	0	0	0	NR
AST > 3 × ULN, N (%)	1 (< 1)	0	0	0	0
ALT > 3 × ULN, N (%)	3 (1)	3 (1)	0	1 (3)	1 (1)

AE = adverse event; ALT = alanine transaminase; AST = aspartate transaminase; b.i.d. = twice daily; CI = confidence interval; HDL = high-density lipoprotein; LDL = low-density lipoprotein; LSM = least squares mean; NR = not reported; SAE = serious adverse events; SD = standard deviation; SE = standard error; ULN = upper limit of normal; WDAE = withdrawal due to adverse event.
[a] Frequency > 5%
[b] Frequency > 1%
[c] The CIs from baseline to three months represent effect of treatment versus placebo.
Note: Outcomes identified as important to the review (see Table 2 for review protocol).

TABLE 29: HARMS FOR STUDY 1064 FROM BASELINE TO THREE MONTHS

	Tofacitinib 5 mg b.i.d. N = 204	Tofacitinib 10 mg b.i.d. N = 201	Adalimumab 40 mg SC q.2w. N = 204	Placebo N = 108
AEs				
Patients with > 0 AEs, N (%)	106 (52)	94 (47)	105 (52)	51 (47)
Most common AEs[a]				
Infections and infestations, N (%)	37 (18)	35 (17)	33 (16)	10 (9)
Gastrointestinal disorders, N (%)	24 (12)	21 (10)	21 (10)	12 (11)
Musculoskeletal and connective tissue disorders, N (%)	15 (7)	8 (4)	12 (6)	7 (7)
Nervous system disorders, N (%)	11 (5)	12 (6)	11 (5)	4 (4)
Investigations, N (%)	13 (6)	10 (5)	11 (5)	3 (3)
Skin and subcutaneous tissue disorders	6 (3)	12 (6)	11 (5)	8 (7)
Respiratory, thoracic, and mediastinal disorders, N (%)	10 (5)	5 (2)	12 (6)	5 (5)
Vascular disorders, N (%)	2 (1)	11 (6)	2 (1)	3 (3)
SAEs				
Patients with > 0 SAEs, N (%)	12 (6)	10 (5)	5 (3)	2 (2)
Most common SAEs[b]				
Infections and infestations, N %	3 (2)	4 (2)	0	1 (1)
WDAEs				
WDAEs, N (%)	14 (7)	10 (5)	10 (5)	3 (3)
Deaths				
Number of deaths, N (%)	1	0	1	0
Notable harms at month 3				
Malignancies, N (%)	1 (1)	0	1 (1)	1 (1)
Gastrointestinal perforations, N (%)	NR	1 (1)[c]	NR	NR
Congestive heart failure, N (%)	3 (2)	5 (3)	3 (2)	3 (3)
Ischemic heart disease, N (%)	1 (1)	5 (3)	2 (1)	1 (1)
Myocardial infarction, N (%)	1 (1)	5 (3)	2 (1)	1 (1)
Cerebrovascular accident, N (%)	NR	NR	NR	NR
Dyslipidemia, N (%)	0	0	1 (1)	0
Neutrophil counts, $10^{-3}/mm^3$, mean change (SE) from baseline	−0.7 (0.1)	−0.8 (0.1)	−1.3 (0.1)	−0.2 (0.2)
Hemoglobin level, g/dL, mean change (SD) from baseline	0.1 (0.9)	−0.1 (2.4)	0.4 (0.8)	0.04 (0.8)
LDL, mean change (SE) from baseline, %	12.2 (2.0)	18.9 (2.0)	3.6 (1.9)	0.3 (2.6)
HDL, mean change (SE) from baseline, %	12.2 (1.5)	11.0 (1.5)	5.6 (1.5)	−1.6 (2.0)
Serum creatinine, mg/dL, mean change (SE) from baseline	0.04 (0.01)	0.1 (0.01)	0.02 (0.01)	0 (0.01)
	N = 186	**N = 183**	**N = 187**	**N = 98**
Mild neutropenia (1,500 to 1,999 cells/mm^3), N (%)	3 (2)	3 (2)	5 (3)	2 (2)
Moderate to severe neutropenia (500 to 1,499	2 (1)	3 (2)	0	0

	Tofacitinib 5 mg b.i.d. N = 204	Tofacitinib 10 mg b.i.d. N = 201	Adalimumab 40 mg SC q.2w. N = 204	Placebo N = 108
cells/mm^3), N (%)				
Potentially life-threatening neutropenia (< 500 cells/mm^3), N (%)	NR	NR	NR	NR
Decreased hemoglobin (–1.0 to –3.0 g/dL)[d], N (%)	15 (8)	15 (8)	10 (5)	9 (9)
	N = 203	**N = 201**	**N = 204**	**N = 105**
AST > 3 × ULN, no. of incidents/total no. (%)[e]	1/2 (50)	0	0	1/2 (50)
ALT > 3 × ULN, no. of incidents/total no. (%)[e]	2/4 (50)	1/4 (25)	0	1/4 (25)

AE = adverse event; ALT = alanine transaminase; AST = aspartate transaminase; b.i.d. = twice daily; HDL = high-density lipoprotein; LDL = low-density lipoprotein; NR = not reported; q.2w. = every two weeks; SAE = serious adverse events; SC = subcutaneous injection; SD = standard deviation; SE = standard error; ULN = upper limit of normal; WDAE = withdrawal due to adverse event.
[a] Frequency > 5%
[b] Frequency > 1%
[c] Serious adverse event leading to permanent study drug discontinuation.
[d] A severe decreased hemoglobin is also defined as an actual level of greater than 7 g/dL but less than 8 g/dL.
[e] Number of incidents in which values were higher than the ULN. More than one incident may be recorded for the same patient.
Note: Outcomes identified as important to the review (see Table 2 for review protocol).

TABLE 30: HARMS FOR STUDY 1064 FROM THREE TO SIX MONTHS

	Tofacitinib 5 mg b.i.d. N = 204	Tofacitinib 10 mg b.i.d. N = 201	Placebo → Tofacitinib 5 mg b.i.d. N = 28	Placebo → Tofacitinib 10 mg b.i.d. N = 21	Adalimumab 40 mg SC q.2w. N = 204	Placebo N = 59
AEs						
Patients with > 0 AEs, N (%)	67 (33)	62 (31)	7 (25)	9 (43)	68 (33)	16 (27)
Most common AEs[a]						
Infections and infestations, N (%)	27 (13)	24 (12)	2 (7)	2 (10)	28 (14)	4 (7)
Influenza, N (%)	0	2 (1)	0	1 (5)	1 (1)	0
Upper respiratory tract infection, N (%)	3 (2)	3 (2)	1 (4)	1 (5)	2 (1)	0
Urinary tract infection, N (%)	4 (2)	0	1 (4)	1 (5)	6 (3)	0
Gastrointestinal disorders, N (%)	6 (3)	8 (4)	0	1 (5)	11 (5)	1 (2)
Nausea, N (%)	1 (1)	3 (2)	0	1 (5)	1 (1)	0
Musculoskeletal and connective tissue disorders, N (%)	11 (5)	7 (4)	1 (4)	2 (10)	13 (6)	1 (2)
Back pain, N (%)	3 (2)	1 (1)	0	1 (5)	5 (3)	0
Tendonitis, N (%)	0	0	0	1 (5)	0	0
Nervous system disorders, N (%)	4 (2)	3 (2)	1 (4)	0	5 (3)	3 (5)
Investigations, N (%)	6 (3)	8 (4)	4 (14)	0	5 (3)	0
Weight increased, N (%)	2 (1)	0	2 (7)	0	0	0

	Tofacitinib 5 mg b.i.d. N = 204	Tofacitinib 10 mg b.i.d. N = 201	Placebo → Tofacitinib 5 mg b.i.d. N = 28	Placebo → Tofacitinib 10 mg b.i.d. N = 21	Adalimumab 40 mg SC q.2w. N = 204	Placebo N = 59
Respiratory, thoracic, and mediastinal disorders, N (%)	4 (2)	5 (3)	3 (11)	1 (5)	0	2 (3)
Cough, n (%)	2 (1)	1 (1)	0	1 (5)	0	1 (2)
SAEs						
Patients with > 0 SAEs, N (%)	10 (5)	7 (4)	0	0	6 (3)	2 (3)
Most common SAEs[b]						
Infections and infestations, N (%)	2 (1)	1 (1)	0	0	2 (1)	0
WDAEs						
WDAEs, N (%)	5 (2.5)	11 (5.5)	1 (3.6)	0	9 (4)	0
Notable harms at month 6						
Malignancies, N (%)	0	1 (1)	0	0	0	0
Gastrointestinal perforations, N (%)	NR	NR	NR	NR	NR	NR
Patients with congestive heart failure AE, N (%)	3 (2)	0	1 (4)	0	1 (1)	0
Patients with ischemic heart disease AE, N (%)	1 (1)	0	0	0	1 (1)	0
Patients with myocardial infarction AE, N (%)	1 (1)	0	0	0	0	0
Cerebrovascular accident, N (%)	NR	NR	NR	NR	NR	NR
Dyslipidemia, N (%)	NR	NR	NR	NR	NR	NR
Neutrophil counts, $10^{-3}/mm^3$, mean change (SE) from baseline	−0.8 (0.1)	−1.0 (0.1)	−0.5 (0.2)	−0.9 (0.3)	−1.4 (0.1)	NA
Hemoglobin level, g/dL, mean change (SD) from baseline	0.2 (0.9)	−0.04 (2.4)	0.3 (0.9)	−0.2 (1.0)	0.3 (1.0)	NA
LDL, mean change (SE) from baseline, %	14.1 (2.0)	21.2 (2.0)	10.1 (3.6)	−1.9 (4.0)	2.4 (2.0)	NA
HDL, mean change (SE) from baseline, %	11.0 (1.5)	13.3 (1.5)	5.9 (2.8)	−2.2 (3.0)	2.8 (1.5)	NA
Serum creatinine, mg/dL, mean change (SE) from baseline	0.05 (0.01)	0.06 (0.01)	0.04 (0.02)	0.03 (0.02)	0.02 (0.01)	NA
	N = 170	**N = 181**	**N = 52**	**N = 42**	**N = 178**	**N = 0**
Mild neutropenia (1,500 to 1,999 cells/mm^3), N (%)	3 (2)	7 (4)	1 (2)	0	9 (5)	NA
Moderate to severe neutropenia (500 to 1,499 cells/mm3), N (%)	0	2 (1)	0	0	1 (< 1)	NA
Potentially life-threatening neutropenia (< 500 cells/mm^3), N (%)	NR	NR	NR	NR	NR	NA

	Tofacitinib 5 mg b.i.d. N = 204	Tofacitinib 10 mg b.i.d. N = 201	Placebo → Tofacitinib 5 mg b.i.d. N = 28	Placebo → Tofacitinib 10 mg b.i.d. N = 21	Adalimumab 40 mg SC q.2w. N = 204	Placebo N = 59
Decreased hemoglobin (–1.0 to –3.0 g/dL)[c], N (%)	17 (10)	21 (12)	3 (6)	7 (17)	12 (7)	NA
	N = 186	**N = 183**	**N = 28**	**N = 20**	**N = 178**	**N = 46**
AST > 3 × ULN, no. of incidents/total no. (%)[d]	2/5 (40)	2/5 (40)	0	0	1/5 (20)	0
ALT > 3 × ULN, no. of incidents/total no. (%)[d]	3/6 (50)	2/6 (33)	0	0	1/6 (17)	0

AE = adverse event; ALT = alanine transaminase; AST = aspartate transaminase; b.i.d. = twice daily; HDL = high-density lipoprotein; LDL = low-density lipoprotein; NR = not reported; q.2w. = every two weeks; SAE = serious adverse events; SC = subcutaneous injection; SD = standard deviation; SE = standard error; ULN = upper limit of normal; WDAE = withdrawal due to adverse event.

[a] Frequency > 5%.

[b] Frequency > 1%

[c] A severe decreased hemoglobin is also defined as an actual level of greater than 7 g/dL but less than 8 g/dL.

[d] Number of incidents in which values were higher than the ULN. More than one incident may be recorded for the same patient.

Note: Outcomes identified as important to the review (see Table 2 for review protocol).

APPENDIX 5: VALIDITY OF OUTCOME MEASURES

Aim

To summarize the validity of the following outcome measures:

- American College of Rheumatology (ACR) 20, ACR 50, and ACR 70
- Disease Activity Score (DAS)28
- Health Assessment Questionnaire–Disability Index (HAQ-DI)
- Modified total Sharp scale (mTSS)
- Functional Assessment of Chronic Illness Therapy–Fatigue scale (FACIT-Fatigue)
- Short-Form (36) Health Survey (SF-36).

Findings

ACR criteria, DAS28, HAQ-DI, mTSS, and FACIT-Fatigue are briefly summarized inTable 31.

TABLE 31: VALIDITY AND MINIMAL CLINICALLY IMPORTANT DIFFERENCE OF OUTCOME MEASURES

Instrument	Type	Evidence of Validity	MCID	References
ACR 20 ACR 50 ACR 70	ACR 20, ACR 50, and ACR 70 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 • Health Assessment Questionnaire • CRP or ESR	Yes	Unspecified	van Riel and van Gestel (2000)[28] Cohen et al. (2006)[29] Bansback et al. (2008)[30] ACR criteria (2007)[31] Chung et al. (2006)[32]
DAS28	DAS28 is an abbreviated version of the DAS, based on a 28-joint count that omits the feet and ankle joints.	Yes	Unspecified	Wells et al. (2009)[33] Crowson et al. (2009)[34]
HAQ-DI	The HAQ-DI is the disability assessment component of the HAQ.	Yes	0.22 points	Bruce and Fries (2003)[35,36]
mTSS	The mTSS is a composite measure of joint erosion and joint space narrowing based on radiograph assessment.	Yes	4.6 units	Bruynesteyn et al. 2002[37]
FACIT-Fatigue	The FACIT-Fatigue scale is a 13-item self-report measure of fatigue.	Yes	3 to 4 points	Cella et al. 2005[38]
SF-36	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 improvement in health status.	Yes	2.5 to 5.0	Gallagher et al. (2001)[39] Hays and Morales (2001)[40] Samsa et al. (1999)[41] Strand and Singh (2008)[42]

ACR = American College of Rheumatology; CRP = C-reactive protein; DAS = Disease Activity Score; ESR = erythrocyte sedimentation rate; FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy fatigue scale; HAQ-DI = Health Assessment Questionnaire–Disability Index; MCID = minimal clinically important difference; MCS = mental component score; mTSS = modified Total Sharp Score; PCS = physical component score; SF-36 = Short-Form (36) Health Survey.

American College of Rheumatology Response Criteria

The ACR criteria for assessing joint status were initially developed for patients with RA.[28] 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-DI
- 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 visual analogue scale (VAS) or Likert scale measurements. ACR 20, 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. This core set of measures included in the ACR response criteria was established through a consensus process of clinical experts. Individual criteria were selected based on their construct validity, face validity, content validity, criterion validity, and discriminant validity.[43] 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 as measured by the HAQ–DI was a strong predictor of mortality 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 such as ESR and CRP.[43] When considering the ability of an outcome measure to detect change, pain assessments, global assessments, tender joint counts, and HAQ-DI scores all had strong discriminant validity.

The ACR 20 is most commonly used as the primary end point in randomized controlled trials evaluating biologics in the treatment of RA. The US Food and Drug Administration (FDA) considers ACR 20 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.[44] ACR 50 and ACR 70 are often reported in clinical trials and are considered more stringent outcome measures.

Chung et al.[32] conducted a meta-analysis of 21 randomized controlled trials of RA therapies published between 1997 and 2004 to compare the discriminant capabilities of the ACR 50 and ACR 20 responses and to determine whether ACR 50 is as informative as ACR 20 in distinguishing between active therapies and control groups. While both measures can distinguish an active therapy from a control therapy, the levels of improvement captured by ACR 20 response do not generally represent an optimal clinical improvement. Furthermore, since the development of the ACR 20 response criteria, much more aggressive therapies in the treatment of RA have been introduced, and larger clinical responses can be expected. This meta-analysis concluded that ACR 20 and ACR 50 are similar in distinguishing between active and control therapies but that ACR 50 represents a more robust clinical response and may be a preferred end point in clinical trials.[32]

ACR 70 is considered even more rigorous than ACR 50. 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 ACR 70 response, maintained over six months, with active therapy compared with the control group.[44]

With 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.[30] Other criticisms of the ACR criteria are that most of its component measures are subjective, that dichotomous measures such as ACR lack sensitivity to change compared with continuous measures of response, and that the ACR 20 response threshold is too low relative to treatment goals applied in clinical practice.[31] 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.[31,45]

Disease Activity Score 28 (DAS28)

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).[46] 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.[47] The DAS components correlate well with each other and with the ACR criteria.[48-51] 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, ESR = erythrocyte sedimentation rate, ln(ESR) = natural logarithm of ESR value, sw28 = swollen joint count of 28 joints, t28 = tender joint count of 28 joints, 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 [GH]) for a panel of RA patients at times when RA is poorly controlled and well controlled.[52] A DAS28 score indicates an absolute level of disease activity, with a score of 5.1 or greater being considered high disease activity, while a DAS28 score lower than 3.2 indicates low disease activity state (LDAS), and a DAS28 score lower than 2.6 indicates remission.[33,53,54]

In recent years, CRP has been used to calculate the DAS28 in place of ESR. The trend of using CRP levels rather than ESR is mainly driven by greater availability, reduced cost, and increased sensitivity of CRP to short-term changes of in disease activity.[33,34] The formula used to calculate the DAS28(CRP) is as follows:

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

where CRP = C-reactive protein, DAS28 = Disease Activity Score 28, ln(CRP + 1) = natural logarithm of (CRP value +1), sw28 = swollen joint count of 28 joints, t28 = tender joint count of 28 joints, GH = general health measured by Patient's Global Assessment of Disease Activity on a VAS of 100 mm.

Overall, the DAS28(CRP) correlates well with DAS28(ESR), and both are validated measures for assessing disease activity in RA.[33,53,55-57] However, studies have shown that the DAS28(CRP) score is usually lower than the DAS28(ESR) score.[53,55-6354-58] The difference (DAS28[CRP] minus DAS28[ESR]) ranges from −0.2[55] to −0.8.[58] 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 for DAS28(ESR).[57] Furthermore, the European League Against Rheumatism (EULAR) recommended that the clinical implications of the DAS28 score (such as good response, moderate response, or no response) should be determined based on the baseline DAS28 scores (see Table 29).[64] Finally, no MCID for change in DAS28 scores exists.

TABLE 32: THE EULAR IMPROVEMENT RESPONSE CRITERIA (DAS28)

Baseline DAS28 Score	DAS28 Improvement Over Time Points		
	> 1.2	0.6 – 1.2	< 0.6
< 3.2	Good response	Moderate response	No response
3.2 – 5.1	Moderate response	Moderate response	No response
> 5.1	Moderate response	No response	No response

DAS28 = Disease Activity Score 28 items; EULAR = European League Against Rheumatism.
Source: Matsui et al. (2007).[57]

Health Assessment Questionnaire and Disability Index

The Health Assessment Questionnaire (HAQ) was originally developed in 1978 at Stanford University.[65] It was one of the first self-reported functional status (disability) measures and has become the dominant instrument in many disease areas, including arthritis.[66] The HAQ has been widely validated in patients with RA.[66] 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.[54]

The HAQ–Disability Index (HAQ-DI) is the disability assessment component of the HAQ. It assesses a patient's level of functional ability. There are 20 questions in eight categories to assess a patient's physical functional status: dressing, arising, eating, walking, hygiene, reach, grip, and common activities.[35,36] 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 randomized controlled trials have demonstrated that the HAQ-DI possesses face validity, content validity, construct validity, predictive validity, and discriminant validity. There is evidence suggesting that baseline HAQ scores are predictive of radiographic damage, work disability, and quality of life.[29,67] A number of investigators have suggested that the MCID is 0.22; however, differences as small as 0.10 have been suggested as clinically important.[35]

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 Score. This method allows for the assessment of two different 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 radiographs of specific joints (typically in the hands and feet) before treatment and at various points after treatment has been initiated. The Sharp scoring system was most recently modified by van der Heijde,[68] who scores erosions as listed in Table 33.

TABLE 33: MODIFIED TOTAL SHARP SCORE

Sharp/van der Heijde[69]	
Erosions	
0	Normal
1	Discrete erosions
2–3	Larger erosions according to surface area involved
4	Erosion extending over the middle of the bone
5	Complete collapse
Joint space narrowing	
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 (the 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.[70] Maximum total scores for both erosion and joint space narrowing are:

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 both erosions and joint space narrowing; and 3) it covers a broad spectrum of joints, providing sensitivity to change.[71]

In the early stages of RA, inflammation appears to be the main contributor to increased disability, rather than actual damage to joints.[72,73] The relationship between radiological and functional changes has been studied. A re-analysis of published data performed by Welsing et al. found that patients must have a certain amount of radiological damage before an increase in damage will affect disability. The authors also found that changes in Sharp scores had a greater impact on disability with advancing age.[74] A study by Sahin 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.[75] 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.[76] Therefore, radiological changes, assessed by Sharp scores, and functional changes, assessed by the HAQ, do not correlate with each other early in RA, but do correlate after several years of disease.

There are several limitations with using radiographs for assessing clinical status in RA. Radiographs tend to change slowly in RA, requiring at least six months to a year to detect changes in a patient. Inter-rater and intra-rater reliability is also a concern because of 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.[77] Given these limitations, beginning in the early 1990s, the use of magnetic resonance imaging (MRI) was being examined as an alternative for assessing disease progression.[78] However, the use of MRI for assessing clinical status of RA is limited because of cost and accessibility.

In a study by Bruynesteyn et al., the authors determined an MCID of 4.6 units for the Sharp/van der Heijde method, using a panel of experts.[37] They defined the MCID as a progression in radiologic joint damage that leads a rheumatologist to change a patient's therapy. This MCID was equal to, or slightly lower than, the smallest detectable difference (SDD) for this scoring system. The SDD represents the smallest change score that can be reliably discriminated from the measurement error of the scoring method.[79] The smallest detectable change (SDC) score is another method of measuring reliability. Similar to the MCID, the SDC score can provide guidance for interpreting whether there has been a real change in patient outcomes over time. A study by Navarro-Compan et al. assessed the level of agreement between two readers on radiographic images from patients with RA.[80] The authors found a SDC 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.[80]

Functional Assessment of Chronic Illness Therapy Fatigue Scale

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 rheumatoid arthritis (RA).[38]

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 the FACIT-Fatigue in RA patients is between three and four points.[38] This MCID was found in a sample of 271 patients (77% female, 81% Caucasian, median age of 56 years [range 28 to 84 years], a median tender joint count of 26 [range 9 to 68], and a median swollen joint count of 15 [range 2 to 43]).[38]

The SF-36 is a generic health assessment questionnaire that has been used in clinical trials to study the impact of chronic disease on health-related quality of life. The SF-36 consists of eight subdomains: physical functioning, pain, vitality, social functioning, psychological functioning, general health perceptions, and role limitations due to physical and emotional problems.[39] The SF-36 also provides two component summaries, the physical component summary (PCS) and the mental component summary (MCS). The eight subdomains are each measured on a scale 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.[40-42]

Conclusion

ACR response, DAS28, HAQ-DI, and mTSS were used as primary efficacy measures in the RA studies included in this analysis. The ACR 20, 50, and 70 indicate a percentage improvement from baseline (but not a final level of disease activity). ACR 20 is most commonly reported in clinical trials; however, ACR 50 or ACR 70 are often cited as evidence of a more robust treatment effect. The DAS28 measures an absolute rather than relative level of disease activity and thus may be preferred to the ACR response rates. The DAS28 components correlate well with each other and with the ACR components. However, it

was reported that DAS28(CRP) underestimates disease activity and overestimates the improvement in disease activity and the remission rate compared with DAS28(ESR). The MCID for a change in DAS28 scores has not been specified. The HAQ is a comprehensive measure of the patient's perception of functional status and has been widely validated in RA. The HAQ-DI is one of five components (the disability component) of the full HAQ. A suggested MCID in patients with RA is 0.22; however, differences as small as 0.10 have also been suggested. The mTSS allows for the assessment of two different 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 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 4.6 units on the mTSS has been suggested. The FACIT is a self-report measure of fatigue that has been validated for use in patients with RA. The MCID for the FACIT-Fatigue scale has been cited as a three- to four-point change in score.

APPENDIX 6: SUMMARY OF EXTENSION STUDIES

Objective

The objective is to summarize the clinical efficacy and harms of seven studies assessing the long-term efficacy and safety of tofacitinib for the treatment of rheumatoid arthritis (RA). All five pivotal studies (1044, 1046, 1064, 1032, and 1045),[4,7,9,12,14] had a double-blind active extension period (18 months, six months, six months, three months, and three months, respectively) following the placebo-controlled portion of the study. Two additional studies, 1024 and 1041,[81] were open-label extension studies available to patients involved in a previous tofacitinib trial, which followed patients for up to five years. For the purposes of assessing the long-term efficacy and safety of tofacitinib, the focus of this summary will be on outcomes collected at six months or later.

Findings

Figure 5 illustrates the key phases of the five pivotal trials and the flow of patients into the long-term open-label extension trials.

Study Design

Double-blind active extension period of pivotal trials (more than six months).

Study 1044[9]

After the placebo-controlled phase of the study (at three months for a non-responder [NR] in the placebo group; at six months for remaining patients) patients were advanced to a double-blind active extension period (up to two years), in which patients received tofacitinib 5 mg or 10 mg twice daily. An NR was defined as less than 20% improvement in tender/painful and swollen joint counts. Patients who were receiving tofacitinib 5 mg or 10 mg twice daily during the study stayed on their study dose. For patients in the placebo group, the randomization procedure at baseline assigned patients to either a placebo switched to tofacitinib 5 mg group or a placebo switched to tofacitinib 10 mg group. Efficacy outcomes were assessed using the full analysis set, which included all patients who had at least one dose of tofacitinib and at least one post-baseline measurement (efficacy or safety end point). For patients who remained in the study but were missing components of their ACR scores, a last observation carried forward (LOCF) mixed components method was used. ACR scores were calculated using a mix of values from the visit and values from the patients' last observation. Patients who withdrew from the study at any time or who were advanced to treatment at month 3 were considered to be NRs.

Studies 1046 and 1064[12,14]

After the placebo-controlled phase of the study (at three months for NR in the placebo group; at six months for remaining patients), patients were advanced to the double-blind active extension period (up to one year), in which patients received tofacitinib 5 mg or 10 mg twice daily. An NR was defined as less than 20% improvement in tender/painful and swollen joint counts. Patients who were receiving active treatment during the study stayed on their study dose. For patients in the placebo group, the randomization procedure at baseline assigned patients to either a placebo switched to tofacitinib 5 mg group or a placebo switched to tofacitinib 10 mg group. Efficacy outcomes were assessed using the full analysis set, which included all patients that had at least one dose of tofacitinib and at least one post-baseline measurement (efficacy or safety end point). For patients who remained in the study but were missing components of their ACR scores, an LOCF mixed components method was used. ACR scores were calculated using a mix of values from the visit and values from the patients' last observation.

Patients who withdrew from the study at any time or who were advanced to treatment at month 3 were considered to be NRs.

Long-Term Open-Label Extension Study[81]

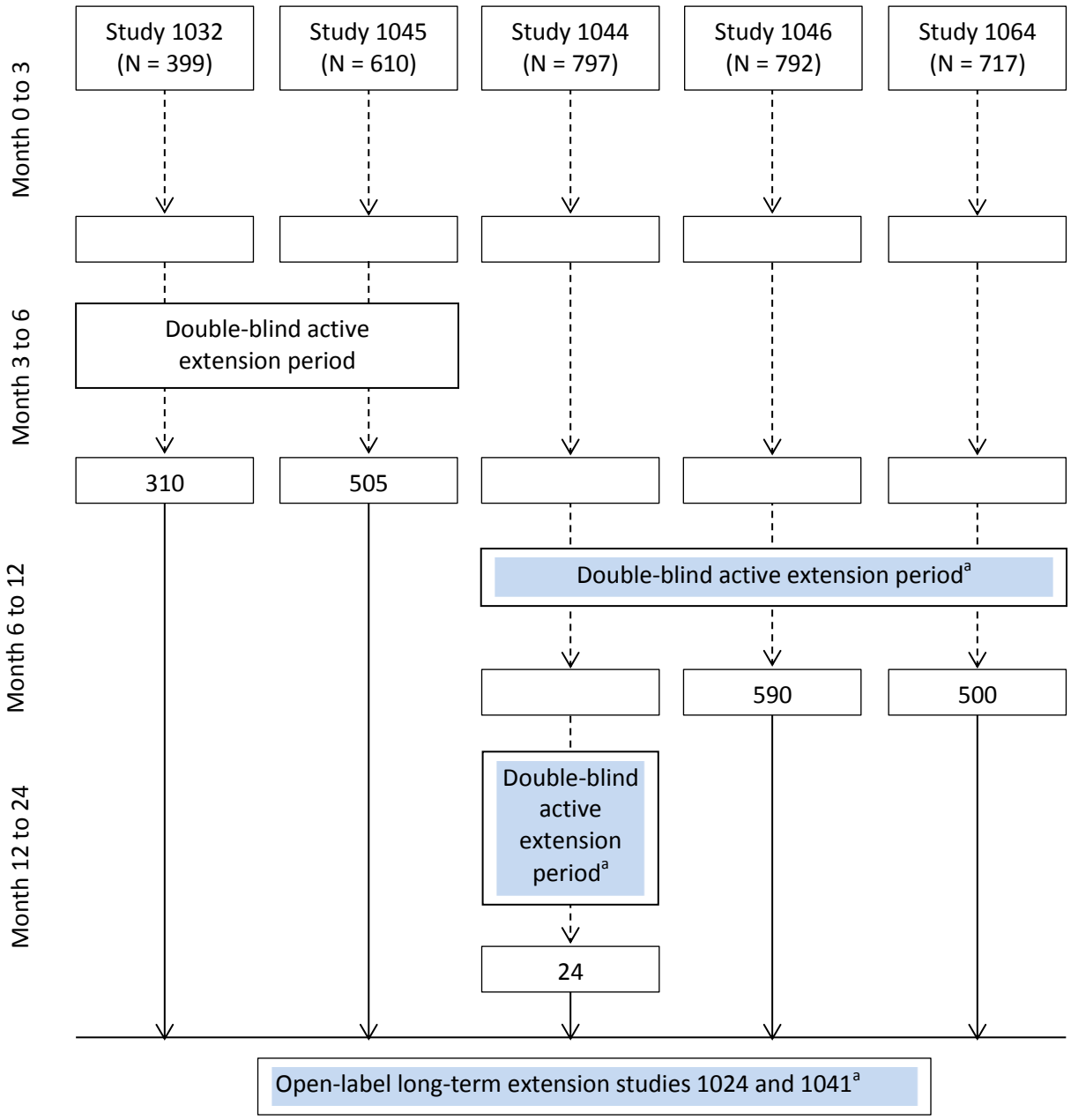
Study 1024

Patients for this long-term extension study had to come from a previous tofacitinib phase 2 or phase 3 trial. All patients were screened for eligibility at the baseline visit, which could correspond with the last visit of the randomized trial. Patients initiated tofacitinib treatment at a dose of 5 mg twice daily until amendment three occurred, at which point all patients initiated treatment at 10 mg (except for patients in China, who initiated treatment at 5 mg). Based on investigator discretion, patients who were receiving tofacitinib 5 mg could increase their dose to 10 mg in cases of inadequate response, and patients who were receiving tofacitinib 10 mg could decrease their dose to 5 mg for safety reasons. Throughout follow-up, drug dosing was also modified at the discretion of the investigator. Patients were followed up at months 1, 2 and 3, and every three months thereafter.

Study 1041

Patients for this long-term extension trial came from Study 1039 or Study 1040, both of which were phase 2 studies based in Japan. Patients initiated tofacitinib treatment at a dose of 5 mg twice daily. Throughout follow-up, modifications to drug dosing were allowed at the discretion of the investigator. Patients were followed up at week 2, week 4 and then every four weeks during the first year of the study. In subsequent years, visits occurred every 8 weeks and then every 12 weeks.

FIGURE 5: FLOW OF PATIENTS INTO LONG-TERM EXTENSION STUDIES



[a] Indicates the safety and efficacy data included in the present long-term exposure summary (i.e., after six months).

Efficacy

Studies 1044, 1046, and 1064[9,12,14]

Efficacy outcomes were assessed using the full analysis set, which included all patients who had at least one dose of tofacitinib and at least one post-baseline measurement (efficacy or safety end point). ACR 20, ACR 50, and ACR 70 results for patients in the tofacitinib 5 mg group and the placebo switched to tofacitinib 5 mg group are presented for months 6 (end of the placebo-controlled phase) to month 12,[12,14] or month 24[9] (end of study) (Table 30). The ACR 20 response rates in the tofacitinib 5 mg twice daily trial group in Study 1044 did show some decline between month 6 (51.46%) and month 24 (41.10%). Otherwise, ACR response rates appear to be generally well maintained over the duration of the studies.

Studies 1024 and 1041[26,81]

Efficacy outcomes for tofacitinib 5 mg twice daily with and without background DMARDs for the two long-term extension studies are displayed in Table 34. ACR 20, ACR 50, and ACR 70 response rates [redacted] in the Japanese population (Study 1041) compared with the global population (Study 1024) [redacted]. ACR response rates [redacted].

TABLE 34: OVERVIEW OF EFFICACY DATA FOR THE DOUBLE-BLIND, ACTIVE EXTENSION PHASE OF STUDIES 1044, 1046, AND 1064 FROM MONTH 6 TO MONTH 24 (FAS POPULATION, NRI)

Time Point	Study 1044 AEP		Study 1046 AEP		Study 1064 AEP	
	TOF 5 mg b.i.d. (N = 316)	Placebo Switched to TOF 5 mg b.i.d. (N = 79)	TOF 5 mg b.i.d. (N = 312)	Placebo Switched to TOF 5 mg b.i.d. (N = 79)	TOF 5 mg b.i.d. (N = 201)	Placebo Switched to TOF 5 mg b.i.d. (N = 56)
ACR 20 response rate, n/N (%)						
Month 6	159/309 (51.46)	21/79 (26.58)	164/311 (52.73)	27/79 (34.18)	101/196 (51.53)	15/56 (26.79)
Month 9	148/309 (47.90)	26/79 (32.91)	157/311 (50.48)	30/79 (37.97)	97/196 (49.49)	17/56 (30.36)
Month 12	150/309 (48.54)	24/79 (30.38)	160/311 (51.45)	25/79 (31.65)	97/196 (49.49)	19/56 (33.93)
Month 15	147/309 (47.57)	23/79 (29.11)	NA	NA	NA	NA
Month 18	139/309 (44.98)	21/79 (26.58)	NA	NA	NA	NA
Month 21	124/309 (40.13)	18/79 (22.78)	NA	NA	NA	NA
Month 24	127/309 (41.10)	20/79 (25.32)	NA	NA	NA	NA
ACR 50 response rate, n/N (%)						
Month 6	100/309 (32.36)	10/79 (12.66)	105/311 (33.76)	27/79 (34.18)	101/196 (51.53)	15/56 (26.79)
Month 9	101/309 (32.69)	18/79 (22.78)	98/311 (31.51)	30/79 (37.97)	97/196 (49.49)	17/56 (30.36)
Month 12	100/309 (32.36)	17/79 (21.52)	104/311 (33.44)	25/79 (31.65)	97/196 (49.49)	19/56 (33.93)
Month 15	99/309 (32.04)	17/79(21.52)	NA	NA	NA	NA
Month 18	96/309 (31.07)	14/79 (17.72)	NA	NA	NA	NA
Month 21	85/309 (27.51)	13/79 (16.46)	NA	NA	NA	NA

Time Point	Study 1044 AEP		Study 1046 AEP		Study 1064 AEP	
	TOF 5 mg b.i.d. (N = 316)	Placebo Switched to TOF 5 mg b.i.d. (N = 79)	TOF 5 mg b.i.d. (N = 312)	Placebo Switched to TOF 5 mg b.i.d. (N = 79)	TOF 5 mg b.i.d. (N = 201)	Placebo Switched to TOF 5 mg b.i.d. (N = 56)
Month 24	89/309 (28.80)	16/79 (20.25)	NA	NA	NA	NA
ACR 70 response rate, n/N (%)						
Month 6	45/309 (14.56)	2/79 (2.53)	105/311 (33.76)	27/79 (34.18)	101/196 (51.53)	15/56 (26.79)
Month 9	56/309 (18.12)	7/79 (8.86)	98/311 (31.51)	30/79 (37.97)	97/196 (49.49)	17/56 (30.36)
Month 12	58/309 (18.77)	8/79 (10.13)	104/311 (33.44)	25/79 (31.65)	97/196 (49.49)	19/56 (33.93)
Month 15	61/309 (19.74)	8/79 (10.13)	NA	NA	NA	NA
Month 18	51/309 (16.50)	11/79 (13.92)	NA	NA	NA	NA
Month 21	50/309 (16.18)	6/79 (7.59)	NA	NA	NA	NA
Month 24	53/309 (17.15)	8/79 (10.13)	NA	NA	NA	NA

ACR = American College of Rheumatology; AEP = active extension phase; b.i.d. = twice daily; FAS = full analysis set; NRI = non-response imputation; NA = not applicable; TOF = tofacitinib.

TABLE 35: OVERVIEW OF EFFICACY DATA FOR THE LONG-TERM EXTENSION STUDIES 1024 AND 1041 FROM MONTH 12 TO MONTH 60

Time Point	5 mg b.i.d. TOF Monotherapy		5 mg b.i.d. TOF + DMARD	
	Study 1024 (Total N = [redacted])	Study 1041 (Total N = [redacted])	Study 1024 (Total N = [redacted])	Study 1041 (Total N = [redacted])
ACR 20 response rate, n/N (%)				
Month 12	[redacted]	[redacted]	[redacted]	[redacted]
Month 24	[redacted]	[redacted]	[redacted]	[redacted]
Month 36	[redacted]	[redacted]	[redacted]	[redacted]
Month 48	[redacted]	[redacted]	[redacted]	[redacted]
Month 60	[redacted]	[redacted]	[redacted]	[redacted]
ACR 50 response rate, n/N (%)				
Month 12	[redacted]	[redacted]	[redacted]	[redacted]
Month 24	[redacted]	[redacted]	[redacted]	[redacted]
Month 36	[redacted]	[redacted]	[redacted]	[redacted]
Month 48	[redacted]	[redacted]	[redacted]	[redacted]
Month 60	[redacted]	[redacted]	[redacted]	[redacted]
ACR 70 response rate, n/N (%)				
Month 12	[redacted]	[redacted]	[redacted]	[redacted]
Month 24	[redacted]	[redacted]	[redacted]	[redacted]
Month 36	[redacted]	[redacted]	[redacted]	[redacted]
Month 48	[redacted]	[redacted]	[redacted]	[redacted]
Month 60	[redacted]	[redacted]	[redacted]	[redacted]

ACR = American College of Rheumatology; DMARD = disease-modifying antirheumatic drug; LTE = long-term extension; NA = not applicable; TOF = tofacitinib.
Source: Manufacturer-provided additional information.[26] Data available up to April 3, 2014, for Study 1024 (ongoing) and up to April 24, 2014, for Study 1041 (completed).

Safety

Studies 1044, 1046, and 1064

Harms evaluations were based on the "safety data set," which was composed of all patients who received at least one dose of tofacitinib. Data were available from month 6 (end of the placebo-controlled phase) to month 12[12,14] or month 24[9] (end of study) (Table 32).

Overall Harms Data: In Study 1044[9] (24 months' duration), patients were exposed to treatment for a median of 709 to 714 days. Five deaths occurred in the tofacitinib 5 mg twice daily group (one due to pneumonia, one due to acute myocardial infarction, one due to cardiac and respiratory arrest, one due to multiple causes [lung cancer, hepatic and renal failure], and one due to cardiac failure and pneumonia) and two in the placebo switched to tofacitinib 5 mg twice daily group (congestive heart failure and cardiac arrest). In studies 1046 and 1064 (12 months' duration), patients were exposed to treatment for a median of 358 to 363 days, and a total of two deaths occurred.

Special Interest Harms Data: In Study 1044[9] (24 months' duration), serious infection occurred in eight (2.5%) patients. There were no cases of tuberculosis reported, and nasopharyngitis was the most common infection. No gastrointestinal perforation was reported. There were a total of 28 cardiac events (congestive heart failure, ischemic heart disease, and myocardial infarction) in the tofacitinib 5 mg twice daily group. In studies 1046 and 1064 (12 months' duration), there were a total of five cases of serious infection across all study groups. There were no cases of tuberculosis, and upper respiratory tract infection was most common type of infection. No gastrointestinal perforation was reported. There were a total of eight cardiac events (congestive heart failure, ischemic heart disease, and myocardial infarction) reported in the tofacitinib 5 mg group in Study 1046, and one case of ischemic heart disease in the placebo to tofacitinib 5 mg group.

Laboratory Data: In Study 1044[9] (24 months' duration), there was evidence of increases in liver enzymes (ALT ≥ 3 × upper limit of normal [ULN] in 5.4% of patients; AST ≥ 3 × ULN in 2.9% of patients), levels of LDL (mean 7% to 19% increase from baseline between month 6 and month 24), and levels of HDL. Between month 6 and month 24, 2.1% or fewer patients had evidence of moderate neutropenia, and at month 24, five (2.4%) patients had a severe decrease in levels of hemoglobin. In studies 1046 and 1064 (12 months' duration), there was evidence of increases in liver enzymes, levels of LDL, and levels of HDL, as well as some evidence of moderate neutropenia and severe hemoglobin decreases.

Studies 1024 and 1041

Harms evaluations were based on the "safety data set," which was composed of all patients who received at least one dose of tofacitinib. Data were available from baseline (end of prior tofacitinib trials) to month 60 (based on the latest information available; Study 1024 is ongoing) (Table 37; Table 38).

Overall Harms Data: In Study 1024 (ongoing), patients had [redacted] patient-years of exposure to tofacitinib 5 mg ([redacted] patient-years of tofacitinib monotherapy and [redacted] patient-years of tofacitinib + DMARD. [redacted] had been reported up to the data cut-off point and [redacted]. In Study 1041 (completed), patients accumulated [redacted] patient-years of exposure to tofacitinib 5 mg ([redacted] patient-years of tofacitinib monotherapy and [redacted] patient-years of tofacitinib + DMARD). [redacted] occurred in this study.

Special Interest Harms Data: [redacted] in Study 1024 for patients receiving tofacitinib monotherapy and tofacitinib +DMARD. [redacted] in Study 1041 ([redacted] for tofacitinib monotherapy and [redacted] for tofacitinib +DMARD). [redacted]

Laboratory Data: Similar to the active extension period in the pivotal trials, there was evidence of increases in levels of LDL and decreases in neutrophil counts from baseline. Data derived from a published report[24] that pooled outcomes from studies 1024 and 1041 showed that a serum creatinine increase of greater than 50% occurred in 3.4% of patients, a decrease in hemoglobin classified as severe occurred in 3.5%, and a decrease in hemoglobin classified as potentially life-threatening occurred in 1.7% of patients over the duration of the study.

Critical Appraisal

Studies 1044, 1046, and 1064

Beyond the six-month mark of studies 1044, 1046, and 1064, all patients either maintained their current treatment dose or switched from placebo to active treatment. Patients remained blinded to their dose of tofacitinib, but no control group was available to make comparisons. Therefore, results are descriptive in nature. In Study 1044, at 24 months, ACR 20 data were available for 71% of patients in the 5 mg tofacitinib group and for 72% of patients in the placebo switched to tofacitinib 5 mg group. ACR 50 and ACR 70 data were available for 68% of patients in both groups. Finally, harms data were not expressed in patient-years, so the incidence of harms is not interpretable based on the amount of exposure to the intervention.

Studies 1024 and 1041

Because of the lack of control group, the results presented herein are descriptive in nature and should be interpreted as such. The most up-to-date information was received from the manufacturer and included data up to April 3, 2014, for Study 1024 (ongoing) and up to April 24, 2014, for Study 1041 (completed).[26] Data are displayed separately for the global and Japanese study, and according to monotherapy or combination therapy, allowing for comparisons between groups. [redacted].

Summary

The long-term safety and efficacy of tofacitinib 5 mg twice daily have been assessed using three double-blind active extension periods (up to 12 and 24 months) and two long-term open-label extension studies (up to 48 months for efficacy and up to 60 months for safety). There is some evidence of decrease in ACR 20 response rates over time (Study 1044); however, efficacy, as assessed by ACR 20/50/70 response rates, was generally maintained. There were a greater number of gastrointestinal perforations and cases of tuberculosis in the long-term extension trial compared with the active extension periods of the pivotal trials.

TABLE 36: OVERVIEW OF SAFETY DATA FOR DOUBLE-BLIND, ACTIVE EXTENSION PHASE OF STUDIES 1044, 1046, AND 1064 FROM MONTH 6 TO MONTH 24 (FAS POPULATION, NRI)

	Study 1044 AEP		Study 1046 AEP		Study 1064 AEP	
	TOF 5 mg b.i.d. (N = 321)	Placebo Switched to TOF 5 mg b.i.d. (N = 81)	TOF 5 mg b.i.d. (N = 315)	Placebo Switched to TOF 5 mg b.i.d. (N = 79)	TOF 5 mg b.i.d. (N = 204)	Placebo Switched to TOF 5 mg b.i.d. (N = 56)
Median duration of treatment, days (range)	709.0 (2 to 742)	714.0 (8 to 736)	358.0 (5 to 387)	361.0 (21 to 372)	363.0 (7 to 385)	363.0 (73 to 368)
Overall harms data						
Deaths, n (%)	5 (1.6)	2 (2.5)	2 (0.6)	0 (0.0)	0 (0.0)	0 (0.0)
SAEs, n of patients (%)	51 (15.9)	6 (7.4)	7 (2.2)	2 (2.5)	10 (4.9)	1 (1.8)
Severe AEs, n of patients (%)	33 (10.3)	5 (6.2)	8 (2.5)	1 (1.3)	5 (2.5)	0 (0.0)
AEs, n	841	177	208	57	162	34
AEs, n of patients (%)	230 (71.7)	48 (59.3)	104 (33.0)	34 (43.0)	89 (43.6)	18 (32.1)
WDAEs, n of patients (%)	32 (10.0)	6 (7.4)	1 (0.3)	0 (0.0)	6 (2.9)	0 (0.0)
Special interest harms data						
Serious infection, n (%)	8 (2.5)	1 (1.2)	1 (0.3)	0 (0.0)	4 (2.0)	0 (0.0)
Tuberculosis (lymph node), n (%)	0 (0.0)	0 (0.0)	NR	NR	NR	NR
Tuberculosis (disseminated), n (%)	0 (0.0)	0 (0.0)	NR	NR	NR	NR
Tuberculosis (pericarditis), n (%)	0 (0.0)	0 (0.0)	NR	NR	NR	NR
Tuberculosis (pulmonary), n (%)	NR	NR	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Bronchitis, n (%)	14 (4.4)	2 (2.5)	3 (1.0)	2 (2.5)	5 (2.5)	2 (3.6)
Herpes zoster, n (%)	15 (4.7)	4 (4.9)	1 (0.3)	0 (0.0)	4 (2.0)	0 (0.0)
Influenza, n (%)	3 (0.9)	1 (1.2)	1 (0.3)	1 (1.3)	2 (1.0)	0 (0.0)
Nasopharyngitis, n (%)	47 (14.6)	7 (8.6)	2 (0.6)	1 (1.3)	8 (3.9)	0 (0.0)
Pharyngitis, n (%)	4 (1.2)	0 (0.0)	2 (0.6)	1 (1.3)	4 (2.0)	1 (1.8)
Pneumonia, n (%)	12 (3.7)	1 (1.2)	5 (1.6)	0 (0.0)	3 (1.5)	0 (0.0)
Sinusitis, n (%)	11 (3.4)	1 (1.2)	5 (1.6)	0 (0.0)	1 (0.5)	1 (1.0)
Upper respiratory tract infection, n (%)	24 (7.5)	6 (7.4)	11 (3.5)	3 (3.8)	8 (3.9)	1 (1.8)
Urinary tract infection, n (%)	13 (4.0)	3 (3.7)	3 (1.0)	0 (0.0)	2 (1.0)	0 (0.0)
Neoplasms (benign, malignant, unspecified), n (%)	5 (1.6)	1 (1.2)	1 (0.3)	1 (0.3)	1 (0.5)	1 (1.8)
Congestive heart failure, n (%)	9 (2.8)	2 (2.5)	2 (0.6)	0 (0.0)	0 (0.0)	0 (0.0)

	Study 1044 AEP		Study 1046 AEP		Study 1064 AEP	
	TOF 5 mg b.i.d. (N = 321)	Placebo Switched to TOF 5 mg b.i.d. (N = 81)	TOF 5 mg b.i.d. (N = 315)	Placebo Switched to TOF 5 mg b.i.d. (N = 79)	TOF 5 mg b.i.d. (N = 204)	Placebo Switched to TOF 5 mg b.i.d. (N = 56)
Ischemic heart disease, n (%)	10 (3.1)	1 (1.2)	3 (1.0)	1 (1.3)	0 (0.0)	0 (0.0)
Myocardial infarction, n (%)	9 (2.8)	1 (1.2)	3 (1.0)	0 (0.0)	0 (0.0)	0 (0.0)
Gastric ulcer perforation, n (%)	0 (0.0)	0 (0.0)	NR	NR	NR	NR
Laboratory Data						
ALT ≥ 3 × ULN, n (%)	15 (5.4)	2 (2.9)	7 (2.6)	0 (0.0)	5 (3.0)	0 (0.0)
AST ≥ 3 × ULN, n (%)	8 (2.9)	1 (1.4)	3 (1.1)	0 (0.0)	2 (1.2)	0 (0.0)
Serum creatinine mg/dL increase > 50% (from baseline), n (%)	3 (< 1.0)	3 (3.7)	0 (0.0)	0 (0.0)	0 (0.0)[a]	0 (0.0)[a]
Moderate neutropenia (ANC ≥ 0.5 to < 1.5 × 10^3/µL), n (%)						
Month 6	3 (1.1)	0 (0.0)	1 (< 1.0)	1(1.4)	0 (0.0)	0 (0.0)
Month 9	1 (< 1.0)	0 (0.0)	2 (< 1.0)	2 (2.8)	0 (0.0)	0 (0.0)
Month 12	1 (< 1.0)	1 (1.5)	2 (< 1.0)	3 (4.3)	0 (0.0)	0 (0.0)
Month 15	5 (2.1)	1 (1.6)	NA	NA	NA	NA
Month 18	1 (< 1.0)	1 (1.8)	NA	NA	NA	NA
Month 21	2 (< 1.0)	1 (1.8)	NA	NA	NA	NA
Month 24	2 (< 1.0)	0 (0.0)	NA	NA	NA	NA
Potential life-threatening neutropenia (ANC < 0.5 × 10^3/µL), n (%)						
Month 6	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Month 9	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Month 12	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Month 15	1 (< 1.0)	0 (0.0)	NA	NA	NA	NA
Month 18	0 (0.0)	0 (0.0)	NA	NA	NA	NA
Month 21	0 (0.0)	0 (0.0)	NA	NA	NA	NA
Month 24	0 (0.0)	0 (0.0)	NA	NA	NA	NA
LDL mg/dL, mean (SD)						
Month 6	128.99 (38.48)	116.99 (40.49)	126.29 (40.21)	124.69 (39.94)	131.93 (43.17)	128.81 (40.65)
Month 9	129.67 (37.89)	123.83 (38.08)	125.42 (41.85)	125.74 (36.29)	130.89 (38.02)	135.04 (37.72)
Month 12	130.48 (38.85)	120.13 (33.04)	127.71 (42.77)	127.29 (39.43)	126.97 (38.65)	134.17 (39.53)
Month 18	125.52 (37.55)	118.43 (33.12)	NA	NA	NA	NA

	Study 1044 AEP		Study 1046 AEP		Study 1064 AEP	
	TOF 5 mg b.i.d. (N = 321)	Placebo Switched to TOF 5 mg b.i.d. (N = 81)	TOF 5 mg b.i.d. (N = 315)	Placebo Switched to TOF 5 mg b.i.d. (N = 79)	TOF 5 mg b.i.d. (N = 204)	Placebo Switched to TOF 5 mg b.i.d. (N = 56)
Month 24	124.57 (38.00)	119.89 (36.37)	NA	NA	NA	NA
LDL mg/dL, % change from baseline, mean per visit (SD)						
Month 6	17.0 (30.17)	7.2 (22.69)	14.7 (26.06)	2.0 (27.53)	13.34 (26.52)	9.44 (19.37)
Month 9	19.3 (34.84)	15.0 (28.94)	12.4 (24.64)	3.2 (19.02)	12.78 (28.43)	17.49 (17.90)
Month 12	17.8 (28.61)	13.3 (29.69)	15.5 (27.61)	5.5 (20.78)	8.88 (26.08)	16.33 (23.01)
Month 18	14.7 (28.39)	10.7 (25.36)	NA	NA	NA	NA
Month 24	15.3 (31.18)	10.6 (29.00)	NA	NA	NA	NA
HDL mg/dL, mean (SD)						
Month 6	66.50 (18.51)	63.23 (19.82)	66.38 (18.32)	61.85 (17.64)	66.16 (18.47)	65.11 (16.60)
Month 9	66.44 (18.96)	68.57 (19.89)	64.03 (17.59)	62.56 (16.69)	66.90 (17.56)	69.66 (18.52)
Month 12	67.64 (19.05)	69.37 (20.27)	68.47 (19.63)	68.93 (19.05)	68.82 (18.18)	71.54 (20.57)
Month 18	69.20 (19.69)	72.25 (20.36)	NA	NA	NA	NA
Month 24	71.00 (21.42)	72.78 (18.88)	NA	NA	NA	NA
HDL mg/dL, % change from baseline, mean per visit (SD)						
Month 6	10.7 (20.15)	4.0 (17.68)	13.55 (22.68)	4.25 (24.23)	12.85 (19.45)	6.18 (20.55)
Month 9	10.8 (23.72)	13.1 (17.07)	9.62 (22.07)	4.48 (20.04)	14.57 (18.84)	14.24 (26.16)
Month 12	13.6 (23.13)	13.5 (17.24)	16.71 (24.75)	14.62 (22.56)	17.56 (20.29)	16.62 (24.55)
Month 18	16.8 (23.85)	17.2 (15.49)	NA	NA	NA	NA
Month 24	18.2 (24.55)	17.1 (18.31)	NA	NA	NA	NA
Severe hemoglobin decrease (> 2 g/dL to < 3 g/dL from baseline)[b], n (%)						
Month 6	1 (< 1.0)	1 (1.4)	1 (< 1.0)	0 (0.0)	0 (0.0)	0 (0.0)
Month 9	1 (< 1.0)	0 (0.0)	4 (1.5)	1 (1.4)	0 (0.0)	0 (0.0)
Month 12	3 (1.2)	0 (0.0)	1 (< 1.0)	0 (0.0)	0 (0.0)	0 (0.0)
Month 15	0 (0.0)	0 (0.0)	NA	NA	NA	NA
Month 18	5 (2.1)	0 (0.0)	NA	NA	NA	NA
Month 21	2 (< 1.0)	0 (0.0)	NA	NA	NA	NA
Month 24	5 (2.4)	0 (0.0)	NA	NA	NA	NA

	Study 1044 AEP		Study 1046 AEP		Study 1064 AEP	
	TOF 5 mg b.i.d. (N = 321)	Placebo Switched to TOF 5 mg b.i.d. (N = 81)	TOF 5 mg b.i.d. (N = 315)	Placebo Switched to TOF 5 mg b.i.d. (N = 79)	TOF 5 mg b.i.d. (N = 204)	Placebo Switched to TOF 5 mg b.i.d. (N = 56)
Potential life-threatening hemoglobin decrease (≥ 3 g/dL from baseline)[c], n (%)						
Month 6	1 (< 1.0)	0 (0.0)	0 (0.0)	1 (1.4)	0 (0.0)	0 (0.0)
Month 9	1 (< 1.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Month 12	1 (< 1.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (< 1.0)	0 (0.0)
Month 15	0 (0.0)	0 (0.0)	NA	NA	NA	NA
Month 18	0 (0.0)	0 (0.0)	NA	NA	NA	NA
Month 21	2 (< 1.0)	1 (1.7)	NA	NA	NA	NA
Month 24	0 (0.0)	0 (0.0)	NA	NA	NA	NA

AE = adverse events; AEP = active extension phase; ALT = alanine transaminase; ANC = absolute neutrophil count; AST = aspartate transaminase; b.i.d. = twice daily; FAS = full analysis set; LDL = low-density lipoprotein; NRI = non-responder imputation; SAE = serious adverse event; SD = standard deviation; TOF = tofacitinib; ULN = upper limit of normal; WDAE = withdrawal due to adverse events.

[a] Serum creatinine > 50% at two sequential visits.

[b] Severe hemoglobin decrease also defined as an actual hemoglobin count of > 7g/dL to < 8g/dL.

[c] Potential life-threatening hemoglobin decrease also defined as an actual hemoglobin count of ≤ 7g/dL.

Source: Manufacturer submitted clinical study reports.[9,12,14]

TABLE 37: OVERVIEW OF SAFETY DATA FOR LONG-TERM EXTENSION STUDIES 1024 AND 1041 FROM BASELINE TO MONTH 60

	5 mg b.i.d. TOF Monotherapy		5 mg b.i.d. TOF + DMARD	
	Study 1024 (N = [redacted])	Study 1041 (N = [redacted])	Study 1024 (N = [redacted])	Study 1041 (N = [redacted])
Total PYs of exposure	[redacted]	[redacted]	[redacted]	[redacted]
Overall harms data, n (events per 100 PYs)[a]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Special interest harms data, n (events per 100 PYs; 95% CI)[a]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Laboratory data				
[redacted]				
Month 12	[redacted]	[redacted]	[redacted]	[redacted]
Month 24	[redacted]	[redacted]	[redacted]	[redacted]
Month 36	[redacted]	[redacted]	[redacted]	[redacted]
Month 48	[redacted]	[redacted]	[redacted]	[redacted]

	5 mg b.i.d. TOF Monotherapy		5 mg b.i.d. TOF + DMARD	
	Study 1024 (N =)	Study 1041 (N =)	Study 1024 (N =)	Study 1041 (N =)
Month 60				
Month 12				
Month 24				
Month 36				
Month 48				
Month 60				

; b.i.d. = twice daily; CI = confidence interval; DMARD = disease-modifying antirheumatic drugs; ; ; PY = patient-year; ; TOF = tofacitinib; ; ; .

[a] Events per 100 patient-years were based on patient-years of exposure that were censored at the time of event.
Source: Manufacturer-provided additional information.[26] Data available up to April 3, 2014, for Study 1024 (ongoing) and up to April 24, 2014, for Study 1041 (completed).

TABLE 38: OVERVIEW OF ADDITIONAL LABORATORY DATA FOR LONG-TERM EXTENSION STUDIES 1024 AND 1041 FROM BASELINE TO MONTH 60

	Pooled Data: LTE Studies 1024 and 1041
	TOF 5 mg b.i.d. ± Background DMARDs (N = 1,421)
Laboratory data	
ALT ≥ 3 × ULN, N (%)	64 (4.5)
AST ≥ 3 × ULN, N (%)	31 (2.2)
Serum creatinine mg/dL increase > 50% (from baseline), N (%)	49 (3.4)
Moderate neutropenia (ANC ≥ 0.5 to < 1.5 × 10^3/µL), N (%)	15 (1.1)
Potential life-threatening neutropenia (ANC < 0.5 × 10^3/µL), N (%)	0 (0.0)
Severe hemoglobin decrease (> 2 g/dL to < 3 g/dL from baseline), N (%)	49 (3.5)
Potential life-threatening hemoglobin decrease (≥ 3 g/dL from baseline), N (%)	24 (1.7)

ALT = alanine transaminase; ANC = absolute neutrophil count; AST = aspartate transaminase; b.i.d. = twice daily; DMARD = disease-modifying antirheumatic drugs; LTE = long-term extension; TOF = tofacitinib.
Source: Wollenhaupt et al. 2014.[82]

APPENDIX 7: SUMMARY AND APPRAISAL OF MANUFACTURER-SUBMITTED MIXED TREATMENT COMPARISON (1 OF 3)

Objective

The objective of this review is to summarize the methods and results, and to conduct a critical appraisal of the manufacturer-provided mixed treatment comparison (MTC) comparing the efficacy of biologic drugs in combination with methotrexate in patients with rheumatoid arthritis (RA) who have previously been treated with methotrexate.[83]

Summary of Mixed Treatment Comparison

Rationale

The manufacturer indicated that the MTC was undertaken because of lack of randomized controlled trials designed to assess the comparative efficacy of tofacitinib with other biologic drugs. Comparative data were needed to inform the health economics model submitted to the CADTH Common Drug Review (CDR).

Methods

Eligibility Criteria

In order to be eligible for inclusion, trials had to include patients with RA, enrol patients with an inadequate response to methotrexate (MTX) or tumour necrosis factor (TNF) inhibitors and have a randomized trial design. There was no restriction on duration of trial, but attempts were made to gather data from a six-month time point.

Intervention and Comparators

The included interventions were tofacitinib, etanercept, infliximab, adalimumab, certolizumab pegol, golimumab, rituximab, abatacept, tocilizumab, and anakinra, in combination with MTX. Studies were excluded if patients were allowed to continue therapy with DMARDs other than MTX. Mean baseline MTX dose was required to be greater than or equal to 15 mg. Studies had to include dosages approved for use in Canada in at least one intervention. The dosage of tofacitinib included in the analysis was 5 mg twice daily (studies 1044 and 1064 and a phase 2 dose-finding study not included in this CDR systematic review).

Outcomes

The outcomes of interest were American College of Rheumatology (ACR) 20, ACR 50, and ACR 70.

Analysis

The manufacturer used Bayesian MTC models to compare relative efficacy of the drugs of interest. It used the data from the trials, likelihood distributions, a model, and prior distributions. WinBUGS 1.4.3 software was used for all MTC analyses. For each of the ACR end points, a network meta-analysis was used to compare the relative effects of the drugs. Both fixed and random effects models were implemented.

Results

Study and Patient Characteristics

There were a total of 17 placebo-controlled trials included in the network meta-analysis (abatacept n = 3, adalimumab n = 4, anakinra n = 2, etanercept n = 2, golimumab n = 1, infliximab n = 2, rituximab n = 1, tocilizumab n = 1, and tofacitinib n = 3). Two of these trials had multiple groups with a head-to-

head biologics comparison (abatacept versus infliximab versus placebo and adalimumab versus tofacitinib versus placebo).
The authors did not report assessment of quality of included studies, and very few details were provided about the studies. No information was provided regarding the study populations, inclusion/exclusion criteria, blinding, or timing of end points.

Results of the Network Meta-Analysis
The end points assessed were ACR 20, ACR 50, and ACR 70. The results of the manufacturer's Bayesian random effects model are summarized in Table 39.

TABLE 39: ODDS RATIO ESTIMATES OF ACHIEVING ACR RESPONSE FOR BIOLOGIC AGENT VERSUS PLACEBO

Drug	Median (95% Credible Interval)		
	ACR 20	ACR 50	ACR 70
Abatacept	3.11 (1.75 to 5.49)	3.66 (1.65 to 8.43)	4.54 (1.87 to 15.57)
Adalimumab	3.79 (2.32 to 6.60)	5.16 (2.54 to 11.17)	5.77 (2.33 to 15.51)
Anakinra	2.20 (1.04 to 4.73)	3.32 (1.16 to 11.41)	5.31 (1.37 to 27.76)
Etanercept	2.49 (1.23 to 6.18)	3.65 (1.40 to 14.49)	4.20 (1.37 to 27.76)
Golimumab	3.85 (1.34 to 11.27)	3.83 (0.87 to 16.97)	4.70 (0.78 to 30.97)
Infliximab	2.68 (1.35 to 5.61)	3.82 (1.45 to 11.38)	5.99 (1.95 to 35.44)
Rituximab	3.11 (1.10 to 8.78)	3.44 (0.78 to 15.04)	5.03 (1.95 to 35.44)
Tocilizumab	3.09 (1.21 to 8.02)	3.57 (0.89 to 14.39)	5.03 (0.80 to 33.12)
Tofacitinib	3.28 (1.81 to 5.78)	4.53 (1.94 to 10.03)	10.3 (3.37 to 30.27)

ACR = American College of Rheumatology.

The authors also ranked the nine biologics using rankograms for ACR 20, ACR 50, and ACR 70. Using the surfaces under the cumulative ranking curve (SUCRA) method to rank the relative probability of achieving ACR response for each biologic agent, tofacitinib ranked fifth for ACR 20, third for ACR 50, and first for ACR 70. No specific statistical comparisons between tofacitinib and other drugs were provided.

The authors assume that "Overlapping 95% credible intervals indicated that no intervention was found to be significantly superior to the others for any of the three end points of interest." The manufacturer did not provide pairwise comparisons for any combination of biologic drugs, so there was no evidence given to support their assumption.

Critical Appraisal of Network Meta-analysis
The quality of the manufacturer's network meta-analysis was assessed according the recommendations provided by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Task Force on Indirect Treatment Comparisons.[84] Details and commentary for each of the relevant items identified by the ISPOR group are provided in Table 40.

Limitations
- The most significant limitation of the study is that the authors did not provide any statistical point estimates for the relative efficacy of tofacitinib versus any other biologic agent.
- There was very little information provided that would allow assessment of internal or external validity of the included studies (e.g., baseline data and demographics).

- The authors discussed heterogeneity, but only from a statistical viewpoint. For example, there were no details regarding the heterogeneity of the time points for ACR response that were pooled by the authors. Because of reporting limitations, this could not be assessed.
- Of the tofacitinib studies that were ≥ 12 months' duration, tofacitinib studies 1044 and 1064 were included, but no data from Study 1046 were used. No explanation was provided for this.
- The search strategy appeared to have been limited to 2011 and 2012, but trials were selected from earlier years. No explanation was given for this.
- No other outcomes, other than ACR response, were reported.
- No indirect comparison statistics were provided. Data were all drug versus placebo.
- No explanation was given for why the certolizumab studies did not meet the inclusion criteria.

Strengths

- ACR 20, ACR 50, and ACR 70 were appropriate outcomes to select for the analysis.
- The biologic drugs selected for the analysis are relevant comparators.
- The authors attempted to reduce clinical heterogeneity by limiting data extraction to ACR response data from the six-month time point, or close to this time point.

TABLE 40: CRITICAL APPRAISAL

Checklist Item	Details and Comments
Are the rationale for the study and the study objectives stated clearly?	• Rationale clearly stated – no head-to-head trials, need to determine the comparative effectiveness and safety for an economic analysis.
Does the methods section include the following? • Description of eligibility criteria • Information sources • Search strategy • Study selection process • Data extraction (validity/quality assessment of individual studies)	• Literature search methods, search terms, and dates presented • Search strategy • Inclusion criteria presented • No critical appraisal performed • Very little information on patient characteristics in the included studies.
Are the outcome measures described?	• No, but this was not necessary since ACR response is a well-understood and standardized outcome.
Is there a description of methods for analysis/synthesis of evidence? Do the methods described include the following? • Description of analyses methods/models • Handling of potential bias/inconsistency • Analysis framework	• Brief description provided for analysis methods and models, description of statistics used • Brief description of how statistical heterogeneity was dealt with using random effects modelling • Bayesian modelling used for MTC.
Are sensitivity analyses presented?	• No sensitivity analyses were provided using selection of different trials. • Various analyses were presented to provide different perspectives on the data: fixed effects, random effects, rankogram, surfaces under the cumulative ranking curve, mean, median.
Do the results include a summary of the studies included in the network of evidence? • Individual study data? • Network of studies?	• Table/list of studies with information regarding study design and patient characteristics presented but very few data were provided • Network of studies presented.

Checklist Item	Details and Comments
Does the study describe an assessment of model fit? Are competing models being compared?	• Authors describe how the total residual favoured the random effects model, but both random effects and fixed effects results were presented.
Are the results of the evidence synthesis (MTC) presented clearly?	• Tables with results were clearly presented. • Point estimates (mean, median) and measure of uncertainty (95% credible intervals) presented.
Sensitivity/scenario analyses	• No sensitivity analyses performed.
Does the discussion include the following? • Description/summary of main findings • Internal validity of analysis • External validity • Implications of results for target audience	• Main findings discussed in general terms • No information was provided to allow a thorough analysis of internal or external validity • Comparisons were made with the 2010 CADTH biologics class review.

ACR = American College of Rheumatology; MTC = mixed treatment comparison.
Source: Manufacturer-submitted MTC.

Summary

Because of the absence of head-to-head trials comparing tofacitinib with other biologic drugs currently used to treat RA, the manufacturer undertook a MTC of randomized controlled trials. No pairwise statistical comparisons were performed; therefore, the relative efficacy of tofacitinib compared with other biologics remains uncertain.

APPENDIX 8: SUMMARY AND APPRAISAL OF MANUFACTURER-SUBMITTED MIXED TREATMENT COMPARISONS (2 OF 3 AND 3 OF 3)

Objective

To summarize and critically appraise two additional network meta-analyses (NMAs) submitted by the manufacturer.

Summary of Mixed Treatment Comparison

Rationale

Randomized controlled trials that evaluate the relative efficacy of tofacitinib compared with all other available biologics for the treatment of RA are not currently available. Three NMAs were undertaken by the manufacturer to obtain evidence of the comparative efficacy of RA biologics. These studies vary according to patient response to previous RA treatment:

1. **MAPI 2011**:[85] To compare the efficacy and safety of tofacitinib 5 mg and 10 mg twice daily (administered alone or in combination with a DMARD) to other biologic response modifiers in adult patients who are inadequate responders to DMARDs (DMARD-IR) or who are inadequate responders to TNF inhibitors (TNF-IR).
2. **MAPI 2013**:[86] To compare the efficacy, safety and tolerability of tofacitinib 5 mg twice daily (administered alone or in combination with a DMARD) to other biologic response modifiers in patients who are TNF-IR.

Methods

Eligibility Criteria

The inclusion criteria for the systematic reviews of the NMAs are listed in Table 41.

TABLE 41: INCLUSION CRITERIA FOR THE SYSTEMATIC REVIEWS OF THE NMAS

	MAPI 2011[85]	MAPI 2013[86]
Population	Adults with RA and an inadequate response to DMARDs or TNF inhibitors.	Adults with moderate to severe RA and an inadequate response to TNF inhibitors.
Intervention	Tofacitinib, etanercept, infliximab, adalimumab, certolizumab pegol, golimumab, rituximab, abatacept, tocilizumab, and anakinra as monotherapy or in combination with a DMARD. **Dose** At dosage levels approved by the FDA	Tofacitinib, adalimumab, etanercept, and tocilizumab as monotherapy, or tofacitinib, adalimumab, etanercept, tocilizumab, infliximab, golimumab, rituximab, abatacept, and anakinra in combination with a DMARD. **Dose** At dosage levels approved in the UK
Comparator	Any active intervention in the previous row or placebo.	Any active intervention in the previous row or placebo.

	MAPI 2011[85]	MAPI 2013[86]
Outcomes	**Clinical Efficacy** ACR response, DAS28, DAS remission, HAQ-DI, EULAR response criteria, patient's assessment of pain, patient's global assessment, physician's global assessment, CRP, erythrocyte sedimentation rate, tender and swollen joint counts, and accepted indices of joint damage. *Assessment time points:* not specified a priori. **Safety** serious adverse events (serious and opportunistic infections, malignancies, tuberculosis, cardiovascular and hematologic events), withdrawals (all causes, due to adverse events, and due to lack of efficacy). *Assessment time point:* not specified a priori.	**Clinical Efficacy** ACR response, HAQ-DI CFB, and DAS28. *Assessment time points:* 12 and 24 weeks. **Safety** withdrawals (all reasons, due to adverse events, and due to lack of efficacy), adverse events, serious adverse events, infections, serious infections. *Assessment time point:* end of trial.
Study Type	RCTs	RCTs

ACR = American College of Rheumatology; CFB = change from baseline; CRP = C-reactive protein; DAS = Disease Activity Score; DMARD = disease-modifying antirheumatic drug; EULAR = European League Against Rheumatism; FDA = Food and Drug Administration; HAQ-DI = Health Assessment Questionnaire –Disability Index; NMA = network meta-analysis; RA = rheumatoid arthritis; RCT = randomized controlled trial; TNF = tumour necrosis factor.
Source: Manufacturer's submission material.[85,86]

Analysis

DMARD-IR and TNF-IR Population:[85] All included studies were assessed using the Jadad quality assessment instrument. The authors stated that they did not plan to use the results of the quality assessment in their analysis, but provided the information for the sake of completeness for interpreting the results. Potential sources of clinical, methodological, or statistical heterogeneity between studies were not stated a priori by the authors and therefore the methods for handling such heterogeneity were not described.

A Bayesian NMA was used to determine the relative efficacy of tofacitinib 5 mg and 10 mg twice daily monotherapy or combination therapy (with MTX) to other biologics (alone or in combination with MTX) at their approved FDA doses. All outcomes were assessed using both a fixed and random effects model. Normal, binomial, and Poisson likelihood distributions were used according to outcome type (continuous, binary, and rate, respectively), and non-informative prior distributions were used for the relative treatment effects in the model. The statistical software WinBUGS was used for the analysis.

TNF-IR Population:[86] All included studies were assessed using the Jadad and the Centre for Reviews and Dissemination quality assessment instruments. The authors stated that they did not plan to use the results of the quality assessment in their analysis, but provided the information for the sake of completeness for interpreting the results. Potential effect modifiers (differences in patient characteristics, concomitant treatments, duration of disease, and baseline disease severity) and the

methods for handling such heterogeneity (subgroup analysis, meta-regression) were stated a priori. Intended analyses also included a comparison of direct and indirect evidence.

A Bayesian NMA was used to determine the efficacy of tofacitinib 5 mg twice daily monotherapy or combination therapy (with DMARDs) relative to other biologics (alone or in combination with DMARDs) at their approved UK doses. All outcomes were planned to be assessed using both a fixed and random effects model. Normal, binomial/ordered, and Poisson likelihood distributions were used according to outcome type (continuous, binary/multinomial, and rate, respectively), and non-informative prior distributions were used for the relative treatment effects in the model. The statistical software WinBUGS and Open BUGS were used for the analysis.

Results

Study and Patient Characteristics

DMARD-IR and TNF-IR Population (MAPI 2011):[85] Thirty-eight relevant trials were identified: 33 involving a DMARD-IR population (Table 42) and five involving a TNF-IR population (Table 43).The DMARD-IR studies included six studies assessing biologic response modifiers as monotherapy versus placebo monotherapy, and 27 looking at combination biologic response modifiers versus combination placebo. Overall, the characteristics of the included studies and their patient populations were generally similar. The included studies were randomized, double-blind, placebo-controlled trials. Most were multi-centre, lasting between 12 and 104 weeks, and assessed ACR 20, ACR 50, ACR 70, HAQ, as well as adverse events, withdrawals due to adverse events, and infections at 12 and 24 weeks. The majority of patients were female, were a mean age of 42 to 57 years old, and had experienced RA for a median of 9 years. Disease severity according to swollen joint count (SJC) and tender joint count (TJC) ranged from a median of 9.7 to a mean of 25 joints, and a median of 12.4 to a mean of 35.5 joints, respectively. All patients in the trials were receiving at least one non-DMARD background therapy, such as nonsteroidal anti-inflammatory drugs (NSAIDs) or low-dose corticosteroids. In studies that were assessing a biologic in combination with a DMARD, the DMARD of choice was generally MTX. Important differences were noted for the following studies:

1. Maini et al. (tocilizumab): included patients with mean baseline **disease duration of 9 to 11 months**, a much shorter duration of disease than other studies
2. ADORE (etanercept): **open-label trial**
3. TEMPO (etanercept): included patients who failed a DMARD other than MTX
4. COMBE (etanercept): the intervention involved the biologic + **sulfasalazine** (instead of MTX)
5. ORAL Sync 1046 (tofacitinib): the intervention involved the biologic + **any DMARD** (not limited to MTX)
6. Fleischmann et al. 2010 (tofacitinib), Van de Putte et al. 2003 (adalimumab), Weinblatt et al. 2003 (abatacept), Keystone et al. 2009 (golimumab), Keystone et al. 2008 (certolizumab pegol), RAPID2 (certolizumab pegol), START (infliximab), OPTION (tocilizumab), TOWARD (tocilizumab), DANCER (rituximab), SERENE (rituximab), Kremer 2010 (tofacitinib), ORAL Scan 1044 (tofacitinib), ORAL Solo 1045 (tofacitinib), ORAL Sync 1046 (tofacitinib), ORAL Standard 1064 (tofacitinib): employed **early escape** prior to study week 24.

The characteristics of the TNF-IR studies and their patient populations were generally similar. Important differences were noted for the following studies:

1. ATTAIN (abatacept): patients were eligible only if they were taking an oral DMARD or anakinra
2. ORAL Step 1032 (tofacitinib): concomitant MTX dose ranged from 7.5 mg to 25 mg weekly
3. REFLEX and RADIATE (rituximab; tocilizumab): concomitant MTX dose ranged from 10 mg to 25 mg weekly

4. GO-AFTER (golimumab): did not require concomitant DMARD therapy
5. GO-AFTER (golimumab), RADIATE (tocilizumab), REFLEX (rituximab), ORAL Step 1032 (tofacitinib): employed **early escape** prior to study week 24.

TNF-IR Population (MAPI 2013):[86] Five trials were identified from the systematic search of the literature: RELFEX, RADIATE, ORAL Step 1032, ATTAIN, and GO-AFTER. Details of the sample sizes, interventions, definition of IR, reassignment schemes, and validity are detailed Table 43. The characteristics of the TNF-IR studies and their patient population included in this NMA were the same as in the MAPI 2011 NMA. The majority of patients were female, with a mean age of 52 to 55 years and a mean disease duration of 11.5 years. Disease severity according to SJC ranged from a median of 14 to a mean of 23 and according to TJC from a median of 26 to 27 to a mean of 33 to 34. The number and type of TNF inhibitors previously used by patients varied according to study. In the ATTAIN study, the number of previous anti-TNF inhibitors used was not available. The predominant previous anti-TNF in ATTAIN was infliximab (67.8% of patients in the active-treatment group and 60.2% in the placebo group). The predominant previous anti-TNF agent in GO-AFTER was adalimumab in the placebo group (55% of patients) and etanercept in the active-treatment group (50% of patients). Likewise, the predominant previous anti-TNF agent in RADIATE was adalimumab in the placebo group (39.4% of patients) and etanercept in the active-treatment group (38.3% of patients). Infliximab was the most predominant in the REFLEX study (81% of patients in the placebo group and 71% of patients in the active-treatment group). Adalimumab was predominant in ORAL Step 1032 (59.1% in the placebo group and 48.9% in the active-treatment group). Mean number of previous drugs was not reported; however, the percentage of patients having one, two, and greater than three anti-TNF drugs ranged between 42% and 65.2%, 27.8% and 44%, and 9% and 18%, respectively, among the six treatment groups among the studies RADIATE, REFLEX, and ORAL Step 1032.

TABLE 42: CHARACTERISTICS OF INCLUDED RANDOMIZED CONTROLLED TRIALS — DMARD-IR POPULATION

Trial	Sample Size (Total of Study Groups)	Intervention	Female (Range of Percentages in Study Groups)	Age (in Years, Range of Means in Study Groups)	Disease Duration (in Years, Range of Means in Study Groups)	Rescue Strategy, Week, Imputation Method
Biologic monotherapy vs. placebo						
ORAL Solo 1045	610	TOF	85.2 to 88.2	49.8 to 52.4	NR	Yes, week 12, LOCF
Fleischmann 2010	222	TOF / ADA	85 to 88	52 to 54	8 to 11	Yes, week 12, LOCF
Fleischmann 2009	220	CTZ	78 to 89	53 to 55	9 to 10	No rescue
Van de Putte 2004	223	ADA	77 to 80	53 to 54	11 to 12	No rescue
Van de Putte 2003	140	ADA	81	50 to 53	9 to 10	Yes, week 12, not reported
Moreland 1999[a]	158	ETN	74 to 76	51 to 53	11 to 12	No rescue
Biologic + MTX vs. placebo + MTX						
A3921025	214	TOF	74 to 81	52 to 56	8 to 9	NR
ORAL Scan 1044	797	TOF	84 to 86	52 to 54	NR	Yes, week 12, BOCF[b]
ORAL Sync 1046[c]	792	TOF	77 to 83	52 to 53	NR	Yes, week 12, BOCF[b]
ORAL Standard 1064	717	TOF / ADA	76 to 85	53 to 54	NR	Yes, week 12, BOCF[b]
Weinblatt 2003	129	ADA	75 to 82	56 to 57	11 to 12	Yes, week 16, unclear
Furst 2003	636	ADA	79 to 80	55 to 56	9 to 12	No[d], week 12, NR
Keystone 2004	407	ADA	73 to 76	56	11	No[d], week 16, NR
Kremer 2003[a]	234	ABT	66 to 75	55 to 56	9 to 10	No rescue
Kremer 2006[a]	652	ABT	78 to 82	50 to 52	9	No rescue
Kay 2008	70	GLB	74 to 86	52 to 57[e]	6 to 8	No rescue
Keystone 2009	222	GLB	81 to 82	52[e]	4.5 to 6.2[e]	Yes, week 16, LOCF
Keystone 2008[a]	592	CTZ	82 to 84	51 to 52	6	Yes, week 16, NRI
Smolen 2009	373	CTZ	84	52	6	Yes, week 16, LOCF
Maini 1999[a]	174	IFX	80 to 81	51 to 54	10 to 11	No rescue
Westhovens 2006	723	IFX	80 to 83	42 to 53	8	Yes, week 22, LOCF
Smolen 2008	409	TCZ	NR	51	8	Yes, week 16, LOCF
Genovese 2008	1216	TCZ	81 to 84	53 to 54	10	Yes, week 16, NR
Maini 2006	99	TCZ	NR	50 to 51	NR	No rescue

Trial	Sample Size (Total of Study Groups)	Intervention	Female (Range of Percentages in Study Groups)	Age (in Years, Range of Means in Study Groups)	Disease Duration (in Years, Range of Means in Study Groups)	Rescue Strategy, Week, Imputation Method
Cohen 2004	501	ANA	75 to 79	56 to 57	10 to 11	No rescue
Emery 2006[a]	341	RTX	80	51	9 to 11	Yes, 16 weeks, LOCF
Emery 2010	344	RTX	81 to 86	51 to 52	7	Yes, 16 weeks, NRI/LOCF
Weinblatt 1999	89	ETN	73 to 90	48 to 53	13	No rescue
Klareskog 2004[a]	682	ETN	74 to 79	53	6 to 7	No rescue
COMBE 2006[a]	254	ETN	79 to 82	51 to 53	6 to 7	No rescue
Schiff 2008	431	ABT	82 to 87	49	7 to 8	No rescue
Edwards 2004[a]	120	RTX	73 to 80	54	9 to 12	No rescue
Van Riel 2006[a]	307	ETN	77 to 79	53 to 54	10	No rescue

ABT = abatacept; ADA = adalimumab; ANA = anakinra; BOCF = baseline observation carried forward; CTZ = certolizumab; DMARD-IR = disease-modifying antirheumatic drugs inadequate responder; ETN = etanercept; GLB = golimumab; IFX = infliximab; IR = inadequate response; LOCF = last observation carried forward; MTX = methotrexate; NR = not reported; NRI = non-response imputation; RCT = randomized controlled trial; RTX = rituximab; TCZ = tocilizumab; TOF = tofacitinib.

[a] Index study, multiple publications.

[b] "Patients who 'advance' to their next assigned treatment are considered to be non-responders." (Appendix A-F)

[c] Background DMARD could be other than MTX.

[d] No rescue strategy, but patients could have received a dose increase for DMARD or corticosteroid therapy.

[e] Median value.

Note: Only the first author of each publication is indicated for brevity

Source: MAPI Report 2011 (Table 5) & MAPI Report 2011 AppendicesA-F[62]

TABLE 43: CHARACTERISTICS OF INCLUDED RANDOMIZED CONTROLLED TRIALS —TNF-IR POPULATION

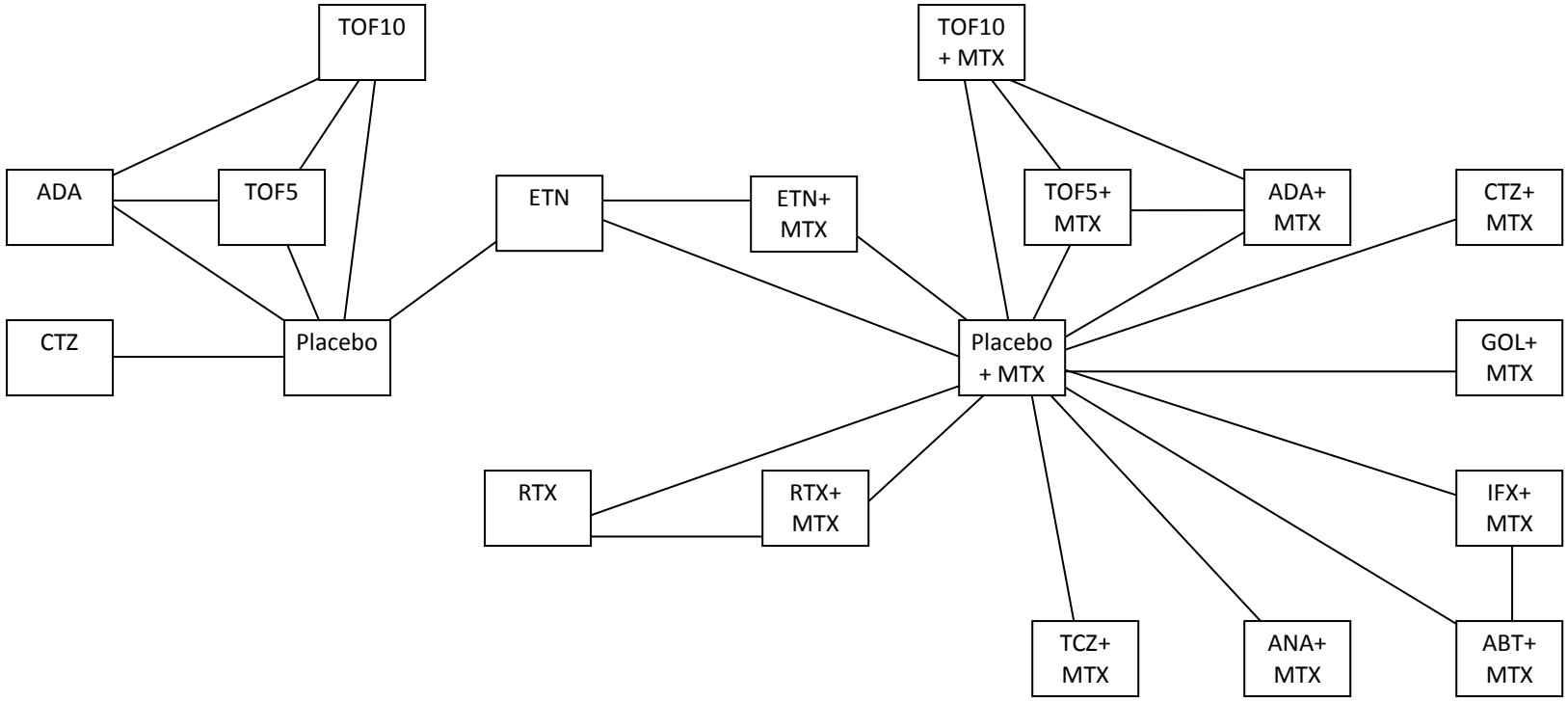
Trial	Sample Size (Total of Study Groups)	Intervention	Female (Range of Percentages in Study Groups)	Age (In Years, Range of Means in Study Groups)	Disease Duration (In Years, Range of Means in Study Groups)	Rescue Strategy, Week, Imputation Method
Biologic + DMARD vs. placebo + DMARD						
Genovese 2005	391	ABT[a]	77 to 80	53	11 to 12	No rescue
Smolen 2009	308	GLB[a]	74 to 85	54 to 55	10	Yes, week 16, LOCF
Emery 2008	335	TCZ	79 to 84	53 to 54	12	Yes, week 16. LOCF
Cohen 2006	517	RTX	81	52 to 53	12	Yes, week 16-24, LOCF
ORAL Step 1032	399	TOF	80 to 87	54 to 55	NR	Yes, week 12, LOCF

ABT = abatacept; DMARD = disease-modifying antirheumatic drug; GLB = golimumab; IR = inadequate response; LOCF = last observation carried forward; RCT = randomized controlled trial; RTX = rituximab; TCZ = tocilizumab; TNF = tumour necrosis factor; TOF = tofacitinib.

[a] Background DMARD could be other than MTX.

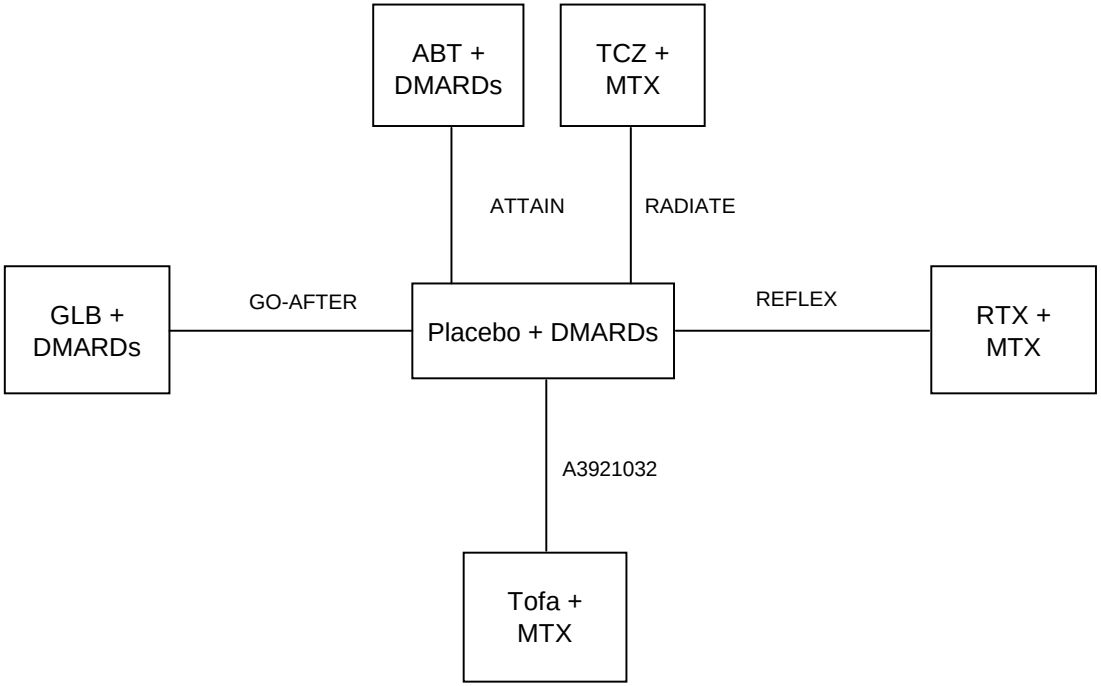
Source: MAPI Report 2011 (Table 5) & MAPI Report 2011 Appendices A to F.[62]

FIGURE 6: NETWORK DIAGRAM FOR DMARD-IR POPULATION



ABT = abatacept; ADA = adalimumab; ANA = anakinra; BOCF = baseline observation carried forward; CTZ = certolizumab; DMARD-IR = disease-modifying antirheumatic drugs inadequate responder; ETN = etanercept; GLB = golimumab; IFX = infliximab; IR = inadequate response; LOCF = last observation carried forward; MTX = methotrexate; NR = not reported; NRI = non-response imputation; RCT = randomized controlled trial; RTX = rituximab; TCZ = tocilizumab; TOF = tofacitinib.
Source: MAPI 2011 Report.[85]

FIGURE 7: NETWORK DIAGRAM FOR TNF-IR POPULATION



ABT = abatacept; ADA = adalimumab; ANA = anakinra; BOCF = baseline observation carried forward; CTZ = certolizumab; DMARD = disease-modifying antirheumatic drug; ETN = etanercept; GLB = golimumab; IFX = infliximab; IR = inadequate response; LOCF = last observation carried forward; MTX = methotrexate; NR = not reported; NRI = non-response imputation; RCT = randomized controlled trial; RTX = rituximab; TCZ = tocilizumab; TOF = tofacitinib.
Source: MAPI 2011 Report.[85]

Results of the Mixed Treatment Comparison
DMARD-IR and TNF-IR Population (MAPI 2011)
a) DMARD-IR (MAPI 2011)

For the purposes of the summary and critical appraisal, only the results related to the Health Canada–approved tofacitinib dose of 5 mg are reported.

ACR 20 (12 and 24 Weeks) (See Table 45***):*** As a monotherapy, tofacitinib was not significantly different from other biologic response modifier monotherapies and combined therapy with respect to patients achieving ACR 20 response at 12 and 24 weeks. As a combination therapy, tofacitinib + MTX was generally not statistically significantly different from other biologic response modifiers + MTX in achieving ACR 20 response at 12 and 24 weeks. However, tofacitinib + MTX had statistically significant greater odds of ACR 20 response than anakinra + MTX and etanercept + MTX at 12 weeks, but statistically significant lower odds of ACR 20 response than certolizumab + MTX at 12 and 24 weeks.

A sensitivity analysis removing the COMBE and A3921046 studies (because of different background DMARDs) overall did not influence 12-week comparisons for ACR 20, but did influence the direction of the results for the comparisons between tofacitinib and adalimumab, and between tofacitinib + MTX and etanercept at 24 weeks. However, the differences between the drugs compared were still not statistically significant. Another sensitivity analysis removing the TEMPO study most often changed the direction, but not significance, of the effect in favour of tofacitinib monotherapy over biologic combination therapy, and in favour of biologic monotherapy over tofacitinib combination therapy at 12 weeks.

ACR 50 (12 and 24 Weeks) (See Table 45***):*** As a monotherapy and combination therapy, tofacitinib was generally not statistically significantly different from other biologic response modifier monotherapy and combined therapy in achieving ACR 50 response at 12 and 24 weeks. A sensitivity analysis removing the COMBE and A3921046 studies changed the direction, but not the significance, of the effect in favour of tofacitinib + MTX over tocilizumab + MTX at 12 weeks and over infliximab + MTX at 24 weeks. The sensitivity analysis removing the TEMPO study changed the direction, but not significance, of the effect in a number of comparisons at 12 and 24 weeks, and resulted in certolizumab + MTX having statistically significant greater odds of ACR 50 response at 24 weeks than tofacitinib + MTX.

ACR 70 (12 and 24 Weeks) (See Table 45***):*** As a monotherapy, tofacitinib was not statistically significantly different from other biologic monotherapy and combined therapy at 12 and 24 weeks. As a combination therapy, tofacitinib + MTX had statistically significantly lower odds of ACR 70 response than certolizumab + MTX at 12 weeks, and was not statistically significantly different from all other biologic monotherapy and combination therapy at 12 and 24 weeks. A sensitivity analysis removing the COMBE and A3921046 studies changed the direction of the effect in favour of tofacitinib in one case (tofacitinib + MTX compared with certolizumab + MTX) and in favour of the comparison biologic in another case (tofacitinib compared with adalimumab + MTX) at 12 weeks and resulted in statistically significant greater odds of ACR 70 response for tofacitinib + MTX compared with etanercept monotherapy at 24 weeks. The sensitivity analysis removing the TEMPO study changed the direction, but not significance, of the effect in a number of cases at 12 and 24 weeks, and resulted in statistically significantly greater odds of ACR 70 response at 24 weeks for tofacitinib + MTX compared with abatacept + MTX.

HAQ (12 and 24 Weeks) (See Table 46***):*** As a monotherapy, there were no statistically significant differences in HAQ scores between tofacitinib and other biologic response modifier monotherapies and combined therapies at 12 and 24 weeks. Tofacitinib + MTX had a statistically significantly greater difference in HAQ score from baseline to 12 weeks than anakinra, but no statistically significant

differences from all other biologic response modifier monotherapies and combination therapies at 12 and 24 weeks. With the removal of the COMBE and A3921046 studies in the sensitivity analysis, the direction of the effect between tofacitinib + MTX and anakinra + MTX changed in favour of anakinra at 12 weeks and had no effects on the 24-week results. The sensitivity analysis removing the TEMPO study had no effect on HAQ results at 12 weeks, but resulted in a statistically significantly greater difference in HAQ scores for tofacitinib + MTX than for anakinra + MTX , and changed the direction, but not significance, of the effect in favour of etanercept + MTX and etanercept monotherapy at 24 weeks.

Withdrawals and Adverse Events (See Table 47***):*** As a monotherapy, there were no statistically significant differences between tofacitinib and all other biologic response modifier monotherapies and combination therapies for the rates of withdrawals due to adverse events, adverse events, and infections. Tofacitinib in combination with MTX had statistically significant greater rates of withdrawals due to adverse events than etanercept + MTX, and a statistically significant greater rate of infection than adalimumab + MTX. There were no statistically significant differences in rates of adverse events between tofacitinib + MTX and all other biologic response modifiers with or without MTX. Comparisons between treatments for rates of serious adverse events were not possible due to a lack of model convergence.

Sensitivity Analysis: As presented in these results, the COMBE and A3921046 studies differed from the other included studies based on the concomitant DMARD provided with the active or placebo intervention. These two studies allowed DMARDs other than MTX, whereas all other studies included only MTX as their concomitant therapy. The TEMPO study included patients who had a previous failure on a DMARD other than MTX, which was systematically different than the other included trials. As a result, the NMAs were rerun to assess the robustness of the model. The authors also presented the results of both the fixed and random effects models.

The authors of the NMA also noted other variables that could have influenced the results, such as the variations in placebo response between trials (especially the particularly low placebo response for the certolizumab pegol trial) and the variations in the reassignment schemes between trials. The authors note that these differences may have led to an overestimation of the effects of certolizumab pegol and a difficulty in interpreting the results of the 24-week outcomes. Given the noted differences in trial designs, the authors undertook a meta-regression analysis to control for placebo discontinuation rates (in combination with the removal of the TEMPO study). The results of this analysis generally supported the findings of the unadjusted analysis for ACR 20 at 12 weeks (Table 48).[87] At 24 weeks, there were noted differences in the point estimates and statistical significance of ACR 20 response rates between the unadjusted and adjusted network meta-analysis (Table 48. In all but one case (tofacitinib + MTX versus etanercept monotherapy), tofacitinib with or without MTX was at least as effective as, or more effective than, comparator biologic therapies with or without MTX.

b) TNF-IR Population (MAPI 2011)

ACR 20 (12 Weeks) (See Table 49***):*** There were no statistically significant differences in ACR 20 response at 12 weeks between tofacitinib + DMARDs and abatacept + DMARDs, golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs.

ACR 50 (12 Weeks) (See Table 49***):*** There were no statistically significant differences in ACR 50 response at 12 weeks between tofacitinib + DMARDs and golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs. Data for abatacept +DMARDs were not available.

ACR 70 (12 Weeks) (See Table 49***):*** There were no statistically significant differences in ACR 70 response at 12 weeks between tofacitinib + DMARDs and golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs. Data for abatacept +DMARDs were not available.

HAQ (12 Weeks) (See Table 50***): T***ofacitinib + DMARDs therapy had a statistically significant greater difference in HAQ score from baseline to 12 weeks than golimumab + DMARDs. There was no statistically significant difference between tofacitinib + DMARD and rituximab + DMARDs. Data for abatacept + DMARDs and golimumab + DMARDs were not available.

Withdrawals and Adverse Events (See Table 51***):*** There were no statistically significant differences in rates of withdrawals due to adverse events, adverse events, serious adverse events, and infections between tofacitinib + DMARDs and abatacept + DMARDs, golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs.

Sensitivity Analysis: No sensitivity analyses were conducted for the TNF-IR population.

TNF-IR Population (MAPI 2013)

ACR 20 (12 and 24 Weeks) (See Table 52***):*** Results were similar to those reported from the MAPI 2011 TNF-IR population at 12 weeks. There were no statistically significant differences in ACR 20 response at 12 weeks between tofacitinib + DMARDs and abatacept + DMARDs, golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs. At 24 weeks, tofacitinib + DMARDs had statistically significant lower odds of ACR 20 response than tocilizumab + DMARDs, and no statistically significant differences between abatacept + DMARDs, golimumab + DMARDs, and rituximab + DMARDs.

ACR 50 (12 and 24 Weeks) (See Table 52***):***Results were similar to those reported from the MAPI 2011 TNF-IR population at 12 weeks. There were no statistically significant differences in ACR 50 response at 12 weeks between tofacitinib + DMARDs and golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs. Data for abatacept at 12 weeks were not available. At 24 weeks, there were no statistically significant differences in ACR 50 response between tofacitinib + DMARDs and golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs.

ACR 70 (12 and 24 Weeks) (See Table 52***):*** Results differed from those reported in the MAPI 2011 TNF-IR population at 12 weeks. Tofacitinib + DMARDs had a greater odds of ACR 70 response than rituximab + DMARDs. There were no statistically significant differences between tofacitinib + DMARDs and golimumab + DMARDs and tocilizumab + DMARDs. Data for abatacept at 12 weeks were not available. At 24 weeks, there were no statistically significant differences in ACR 70 response between tofacitinib + DMARDs and abatacept + DMARDs, golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs.

HAQ (12 and 24 Weeks) (See Table 53***):*** Results differed from those reported in the MAPI 2011 TNF-IR population at 12 weeks. There were no statistically significant differences in change in HAQ scores from baseline to 12 weeks between tofacitinib + DMARDs and golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs. Likewise, at 24 weeks, there were no statistically significant differences between tofacitinib + DMARDs and abatacept + DMARDs, golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs.

DAS28 (12 and 24 Weeks): Authors of the NMA suggest that DAS28 outcome measures not be reported because of missing data.

Withdrawals and Adverse Events (See Table 54***):*** Results were similar to those reported in the MAPI 2011 TNF-IR population. There were no statistically significant differences in rates of withdrawals due to adverse events, adverse events, and serious adverse events between tofacitinib + DMARDs and abatacept + DMARDs, golimumab + DMARDs, tocilizumab + DMARDs, and rituximab + DMARDs. Rates of serious infection were not available.

Sensitivity Analysis: Given the limited number of studies available for the NMA, authors noted that they were unable to determine the effects of the differences in the potential effect modifiers across studies. The authors also highlighted the limitations of the variations in reassignment schemes among studies. In particular, they note that the manufacturer-submitted study, A3921032, deviated substantially from the other studies in the rescue scheme imposed on patients. All patients in the placebo group of this study were reassigned to tofacitinib treatment after 12 weeks regardless of response. Efficacy and harms outcomes at 24 weeks were then assessed based on last observation carried forward (LOCF) in the placebo group and based on 24-week outcome data for the active-treatment group. Because of the limitations of this analysis, the NMA for the 24-week ACR 20/50/70 response was rerun using 12-week data carried forward to week 24 for both the placebo and active groups for trial A3921032. All results were consistent with the main findings, except for tofacitinib combination therapy compared with rituximab combination therapy, for which tofacitinib had statistically significant lower odds of ACR 20 response at 24 weeks. Another sensitivity analysis was conducted to account for the differences in populations used for the withdrawal and adverse events outcomes. In the A3921032 trial, results were available for withdrawal and adverse event outcomes based on both the safety population and the full analysis set (defined as those patients that receive both at least one treatment and at least one assessment). The main analysis was based on the safety population, and there were no differences in outcomes from those presented in a supplementary analysis with the full analysis set.

Critical Appraisal of Network Meta-analysis

The quality of the NMAs submitted by the manufacturer was assessed according to the checklist provided by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Task Force on Indirect Treatment Comparisons (Table 44).

Limitations

DMARD-IR and TNF-IR Population (MAPI 2011)

General limitations of the 2011 NMA include no a priori description of how potential biases and inconsistencies in trial methodology or patient characteristics would be handled, no presentation of the traditional pairwise meta-analyses results so that consistency between direct and indirect evidence could be assessed, and no primary efficacy or safety outcomes stated.

A key limitation of the 2011 NMA for the DMARD-IR population was the limited number of trials available for the meta-analyses. This is especially apparent for the network link between combination therapies and monotherapies. This link was based on a single trial that compared etanercept with etanercept plus MTX, placebo, and placebo plus MTX. Similarly, for the TNF-IR population, only one study was available for each of the active interventions compared with placebo.

A second key limitation was the heterogeneity between trials. For the DMARD-IR population, sensitivity analyses were conducted for three trials (COMBE, A3921046, and TEMPO) which differed from other trials in the concomitant DMARD used, and in previous DMARD failure being defined as failure on DMARDs other than MTX; however, there were other differences between trials that may have acted as effect modifiers. These variables include the dropout rates in the placebo groups, doses of concomitant

MTX, and the reassignment schemes imposed for patients after 12 weeks of treatment (i.e., early escape design). The variability in dropout rates between placebo groups, if higher in one study, may artificially inflate the results of the competing intervention, and the authors suggest this was a reason for the relatively greater outcomes for certolizumab pegol. Variations in dosing of concomitant MTX, if greater in one trial or intervention group relative to another, may enhance the effects of the intervention, a result that may be mistaken for an increase or decrease in efficacy and safety outcomes owing to the biologic of interest. Finally, the reassignment schemes were variable between studies. Twenty-one (55%) of the trials for the DMARD-IR population incorporated some form of reassignment based on a criterion of non-response at a certain point in time. Trials imposed reassignment at 12 weeks, 16 weeks, or (in one case) 22 weeks, and used either LOCF, BCOF, or non-responder imputation to obtain 24-week data. Given this variation, 24-week efficacy and harm outcomes become difficult to interpret. Twenty-four week outcomes were not reported for the TNF-IR population.

TNF-IR Population (MAPI 2013)

Key limitations, similar to those reported for the 2011 NMA and highlighted by the authors, are important to note. First, as cited in the TNF-IR population of the 2011 NMA, only one study was available for each of the active interventions compared with placebo. This raises uncertainty regarding the reliability of results from an NMA with a limited number of trials. As well, additional analysis, including sensitivity analysis and meta-regressions for heterogeneity between study methodology and baseline patient characteristics, was not possible. Second, because this NMA reported efficacy and safety outcomes at 24 weeks, the 24-week results have the same limitations as those reported for the DMARD-IR population from the 2011 NMA. The ATTAIN trial did not use reassignment, while the other trials did. The RADIATE, GO-AFTER, and REFLEX trials defined withdrawals as any patient who withdrew before reassignment, while the A3921032 trial identified withdrawals at week 12. All four trials therefore have potentially underestimated withdrawal and adverse event outcomes. Efficacy outcomes are similarly difficult to interpret.

Strengths

DMARD-IR and TNF-IR Population (MAPI 2011)

The rationale, objectives, methods, and results of the 2011 NMA were very well reported. The report carried out a systematic review of the literature with a search strategy that was clearly described. Study selection was performed independently by two reviewers, and individual trials were assessed for validity using the Jadad scale. There was detailed reporting of individual study characteristics, including the study methodology and patient population, and the reporting of the NMA methodology and corresponding network figure were well documented. Additionally, results of the fixed and random effects models were reported, and a sensitivity analysis that excluded studies that were believed to have substantial heterogeneity was conducted.

TNF-IR Population (MAPI 2013)

The rationale, objectives, methods, and results of the 2011 NMA were very well reported. The report carried out a systematic review of the literature with a search strategy that was clearly described. Study selection was performed independently by two reviewers, and individual trials were assessed for validity using the Jadad scale and the Centre for Reviews and Dissemination quality assessment of studies tool. There was detailed reporting of individual study characteristics, including the study methodology and patient population, and the reporting of the NMA methodology and corresponding network figure were well documented.

Summary

The results of the NMA for the DMARD-IR population were based on a strong network of 33 studies, consisting of a generally homogenous patient population and reasonably conducted trials. The results remained relatively stable following sensitivity analyses that involved removing studies that were deemed to be different from other studies (DMARD use other than MTX). The resulting lack of statistically significant differences in the efficacy and safety of tofacitinib compared with other biological response modifiers at week 12 seems reasonable to conclude, especially when comparing combination therapies. The results of the comparisons between monotherapies and combination therapies have greater limitations to interpretation because the link between these populations was based on a single study. Finally, given the differences in rescue therapy protocols imposed on patients between the studies, 24-week results should be interpreted more cautiously.

The results of the NMA for the TNF-IR population were based on a smaller network of five studies consisting of a generally homogeneous patient population, with a fairly high degree of methodological heterogeneity between studies. Because of the limited number of studies available in the network, sensitivity analyses were not conducted. As a result, the impact of the differences between studies on the efficacy and safety outcomes is unknown. This is especially concerning for 24-week efficacy and safety outcomes.

TABLE 44: CRITICAL APPRAISAL OF NMAS USING THE ISPOR CHECKLIST

ISPOR Checklist Item		Details and Comments (MAPI 2011)	Details and Comments (MAPI 2013)
1.	Are the rationale for the study and the objectives stated clearly?	• Rationale and objectives are clearly described.	• Rationale and objectives are clearly described.
2.	• Does the methods section include the following? • Eligibility criteria • Information sources • Search strategy • Study selection process • Data extraction • Validity of individual studies	• Databases searched are listed (MEDLINE and Embase). • Search strategy and search terms are described (full search detailed in the Appendix). • Study eligibility criteria are outlined according to the PICO model. • Study selection process is described and involved two reviewers. • The validity of the individual trials was assessed using the Jadad instrument.	• Databases searched are listed (MEDLINE and Embase). • Search strategy and search terms are described (full search detailed in the Appendix). • Study eligibility criteria are outlined according to the PICO model. • Study selection process is described and involved two reviewers. • The validity of the individual trials was assessed using the Jadad and CRD quality appraisal instruments.
3.	Are the outcome measures described?	• A list of relevant outcome measures is provided. Justification for these outcome measures is not provided. • Primary outcome not stated.	• A list of relevant outcome measures is provided. Justification for these outcome measures is not provided. • Primary outcome not stated.

ISPOR Checklist Item		Details and Comments (MAPI 2011)	Details and Comments (MAPI 2013)
4.	Is there a description of methods for analysis/synthesis of evidence? • Description of analyses methods/models • Handling of potential bias/inconsistency • Analysis framework	• No description of the pairwise synthesis of evidence was provided. • A full description of the methods for the NMAs was provided. • No a priori description of how potential biases and inconsistencies would be handled. • The validity of the studies was reported, but the results of the assessment were not considered when conducting the statistical analyses.	• Methods describe a planned analysis to assess the consistency between direct and indirect evidence. • Methods describe that potential biases and inconsistencies between studies would be handled using meta-regressions or subgroup analyses. • Methods describe how missing data would be handled. • A full description of the methods for the NMAs was provided. • Gelman-Rubin statistics were used to assess the convergence of the model.
5.	Are sensitivity analyses presented?	• Results of the fixed and random effects NMAs models are presented. • Separate results were presented for an analysis that included all trials and for analyses that excluded studies with clinical and methodological heterogeneity. • No sensitivity analysis was conducted using a different prior distribution to inform the model.	• Potential effect modifiers (differences in patient characteristics, concomitant treatments, duration of disease and baseline disease severity) were stated a priori. • No analysis was carried out to assess the effects of potential effect modifiers because of the limited number of studies available. Authors state that if there are < 10 studies, meta-regressions should not be carried out. • No sensitivity analysis was conducted using a different prior distribution to inform the model.
6.	Do the results include a summary of the studies included in the network of evidence? • Individual study data? • Network of studies?	• A summary of the individual studies included in the NMA was provided and included potential effect modifiers such as age, sex, and disease duration. • NMA figures outlining the network of studies according to outcome measure was provided.	• A summary of the individual studies included in the network meta-analysis was provided and included potential effect modifiers such as treatment type, inclusion and exclusion criteria, definition of TNF-IR, concomitant treatments, patient age, sex, disease duration, and disease severity at baseline. • NMA figures outlining the network of studies according to outcome measure were provided.
7.	Does the study describe an assessment of model fit? Are competing models being compared?	• Model fit was assessed using the deviance information criterion (DIC). A fixed or random effects model was selected based on the model with the lowest DIC. The results of both models are presented.	• The methods describe performing the analyses using a fixed and random effects approach and then assessing for model fit using the DIC. Only the fixed effects results are presented.

ISPOR Checklist Item		Details and Comments (MAPI 2011)	Details and Comments (MAPI 2013)
8.	Are the results of the evidence synthesis presented clearly?	• The results of pairwise meta-analyses were not presented. • The results of the evidence synthesis (NMA) were presented clearly.	• Only one study per group was available for the NMA, so pairwise meta-analyses were not necessary. • The results of the evidence synthesis (NMA) were presented clearly.

CRD = Centre for Research and Dissemination; DIC = deviance information criterion; IR = inadequate response; ISPOR = International Society for Pharmacoeconomics Research; NMA = network meta-analysis; PICO = population, intervention, comparator, outcome; TNF = tumour necrosis factor.

TABLE 45: DMARD-IR POPULATION (MAPI 2011): ACR 20/50/70 OUTCOME DATA FOR PATIENTS WITH INADEQUATE RESPONSE TO DMARDS (MTX OR OTHERS)

	ACR 20, OR (95% CrI)		ACR 50, OR (95% CrI)		ACR 70, OR (95% CrI)	
	12 Weeks	24 Weeks	12 Weeks	24 Weeks	12 Weeks	24 Weeks
Tofacitinib Monotherapy vs. Biologic Monotherapy						
Tofacitinib 5 mg vs. Adalimumab	1.40 (0.80 to 2.51)	1.04 (0.25 to 4.48)[a]	1.10 (0.40 to 2.91)	0.91 (0.18 to 4.92)[b]	1.52 (0.33 to 5.84)	0.26 (0.03 to 2.21)
Tofacitinib 5 mg vs. Certolizumab	0.60 (0.22 to 1.56)	0.46 (0.09 to 2.50)	0.35 (0.03 to 2.38)	0.63 (0.09 to 4.37)	0.65 (0.01 to 8.19)	0.32 (0.01 to 5.27)
Tofacitinib 5 mg vs. Etanercept	1.04 (0.41 to 2.64)	0.66 (0.15 to 2.84)[c]	0.72 (0.15 to 3.50)	0.77 (0.14 to 4.07)	1.13 (0.13 to 8.87)[b]	0.52 (0.06 to 4.13)
Tofacitinib 5 mg vs. Rituximab	NA	0.77 (0.12 to 5.02)[b]	NA	0.78 (0.10 to 6.14)[b]	NA	0.55 (0.04 to 7.09)[b]
Tofacitinib Combination Therapy vs. Biologic Combination Therapy						
Tofacitinib 5 mg + MTX vs. Adalimumab + MTX	0.88 (0.60 to 1.29)	0.83 (0.44 to 1.52)	1.17 (0.54 to 2.48)	0.91 (0.45 to 1.74)	1.34 (0.59 to 3.28)	1.61 (0.71 to 3.42)
Tofacitinib 5 mg + MTX vs. Certolizumab + MTX	0.39 (0.23 to 0.64)	0.24 (0.10 to 0.58)	0.54 (0.91 to 1.54)	0.39 (0.14 to 1.02)[c]	0.23 (0.03 to 0.98)	0.66 (0.17 to 2.29)
Tofacitinib 5 mg + MTX vs. Etanercept + MTX	1.57 (0.92 to 2.49)	0.71 (0.31 to 1.52)[c]	1.11 (0.37 to 2.64)[b]	1.01 (0.39 to 2.26)[b]	1.11 (0.32 to 3.20)[b]	1.93 (0.55 to 5.10)[b]
Tofacitinib 5 mg + MTX vs. Golimumab + MTX	1.47 (0.79 to 2.76)	0.81 (0.31 to 2.13)	1.07 (0.34 to 3.33)	0.79 (0.27 to 2.21)	0.77 (0.16 to 3.20)	1.88 (0.49 to 6.79)
Tofacitinib 5 mg + MTX vs. Infliximab + MTX	1.17 (0.64 to 2.11)	0.84 (0.39 to 1.78)	0.80 (0.18 to 3.59)	0.99 (0.42 to 2.22)[a,b]	0.86 (0.15 to 4.27)[b]	1.85 (0.62 to 4.92)
Tofacitinib 5 mg + MTX vs. Tocilizumab + MTX	0.81 (0.52 to 1.26)	0.61 (0.26 to 1.42)	0.94 (0.40 to 2.31)[a]	0.60 (0.24 to 1.48)	1.15 (0.43 to 3.25)	0.76 (0.24 to 2.29)
Tofacitinib 5 mg + MTX vs. Abatacept + MTX	1.29 (0.81 to 2.05)	0.84 (0.40 to 1.78)	1.17 (0.43 to 3.26)	1.10 (0.49 to 2.45)	1.03 (0.27 to 3.50)	2.05 (0.71 to 5.26)[c]
Tofacitinib 5 mg + MTX vs. Anakinra + MTX	2.14 (1.16 to 3.94)	1.23 (0.40 to 3.67)	NA	1.58 (0.47 to 5.27)	NA	2.48 (0.47 to 11.65)
Tofacitinib 5 mg + MTX vs. Rituximab + MTX	NA	0.77 (0.34 to 1.70)	NA	1.00 (0.41 to 2.37)	NA	2.34 (0.76 to 6.73)
Tofacitinib Combination Therapy vs. Biologic Monotherapy						
Tofacitinib 5 mg + MTX vs. Adalimumab	2.50 (0.86 to 6.99) [b]	1.28 (0.49 to 3.32)	2.06 (0.29 to 12.04)[b]	1.60 (0.55 to 4.56)	1.97 (0.18 to 19.99)	1.61 (0.46 to 5.59)

	ACR 20, OR (95% CrI)		ACR 50, OR (95% CrI)		ACR 70, OR (95% CrI)	
	12 Weeks	24 Weeks	12 Weeks	24 Weeks	12 Weeks	24 Weeks
Tofacitinib 5 mg + MTX vs. Certolizumab	1.06 (0.28 to 3.71)[b,c]	0.57 (0.13 to 2.42)	0.65 (0.04 to 6.71)	1.11 (0.19 to 6.02)[b]	0.84 (0.01 to 17.07)[a]	1.96 (0.06 to 27.77)
Tofacitinib 5 mg + MTX vs. Etanercept	1.87 (1.07 to 3.01)[b]	0.82 (0.35 to 1.75)[a,c]	1.36 (0.46 to 3.32)[b]	1.37 (0.54 to 3.08)[b]	1.50 (0.41 to 4.35)[b]	3.25 (0.91 to 8.76)[b,d]
Tofacitinib 5 mg + MTX vs. Rituximab	NA	0.96 (0.27 to 3.37)[b]	NA	1.36 (0.35 to 5.25)	NA	3.36 (0.62 to 17.84)
Tofacitinib Monotherapy vs. Biologic Combination Therapy						
Tofacitinib 5 mg vs. Adalimumab + MTX	0.49 (0.17 to 1.51)[b]	0.67 (0.16 to 2.91)	0.62 (0.10 to 4.49)[b]	0.51 (0.10 to 2.80)	1.03 (0.10 to 11.77)[a]	0.26 (0.03 to 2.14)
Tofacitinib 5 mg vs. Certolizumab + MTX	0.22 (0.07 to 0.68)[b]	0.19 (0.04 to 0.99)	0.29 (0.04 to 2.28)	0.22 (0.03 to 1.41)	0.17 (0.01 to 2.44)	0.11 (0.01 to 1.10)
Tofacitinib 5 mg vs. Etanercept + MTX	0.87 (0.32 to 2.36)[b]	0.57 (0.13 to 2.57)	0.59 (0.11 to 3.18)	0.57 (0.10 to 3.08)	0.83 (0.09 to 7.37)	0.31 (0.03 to 2.53)
Tofacitinib 5 mg vs. Golimumab + MTX	0.82 (0.26 to 2.76)[b]	0.65 (0.13 to 3.54)	0.57 (0.08 to 4.74)[b]	0.45 (0.07 to 2.99)	0.58 (0.04 to 7.73)	0.30 (0.03 to 3.34)
Tofacitinib 5 mg vs. Infliximab + MTX	0.65 (0.21 to 2.14)[b]	0.68 (0.15 to 3.24)	0.43 (0.05 to 4.50)	0.56 (0.10 to 3.34)	0.64 (0.04 to 11.37)[b]	0.30 (0.03 to 2.70)
Tofacitinib 5 mg vs. Tocilizumab + MTX	0.45 (0.16 to 1.37)[b]	0.49 (0.10 to 2.48)	0.50 (0.08 to 3.75)[b]	0.34 (0.06 to 2.16)	0.86 (0.09 to 10.68)[b]	0.12 (0.01 to 1.19)
Tofacitinib 5 mg vs. Abatacept + MTX	0.71 (0.25 to 2.21)[b]	0.68 (0.15 to 3.28)	0.62 (0.10 to 4.91)[b]	0.63 (0.11 to 3.69)	0.77 (0.07 to 10.93)[b]	0.33 (0.04 to 3.04)
Tofacitinib 5 mg vs. Anakinra + MTX	1.19 (0.38 to 3.99)	0.99 (0.18 to 5.77)[b]	NA	0.90 (0.13 to 6.78)[b]	NA	0.40 (0.03 to 5.17)
Tofacitinib 5 mg vs. Rituximab + MTX	NA	0.62 (0.13 to 3.03)	NA	0.57 (0.10 to 3.46)	NA	0.38 (0.04 to 3.62)

ACR = American College of Rheumatology; CrI = credible interval; CSR = Clinical Study Report; DMARDs = disease-modifying antirheumatic drugs; IR = inadequate response; MTX = methotrexate; NA = not available; NMA = network meta-analysis; OR = odds ratio.

[a] Change in direction (with removal of COMBE/CSR1046).

[b] Change in direction (with removal of TEMPO Study).

[c] Change in significance (with removal of TEMPO Study).

[d] Change in significance (with removal of COMBE/CSR1046).

Note: Random effects NMA.

TABLE 46: DMARD-IR POPULATION (MAPI 2011): HAQ OUTCOME DATA FOR PATIENTS WITH INADEQUATE RESPONSE TO DMARDS (MTX OR OTHERS)

	HAQ 12 Weeks Difference (95% CrI)	HAQ 24 Weeks Difference (95% CrI)
Tofacitinib monotherapy vs. biologic monotherapy		
Tofacitinib 5 mg vs. adalimumab	–0.02 (–0.19 to 0.14)	0.13 (–0.17 to 0.44)
Tofacitinib 5 mg vs. certolizumab	0.16 (–0.17 to 0.47)	0.36 (–0.08 to 0.81)
Tofacitinib 5 mg vs. etanercept	0.07 (–0.52 to 0.67)	0.47 (–0.43 to 1.34)
Tofacitinib 5 mg vs. rituximab	NA	0.70 (–0.23 to 1.58)
Tofacitinib combination therapy vs. biologic combination therapy		
Tofacitinib 5 mg + MTX vs. adalimumab + MTX	–0.03 (–0.16 to 0.08)	–0.03 (–0.16 to 0.11)
Tofacitinib 5 mg + MTX vs. certolizumab + MTX	0.11 (–0.15 to 0.36)	0.08 (–0.10 to 0.26)
Tofacitinib 5 mg + MTX vs. etanercept + MTX	–0.06 (–0.19 to 0.08)[a]	–0.03 (–0.18 to 0.14)[b]
Tofacitinib 5 mg + MTX vs. golimumab + MTX	NA	–0.05 (–0.28 to 0.18)
Tofacitinib 5 mg + MTX vs. infliximab + MTX	–0.06 (–0.26 to 0.13)	–0.18 (–0.41 to 0.05)
Tofacitinib 5 mg + MTX vs. tocilizumab + MTX	0.04 (–0.19 to 0.28)	–0.03 (–0.22 to 0.15)
Tofacitinib 5 mg + MTX vs. abatacept + MTX	–0.15 (–0.34 to 0.04)	–0.03 (–0.34 to 0.30)
Tofacitinib 5 mg + MTX vs. anakinra + MTX	–0.20 (–0.36 to –0.04)[a]	–0.19 (–0.39 to 0.01)[c]
Tofacitinib 5 mg + MTX vs. rituximab + MTX	–0.04 (–0.26 to 0.16)	–0.13 (–0.31 to 0.05)
Tofacitinib combination therapy vs. biologic monotherapy		
Tofacitinib 5 mg + MTX vs. adalimumab	–0.23 (–0.85 to 0.37)	–0.48 (–1.32 to 0.40)
Tofacitinib 5 mg + MTX vs. certolizumab	–0.05 (–0.73 to 0.61)	–0.26 (–1.12 to 0.68)
Tofacitinib 5 mg + MTX vs. etanercept	–0.14 (–0.29 to 0.02)	–0.15 (–0.31 to 0.05)[b]
Tofacitinib 5 mg + MTX vs. rituximab	NA	0.09 (–0.13 to 0.30)
Tofacitinib monotherapy vs. biologic combination therapy		
Tofacitinib 5 mg vs. adalimumab + MTX	0.17 (–0.44 to 0.79)	0.59 (–0.33 to 1.46)
Tofacitinib 5 mg vs. certolizumab + MTX	0.31 (–0.34 to 0.96)	0.69 (–0.23 to 1.57)
Tofacitinib 5 mg vs. etanercept + MTX	0.14 (–0.45 to 0.76)	0.59 (–0.31 to 1.46)
Tofacitinib 5 mg vs. golimumab + MTX	NA	0.56 (–0.37 to 1.45)
Tofacitinib 5 mg vs. infliximab + MTX	0.14 (–0.48 to 0.78)	0.44 (–0.50 to 1.32)
Tofacitinib 5 mg vs. tocilizumab + MTX	0.25 (–0.40 to 0.90)	0.58 (–0.35 to 1.46)
Tofacitinib 5 mg vs. abatacept + MTX	0.06 (–0.56 to 0.69)	0.59 (–0.39 to 1.50)
Tofacitinib 5 mg vs. anakinra + MTX	0.01 (–0.31 to 0.64)	0.42 (–0.50 to 1.31)
Tofacitinib 5 mg vs. rituximab + MTX	0.17 (–0.46 to 0.80)	0.49 (–0.44 to 1.36)

CrI = credible interval; DMARDs = disease-modifying antirheumatic drugs; HAQ = Health Assessment Questionnaire; IR = inadequate response; MTX = methotrexate; NA = not available; NMA = network meta-analysis; vs.=versus.

[a] Change in direction (with removal of COMBE/Clinical Study Report 1046).

[b] Change in direction (with removal of TEMPO Study).

[c] Change in significance (with removal of TEMPO Study).

Note: Random effects NMA.

TABLE 47: DMARD-IR POPULATION (MAPI 2011): WITHDRAWALS AND ADVERSE EVENTS OUTCOME DATA FOR PATIENTS WITH INADEQUATE RESPONSE TO DMARDS (MTX OR OTHERS)

	Withdrawals Due to Adverse Events RR (95% CrI)	Adverse Events RR (95% CrI)	Infections RR (95% CrI)
Tofacitinib monotherapy vs. biologic monotherapy			
Tofacitinib 5 mg vs. adalimumab	0.07 (0.01 to 0.26)	1.05 (0.76 to 1.42)	5.61 (0.87 to 241.80)
Tofacitinib 5 mg vs. certolizumab	0.13 (0.01 to 1.28)	0.84 (0.54 to 1.30)	NA
Tofacitinib 5 mg vs. etanercept	0.50 (0.05 to 4.78)	0.95 (0.61 to 1.50)	0.47 (0.17 to 1.33)
Tofacitinib 5 mg vs. rituximab	0.09 (0.00 to 2.44)	0.99 (0.56 to 1.68)	NA
Tofacitinib combination therapy vs. biologic combination therapy			
Tofacitinib 5 mg + MTX vs. adalimumab + MTX	1.05 (0.60 to 1.93)	0.94 (0.72 to 1.21)	0.94 (0.58 to 1.54)
Tofacitinib 5 mg + MTX vs. certolizumab + MTX	0.53 (0.15 to 1.61)	0.93 (0.64 to 1.30)	0.91 (0.45 to 1.93)
Tofacitinib 5 mg + MTX vs. etanercept + MTX	2.26 (1.03 to 5.01)	0.95 (0.71 to 1.26)	1.23 (0.76 to 2.06)
Tofacitinib 5 mg + MTX vs. golimumab + MTX	2.44 (0.61 to 11.41)	0.88 (0.62 to 1.22)	1.20 (0.63 to 2.33)
Tofacitinib 5 mg + MTX vs. infliximab + MTX	0.82 (0.35 to 1.87)	0.97 (0.76 to 1.23)	1.00 (0.54 to 2.00)
Tofacitinib 5 mg + MTX vs. tocilizumab + MTX	0.85 (0.37 to 1.84)	0.86 (0.68 to 1.08)	1.09 (0.66 to 1.85)
Tofacitinib 5 mg + MTX vs. abatacept + MTX	1.61 (0.70 to 3.65)	0.99 (0.78 to 1.25)	0.69 (0.19 to 2.05)
Tofacitinib 5 mg + MTX vs. anakinra + MTX	2.63 (0.34 to 25.61)	0.89 (0.68 to 1.18)	1.00 (0.52 to 1.93)
Tofacitinib 5 mg + MTX vs. rituximab + MTX	0.45 (0.10 to 1.59)	0.95 (0.69 to 1.25)	1.28 (0.76 to 2.22)
Tofacitinib combination therapy vs. biologic monotherapy			
Tofacitinib 5 mg + MTX vs. adalimumab	0.26 (0.03 to 2.56)	1.04 (0.70 to 1.58)	15.39 (1.67 to 763.80)
Tofacitinib 5 mg + MTX vs. certolizumab	0.55 (0.04 to 7.82)	0.88 (0.50 to 1.41)	NA
Tofacitinib 5 mg + MTX vs. etanercept	2.05 (0.92 to 4.46)	0.96 (0.71 to 1.27)	1.26 (0.75 to 2.11)
Tofacitinib 5 mg + MTX vs. rituximab	0.33 (0.02 to 4.03)	1.01 (0.61 to 1.65)	NA
Tofacitinib monotherapy vs. biologic combination therapy			
Tofacitinib 5 mg vs. adalimumab + MTX	0.25 (0.02 to 3.00)	0.92 (0.58 to 1.52)	0.35 (0.11 to 1.18)
Tofacitinib 5 mg vs. certolizumab + MTX	0.13 (0.01 to 1.69)	0.91 (0.54 to 1.47)	0.34 (0.10 to 1.25)
Tofacitinib 5 mg vs. etanercept + MTX	0.55 (0.06 to 5.74)	0.94 (0.62 to 1.49)	0.46 (0.16 to 1.40)
Tofacitinib 5 mg vs. golimumab + MTX	0.62 (0.03 to 8.46)	0.87 (0.54 to 1.39)	0.44 (0.14 to 1.53)
Tofacitinib 5 mg vs. infliximab + MTX	0.20 (0.02 to 2.20)	0.95 (0.64 to 1.49)	0.38 (0.12 to 1.34)
Tofacitinib 5 mg vs. tocilizumab + MTX	0.21 (0.02 to 2.26)	0.84 (0.57 to 1.29)	0.41 (0.13 to 1.34)
Tofacitinib 5 mg vs. abatacept + MTX	0.39 (0.04 to 4.50)	0.96 (0.64 to 1.50)	0.25 (0.05 to 1.20)
Tofacitinib 5 mg vs. anakinra + MTX	0.65 (0.03 to 16.97)	0.88 (0.56 to 1.40)	0.37 (0.11 to 1.30)
Tofacitinib 5 mg vs. rituximab + MTX	0.11 (0.01 to 1.43)	0.93 (0.66 to 1.33)	0.48 (0.15 to 1.57)

CrI = credible interval; DMARDs = disease-modifying antirheumatic drugs; IR = inadequate response; MTX = methotrexate; NA = not available; NMA = network meta-analysis; RR = rate ratio; vs. = versus.
Note: Random effects NMA.

TABLE 48: DMARD-IR POPULATION (MAPI 2011): UNADJUSTED AND ADJUSTED NMA MODELS FOR ACR 20 AT 12 AND 24 WEEKS

	ACR 20 — OR (95% CrI) 12 Weeks			ACR 20 — OR (95% CrI) 24 Weeks		
	Unadjusted Model	Adjusted Model[a]	Adjusted Model[b]	Unadjusted Model	Adjusted Model[a]	Adjusted Model[b]
Tofacitinib monotherapy vs. biologic monotherapy						
Tofacitinib 5 mg vs. adalimumab	1.40 (0.80 to 2.51)	1.42 (0.83 to 2.42)	1.56 (0.93 to 2.90)	1.04 (0.25 to 4.48)	1.14 (0.45 to 2.97)	**1.80 (1.51 to 2.49)[c]**
Tofacitinib 5 mg vs. certolizumab	0.60 (0.22 to 1.56)	0.61 (0.24 to 1.49)	1.23 (0.50 to 3.45)	0.46 (0.09 to 2.50)	0.47 (0.15 to 1.40)	**1.61 (1.60 to 2.56)[c]**
Tofacitinib 5 mg vs. etanercept	1.04 (0.41 to 2.64)	1.02 (0.42 to 2.44)	0.93 (0.42 to 2.23)	0.66 (0.15 to 2.84)	**0.35 (0.13 to 0.95)[c]**	**1.56 (1.20 to 1.56)[c]**
Tofacitinib 5 mg vs. rituximab	NA	NA	NA	0.77 (0.12 to 5.02)	1.02 (0.29 to 3.62)	NA
Tofacitinib combination therapy vs. biologic combination therapy						
Tofacitinib 5 mg + MTX vs. adalimumab + MTX	0.88 (0.60 to 1.29)	0.88 (0.63 to 1.25)	0.97 (0.72 to 1.34)	0.83 (0.44 to 1.52)	0.92 (0.70 to 1.19)	1.30 (0.96 to 1.62)
Tofacitinib 5 mg + MTX vs. certolizumab + MTX	**0.39 (0.23 to 0.64)**	**0.39 (0.24 to 0.62)**	0.55 (0.25 to 1.31)[c]	**0.24 (0.10 to 0.58)**	**0.26 (0.16 to 0.40)**	0.81 (0.74 to 1.08)[c]
Tofacitinib 5 mg + MTX vs. etanercept + MTX	1.57 (0.92 to 2.49)	1.08 (0.56 to 2.03)	0.94 (0.56 to 1.55)	0.71 (0.31 to 1.52)	**0.38 (0.21 to 0.66)[c]**	1.01 (0.80 to 1.20)
Tofacitinib 5 mg + MTX vs. golimumab + MTX	1.47 (0.79 to 2.76)	1.48 (0.83 to 2.68)	1.14 (0.72 to 1.81)	0.81 (0.31 to 2.13)	0.79 (0.46 to 1.35)	**1.22 (1.10 to 1.84)[c]**
Tofacitinib 5 mg + MTX vs. infliximab + MTX	1.17 (0.64 to 2.11)	1.18 (0.68 to 2.05)	1.21 (0.73 to 1.96)	0.84 (0.39 to 1.78)	0.83 (0.60 to 1.15)	**1.54 (1.10 to 1.54)[c]**
Tofacitinib 5 mg + MTX vs. tocilizumab + MTX	0.81 (0.52 to 1.26)	0.82 (0.55 to 1.21)	0.85 (0.62 to 1.09)	0.61 (0.26 to 1.42)	**0.61 (0.44 to 0.83)[c]**	1.16 (0.88 to 1.23)
Tofacitinib 5 mg + MTX vs. abatacept + MTX	1.29 (0.81 to 2.05)	1.30 (0.83 to 1.98)	1.00 (0.70 to 1.49)	0.84 (0.40 to 1.78)	0.84 (0.61 to 1.17)	1.22 (0.96 to 1.25)
Tofacitinib 5 mg + MTX vs. anakinra + MTX	**2.14 (1.16 to 3.94)**	**2.14 (1.24 to 3.68)**	**2.51 (1.67 to 3.72)**	1.23 (0.40 to 3.67)	1.26 (0.80 to 1.98)	**2.84 (2.47 to 3.62)[c]**
Tofacitinib 5 mg + MTX vs. rituximab + MTX	NA	NA	NA	0.77 (0.34 to 1.70)	0.81 (0.55 to 1.21)	NA

	ACR 20 — OR (95% CrI) 12 Weeks			ACR 20 — OR (95% CrI) 24 Weeks		
	Unadjusted Model	Adjusted Model[a]	Adjusted Model[b]	Unadjusted Model	Adjusted Model[a]	Adjusted Model[b]
Tofacitinib combination therapy vs. biologic monotherapy						
Tofacitinib 5 mg + MTX vs. adalimumab	2.50 (0.86 to 6.99)	0.94 (0.50 to 1.73)	1.51 (0.62 to 3.40)	1.28 (0.49 to 3.32)	1.12 (0.66 to 1.87)	1.05 (0.83 to 1.42)
Tofacitinib 5 mg + MTX vs. certolizumab	1.06 (0.28 to 3.71)	0.40 (0.16 to 0.97)	1.17 (0.43 to 3.15)	0.57 (0.13 to 2.42)	0.46 (0.18 to 1.10)	0.93 (0.93 to 1.36)
Tofacitinib 5 mg + MTX vs. etanercept	**1.87 (1.07 to 3.01)**	0.84 (0.48 to 1.46)[c]	0.89 (0.51 to 1.60)[c]	0.82 (0.35 to 1.75)	**0.34 (0.19 to 0.59)[c]**	**0.88 (0.67 to 0.91)[c]**
Tofacitinib 5 mg + MTX vs. rituximab	NA	NA	NA	0.96 (0.27 to 3.37)	1.00 (0.42 to 2.30)	NA
Tofacitinib monotherapy vs. biologic combination therapy						
Tofacitinib 5 mg vs. adalimumab + MTX	0.49 (0.17 to 1.51)	0.91 (0.31 to 2.79)	1.02 (0.40 to 2.83)	0.67 (0.16 to 2.91)	0.94 (0.36 to 2.48)	**2.22 (1.71 to 2.84)[c]**
Tofacitinib 5 mg vs. certolizumab + MTX	**0.22 (0.07 to 0.68)**	0.40 (0.13 to 1.26)[c]	0.57 (0.14 to 2.63)[c]	**0.19 (0.04 to 0.99)**	**0.26 (0.10 to 0.75)**	**1.39 (1.35 to 1.93)[c]**
Tofacitinib 5 mg vs. etanercept + MTX	0.87 (0.32 to 2.36)	1.11 (0.42 to 2.84)	0.98 (0.41 to 2.50)	0.57 (0.13 to 2.57)	0.39 (0.14 to 1.09)	**1.73 (1.61 to 2.12)[c]**
Tofacitinib 5 mg vs. golimumab + MTX	0.82 (0.26 to 2.76)	1.51 (0.47 to 5.07)	1.20 (0.46 to 3.31)	0.65 (0.13 to 3.54)	0.81 (0.28 to 2.39)	**2.10 (1.89 to 3.32)[c]**
Tofacitinib 5 mg vs. infliximab + MTX	0.65 (0.21 to 2.14)	1.21 (0.39 to 3.92)	1.25 (0.45 to 3.96)	0.68 (0.15 to 3.24)	0.85 (0.32 to 2.31)	**2.64 (1.97 to 2.64)[c]**
Tofacitinib 5 mg vs. tocilizumab + MTX	0.45 (0.16 to 1.37)	0.84 (0.29 to 2.58)	0.88 (0.37 to 2.24)	0.49 (0.10 to 2.48)	0.62 (0.24 to 1.68)	**2.10 (1.51 to 2.11)[c]**
Tofacitinib 5 mg vs. abatacept + MTX	0.71 (0.25 to 2.21)	1.32 (0.46 to 4.08)	1.03 (0.40 to 3.19)	0.68 (0.15 to 3.28)	0.86 (0.33 to 2.33)	**2.10 (1.76 to 2.32)[c]**
Tofacitinib 5 mg vs. anakinra + MTX	1.19 (0.38 to 3.99)	2.19 (0.71 to 7.11)	**2.63 (1.03 to 7.21)[c]**	0.99 (0.18 to 5.77)	1.28 (0.46 to 3.65)	**4.86 (4.49 to 6.58)[c]**
Tofacitinib 5 mg vs. rituximab + MTX	NA	NA	NA	0.62 (0.13 to 3.03)	0.83 (0.31 to 2.32)	NA

ACR = American College of Rheumatology; CrI = credible interval; MTX = methotrexate; NA = not available; NMA = network meta-analysis; OR = odds ratio; vs. = versus.

[a] With removal of the TEMPO study.

[b] With removal of TEMPO and adjusted for placebo discontinuation.

[c] Change in significance relative to the unadjusted model.

Notes: **Bold** = statically significant. Random effects NMA.

Source: MAPI report 2011.[85,87]

TABLE 49: TNF-IR POPULATION (MAPI 2011): ACR 20/50/70 12 WEEK OUTCOME DATA FOR PATIENTS WITH INADEQUATE RESPONSE TO TNF INHIBITOR

	ACR 20 OR (95% CrI)	ACR 50 OR (95% CrI)	ACR 70 OR (95% CrI)
Tofacitinib combination therapy vs. biologic combination therapy			
Tofacitinib 5 mg + DMARDs vs. abatacept + DMARDs	0.57 (0.27 to 1.20)	NA	NA
Tofacitinib 5 mg + DMARDs vs. golimumab + DMARDs	1.08 (0.51 to 2.27)	1.41 (0.48 to 4.06)	1.82 (0.24 to 20.24)
Tofacitinib 5 mg + DMARDs vs. tocilizumab + DMARDs	0.56 (0.27 to 1.15)	0.61 (0.19 to 1.79)	0.55 (0.02 to 8.87)
Tofacitinib 5 mg + DMARDs vs. rituximab + DMARDs	0.57 (0.29 to 1.11)	0.91 (0.34 to 2.45)	5.00 (0.95 to 47.38)

ACR = American College of Rheumatology; CrI = credible limits; DMARDs = disease-modifying antirheumatic drugs; IR = inadequate response; NA = not available; NMA = network meta-analysis; OR = odds ratio; TNF = tumour necrosis factor; vs. = versus.
Note: Fixed effects NMA.

TABLE 50: TNF-IR POPULATION (MAPI 2011): HAQ 12-WEEK OUTCOME DATA FOR PATIENTS WITH INADEQUATE RESPONSE TO TNF INHIBITOR

	HAQ Difference (95% CrI)
Tofacitinib combination therapy vs. biologic combination therapy	
Tofacitinib 5 mg + DMARDs vs. abatacept + DMARDs	NA
Tofacitinib 5 mg + DMARDs vs. golimumab + DMARDs	–0.26 (–0.49 to –0.03)
Tofacitinib 5 mg + DMARDs vs. tocilizumab + DMARDs	NA
Tofacitinib 5 mg + DMARDs vs. rituximab + DMARDs	–0.04 (–0.29 to 0.21)

CrI = credible interval; DMARDs = disease-modifying antirheumatic drugs; HAQ = Health Assessment Questionnaire; IR = inadequate response; NA = not available; NMA = network meta-analysis; TNF = tumour necrosis factor.
Note: Fixed effects NMA.

TABLE 51: TNF-IR POPULATION (MAPI 2011): WITHDRAWALS AND ADVERSE EVENTS OUTCOME DATA FOR PATIENTS WITH INADEQUATE RESPONSE TO TNF INHIBITOR

	Withdrawals Due to Adverse Events RR (95% CrI)	Adverse Events RR (95% CrI)	Serious Adverse Events RR (95% CrI)	Infections RR (95% CrI)
Tofacitinib combination therapy vs. biologic combination therapy				
Tofacitinib 5 mg + DMARDs vs. abatacept + DMARDs	1.61 (0.37 to 7.11)	0.84 (0.56 to 1.27)	0.32 (0.04 to 1.76)	0.82 (0.42 to 1.63)
Tofacitinib 5 mg + DMARDs vs. golimumab + DMARDs	4.04 (0.97 to 20.71)	1.03 (0.68 to 1.56)	0.39 (0.04 to 2.31)	0.86 (0.44 to 1.72)
Tofacitinib 5 mg + DMARDs vs. tocilizumab + DMARDs	1.52 (0.44 to 5.57)	0.89 (0.60 to 1.34)	0.53 (0.06 to 3.09)	0.76 (0.40 to 1.45)
Tofacitinib 5 mg + DMARDs vs. rituximab + DMARDs	0.49 (0.06 to 2.83)	0.97 (0.66 to 1.42)	0.40 (0.05 to 2.11)	0.84 (0.45 to 1.58)

CrI = credible interval; DMARDs = disease-modifying antirheumatic drugs; IR = inadequate response; NMA = network meta-analysis; RR = rate ratio; TNF = tumour necrosis factor.
Note: Fixed effects NMA.

TABLE 52: TNF-IR POPULATION (MAPI 2013): ACR 20/50/70 OUTCOME DATA FOR PATIENTS WITH INADEQUATE RESPONSE TO A TNF INHIBITOR

	ACR 20 OR (95% CrI)		ACR 50 OR (95% CrI)		ACR 70 OR (95% CrI)	
	12 Weeks	24 Weeks	12 Weeks	24 Weeks	12 Weeks	24 Weeks
Tofacitinib combination therapy vs. biologic combination therapy						
Tofacitinib 5 mg + DMARDs vs. abatacept + DMARDs	0.56 (0.27 to 1.17)	0.78 (0.38 to 1.60)	NA	0.97 (0.27 to 3.06)	NA	1.63 (0.15 to 17.15)
Tofacitinib 5 mg + DMARDs vs. golimumab + DMARDs	0.89 (0.42 to 1.89)	1.29 (0.60 to 2.74)	1.39 (0.47 to 4.09)	1.57 (0.50 to 4.68)	1.89 (0.25 to 17.45)	3.46 (0.59 to 31.29)
Tofacitinib 5 mg + DMARDs vs. yocilizumab + DMARDs	0.52 (0.25 to 1.09)	0.36 (0.16 to 0.80)	0.43 (0.12 to 1.39)	0.61 (0.18 to 1.87)	0.20 (0.00 to 4.64)	1.17 (0.12 to 12.31)
Tofacitinib 5 mg + DMARDs vs. rituximab + DMARDs	0.59 (0.30 to 1.16)	0.69 (0.35 to 1.37)	0.96 (0.37 to 2.55)	0.92 (0.33 to 2.50)	6.06 (1.19 to 47.50)	0.87 (0.08 to 9.79)

CrI = credible interval; DMARDs = disease-modifying antirheumatic drugs; IR = inadequate response; NA = not available; NMA = network meta-analysis; OR = odds ratio; TNF = tumour necrosis factor.
Note: Fixed effects NMA.

TABLE 53: TNF-IR POPULATION (MAPI 2013): HAQ OUTCOME DATA FOR PATIENTS WITH INADEQUATE RESPONSE TO A TNF INHIBITOR

	HAQ 12 Weeks Difference (95% CrI)	HAQ 24 Weeks Difference (95% CrI)
Tofacitinib Combination Therapy vs. Biologic Combination Therapy		
Tofacitinib 5 mg + DMARDs vs. abatacept + DMARDs	NA	0.01 (–0.15 to 0.17)
Tofacitinib 5 mg + DMARDs vs. golimumab + DMARDs	–0.04 (–0.19 to 0.11)	–0.12 (–0.28 to 0.04)
Tofacitinib 5 mg + DMARDs vs. tocilizumab + DMARDs	0.10 (–0.05 to 0.25)	0.01 (–0.16 to 0.18)
Tofacitinib 5 mg + DMARDs vs. rituximab + DMARDs	–0.01 (–0.15 to 0.13)	0.04 (–0.11 to 0.19)

CrI = credible interval; DMARDs = disease-modifying antirheumatic drugs; HAQ = Health Assessment Questionnaire; IR = inadequate response; NA = not available; NMA = network meta-analysis; TNF = tumour necrosis factor
Note: Fixed effects NMA.

TABLE 54: TNF-IR POPULATION (MAPI 2013): WITHDRAWALS AND ADVERSE EVENTS OUTCOME DATA FOR PATIENTS WITH INADEQUATE RESPONSE TO A TNF INHIBITOR

	Withdrawals Due to Adverse Events RR (95% CrI)	Adverse Events RR (95% CrI)	Serious Adverse Events RR (95% CrI)	Infections RR (95% CrI)
Tofacitinib combination therapy vs. biologic combination therapy				
Tofacitinib 5 mg + DMARDs vs. abatacept + DMARDs	1.26 (0.26 to 5.81)	0.90 (0.60 to 1.34)	0.33 (0.04 to 1.77)	NA
Tofacitinib 5 mg + DMARDs vs. golimumab + DMARDs	4.18 (0.87 to 23.31)	1.10 (0.72 to 1.68)	0.41 (0.04 to 2.46)	NA
Tofacitinib 5 mg + DMARDs vs. tocilizumab + DMARDs	1.65 (0.40 to 6.80)	0.93 (0.62 to 1.37)	0.54 (0.06 to 3.09)	NA
Tofacitinib 5 mg + DMARDs vs. rituximab + DMARDs	0.56 (0.06 to 3.61)	1.15 (0.79 to 1.66)	0.46 (0.05 to 2.42)	NA

CrI = credible interval; DMARDs = disease-modifying antirheumatic drugs; IR = inadequate response; NA = not available; NMA = network meta-analysis; RR = rate ratio; TNF = tumour necrosis factor.
Note: Fixed effects NMA.

APPENDIX 9: SUMMARY OF OTHER STUDIES

Objective

The objective of this review is to summarize the results of a study that was excluded from the main analysis because the patient population did not meet the inclusion criteria specified in the protocol. The patients in this study were naive to MTX. This summary is derived from an unpublished report submitted by the manufacturer (1069),[88] and a published study by Lee et al. 2014.[89]

Findings

Study Design

This study was a two-year, double-blind, randomized controlled trial for patients with moderate to severe RA. Patients were randomized in a 2:2:1 ratio into one the following three study groups:

1. Tofacitinib 5 mg twice daily
2. Tofacitinib 10 mg twice daily
3. MTX 10 mg/week, titrated to 20 mg/week as tolerated by week 8.

Patients must have been MTX naive prior to study entry. If a patient had received more than three weekly doses of MTX, or discontinued the use of MTX before reaching three doses due to a treatment-related adverse event, the patient was excluded from the study. If receiving three or fewer doses prior to the study, patients must have not been receiving MTX for at least four weeks before receiving the first dose of the study drug. Of the 186 patients assigned to the MTX group, 130 patients had reached a dose of 20 mg/week, 26 patients had reached 15 mg/week and 13 patients had reached 10 mg/week, with an overall mean dose of methotrexate of 18.5 mg/week.

The co-primary end points of the study were the mTSS and ACR 70 at six months. Secondary outcomes included mTSS and ACR 70 at 12 and 24 months; and ACR 20, ACR 50, and other disease, quality of life, work-related, and fatigue scales at six, 12, and 24 months. Safety outcomes were assessed every three months and included overall event outcomes (adverse events, WDAEs, SAEs, and serious infections), harms of special interest (cardiovascular, infections, malignancies, gastrointestinal) and laboratory data. Background therapy for pain or other associated RA therapy was permissible during the study, provided it remained at stable doses and was in accordance with study defined protocol.

The FAS (patients who received at least one dose, baseline assessment, and at least one post-baseline assessment) was used to assess the primary and secondary efficacy outcomes. Safety outcomes were assessed using the safety analysis set (patients who received at least one dose of the drug). Laboratory outcomes included only patients who had laboratory results available.

Results

Nine hundred and fifty-eight patients were randomized in the study (N = 373, tofacitinib 5 mg; N = 397, tofacitinib 10 mg; N = 186, MTX). The population was 77% to 78% female, 64% to 68% Caucasian, a mean of 49 to 50 years of age, had RA for a mean duration of 2.7 to 2.9 years, a mean of 25 to 26 tender joints, a mean of 16 to 17 swollen joints, and 37% to 41% had received prior non-biologic DMARDs other than MTX.

For the primary efficacy outcomes, the tofacitinib 5 mg twice daily group had a statistically significantly greater mTSS and ACR 70 response rate at six months compared with the MTX group. Table 52 and Table 53 show the results of the secondary efficacy outcomes. All outcomes listed favoured the

tofacitinib 5 mg twice daily group and the observed differences were statistically significant. Harms data are listed in Table 54. The tofacitinib 5 mg group had greater increases in LDL and serum creatinine levels from baseline to six, 12, and 24 months compared with MTX.

TABLE 55: EFFICACY OUTCOMES (MTSS, DAS-4, HAQ-DI)

	Tofacitinib 5 mg b.i.d. (N = 371)	MTX[a] (N = 186)
mTSS, LS mean change[b] from baseline (SE)		
Month 6	0.18[c] (0.12)[d]	0.84 (0.16)
Month 12	0.36[c] (0.14)	1.18 (0.20)
Month 24	0.55[c] (0.25)	2.08 (0.34)
DAS-4 (ESR) < 2.6, % of patients (SE)		
Month 6	14.6[e] (1.91)	7.6 (2.02)
Month 12	18.7[e] (2.10)	11.7 (2.45)
Month 24	20.8[c] (2.19)	9.9 (2.28)
HAQ-DI, LS mean change¥ from baseline (SE)		
Month 6	–0.8[c] (0.03)	–0.6 (0.04)
Month 12	–0.9[c] (0.03)	–0.7 (0.04)
Month 24	–0.9[c] (0.03)	–0.7 (0.05)

b.i.d. = twice daily; DAS = Disease Activity Score; ESR = erythrocyte sedimentation rate; HAQ-DI = Health Assessment Questionnaire–Disability Index; LS = least squares; mTSS = modified Total Sharp Score; MTX = methotrexate; N = sample size; SE = standard error.

[a] Mean MTX dose at the end of week 12 was 18.5 mg/week.

[b] Controlling for baseline score and time since diagnosis.

[c] $P < 0.001$.

[d] Primary outcome.

[e] $P < 0.05$.

Source: Lee et al.[89] and manufacturer's submission.[88]

TABLE 56: EFFICACY OUTCOMES (ACR RESPONSE)

	ACR 20 Response		ACR 50 Response		ACR 70 Response	
	TOF 5 mg b.i.d. (N = 371)	MTX[a] (N = 186)	TOF 5 mg b.i.d. (N = 371)	MTX[a] (N = 186)	TOF 5 mg b.i.d. (N = 371)	MTX[a] (N = 186)
Percentage of patients (SE)						
Month 6	71.3[b] (2.35)	50.5 (3.68)	46.6[b] (2.59)	26.6 (3.25)	25.5[b] (2.26)[c]	12.0 (2.39)
Month 12	67.8[b] (2.43)	51.1 (3.68)	49.9ŧ (2.60)	33.7 (3.48)	28.7[b] (2.35)	15.2 (2.64)
Month 24	64.2[b] (2.49)	42.4 (3.64)	49.3[b] (2.60)	28.3 (3.31)	34.4[b] (2.47)	15.2 (2.64)

ACR = American College of Rheumatology; b.i.d. = twice daily; MTX = methotrexate; N = sample size; SE = standard error; TOF = tofacitinib.

[a] Mean MTX dose at the end of week 12 was 18.5 mg/week.

[b] $P < 0.001$ compared with MTX

[c] Primary outcome

Source: Lee et al.[89] and manufacturer's submission.[88]

TABLE 57: SAFETY OUTCOMES

	Tofacitinib 5 mg b.i.d. (N = 371)	MTX[a] (N = 186)
AEs		
Deaths, N	3	0
AEs, N	1097	561
AEs, N of patients, %	297 (79.6)	147 (79.0)
SAEs, N of patients, %	40 (10.7)	22 (11.8)
Serious infection, N of patients, %	11 (3.0)	5 (2.7)
WDAEs, N of patients, %	40 (10.7)	25 (13.4)
Harms of special interest[b]		
Cancer (confirmed)	2	1
Cardiac disorders	5	3
Gastrointestinal disorders¥	6	4
Infections and Infestations¥	10	5
Laboratory data		
LDL, LS mean change from baseline, % change (SE)		
Month 6	14.47 (1.54)	2.12 (2.18)
Month 12	16.77 (1.58)	2.24 (2.29)
Month 24	18.57 (1.65)	3.91 (2.47)
Serum creatinine, LS mean change from baseline, mg/dL (SE)		
Month 6	0.08 (0.01)	0.03 (0.01)
Month 12	0.09 (0.01)	0.04 (0.01)
Month 24	0.10 (0.01)	0.04 (0.01)
Moderate to severe neutropenia ANC ≥ 0.5 to < 1.5 x 10^3/µL, n (%)		
Month 6	3 (< 1.0)	0 (0.0)
Month 12	3 (< 1.0)	0 (0.0)
Month 24	1 (< 1.0)	0 (0.0)
Potentially life-threatening neutropenia ANC < 0.5 × 10^3/µL, n (%)		
Month 6	0 (0.0)	0 (0.0)
Month 12	0 (0.0)	0 (0.0)
Month 24	0 (0.0)	0 (0.0)
AST ≥ 3 x ULN, n (%)	6 (1.6)	6 (3.3)
ALT ≥ 3 x ULN, n (%)	11 (3.0)	13 (7.1)

AEs = adverse events; ALT = alanine transaminase; ANC = absolute neutrophil count; AST = aspartate transaminase; b.i.d. = twice daily; LDL = low-density lipoprotein; LS = least squares; N = sample size; N = number; SAEs = serious adverse events; SE = standard error; ULN = upper limit of normal; WDAEs = withdrawals due to adverse events.

[a] Mean MTX dose at the end of week 12 was 18.5 mg/week

[b] Medical Dictionary for Regulatory Activities (MedDRA) System Organ Class.

Source: Lee et al.[89] and manufacturer's submission.[88]

Critical Appraisal

Adequate blinding, randomization, and allocation concealment were applied in the present study. Baseline characteristics were similar between the groups; however, the duration of disease for the sample was lower than that found in other similar studies (approximately three years compared with six years or more). The dose of MTX should also be considered when interpreting the results of the present study. The maximum dose of MTX, 20 mg/week, was reached by 70% of patients.

Summary

This study compared the efficacy and safety of tofacitinib 5 mg, tofacitinib 10 mg, and MTX for patients with moderate to severe RA who were defined as MTX naive. Overall, the mean dose for the MTX group was 18.5 mg/week. Relative to the MTX group, tofacitinib 5 mg twice daily had superior efficacy outcomes according to mTSS, ACR 20/50/70 response rates, HAQ-DI, and DAS-4(ESR). There was a numerically greater increase in LDL and serum creatinine for 5 mg tofacitinib group compared with the MTX group.

REFERENCES

1. Xeljanz (tofacitinib) tablets 5 mg tofacitinib (as tofacitinib citrate) Tablets for oral administration [product monograph]. Kirkland (QC): Pfizer; 2014.
2. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, III, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum [Internet]. 2010 Sep [cited 2014 Oct 14];62(9):2569-81. Available from: http://onlinelibrary.wiley.com/doi/10.1002/art.27584/pdf
3. European Medicines Agency. Refusal of the marketing authorisation for Xeljanz (tofacitinib). London UK: 2014 Jun 25. [cited 2014 May 4]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/002542/WC500142485.pdf
4. Clinical Study Report: A3921032. Phase 3, randomized, double-blind, placebo-controlled study of the safety and efficacy of 2 doses of CP-690,550 in patients with active rheumatoid arthritis on background methotrexate with inadequate response to TNF inhibitors [**CONFIDENTIAL** internal manufacturer's report]. Kirkland (QC): Pfizer, Inc.; 2012 Jan 26.
5. Burmester GR, Blanco R, Charles-Schoeman C, Wollenhaupt J, Zerbini C, Benda B, et al. Tofacitinib (CP-690,550) in combination with methotrexate in patients with active rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitors: a randomised phase 3 trial. Lancet. 2013 Feb 9;381(9865):451-60.
6. Strand V, Burmester GR, Zerbini CA, Mebus CA, Zwillich SH, Gruben D, et al. Tofacitinib with methotrexate in third-line treatment of patients with active rheumatoid arthritis: Patient-reported outcomes from a Phase 3 trial. Arthritis Care Res (Hoboken). 2014 Sep 3.
7. Clinical Study Report: A3921045. Phase 3, randomized, double blind, placebo controlled study of the efficacy and safety of 2 doses of CP-690,550 monotherapy in patients with active rheumatoid arthritis [**CONFIDENTIAL** internal manufacturer's report]. Kirkland (QC): Pfizer, Inc.; 2012 Jan 27.
8. Fleischmann R, Kremer J, Cush J, Schulze-Koops H, Connell CA, Bradley JD, et al. Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis. N Engl J Med. 2012 Aug 9;367(6):495-507.
9. Clinical study report: A3921044. Phase 3 randomized, double blind, placebo controlled study of the efficacy and safety of 2 doses of CP-690,550 in patients with active rheumatoid arthritis on background methotrexate (1-year analysis) [**CONFIDENTIAL** internal manufacturer's report]. New York (NY): Pfizer, Inc; 2012 Jan 25.
10. Clinical study report: A3921044. Phase 3 randomized, double blind, placebo controlled study of the efficacy and safety of 2 doses of CP-690,550 in patients with active rheumatoid arthritis on background methotrexate [**CONFIDENTIAL** internal manufacturer's report]. [New York (NY)]: Pfizer, Inc.; 2013 Jan 25.
11. van der Heijde D, Tanaka Y, Fleischmann R, Keystone E, Kremer J, Zerbini C, et al. Tofacitinib (CP-690,550) in patients with rheumatoid arthritis receiving methotrexate: twelve-month data from a twenty-four-month phase III randomized radiographic study. Arthritis Rheum. 2013 Mar;65(3):559-70.
12. Clinical Study Report: A3921046 Amended. Phase 3, randomized, double-blind, placebo-controlled study of the safety and efficacy of 2 doses of CP-690,550 in patients with active rheumatoid arthritis on background DMARDs [**CONFIDENTIAL** internal manufacturer's report]. Kirkland (QC): Pfizer, Inc; 2012 Jan 27.
13. Kremer J, Li ZG, Hall S, Fleischmann R, Genovese M, Martin-Mola E, et al. Tofacitinib in combination with nonbiologic disease-modifying antirheumatic drugs in patients with active rheumatoid arthritis: a randomized trial. Ann Intern Med. 2013 Aug 20;159(4):253-61.

14. Clinical Study Report: A3921064 Amended. Phase 3 randomized, double-blind, active comparator, placebo-controlled study of the efficacy and safety of 2 doses of CP-690,550 in patients with active rheumatoid arthritis on background methotrexate [**CONFIDENTIAL** internal manufacturer's report]. Kirkland (QC): Pfizer, Inc; 2012 Jan 27.

15. van Vollenhoven RF, Fleischmann R, Cohen S, Lee EB, Garcia Meijide JA, Wagner S, et al. Tofacitinib or adalimumab versus placebo in rheumatoid arthritis. N Engl J Med. 2012 Aug 9;367(6):508-19.

16. Center for Drug Evaluation and Research, U.S. Food and Drug Administration. Medical review(s). In: Xeljanz (tofacitinib). Pfizer. Application No. 203214Orig1s000. Approval date: 11/06/2012 [Internet]. Rockville (MD): FDA; 2012 Sep 26 [cited 2014 Sep 1]. (FDA drug approval package). Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/203214Orig1s000MedR.pdf

17. The impact of arthritis in Canada: today and over the next 30 years [Internet]. Toronto: Arthritis Alliance of Canada; 2011. [cited 2014 Oct 14]. Available from: http://www.arthritisalliance.ca/en/initiativesen/impact-of-arthritis

18. Bykerk VP, Akhavan P, Hazlewood GS, Schieir O, Dooley A, Haraoui B, et al. Canadian Rheumatology Association recommendations for pharmacological management of rheumatoid arthritis with traditional and biologic disease-modifying antirheumatic drugs. J Rheumatol. 2012 Aug;39(8):1559-82.

19. Vliet Vlieland TP. Non-drug care for RA--is the era of evidence-based practice approaching? Rheumatology (Oxford) [Internet]. 2007 Sep [cited 2014 Oct 14];46(9):1397-404. Available from: http://rheumatology.oxfordjournals.org/content/46/9/1397.full.pdf+html

20. [Pr]Actemra© (tocilizumab) 20 mg/mL concentrate solution for infusion 162 mg/ 0.9 mL solution for injection [product monograph]. Mississauga (ON): Hoffmann-La Roche Limited; 2014 Oct 7.

21. Cividino AA. Use of biologic monotherapy for rheumatoid arthritis. CRA e-newsletter [Internet]. 2012 [cited 2014 Oct 14];(8). Available from: http://rheum.ca/images/documents/Issue_08_Biologic_Monotherapy_EN.pdf

22. [Pr]Orencia© (abatacept) Intravenous Infusion, 250 mg / 15 mL vial Solution for Subcutaneous Injection, 125 mg/ mL [product monograph]. Montreal (QC): Bristol-Myers Squibb Canada; 2013 Oct 7.

23. [Pr]Rituxan© (rituximab) 10 mg/mL intravenous infusion [product monograph]. Mississauga (ON): Hoffmann-La Roche Limited; 2014 May 29.

24. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Office of Translational Sciences, Office of Biostatistics. Statistical review and evaluation. Clinical studies (addendum): Xeljanz (tofacitinib) 5 mg and 10mg tablets [Internet]. Silver Spring (MD): FDA; 2012 Jun 11. [cited 2014 Nov 19]. Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/203214Orig1s000StatR.pdf Addendum for NDA 203214.

25. CDR submission: XELJANZ™ Tofacitinib tablets 5 mg tofacitinib (as tofacitinib citrate) Tablets for oral administration. Pfizer: Kirkland, Quebec [**CONFIDENTIAL** manufacturer's submission]. Kirkland (QC): Pfizer; 2014 Apr 5.

26. Pfizer Canada response to Nov 6, 10, 24 2014 CDR request for additional information: non-responder imputation (NRI) method for all patients that withdrew for any reason **[CONFIDENTIAL** additional manufacturer's information]. Kirkland (QC): Pfizer, Inc; 2015 Jan 9.

27. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis [Internet]. 2014 Mar [cited 2014 Oct 22];73(3):492-509. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3933074/pdf/annrheumdis-2013-204573.pdf

28. van Riel PL, van Gestel AM. Clinical outcome measures in rheumatoid arthritis. Ann Rheum Dis [Internet]. 2000 Nov [cited 2014 Oct 14];59 Suppl 1:i28-i31. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1766622

29. Cohen JD, Dougados M, Goupille P, Cantagrel A, Meyer O, Sibilia J, et al. Health assessment questionnaire score is the best predictor of 5-year quality of life in early rheumatoid arthritis. J Rheumatol. 2006 Oct;33(10):1936-41.

30. 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.

31. 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 Mar 15;57(2):193-202.

32. 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):1602-7.

33. Wells G, Becker JC, Teng J, Dougados M, Schiff M, Smolen 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 Jun;68(6):954-60.

34. 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 Aug;36(8):1606-10.

35. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: dimensions and practical applications. Health Qual Life Outcomes. 2003;1:20.

36. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: a review of its history, issues, progress, and documentation. J Rheumatol. 2003 Jan;30(1):167-78.

37. Bruynesteyn K, van der Heijde D, Boers M, Saudan A, Peloso P, Paulus H, 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 Apr;46(4):913-20.

38. 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):811-9.

39. 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 Dec;38(6):633-8. Temp ID 74.

40. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. Ann Med. 2001 Jul;33(5):350-7. Temp ID 75.

41. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. PharmacoEconomics. 1999 Feb;15(2):141-55. Temp ID 76.

42. Strand V, Singh JA. Improved health-related quality of life with effective disease-modifying antirheumatic drugs: evidence from randomized controlled trials. Am J Manag Care. 2008 Apr;14(4):239-53. Temp ID 77.

43. Cohen S, Hurd E, Cush J, Schiff M, Weinblatt ME, Moreland LW, 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 Mar;46(3):614-24.

44. 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) [Internet]. Rockville (MD): Food and Drug Administration; 1999. [cited 2014 Oct 14]. Available from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071579.pdf

45. Felson DT, Furst DE, Boers M. Rationale and strategies for reevaluating the ACR20. J Rheumatol. 2007 May;34(5):1184-7.

46. Fransen J, Antoni C, Mease PJ, Uter W, Kavanaugh A, Kalden JR, 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 [Internet]. 2006 Apr 27 [cited 2014 Oct 14];65(10):1373-8. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1798317/pdf/1373.pdf

47. 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.

48. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, 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 Nov;49(11):916-20.

49. 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 Feb;51(2):177-81.

50. 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 Jan;38(1):44-8.

51. van Gestel AM, Anderson JJ, van Riel PL, Boers M, Haagsma CJ, Rich B, et al. ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. J Rheumatol. 1999 Mar;26(3):705-11.

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

53. 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 [Internet]. 2007 Mar [cited 2014 Oct 14];66(3):407-9. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1856019/pdf/407.pdf

54. Anderson J, Caplan L, Yazdany J, Robbins ML, Neogi T, Michaud K, et al. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. Arthritis Care Res (Hoboken). 2012 May;64(5):640-7.

55. 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):783-9.

56. Fernandez-Montero JV, Barreiro P, Soriano V. HIV protease inhibitors: recent clinical trials and recommendations on use. Expert Opin Pharmacother. 2009;10(10):1615-29.

57. Matsui T, Kuga Y, Kaneko A, Nishino J, Eto Y, Chiba N, 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 [Internet]. 2007 Sep [cited 2014 Oct 14];66(9):1221-6. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1955164/pdf/1221.pdf

58. Tamhane A, Redden DT, McGwin G, Jr., Brown EE, Westfall AO, Reynolds RJ, 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 Nov;40(11):1812-22.

59. Canadian Agency for Drugs and Technology in Health. CADTH Therapeutic Review. Clinical and economic overview: biological response modifier agents for adults with rheumatoid arthritis [Internet]. Ottawa (ON): CADTH; 2010. [cited 2015 Jan 15]. Available from: http://www.cadth.ca/media/pdf/TR_RA_Clinical_and_Economic_Overview_e.pdf

60. McKesson Pharmaclick [Internet]. Saint-Laurent (QC): McKesson Canada. 2014 [cited 2014 Oct 30]. Available from: https://www.mckesson.ca/ Subscription required.

61. Mapi Values. Efficacy and safety of tofacitinib versus biological treatments for rheumatoid arthritis patients who had an inadequate response with dmards: Appendix G Additional results. In: CDR submission: XELJANZ™ Tofacitinib tablets 5 mg tofacitinib (as tofacitinib citrate) Tablets for oral administration. Pfizer: Kirkland, Quebec [**CONFIDENTIAL** manufacturer's submission]. Kirkland (QC): Pfizer; 2014 Apr 5. Report PI6169D - Version 6. Prepared November 2011.

62. Mapi Values. Efficacy and safety of tofacitinib versus biological treatments for rheumatoid arthritis patients who had an inadequate response with dmards: Appendix A to F. In: CDR submission: XELJANZ™ Tofacitinib tablets 5 mg tofacitinib (as tofacitinib citrate) Tablets for oral administration. Pfizer: Kirkland, Quebec [**CONFIDENTIAL** manufacturer's submission]. Kirkland (QC): Pfizer; 2014 Apr 5. Report PI6169D - Version 6. Prepared November 2011.

63. Ontario Ministry of Health and Long-Term Care. Ministry of Health and Long Term Care Exceptional Access Program (EAP) EAP Reimbursement criteria for frequently requested drugs; 2014 Feb 1 [cited 2014 Oct 30]. Available from: http://www.health.gov.on.ca/en/pro/programs/drugs/pdf/frequently_requested_drugs.pdf

64. Disease activity score in 28 joints (DAS28) [Internet]. Philadelphia: The Institute for Continuing Healthcare Education; 2009. [cited 2014 Oct 14]. Available from: http://www.iche.edu/newsletter/DAS28.pdf

65. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980 Feb;23(2):137-45.

66. Ramey DR, Fries JF, Singh G. The Health Assessment Questionnaire 1995 -- status and review. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1996. p. 227-37.

67. 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 Aug;35(8):1528-37.

68. Boini S, Guillemin F. Radiographic scoring methods as outcome measures in rheumatoid arthritis: properties and advantages. Ann Rheum Dis. 2001 Sep;60(9):817-27.

69. 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 May;32(5):778-86.

70. Sokka T. Radiographic scoring in rheumatoid arthritis: a short introduction to the methods. Bull NYU Hosp Jt Dis. 2008;66(2):166-8.

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

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

73. 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 Sep;42(9):1854-60.

74. Welsing PM, Landewe RB, van Riel PL, Boers M, van Gestel AM, van der LS, et al. The relationship between disease activity and radiologic progression in patients with rheumatoid arthritis: a longitudinal analysis. Arthritis Rheum. 2004 Jul;50(7):2082-93.

75. 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 Sep;24(5):540-5.

76. Clarke AE, St-Pierre Y, Joseph L, Penrod J, Sibley JT, Haga M, et al. Radiographic damage in rheumatoid arthritis correlates with functional disability but not direct medical costs. J Rheumatol. 2001 Nov;28(11):2416-24.

77. 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 Dec;31(3):176-90.

78. Foley-Nolan D, Stack JP, Ryan M, Redmond U, Barry C, Ennis J, et al. Magnetic resonance imaging in the assessment of rheumatoid arthritis--a comparison with plain film radiographs. Br J Rheumatol. 1991 Apr;30(2):101-6.

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

80. 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 Jun;73(6):1067-70.

81. Clinical Study Report: A3921024 and A3921041 Amended. Summary of long-term extension studies study A3921024: a long-term, open-label follow-up study of tasocitinib (CP-690,550) for treatment of rheumatoid arthritis [and] study A3921041: a long-term, open-label study of CP-690,550 to confirm the safety following long term administration of CP-690,550 in the treatment of rheumatoid arthritis [**CONFIDENTIAL** internal manufacturer's report]. Kirkland (QC): Pfizer, Inc; 2012 Jan 31.

82. Wollenhaupt J, Silverfield J, Lee EB, Curtis JR, Wood SP, Soma K, et al. Safety and efficacy of tofacitinib, an oral janus kinase inhibitor, for the treatment of rheumatoid arthritis in open-label, longterm extension studies. J Rheumatol. 2014 May;41(5):837-52.

83. Tran T, van Gaalen R, Woolcott J. Supplementary report: mixed treatment comparison of the efficacy of biologic agents in combination with methotrexate in patients with rheumatoid arthritis who have been previously treated with methotrexate . Kanata (ON): Pfizer Canada; 2014 Apr 29.

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

85. Mapi Values. Efficacy and safety of tofacitinib versus biological treatments for rheumatoid arthritis patients who had an inadequate response with dmards. In: CDR submission: XELJANZ™ Tofacitinib tablets 5 mg tofacitinib (as tofacitinib citrate) Tablets for oral administration. Pfizer: Kirkland, Quebec [**CONFIDENTIAL** manufacturer's submission]. Kirkland (QC): Pfizer; 2014 Apr 5. Report PI6169D - Version 6. Prepared November 2011.

86. Mapi Values. Efficacy and safety of tofacitinib versus biologic treatments for rheumatoid arthritis patients who have failed treatment with TNF inhibitors. In: CDR submission: XELJANZ™ Tofacitinib tablets 5 mg tofacitinib (as tofacitinib citrate) Tablets for oral administration. Pfizer: Kirkland, Quebec [**CONFIDENTIAL** manufacturer's submission]. Kirkland (QC): Pfizer; 2013 Dec 9. Report PI15250A - Version 1.0 December 9, 2013.

87. Pfizer Canada response to Nov 6, 11 2014 CDR request for additional information:updated analysis referred to as Appendix I in the submitted Network Meta-Analysis (NMA) report PI6169D, da Silva, Mapi Values 2011, on p.371. **[CONFIDENTIAL** additional manufacturer's information]. Kirkland (QC): Pfizer, Inc; 2015 Jan 9.

88. Clinical Study Report: A3921069 Phase 3 randomized, double-blind study of the efficacy and safety of 2 doses of CP-690,550 compared to methotrexate in methotrexate-naïve patients with rheumatoid arthritis (1-Year analysis) [**CONFIDENTIAL** internal manufacturer's report]. Kirkland (QC): Pfizer, Inc; 2012 Oct 18.

89. Lee EB, Fleischmann R, Hall S, Wilkinson B, Bradley JD, Gruben D, et al. Tofacitinib versus methotrexate in rheumatoid arthritis. N Engl J Med. 2014 Jun 19;370(25):2377-86.