A Description of the IVI-RA Model v2.0 *†

Devin Incerti‡ Jeroen P. Jansen§

December 16, 2019

Contents

*We thank Melody Owen, Emily Glowienka, Ian McGovern, Zarmina Khankhel, and Maria Lorenzi for conducting the systematic literature review and running the network meta-analysis for this version of the IVI-RA model. We thank Ming Xu for support with model adaptation. We thank Darius Lakdawalla, Jason Shafrin, Mark Linthicum, Jeffrey Curtis, Carole Wiedmeyer, and members of the Innovation and Value Initiative's scientific advisory panel for reviewing an earlier version of this document. Finally, we thank Sam Norton for help with the latent class growth model.

†This report should be referenced as follows: Incerti, D, Jansen, JP. A Description of the IVI-RA Model v2.0. 2020; last updated January 2020.

‡Previously at the Innovation and Value Initiative and Precision Medicine Group

§Innovation and Value Initiative and Precision Medicine Group

List of Figures

List of Tables

Executive summary

This document describes version 2.0 of the Innovation and Value Initiative's (IVI's) individual patient simulation model for rheumatoid arthritis (RA) (the IVI-RA model). The model simulates the costs, health outcomes, and risks associated with disease-modifying anti-rheumatic drugs (DMARDs) including conventional DMARDs (cDMARDs), biologic DMARDs (bDMARDs), and Janus kinase/signal transducers and activators of transcription (JAK/STAT) inhibitors for patients with moderate to severe rheumatoid arthritis (RA) who have previously failed treatment with cDMARDs. The model is intended to help decision-makers assess the value of treatments for a population of patients with RA.

Open-Source Value Project

The IVI-RA model is part of IVI's Open Source Value Project (OSVP), which is building an open, collaborative, and consensus-based process for the development of tools for value assessment. Models developed by the OSVP process are iterative, evolving as the science of value assessment advances and as new evidence becomes available.

OSVP models are released and updated using a four step process:

1. Public release of the model.
2. Invite feedback and suggested improvements to the model in a public comment period.
3. A panel of experts determines which of the evidence-based suggestions for improvement suggested in Step 2 should be implemented by means of peer-review and a formal voting process.
4. Revise the model based on the feedback from the technical expert panel in Step 3.

To provide a starting point for debate, the initial release of each OSVP model (i.e., version 1.0) must be flexible and allow users to choose from a large number of plausible model structures and approaches based on clinical practice and previous modeling efforts. The four-step process is designed to be repeated many times so that the scientific approach and evidence considered can be refined over time. Over time, the number of model structures may shrink as the OSVP process moves toward scientific consensus. To be sure, the OSVP process will not eliminate all the variation in results of value assessment since perspectives on value will vary and disagreements about relevant clinical evidence may persist. But the consensus-based approach will allows users to better understand legitimate and intrinsic reasons why value estimates vary.

Contents of the IVI-RA model

Version 2.0 is IVI's second release of the IVI-RA model. The model is very flexible and allows users to choose from a large number of the plausible model structures supported by clinical practice and prior decision-analytic modeling research in RA. The IVI-RA model is a collaborative multistakeholder effort that produces tools to help decision-makers evaluate the value of pharmaceutical treatments for RA. To facilitate transparency, understanding, and debate among diverse stakeholders, the IVI-RA model consists of the following components:

- **Source code**: R and C++ code for the model is available in our IVI GitHub repository. Modelers and programmers may adapt the source code for their own purposes or collaborate with IVI to improve the code.

- **R package**: The IVI-RA model is released as an R package with documentation available online. Researchers can use the package to run the IVI-RA model for custom analyses. Use of the R package is recommended when peforming analyses for academic publication.
- **Model documentation**: This document provides provides technical details on the model structure, statistical methods for parameter estimation, and source data.
- **IVI-RA Model Interface**: For users not be well-versed in the R programming language, we provide a web application for running the model online. The web application is designed for custom analyses and allows users full control over the treatments, patient population, model structures, parameter values, and simulation settings.
- **The IVI-RA Value Tool**: An important aim of the OSVP project is to obtain feedback from as many relevant stakeholders as possible. The IVI-RA Value Tool is a general audience web-application allowing those who are not experts in modeling, health economics, or RA to interact with the IVI-RA model.

Intended use of the IVI-RA model

The IVI-RA model is not a value assessment framework but a model that simulates the costs, health outcomes, and risks associated with treatments for RA. It can therefore be used with any value framework preferred by the user. Currently, our online tools support both cost-effectiveness analysis (CEA) and multi-criteria decision-analysis (MCDA). IVI has also developed an R package, `hesim`, for health-economic simulation modeling and decision analysis that can be used to perform individualized CEA (Basu and Meltzer 2007; Ioannidis and Garber 2011; Espinoza et al. 2014) on simulation output from the IVI-RA model.

About the IVI-RA model

Overview

The IVI-RA model is a discrete-time individual patient simulation that simulates outcomes for individual patients. Model cycles are 6-months long, which is consistent with clinical trial evidence. The model simulates the progression of the health assessment questionnaire disability index (HAQ), a measure of functional status in RA.

Serious infection rates and changes in HAQ score during the first 6 months from baseline are based on clinical trial evidence. The change in HAQ can be modeled indirectly as a function of the American College of Rheumatology (ACR) response to treatment, the European League Against Rheumatism (EULAR) response to treatment, or directly as a function of the treatment. Patients switch treatment during the initial 6 months if they have a serious infection. Additionally, the user can chose whether treatment switching should be based on disease activity level or treatment response.

After the first 6 months on a new treatment, the HAQ score progresses over time at a rate based on observational data. Progression can either be assumed to be linear (Wolfe and Michaud 2010; Michaud et al. 2011) or modeled using a non-linear mixture model (Norton et al. 2014).

Patients remain on treatment until treatment discontinuation or death. Time to treatment discontinuation is based on parametric survival analyses of real-world data. Seven possible distributions (exponential, Weibull, Gompertz, log-logistic, lognormal, and generalized gamma) can be chosen

by the user. Male and female mortality is based on US lifetables and increases with the HAQ score at baseline and the change in the HAQ score from baseline.

Health care sector costs consist of drug acquisition and administration costs, hospital costs (which increase with the HAQ score), general management costs, and costs caused by serious infections. Non-health care sector costs are those due to lost wages.

Users wishing to calculate utility for CEA can map HAQ and individual characteristics to utility using the logistic regression algorithm of Wailoo et al. (2006) or the Hernández-Alava et al. (2013) mixture model. With both the Wailoo et al. (2006) and Wailoo et al. (2006) mappings, utility is calculated as a function of the HAQ and individual patient characteristic mapping, serious infections, and preferences for treatment attributes unrelated to safety and efficacy. QALYs combine life expectancy with per cycle utility.

Patient preferences and heterogeneity

The IVI-RA model is desgned to capture differences in individual characteristics, preferences, circumstances, and response to treatment. First, progression of disease, mortality, and preferences for treatment vary according to individual characteristics. Second, although current evidence is scarce, users can adapt the model so that treatment effects vary across patients (e.g., as a function of patient characteristics or prognostic factors). Third, the IVI-RA model incorporates preferences for treatment attributes unrelated to safety and efficacy—such as mode of administration and the time a medication has been on the market—that are not typically included in decision-analytic models for value assessment.

Uncertainty analysis

Since there will always be gaps in the available evidence and the appropriate scientific assumptions, it is important to quantify uncertainty. The IVI-RA model consequently contains 384 possible model structures, which can be used to quantify structural uncertainty or to evaluate the implications of different modeling assumptions. Parameter uncertainty is quantified using probabilistic sensitivity analysis (PSA).

We have found that model outcomes are especially sensitive to certain parameters and model structures, which highlights the importance of a flexible and consensus-based model. Primary sources of uncertainty include:

- The effect of treatment on the change in HAQ from baseline during the first 6 months of treatment
- The long-term progression of HAQ
- The reduction in treatment response after previous treatment failures
- The extent to which the HAQ score "rebounds" to its initial level after failing treatment
- Time on biologic treatment
- The relationship between HAQ and quality of life

Real-world evidence

To ensure that simulated clinical and economic outcomes reflect outcomes in routine practice, we model "baseline event rates" (i.e., disease progression, mortality, time on treatment), patient preferences, and costs using real-world data. To minimize bias, relative treatment effects (i.e., differences in safety and efficacy across treatments) are, when possible, based on randomized clinical trials (RCTs), and then applied to the baseline event rates.

Perspective of the decision-maker

Models should be flexible enough to meet the specific needs (e.g., a specific patient population) and perspectives (e.g., relevant sources of value) of different decision-makers. The current model is suitable for decision-makers making decisions for specific populations or subpopulations (e.g., policymakers, insurers, provider groups) but is not suitable for making predictions at the indiviudal level. Future iterations of the model may expand its use so that that it can be used for patients making resource allocation decisions (e.g., individualized cost-effectiveness analysis).

Cost components included in the model are based on the framework suggested by the Second Panel on Cost-Effectiveness in Health and Medicine (Sanders et al. 2016). Analyses based on a health care sector perspective can be performed by only incorporating health care sector costs. Analyses based on a (limited) societal perspective would include lost wages in addition to health care sector costs.

Value to the healthy

Conventional value assessments focus on value to the sick, but recent research provides a framework for valuing technology for the healthy (i.e., "insurance value") as well Lakdawalla et al. (2017). The IVI-RA model allows users to optionally incorporate insurance value, but we note that it is less well established than conventional approaches.

Version 2.0

IVI released Version 1.0 the IVI-RA model in November 2017, after which IVI invited public comment through February 16, 2018. Upon the conclusion of the public comment period, IVI engaged a third-party Technical Expert Panel (TEP) comprised of leaders in health economics, epidemiology, rheumatology, and patient communities to review the public comments and establish priorities for model improvement through a teleconference and a two-part modified Delphi survey. Several priorities emerged from TEP deliberation as described in the following `report`.Version 2.0 of the IVI-RA model, as described in this report, incorporates additional treatment options and uses new 6-month relative treatment effects based on an updated systematic literature review and network meta-analysis. In addition, drug acquisition and resource use cost estimates have been updated to 2019. It is envisoned that other recommendations by the TEP, such as incorporating long-term heterogeneous treatment effects, will be incorporated in the next iteration of the IVI-RA model.

1 Open-source consensus-based models for value assessment

The continuing increase in US health care costs has stimulated the introduction of initiatives to promote the use of high-value care. Decision-analytic models can be used to inform efficient use of health care resources, but are only relevant when deemed credible by different stakeholders, are representative of the local context and patient population, and can be easily updated without duplication of effort.

The nature of simulation modeling often leads to scientific disagreements and mistrust among decision-makers. Models are typically complex and difficult to understand. Even modeling experts may not be able to fully understand a model without public source code and detailed model documentation. Furthermore, efforts to make models accessible to non-experts are lacking. Models also become quickly outdated as new evidence arises or new scientific approaches are developed, which means that previous finding quickly become irrelevant to decision-makers.

The OSVP aims to increase understanding and relevance to diverse stakeholders by developing open-source consensus-based models. The hope is that these efforts can increase confidence in efforts to base reimbursement and policy decisions on value.

OSVP models are released and updated using a four step process:

1. Public release of the model.
2. Invite feedback and suggested improvements to the model in a public comment period.
3. A panel of experts determines which of the evidence-based suggestions for improvement suggested in Step 2 should be implemented by means of peer-review and a formal voting process.
4. Revise the model based on the feedback from the technical expert panel in Step 3.

The four-step process is designed to be repeated many times so that the scientific approach and evidence considered can be refined over time.

2 Overview of the IVI-RA model

2.1 Why IVI is modeling rheumatoid arthritis

Treatment for rheumatoid arthritis (RA) is well suited for the OSVP approach for three reasons. First, modeling methods and assumptions vary considerably across existing simulation models (Brennan et al. 2003; Wailoo et al. 2008; Tosh et al. 2011; Carlson et al. 2015; Stephens et al. 2015; Athanasakis et al. 2015; Stevenson et al. 2016; Institute for Clinical and Economic Review 2017; Stevenson et al. 2017). Predicting disease progression is complex and there are a number of different measures of treatment response and morbidity (Madan et al. 2015). Analyses have, not surprisingly, been performed using different modeling approaches and have reached different conclusions about the cost-effectiveness of treatments for RA.

Second, RA is an area of significant innovation. There have been important advancements in the treatment of RA over the past decade, which suggests that there is an increasing need for tools to assess the cost-effectiveness of these treatments.

Third, not only have new treatments come to market recently, but evidence on existing RA treatments is growing rapidly. Thus, there is a strong need for models that can be updated in a straightforward manner as the evidence base evolves.

2.2 Contents

To facilitate transparency, understanding, and debate among diverse stakeholders, the IVI-RA model consists of the following components:

- **Source code**: R and C++ code for the model is available in our IVI GitHub repository. Modelers and programmers may adapt the source code for their own purposes or collaborate with IVI to improve the code.
- **R package**: The IVI-RA model is released as an R package with documentation available online. Researchers can use the package to run the IVI-RA model for custom analyses. Use of the R package is recommended when peforming analyses for academic publication.
- **Model documentation**: This document provides provides technical details on the model structure, statistical methods for parameter estimation, and source data.
- **IVI-RA Model Interface**: For users not be well-versed in the R programming language, we provide a web application for running the model online. The web application is designed for custom analyses and allows users full control over the treatments, patient population, model structures, parameter values, and simulation settings.
- **The IVI-RA Value Tool**: An important aim of the OSVP project is to obtain feedback from as many relevant stakeholders as possible. The IVI-RA Value Tool is a general audience web-application allowing those who are not experts in modeling, health economics, or RA to interact with the IVI-RA model.

These components along with the OSVP process are designed to encourage collaboration among stakeholders. Stakeholders may collaborate with IVI in at least two ways. First, they can provide feedback on any of the components during the public comment period. Second, programmers can make direct changes to the source code by making a "pull request" on GitHub. IVI will review the proposed changes. Code modifications that affect the scientific approach or evidence considered will only be incorporated after a review by the technical panel but other changes such as bug fixes or performance improvements may be immediately accepted.

2.3 About

The IVI-RA model is a discrete-time individual patient simulation (IPS) with 6 month cycles that simulates patients one at a time. The model accounts for both parameter and structural uncertainty. Since the range of defensible scientific approaches is large, the IVI-RA model consists of 384 possible model structures. Structural uncertainty can be quantified by estimating cost-effectiveness across these different model structures and parameter uncertainty is quantified using probabilistic sensitivity analysis (PSA). (Note that the simulation was primarily written in C++ so that PSAs and analyses of structural uncertainty can be run in a reasonable amount of time.)

To ensure that simulated outcomes reflect outcomes in routine practice, we model "baseline event rates" (i.e., disease progression, mortality, time on treatment), patient preferences, and costs using real-world data. To minimize bias, relative treatment effects (i.e., differences in safety and efficacy across treatments) are, when possible, based on randomized clinical trials (RCTs), and then applied to the baseline event rates.

The IPS approach allows us to take an "individualized" modeling approach that captures both observable and unobservable patient heterogeneity. Disease progression, mortality, and preferences all vary across patients. In addition, although the evidence base is limited, users of the R package can model treatment effects as a function of any combination of patient characteristics (e.g., demographics, prognostic factors). Finally, the model incorporates preferences for treatment attributes unrelated to safety and efficacy.

As recommended by the Second Panel on Cost-Effectiveness in Health and Medicine (Sanders et al. 2016), costs are simulated from both a health care sector perspective and a societal perspective. Productivity losses from lost earnings are included in the societal perspective but not the health care sector perspective. As discussed below (Section 2.4), our individualized approach implies that future iterations of the model could be tailored to fit the perspective of a patient or provider.

2.4 Intended use

The model simulates the costs, health outcomes and risks associated with treatments for RA for each individual in a given population (see Section 5). As described in Section 6 users can model any sequence of biologic treatments and conventional disease-modifying antirheumatic drugs (cDMARDs).

The model can therefore be used for a number of purposes, conditional on the population of interest and the perspective of the decision maker. Here we describe a few possibilities.

The first and most obvious use of the model is for value assessment. Two approaches, cost-effectiveness analysis (CEA) and multi-criteria decision analysis (MCDA), are discussed in more detail in Section 3. Within the CEA approach, cost-effectiveness can be evaluated from the conventional perspective of a sick individual or from the perspective of a healthy individual using the "insurance value" framework developed by Lakdawalla et al. (2017).

Second, the model can be used to evaluate the consequences of clinical guidelines such as the current treat-to-target guidelines in the US (Singh et al. 2016) or guidelines based on treatment response like in the UK (Deighton et al. 2010). Unlike most previous models, our flexible framework allows treatment switching decisions to depend on disease activity level or treatment response, so outcomes under different decision rules can be simulated.

Third, although the model is currently designed for population level decision-making, it could, in principle, be used to predict long-term health and economic consequences for patients. The predicted outcomes could, for example, be used to inform patient and providers decision making. For instance, Ioannidis and Garber (2011) argue that cost-effectiveness has relevance to patients spending their own money on health care services, particularly as out-of-pocket costs grow. Likewise, providers have a growing interest in cost-effectiveness models to demonstrate the value of their care whether through participation in Accountable Care Organizations (ACOs), to ensure coverage of medical interventions for their patients, or to reduce unwanted variability in management.

2.5 Version 2.0

IVI released Version 1.0 the IVI-RA model in November 2017, after which IVI invited public comment through February 16, 2018. Upon the conclusion of the public comment period, IVI engaged a third-party Technical Expert Panel (TEP) comprised of leaders in health economics, epidemiology, rheumatology, and patient communities to review the public comments and establish priorities for model improvement through a teleconference and a two-part modified Delphi survey.

Details about the proces, findings and emerged priorities for next iterations of the model are described in the following `report`.

Updates with Version 2.0 of the IVI-RA model:

- **Treatment options**: Triple cDMARD therapy, sarilumab, baricitinib, upadacitinib, biosimilars
- **Evidence base**: Updated systematic literature review and network meta-analysis to estimate 6-month relative treatment effects regarding ACR 20/50/70, DAS28, and HAQ-DI based on randomized controlled trial evidence.
- **Unit costs**: Drug acquisition costs are updated to reflect 2019 costs. Costs related to other resource use have been updated based on 2019 consumer price index figures.

It is envisoned that other recommendations by the TEP, such as incorporating long-term heterogeneous treatment effects, will be incorporated in the next iteration of the IVI-RA model.

3 Value assessment

The IVI-RA model simulates clinical and economic outcomes for each individual in a given population of interest. Outcomes can be simulated over a particular time horizon or over a lifetime.

Although simulation output can be used with any value assessment framework, IVI tools currently support two methodologies for decision analysis: CEA and MCDA. Cost-effectiveness results and MCDA value scores are automatically generated when users run IVI's web-based user interfaces. In addition, IVI has developed an R package, `hesim`, for health-economic simulation modeling and decision analysis that can be used to perform individualized CEA (Basu and Meltzer 2007; Ioannidis and Garber 2011; Espinoza et al. 2014).

3.1 Cost-effectiveness analysis

CEA is a well-established approach for value assessment grounded in economic theory and widely used in the scientific literature (Briggs et al. 2006; Meltzer et al. 2011; Drummond et al. 2015). In general, CEA can be thought of as a methodology for maximizing health or well being subject to a resource constraint (Garber and Phelps 1997). The total value of a new health technology relative to a comparator is typically assessed using the incremental net monetary benefit (INMB),

$$INMB = k \cdot \Delta e - \Delta p, \tag{1}$$

where $e = e_1 - e_0$ is a measure of the incremental health benefits from the new technology relative to the comparator, $p = p_1 - p_0$ is a measure of the incremental cost of the new technology, and k is the willingness to pay for a one-unit health gain. The new technology can be deemed cost-effective if the $INMB > 0$, or equivalently, in terms of the incremental cost-effectiveness ratio (ICER), if,

$$\frac{\Delta p}{\Delta e} < k. \tag{2}$$

Incremental health benefits are typially measured in terms of health gains or patient well-being. Since treatments can affect both morbidity and mortality, CEAs typically use the quality-adjusted life-year (QALY). Since costs and benefits vary across patients, some researchers have argued for individualized CEA (Basu and Meltzer 2007; Ioannidis and Garber 2011; Espinoza et al. 2014) so that INMBs and ICERs are calculated separately for different subpopulations. It can be shown that if treatment response varies across the population, then making separate decisions in different populations will increase social welfare (Basu and Meltzer 2007).

In practice, costs and health benefits are subject to statistical uncertainty. We quantify this uncertainty using probabilistic sensitivity analysis (PSA) and structural uncertainty analysis, which is described in more detail in Section 9. This approach allows us to generate standard measures of uncertainty in CEA including cost-effectiveness planes (Black 1990; Barton et al. 2008), cost-effectiveness acceptability curves (CEACs) (Van Hout et al. 1994; Briggs et al. 1999; Fenwick et al. 2001; Barton et al. 2008), the cost-effectiveness acceptability frontier (CEAF) (Barton et al. 2008), and estimates of the expected value of perfect information (EVPI) (Fenwick et al. 2001; Barton et al. 2008).

3.2 Multi-criteria decision-analysis

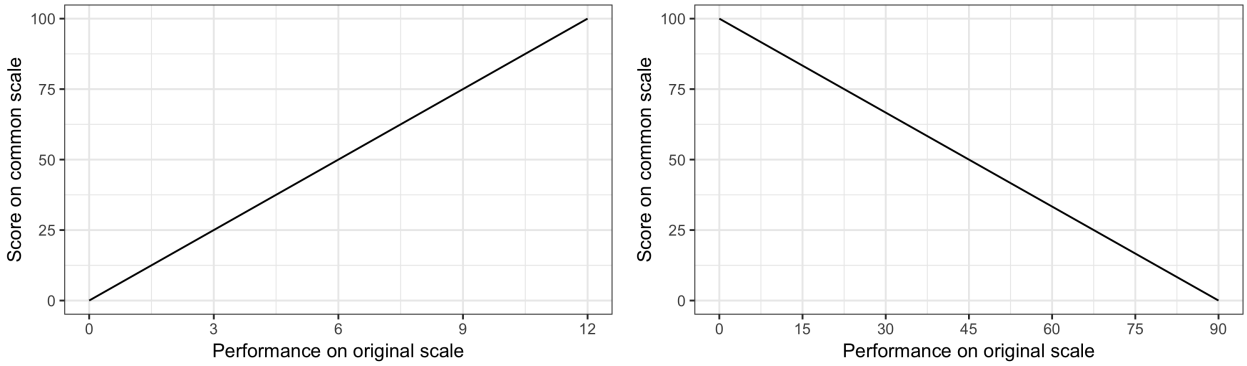
An alternative approach to CEA is MCDA. Keeney and Raiffa (1993) define MCDA as "an extension of decision theory that covers any decision with multiple objectives. A methodology for appraising alternatives on individual, often conflicting criteria, and combining them into one overall appraisal..." We use a similar approach, which implies that separate criteria are aggregated into a single measure of value.

There are many approaches to MCDA; here, we discuss the approach used by IVI in the web-based user interface, which is based on the discussion in Thokala et al. (2016). First, decision-makers must select the relevent criteria for the analysis. These criteria are based on the costs, health outcomes, and risks simulated from the underlying health-economic model. We discuss the criteria relevant to the IVI-RA model in Section 7.8.

Since different criteria may be measured using different units, performance on each criterion is converted into a common scale, for instance, ranging from 0 to 100. There are a number of techniques for creating a common scale; we use a simple linear partial value function to translate scores, which assumes a linear relationship between performance on the original scale of a given criterion and the common scale. To illustrate, Figure 1 demonstrates two mappings between the original scale and the common scale.

Performance on the first criterion, shown in Figure 1a, ranges from 0 to 12 on the original scale, with higher scores denoting better performance. In contrast, performance on the second criterion, shown in Figure 1a, ranges from 0 to 90, with lower scores denoting better performance. The relationship between performance on the original scale and the score on the common scale is therefore positive for the first criterion and negative for the second criterion. In both cases, the relationship follows a straight line because we assume a linear relationship.

Each criterion is assigned points, say ranging from 0 to 10, by the decision maker, and weighted by dividing each criterion's points by the sum of points across all criteria. For example, if there were 3 criteria and each criterion was given a score of 5, then each criterion would receive a weight of 1/3. If, on the other hand, the three criteria were given scores of 2.5, 5, and 7.5, then they would be given weights of .167, .33, and .5, respectively.



(a) Criterion where high performance is better (b) Criterion where low performance is better

Figure 1: Linear partial value functions

To aggregate results, we assume an additive model. In other words, the total score for a given treatment sequence is calculated by multiplying each criterion by the simulated standardized score and summing across criteria.

As with CEA, MCDA results are subject to statistical uncertainty. In our web applications, users choose a single model structure at a time, so uncertainty in MCDA outcomes is quantified using PSA. This produces a probability distribution around the simulated total score for each treatment sequence, which can be used to derive quantities of interest such as Bayesian credible intervals around the total score or the probability that each treatment sequence obtains a particular ranking among relevant treatment sequences.

4 Broader concepts of value

Garrison et al. (2017) suggest five concepts of value that researchers should consider adding to the standard cost per QALY based CEA: (1) a reduction in uncertainty from a diagnostic test; (2) insurance value for healthy patients due to reduction against physical risk; (3) the value of hope for individuals who become risk-loving and would rather pay for a therapy with a long right survival tail than a therapy with a shorter right survival tail but an equivalent (or shorter) expected life-expectancy; (4) real option value when a therapy allows an individual to benefit from future medical innovations; and (5) scientific spillovers when the benefits of an innovation cannot be entirely appropriated by the innovator.

The concept that is arguably most salient to RA is insurance value, which focuses on valuing morbidity-reducing innovations and has the largest effects relative to conventional CEA on treatments for severe diseases where the burden of illness is the greatest. The IVI-RA model allows users to incorporate insurance value into their analyses, while noting that the approach is less well-established than conventional CEA.

Other concepts of value may be incorporated in the future, but likely in future disease areas. For example, real option value is most relevant for innovations that increase longevity and might be particularly well suited to analyses of treatments in oncology. Likewise, survey evidence for the value of hope is based on technologies that increase survival Lakdawalla et al. (2012) rather than those that affect morbidity. Reductions in uncertainty from diagnostic tests are clearly most relevant to diagnostics and scientific spillovers are most relevant to diseases with large externalities

such infectious diseases.

Lakdawalla et al. (2017) provide a general mathematical framework for incorporating the effects of medical innovation on physical and financial risk. Conceptually, innovation can lower physical risk to healthy patients who might get sick in the future. New medical technologies act like "insurance policies" that protect a healthy person from all or part of the costs of falling ill. And while innovation certainly increases financial risk, this increase in financial risk can be mitigated by health care insurance.

The insurance value framework is an extension of the conventional CEA approach from the perspective of a healthy individual deriving utility from non-health consumption, c and health, h, according to $u(c, h)$. The individual is sick with probability π and well with probability $1 - \pi$. Health when well is h^w and health when sick is $h^s < h^w$. Income is y^w when well and $y^s < y^w$ when sick. The marginal utility of good $j \in c, h$ in state $i \in s, w$ is denoted by u^i_j.

The value of a technology to a healthy consumer (with no health insurance), V^{NHI} is derived implicitly by,

$$\pi u(y^s - p - V^{NHI}, h^s + \delta h) + (1 - \pi)u\left(y^w - V^{NHI}, h^w\right) = \pi u(y^s, h^s) + (1 - \pi)u(y^w, h^w). \tag{3}$$

The marginal value of the technology, dV^{NHI}, can be shown to be,

$$dV^{NHI} = \pi(k \cdot dh - dp) + \pi(1 - \pi)(k \cdot dh - dp)\left(\frac{u^s_c - u^w_c}{\pi u^s_c + (1 - \pi)u^w_c}\right) \tag{4}$$

$$= [k \cdot dh - dp]\left[\pi + \pi(1 - \pi)\left(\frac{u^s_c/u^w_c - 1}{\pi u^s_c/u^w_c + 1 - \pi}\right)\right], \tag{5}$$

where $k = \partial u^s_h / \partial u^s_c$ is the marginal value of a one unit health gain in dollar terms, dh is the marginal health gain from the technology, and dp is the marginal cost of the technology. The term $k \cdot dh - dp$ is equivalent to the INMB in conventional CEA. The insurance value framework can therefore be implemented with knowledge of only two additional parameters beyond those in conventional CEA: the probability of illness, π, and the marginal rate of substitution between the sick and the well states, u^s_c/u^w_c.

The probability of illness can be estimated using incidence of disease in the population of interest (e.g., in the RA population). The second term, u^s_c/u^w_c, is harder to estimate, but we allows users to specify it directly in our model and web-based user interfaces. Intuitively, this term reflects the amount of money the consumer would give up when healthy in exchange for gaining an additional dollar when sick. It rises when the consumer faces greater risks from illness.

It is worth emphasizing that insurance value is only larger than conventional value if the consumer is willing to give up more than \$1 in the well state in exchange for an additional \$1 in the sick state (i.e., $u^s_c/u^w_c > 1$). This is likely to be true, because if the demand for health care insurance is positive, then $u^s_c/u^w_c > 1$.

The difference between the insurance value of a technology and its conventional value is even larger when individuals can purchase health insurance. For example, consider an actuarially fair insurance

contract that pays the consumer I(p) when she falls sick. In this case, the insurance value of a health technology can be shown to be:

$$dV^{WHI} = dV^{NHI} + \pi(1-\pi)\left(\frac{u_c^s/u_c^w - 1}{\pi u_c^s/u_c^w + 1 - \pi}\right)\frac{dI}{dp}dp. \tag{6}$$

The term dI/dp is the marginal payment made to the insuree per 1 dollar spent on health care. In the extreme case where there is no cost-sharing so that $I(p) = p$ and $dI/dp = 1$. Here, health insurance completely eliminates spending risk the value of a technology is equal to its conventional value plus the value of physical risk reduction. More generally, $dI/dp < 1$ and the value of a health technology with health insurance is equal to the sum of its conventional value, the insurance value absent health insurance, and the value of health insurance made possible by the technology.

5 Populations

To run the IPS, a patient population must be specified. The model is designed for patients who are cDMARD experienced. The patient characteristics that must be included in the analysis are age, HAQ, gender, weight, the number of previous DMARDs, and disease activity. These variables are measured at the start of the simulation (i.e., model cycle 0).

Two default options for the patient population are available. First, a homogeneous cohort of men and women with gender-specific weights but otherwise identical characteristics can be used. Second, a heterogeneous cohort of patients with gender-specific weights but varying across all other characteristics can be specified. Other populations (i.e., for certain subgroups or based on registry data) can be used as well but are not prespecified in our R package.

Our default population consists of individuals that, on average, have high disease activity. The proportion that is female, age, the number of previous DMARDs, baseline HAQ, and DAS28 are based on the values reported in Curtis et al. (2010). Mean values for the SDAI and CDAI are from the US301 clinical trial—which had a DAS28 score similar to the value from Curtis et al. (2010)—summarized in Smolen et al. (2003). Summaries of each variable are reported in Table 1. Details on the algorithm for simulating heterogeneous patients are described in Appendix B.

Table 1: Default patient population

	Mean	Standard deviation	Minimum	Maximum
Age	55.00	13.00	18	85
Male	0.21	-	-	-
Female weight (kg)	75.00	-	-	-
Male weight (kg)	89.00	-	-	-
Previous DMARDs	3.28	1.72	0	-
DAS28	6.00	1.20	0	9.4
SDAI	43.00	13.00	0	86
CDAI	41.00	13.00	0	76
HAQ	1.50	0.70	0	3

6 Treatment strategies

Since patients typically use multiple treatments over a lifetime, the model is capable of simulating a treatment sequence of any arbitrary length. Treatments that can be included in a sequence include conventional disease-modifying anti-rheumatic drugs (cDMARDs), biologic DMARDs (bDMARDs), and Janus kinase/STAT (JAK/STAT) pathway inhibitors. The bDMARDs and JAK/STAT inhibitors, which we refer to collectively as targeted DMARDs (tDMARDs), included in the current version of the model are:

- **Tumor necrosis factor (TNF) inhibitors**: etanercept, adalimumab, infliximab, certolizumab, golimumab
- **Non-TNF inhibitors**: abatacept, anakinra, rituximab, tocilizumab, sarilumab
- **Janus kinase/signal transducers and activators of transcription (JAK/STAT) inhibitors**: tofacitinib, baricitinib, upadacitinib
- **Biosimilars**: Biosimilars of etanercept, adalimumab, and infliximab
- **Triple therapy**: sulfasalazine + hydroxychloroquine + methotrexate

At the end of a sequence, patient switch to non-biologic therapy (NBT), which encompasses a range of therapies that clinicians may feel is appropriate for all patients such as methotrexate and sulfasalazine (Stevenson et al. 2016, 2017).

7 Competing model structures

The IVI-RA model is a discrete-time IPS with 6 month cycles that can be run using a number of different model structures. Like most decision-analytic models in RA, version 1 of the model measures changes in disease severity using the Health Assessment Questionnaire (HAQ) Disability Index score (Brennan et al. 2003; Wailoo et al. 2008; Tosh et al. 2011; Carlson et al. 2015; Stephens et al. 2015; Athanasakis et al. 2015; Stevenson et al. 2016; Institute for Clinical and Economic Review 2017; Stevenson et al. 2017). At the start of the simulation, each patient is assigned a baseline HAQ score. Subsequently, the impact of the disease measured by the HAQ trajectory over time is modeled as a function of a sequence of treatments (Figure 2). In the absence of treatment, HAQ deteriorates at a certain rate as depicted by the dashed line in the figure. For each treatment in a treatment sequence, treatment is separated into two distinct phases: an initial phase of up to 6 months, consistent with data reported from randomized controlled trials (RCTs), and a maintenance phase thereafter until discontinuation.

7.1 Initial treatment phase

During the initial treatment phase HAQ is modeled as a change from baseline.

- **H1**: Treatment → ACR → HAQ
- **H2**: Treatment → ACR → EULAR → HAQ
- **H3**: Treatment → HAQ

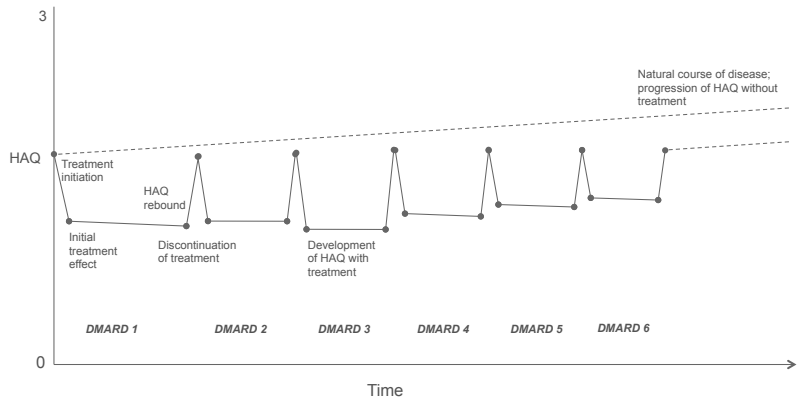


Figure 2: Model structure regarding development of HAQ with sequential biologic treatment

In **H1**, treatment influences HAQ through its effect on the American College of Rheumatology (ACR) response criteria, which is similar to the structure used in US based cost-effectiveness models (e.g. Carlson et al. 2015; Institute for Clinical and Economic Review 2017). ACR 20/50/70 response is defined as at least a 20/50/70% improvement. In the simulation, we convert these overlapping ACR categories to four mutually exclusive categories: no response (defined as less than 20% improvement), ACR 20% to <50% improvement, ACR 50% to <70% improvement, and ACR 70% improvement or greater. The rationale for using ACR response rather than HAQ directly is that the evidence base relating treatment to ACR response is larger than the evidence based relating treatment to HAQ. **H2** follows the National Institute for Health and Care Excellence (NICE) cost-effectiveness model (Stevenson et al. 2016, 2017) and models the effect of treatment on HAQ indirectly through its effect on ACR response and, in turn, the three categories of the European League Against Rheumatism (EULAR) response (no response, moderate response, or good response). Finally, since modeling the effect of treatment on HAQ through intermediary variables may mediate treatment response, in **H3**, treatment impacts HAQ directly.

Treatment switching during the initial treatment phase is modeled using 6 different pathways **S1-S6**.

- **S1:** Treatment → ACR → Switch
- **S2:** Treatment → ACR → ΔDAS28 → DAS28 → Switch
- **S3:** Treatment → ACR → ΔSDAI → SDAI → Switch
- **S4:** Treatment → ACR → ΔCDAI → CDAI → Switch
- **S5:** Treatment → ΔDAS28 → DAS28 → Switch
- **S6:** Treatment → ACR → EULAR → Switch

S1 follows a common approach where ACR non-responders discontinue treatment (e.g. Carlson et al. 2015; Institute for Clinical and Economic Review 2017). One drawback of this approach is that it is not consistent with current treat-to-target guidelines in the United States (Singh et al. 2016). In **S2-S5**, treatment switching consequently depends on disease activity (remission, low, moderate, high) (Anderson et al. 2012). In **S2-S4**, ACR response predicts the change in disease activity from baseline, which along with baseline disease activity, predicts absolute disease activity. Patients with moderate or high disease switch treatment while patients with low disease activity or in remission continue treatment. Disease activity is measured using either the Disease Activity Score with 28-joint counts (DAS28) (Prevoo et al. 1995), Simplified Disease Activity Index (SDAI) (Smolen et al. 2003; Aletaha and Smolen 2005), or the Clinical Disease Activity Index (CDAI) (Aletaha et al. 2005).

S5 is similar to **S2-S4**, but models the effect of treatment on changes in DAS28 directly, rather than indirectly through ACR response. We also aimed to model the direct effect of treatment on SDAI and CDAI, but sufficient clinical trial data are not available. Finally, since in the UK, the British Society for Rheumatology and the British Health Professionals in Rheumatology recommends using the EULAR response (Deighton et al. 2010), treatment switching in **S6** depends on EULAR response. In particular, following the NICE model, we assume that EULAR non-responders discontinue treatment while moderate and good responders continue treatment (Stevenson et al. 2016). The reasoning is that rules stipulated by NICE require a DAS28 improvement of more than 1.2 to continue treatment which is associated with moderate or good EULAR response.

Not all pathways **S1-S6** can be used with each of **H1-H3**. If **H1** is used, then **S1-S5** are available, but **S6** is not because EULAR response is not simulated. In **H2**, **S1-S6** are all available while in **H3** only **S5** can be used since ACR response is not simulated. The 12 possible combinations are outlined in Table 2.

Table 2: Model structures for initial treatment phase

	S1	S2	S3	S4	S5	S6
H1	1	2	3	4	5	-
H2	6	7	8	9	10	11
H3	-	-	-	-	12	-

Notes: Rows denote the pathway used to relate treatment to HAQ and columns denote the pathway used to determine treatment switching. Each number denotes a unique combination of pathways (i.e., 1 corresponds to H1 and S1, and 8 corresponds to H2 and S3) and the "-" denotes a combination of pathways that is not possible. There are 12 possible model structures for the initial treatment phase.

7.2 Maintenance phase

In the maintenance phase, the long-term progression of HAQ can be modeled in two ways. First, as is common in cost-effectiveness analyses (CEAs) of therapies for RA, HAQ is assumed to progress at a constant linear rate over time (see Tosh et al. 2011; Wailoo et al. 2008). However, since emerging evidence suggests that the rate of HAQ progression is non-linear and varies across patients (Gibson et al. 2015), our second scenario simulates HAQ progression using a latent class growth model (LCGM) (Norton et al. 2014) with 4 distinct HAQ trajectories and a rate of HAQ progression that decreases over time within each trajectory. Upon discontinuation of treatment, the HAQ score rebounds by a proportion of the improvement experienced at the end of the initial 6-month period with that treatment.

The duration of the maintenance phase (i.e., time to discontinuation of maintenance treatment) is simulated using parametric time-to-event distributions. When **S1** is used, time to treatment discontinuation is simulated using a single time-to-event curve because we have been unable to obtain curves stratified by ACR response categories. In contrast, when **S2-S5** are selected, the time-to-event curves are a function of disease activity level so patients with lower disease activity at the end of the initial treatment phase stay on treatment longer, on average. Likewise, when structure **S6** is used, the time-to-event distributions are stratified by EULAR response category and patients with good response at the end of the initial treatment phase tend to stay on treatment longer than patients with a moderate response. In each case, time to discontinuation can be simulated using one of 7 possible distributions (exponential, Weibull, Gompertz, gamma, log-logistic, lognormal, generalized gamma).

7.3 Adverse events

In line with Stevenson et al. (2016) the adverse events included in the model are limited to serious infections; we assume that only serious infections have a significant cost impact and increased risk over background rates to be meaningful to include (Ramiro et al. 2017). During the initial treatment phase, a patient immediately stops treatment if a serious infection occurs; during the maintenance phase, time on treatment depends on the sampled time to treatment discontinuation and a patient experiences a serious infection if the individual's sampled time to the adverse event is shorter than the sampled time to treatment discontinuation.

7.4 Mortality

Baseline HAQ scores (and changes in HAQ scores from baseline) are used to determine mortality relative to age/sex specific rates for the US general population (assumed to have a HAQ score of 0). Treatment, therefore, has an indirect effect on mortality through its effect on HAQ.

7.5 Utility

Individual HAQ scores at a particular point in time were also used to simulate EuroQol five dimensions questionnaire (EQ-5D) utility scores (0-1 range), which, in turn, are used to simulate quality-adjusted life-years (QALYs). However, since a number of different methods have been used to convert HAQ into utility, our model contains two different possible mapping algorithms. Our preferred algorithm is the Hernández-Alava et al. (2013) mixture model, which uses a much larger sample size than other statistical models and has been shown to have better predictive accuracy. Other algorithms are typically estimated using clinical trial data (e.g. Carlson et al. 2015; Stephens et al. 2015) and consequently have limited generalizability. The second utility algorithm available within our model is based on a linear regression analysis of real-world data by Wailoo et al. (2006) that has been used in a few previous CEAs (e.g. Wailoo et al. 2008; Institute for Clinical and Economic Review 2017).

7.6 Costs

Annual hospitalization days and productivity losses are simulated as a function of HAQ. Health sector costs considered in the models are related to drug acquisition and administration, adverse events, general management of RA, and hospitalization. Non-health sector costs are limited to work-related productivity loss.

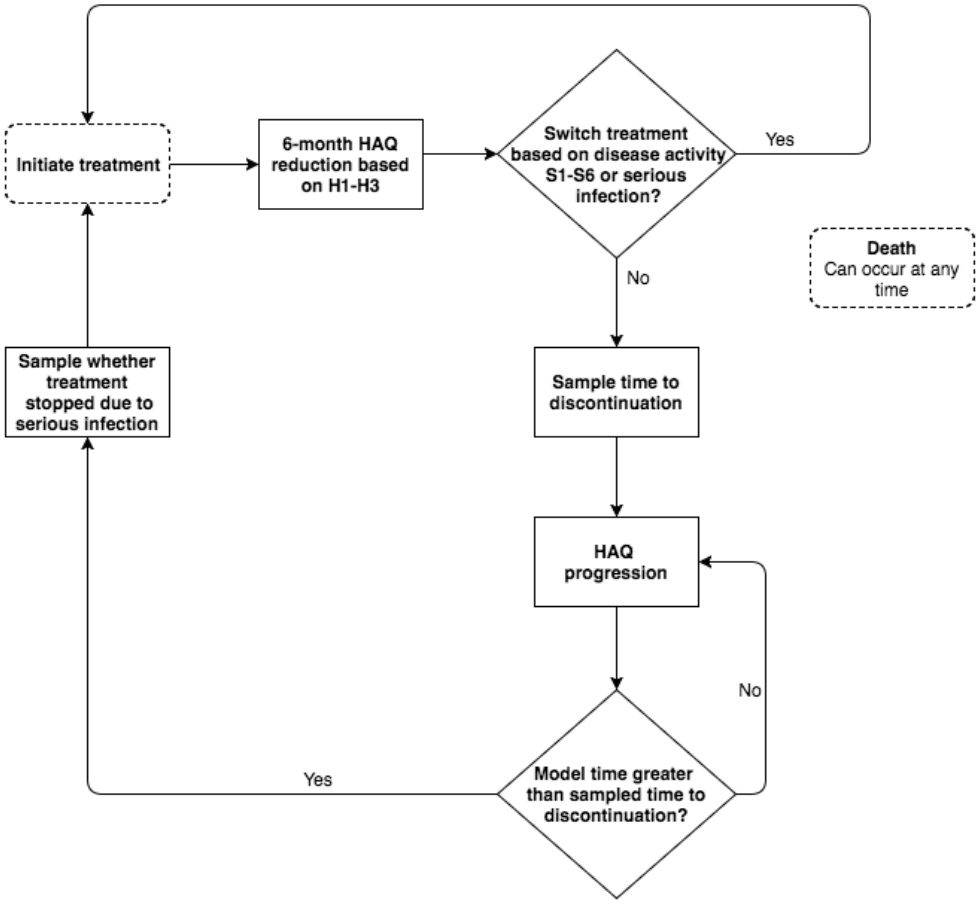


Figure 3: Flow diagram of the simulation for a single patient

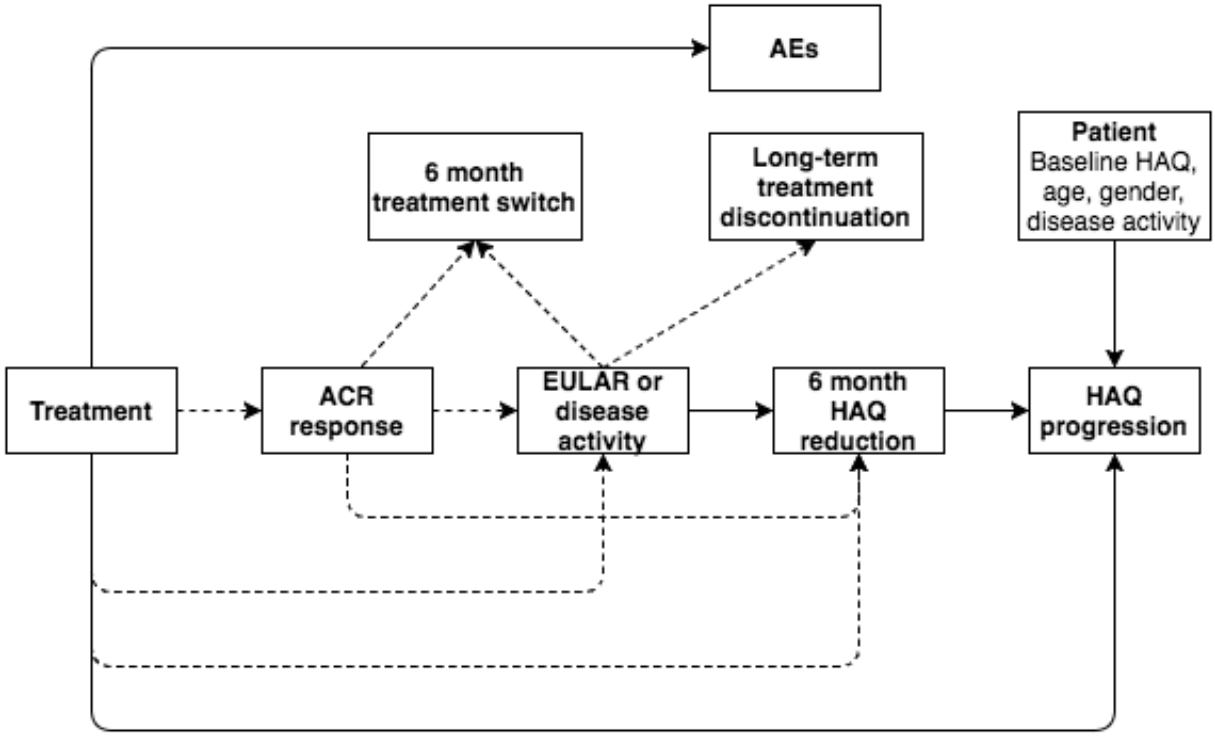
Notes: Rectangles represent "processes" determining the effect of treatment on disease progression, Diamonds represent "decisions" that determine whether a patient will switch to a new treatment. Dotted lines denote start of a new treatment or the end of the simulation.

7.7 Summary of simulation

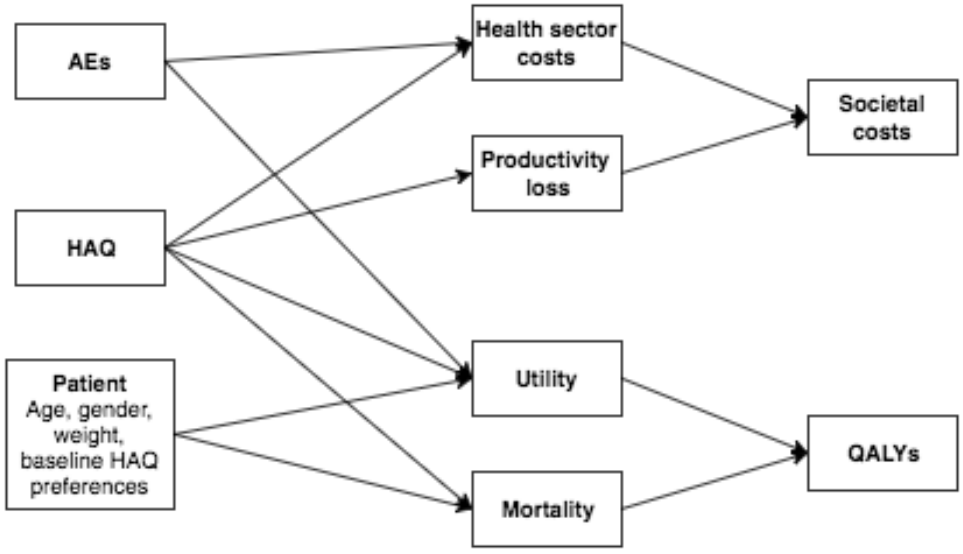
The flow diagram in Figure 3 describes the flow of a single patient through the simulation. The simulation runs for a patient's entire lifespan beginning with treatment initiation and ending in death. The rectangles in the figure represent "processes" determining the effect of treatment on disease progression and the diamonds represent "decisions" that determine whether a patient will switch to a new treatment.

The influence diagram in Figure 4 summarizes the assumed relationships among different variables in the model. Each arrow represents the direct effect of one parameter on another. Dashed lines represent relationships that depend on the structural assumptions used. Figure 4a focuses on the effect of treatment on disease progression and adverse events while Figure 4b looks at the relationships between the health and cost outcome variables.

The model accounts for patient heterogeneity in two ways. First, baseline event rates vary across patients by both observable and unobservable factors. For example, long-term HAQ progression, mortality, and utility depend on patient specific variables including age, gender, and baseline disease level. Moreover, unobserved differences in long-term HAQ progression and utility across patients are



(a) Treatment effects



(b) Long-term model outcomes

Figure 4: Influence diagram outlining structural relationships

Notes: ACR: American College of Rheumatology; EULAR: European League Against Rheumatism; HAQ: Health Assessment Questionnaire: AEs: adverse events; QALYs: quality-adjusted life-years; WTP: willingness to pay. Disease activity refers to the Disease Activity Score with 28-joint counts (DAS28), the Simplified Disease Activity Index (SDAI), or the Clinical Disease Activity Index (CDAI).

modeled using mixture models. Second, relative treatment effects for ACR response, the change in HAQ at 6 months, and the change in DAS28 at 6 months, can be modeled as a function of explanatory variables in the R package.

7.8 Model outcomes

7.8.1 Benefits, costs, and risks

The model simulates the health outcomes, costs, and risks associated with treatment. Depending on the model structure, model outcomes include the following:

- **Clinical outcomes during initial treatment phase**: ACR response, EULAR response, DAS28, SDAI, CDAI
- **Long-term clinical outcomes**: HAQ, QALYs
- **Adverse events**: number of serious infections
- **Health care sector costs**: drug acquisition and administration costs, general management and monitoring costs, adverse event costs, hospitalization costs
- **Non-health care sector costs**: productivity losses

7.8.2 Outcomes for value assessment

If CEA is used for value assessment, then the value of treatment is estimated using the NMB, as described in Section 3.1. CEA from a societal perspective would include productivity losses while analyses from a health care sector perspective would not.

Any combination of simulated model outcomes can be used for MCDA. In IVI's web interfaces, the MCDA is currently based on the following criteria: (i) QALYs, (ii) total health care sector costs, (iii) productivity losses, (iv) number of serious infections, (v) route of administration (oral/injection/infusion) and (vi) time the medication has been on the market. We measure performance for each route of administration by calculating the percentage of total life-years that were spent using that particular route of administration. If a combination therapy is used during the treatment sequence, we allocate time equally among all routes of administration within the combination therapy (i.e., during a time period in which tofacitinib citrate is used with methotrexate, we allocate half of the time to oral admnistration and half to admnistration by injection). Performance on the time since the medication has been on the market criterion is a weighted average of time since FDA approval for each treatment in a treatment sequence, where weights are equal to the number of life-years spent using a particular treatment within the sequence. In the web interfaces users can input their own weights for each of the criteria, but it is important to note that we have not conducted the surveys required to elicit weights in a representative sample of patients.

When analyzing value to healthy individuals—rather than sick patients—we use the framework described in Section 4. Following Lakdawalla et al. (2015) we calculate annual value for patients (e.g., benefits to an insurance enrollee during a plan year) by annualizing lifetime health gains (i.e., QALYs) and costs (see Appendix H for more details). To calculate the conventional value of a treatment to a healthy individual (i.e., $\pi(k \cdot dh - dp)$ from Equation 5), we estimate dh using annualized incremental QALYs, dp using annualized incremental costs, k using willingness to pay thresholds, and π as the probability of obtaining RA within the next year.

8 Source data and parameter estimation

8.1 Treatment effects at 6 months

The effect of treatment on ACR response, DAS28, and HAQ at 6 months for tDMARD naive patients are estimated using Bayesian network meta-analyses (NMA) of published randomized controlled trials (RCTs). Primary outcomes were ACR response, change in DAS28 from baseline at 6 months, and the change in HAQ from baseline at 6 months. Results from the NMA are shown in Table 3. Details of the systematic literature review and the statistical methodology are provided in the Appendix (Section I.2).

Its important to note that treatment effects for each tDMARD were estimated relative to cDMARDs and then applied to the average response for patients using cDMARDs. A limitation of our current approach is that that the average response for patients using cDMARDs is estimated using data from the clinical trials include in the NMA, and may not reflect outcomes seen in routine practice. Future versions of the model could consider using real-world data instead of clinicial trial evidence to estimate this average response.

Given that there is limited evidence that treatment effects vary across patients in the published literature, treatment response at 6 months for a given treatment does not vary according to patient characteristics. Nonetheless, in our R package, treatment effects for each simulated patient can be modeled as a function of any variables chosen by the user. Our approach to modeling treatment effect heterogeneity is described in Section I.2.

Treatment effects for tDMARD experienced patients are reduced by multiplying treatment effects for tDMARD naive patients by a constant γ. Based on evidence reported in Carlson et al. (2015), we assume that γ is uniformly distributed and ranges between .75 and .92, implying that (rounding up) the average value of γ is .84. In other words, reductions in DAS28 and HAQ scores for tDMARD experienced patients are, on average, 84% of the reduction in DAS28 and HAQ scores for tDMARD naive patients, and an ACR response of 60/40/20 for tDMARD naive patients would, on average, be reduced to 50/33/16 for tDMARD experienced patients.

In the simulation, treatment response depends on the line of therapy and whether a patient is tDMARD naive or tDMARD experienced at baseline. For tDMARD naive patients, first line treatment response is based on the NMA results for tDMARD naive patients while response for all other treatments in a treatment sequence is reduced using the constant γ. For tDMARD experienced patients, treatment response is reduced using γ at each line of therapy including the first line. One limitation of this approach is that we are unable to model the relationship between line of therapy and γ; that is, treatment response for a patient who has failed at least one biologic is assumed to be reduced by, on average, .84, regardless of line of therapy.

8.2 Treatment switching at 6 months

The data required to determine treatment switching at 6 months depends on the selected model structure. If **S1** is selected, then treatment switching depends on the simulated ACR response; likewise, if **S5** is selected, then treatment switching depends on the simulated level of DAS28 at 6 months. When **S2-S4** are used, treatment switching is determined by the relationship between ACR response and the change in disease activity, and in **S6**, switching is based on the relationship between ACR response and EULAR response. Details of the mapping between ACR response and change in disease activity and between ACR response and EULAR response are provided below.

Table 3: Network meta-analysis estimates of ACR response, change in DAS28, and change in HAQ for tDMARD naive patients

	ACR response				
	ACR20	ACR50	ACR70	ΔDAS28	ΔHAQ
cDMARDs	0.291 (0.277, 0.306)	0.120 (0.111, 0.130)	0.040 (0.036, 0.044)	-0.992 (-1.046, -0.937)	-0.233 (-0.275, -0.189)
ABT IV + MTX	0.636 (0.546, 0.720)	0.394 (0.306, 0.485)	0.199 (0.139, 0.269)	-2.331 (-2.541, -2.123)	-0.464 (-0.590, -0.347)
ABT SC + MTX	0.632 (0.486, 0.760)	0.392 (0.258, 0.537)	0.200 (0.109, 0.311)	-2.282 (-2.617, -1.985)	-0.452 (-0.624, -0.279)
ADA + MTX	0.588 (0.495, 0.669)	0.346 (0.263, 0.426)	0.166 (0.113, 0.222)	-2.183 (-2.494, -1.871)	-0.554 (-0.673, -0.431)
ADA	0.501 (0.334, 0.645)	0.271 (0.145, 0.399)	0.120 (0.052, 0.202)	-1.377 (-1.863, -0.875)	-0.395 (-0.556, -0.242)
ADA BWWD + MTX	0.585 (0.369, 0.791)	0.352 (0.169, 0.574)	0.175 (0.063, 0.347)	-2.234 (-2.729, -1.744)	-
ANA + MTX	0.460 (0.243, 0.683)	0.243 (0.092, 0.440)	0.105 (0.028, 0.234)	-	-0.343 (-0.499, -0.197)
BCT	0.599 (0.172, 0.924)	0.389 (0.059, 0.794)	0.218 (0.016, 0.590)	-	-
BCT + MTX	0.554 (0.345, 0.760)	0.321 (0.154, 0.535)	0.153 (0.055, 0.308)	-	-
CZP	0.581 (0.286, 0.832)	0.355 (0.116, 0.634)	0.181 (0.038, 0.406)	-	-0.546 (-0.919, -0.157)
CZP + MTX	0.737 (0.639, 0.821)	0.507 (0.394, 0.616)	0.289 (0.198, 0.390)	-3.006 (-3.315, -2.713)	-0.619 (-0.723, -0.518)
ETN	0.598 (0.493, 0.706)	0.356 (0.257, 0.469)	0.173 (0.109, 0.256)	-2.502 (-2.974, -1.999)	-0.381 (-0.577, -0.176)
ETN + MTX	0.584 (0.466, 0.690)	0.343 (0.240, 0.453)	0.165 (0.100, 0.242)	-2.567 (-2.911, -2.226)	-0.569 (-0.748, -0.390)
ETN SZZS + MTX	0.499 (0.263, 0.742)	0.276 (0.104, 0.511)	0.126 (0.033, 0.294)	-2.564 (-3.117, -1.968)	-
ETN YKRO + MTX	0.612 (0.379, 0.820)	0.378 (0.176, 0.618)	0.194 (0.065, 0.390)	-2.667 (-3.220, -2.141)	-0.474 (-0.817, -0.136)
GOL + MTX	0.615 (0.482, 0.744)	0.375 (0.252, 0.513)	0.187 (0.106, 0.292)	-2.457 (-3.002, -1.937)	-0.578 (-0.685, -0.464)
IFX + MTX	0.585 (0.481, 0.701)	0.344 (0.253, 0.460)	0.165 (0.107, 0.253)	-1.922 (-2.363, -1.465)	-0.446 (-0.619, -0.277)
IFX QBTX + MTX	-	-	-	-	-0.474 (-0.743, -0.207)
Placebo	0.183 (0.088, 0.299)	0.065 (0.024, 0.125)	0.019 (0.005, 0.042)	-0.545 (-1.107, -0.023)	-0.062 (-0.254, 0.142)
RTX	0.486 (0.276, 0.713)	0.264 (0.113, 0.477)	0.118 (0.036, 0.261)	-1.735 (-2.296, -1.143)	-
RTX + MTX	0.560 (0.422, 0.704)	0.323 (0.205, 0.466)	0.152 (0.080, 0.252)	-1.978 (-2.283, -1.653)	-0.479 (-0.896, -0.094)
SAR	0.645 (0.373, 0.851)	0.415 (0.175, 0.664)	0.223 (0.064, 0.440)	-2.274 (-2.935, -1.616)	-0.581 (-0.829, -0.339)
SAR + MTX	0.617 (0.423, 0.801)	0.381 (0.206, 0.591)	0.195 (0.080, 0.364)	-	-0.483 (-0.630, -0.328)
SSZ + HCQ + MTX	0.519 (0.279, 0.752)	0.294 (0.112, 0.524)	0.138 (0.037, 0.299)	-2.311 (-2.912, -1.747)	-0.505 (-0.782, -0.216)
TCZ	0.685 (0.554, 0.798)	0.447 (0.313, 0.584)	0.241 (0.142, 0.358)	-2.779 (-3.142, -2.421)	-0.482 (-0.618, -0.351)
TCZ + MTX	0.667 (0.562, 0.761)	0.427 (0.321, 0.535)	0.224 (0.148, 0.313)	-2.928 (-3.173, -2.682)	-0.473 (-0.575, -0.373)
TOF + MTX	0.586 (0.453, 0.704)	0.346 (0.229, 0.466)	0.167 (0.093, 0.253)	-1.937 (-2.464, -1.469)	-0.637 (-0.799, -0.484)
TOF	0.498 (0.332, 0.684)	0.271 (0.144, 0.441)	0.121 (0.050, 0.235)	-1.702 (-2.161, -1.270)	-0.546 (-0.730, -0.367)
UPA + MTX	0.569 (0.369, 0.764)	0.335 (0.168, 0.540)	0.162 (0.062, 0.313)	-2.213 (-2.631, -1.784)	-0.587 (-0.820, -0.354)

Notes: ACR20/50/70 categories are the probability of at least a 20/50/70% improvement. 95% credible intervals are in parentheses. Estimates are based on 1,000 random draws of the NMA parameters. ΔDAS28 and ΔHAQ are changes in the DAS28 and HAQ score from their baseline scores respectively; negative numbers denote reductions in baseline values. cDMARDs = conventional disease-modifying antirheumatic drugs; MTX = methotrexate; ABT IV = abatacept intravenous; ABT SC = abatacept subcutaneous; ADA = adalimumab; ADA BWWD = adalimumab-bwwd (biosimilar Samsung Bioepis); ANA = anakinra; BCT = baricitinib; CZP = certolizumab pegol; ETN = etanercept; ETN SZZS = etanercept-szzs (biosimilar Sandoz); ETN YKRO = etanercept-ykro (biosimilar Samsung Bioepis); GOL = golimumab; HCQ = hydroxychloroquine sulfate; IFX = infliximab; IFX QBTX = infliximab-qbtx (biosimilar Pfizer); RTX = rituximab; SAR = sarilumab; SSZ = sulfazalazine; TCZ = tocilizumab; TOF = tofacitinib; UPA = upadacitinib; ACR = American College of Rheumatology.

8.2.1 ACR response and change in disease activity

There are currently no established mappings between mutually exclusive ACR response categories and DAS28, SDAI, or CDAI (Madan et al. 2015). However, Aletaha and Smolen (2005) provides evidence on the relationship between overlapping ACR response categories (ACR 20/50/70) and mean changes in each of the three disease activity measures. Results are reported for three cohorts—the Leflunomide datasets, the inception cohort, and the routine cohort—with 1,839, 91, and 279 patients, respectively. We transformed mean changes by overlapping ACR response categories to mean changes by mutually exclusive ACR response categories by using the number of patients in each mutually exclusive ACR response category as described in Appendix C. Smolen et al. (2003) provided the number of patients in each ACR response category in the Leflunomide dataset and Aletaha et al. (2005) provided the number of patients in the inception cohort. Mean changes in disease activity in each mutually exclusive ACR response category are shown in Table 4. However, note that the referenced publications did not report mean outcomes, so we were unable to generate standard errors for the estimates. We consequently assume allow the estimates to vary by 20% in either direction.

Table 4: Relationship between ACR response and change in disease activity measures

ACR response	Mean change at 6 months			
	Leflunomide dataset	Inception cohort		
	SDAI	SDAI	CDAI	DAS28
<20	0.000	0.000	0.000	0.000
20 to <50	-30.284	-13.700	-11.300	-1.550
50 to <70	-35.234	-14.882	-12.873	-1.543
≥70	-41.000	-30.100	-27.600	-3.310

Sources: Aletaha and Smolen (2005), Smolen et al. (2003), and Aletaha et al. (2005)

We did not include estimates from the routine cohort for two reasons. First, we were unable to find information on the number of patients in each ACR response category. Second, patients in the routine cohort had considerably lower disease activity levels (Aletaha and Smolen 2005; Aletaha et al. 2005) and our default population (see Section 5) consists of patients with high disease activity at baseline. Mean DAS28 in the inception cohort and routine cohort were 5.62 and 4.09, respectively, while the mean DAS 28 ranged from 6.3 to 7 across the clinical trials making up the Leflunomide dataset.

8.2.2 ACR response and change in EULAR response

ACR responses were translated into EULAR response probabilities based on evidence of their relationship reported in Stevenson et al. (2016) and obtained from the US Veterans Affairs Rheumatoid Arthritis (VARA) registry (Table 5).

8.3 Change in HAQ at 6 months

In model structures including **H1**, the impact of treatment on changes in HAQ at 6 months is modeled by first estimating the effect of treatment on ACR response and then mapping ACR response to a change in HAQ. As in Institute for Clinical and Economic Review (2017), ACR responses from the NMA were translated into HAQ scores based on evidence from the adalimumab

Table 5: Relationship between ACR response and EULAR response

	EULAR response		
ACR response	None	Moderate	Good
<20	755	136	57
20 to <50	4	27	26
50 to <70	2	2	10
≥70	0	2	2

Notes: Obtained from the US Veterans Affairs Rheumatoid Arthritis (VARA) registry by Stevenson et al. (2016). The VARA registry is a multicentre, US database of veterans age 19 and older. Each cell represents the number of patients in the database in a given category.

monotherapy for treatment of rheumatoid arthritis (ADACTA) trial reported in Carlson et al. (2015) (Table 6).

Table 6: Relationship between ACR response and change in HAQ at 6 months

	HAQ change	
ACR response	Mean	Standard error
<20	-0.11	0.06765
20 to <50	-0.44	0.05657
50 to <70	-0.76	0.09059
≥70	-1.07	0.07489

Source: Carlson et al. (2015)

The relationship between EULAR response and HAQ is based on analyses conducted by Stevenson et al. (2016) using the BSRBR database. Their analysis is based on predictions from a mixture model with covariates set to sample means. Moderate and good EULAR responses are associated with -0.317 (SE = 0.048) and -0.672 (SE = 0.112) changes in HAQ scores respectively (Table 7).

Table 7: Relationship between EULAR response and change in HAQ at 6 months

EULAR response	Mean	Standard error
None	0.000	0.000
Moderate	-0.317	0.048
Good	-0.672	0.112

Notes: Based on an analysis of the BSRBR database by Stevenson et al. (2016).

Table 8 compares the impact of treatment on HAQ when using **H1-H3**. Results were estimated by simulating 1,000 patients for 6 months and randomly sampling 1,000 parameter sets. For each randomly sampled parameter set, we calculated the average decrease in HAQ at 6 months across the 1,000 patients. Estimates reported in the table are the mean and 95% credible interval of the mean decrease in HAQ at 6 months. To maintain consistency across **H1-H3**, we did not allow HAQ scores for patients who might have otherwise switched treatments accoring to **S1-S6** to rebound back to their baseline levels (i.e., levels at the start of the simulation) at the end of the 6 month period.

Table 8: Simulated mean change in HAQ at 6 months under different model structures

	H1	H2	H3
cDMARDs	-0.26 (-0.36, -0.17)	-0.20 (-0.25, -0.15)	-0.23 (-0.28, -0.19)
ABT IV + MTX	-0.51 (-0.62, -0.40)	-0.34 (-0.46, -0.24)	-0.46 (-0.59, -0.34)
ABT SC + MTX	-0.51 (-0.65, -0.37)	-0.34 (-0.46, -0.24)	-0.46 (-0.63, -0.29)
ADA + MTX	-0.47 (-0.58, -0.36)	-0.32 (-0.42, -0.23)	-0.56 (-0.68, -0.43)
ADA	-0.40 (-0.54, -0.26)	-0.29 (-0.39, -0.19)	-0.39 (-0.56, -0.23)
ADA BWWD + MTX	-0.46 (-0.65, -0.27)	-0.32 (-0.45, -0.20)	-
ANA + MTX	-0.37 (-0.56, -0.20)	-0.27 (-0.40, -0.17)	-0.34 (-0.49, -0.20)
BCT	-0.52 (-0.89, -0.18)	-0.34 (-0.53, -0.15)	-
BCT + MTX	-0.45 (-0.62, -0.28)	-0.31 (-0.43, -0.21)	-
CZP	-0.47 (-0.74, -0.24)	-0.32 (-0.47, -0.18)	-0.55 (-0.93, -0.16)
CZP + MTX	-0.61 (-0.73, -0.49)	-0.38 (-0.52, -0.28)	-0.62 (-0.73, -0.51)
ETN	-0.47 (-0.59, -0.36)	-0.32 (-0.43, -0.23)	-0.39 (-0.57, -0.19)
ETN + MTX	-0.46 (-0.59, -0.35)	-0.32 (-0.43, -0.23)	-0.56 (-0.73, -0.39)
ETN SZZS + MTX	-0.40 (-0.62, -0.21)	-0.29 (-0.43, -0.18)	-
ETN YKRO + MTX	-0.49 (-0.71, -0.29)	-0.33 (-0.48, -0.21)	-0.47 (-0.78, -0.14)
GOL + MTX	-0.49 (-0.63, -0.36)	-0.33 (-0.45, -0.23)	-0.58 (-0.69, -0.46)
IFX + MTX	-0.47 (-0.59, -0.34)	-0.32 (-0.44, -0.23)	-0.45 (-0.62, -0.27)
IFX QBTX + MTX	-	-	-0.47 (-0.74, -0.20)
Placebo	-0.20 (-0.32, -0.08)	-0.16 (-0.22, -0.11)	-0.06 (-0.26, 0.13)
RTX	-0.40 (-0.58, -0.21)	-0.28 (-0.42, -0.17)	-
RTX + MTX	-0.45 (-0.58, -0.31)	-0.31 (-0.42, -0.21)	-0.48 (-0.89, -0.09)
SAR	-0.52 (-0.76, -0.30)	-0.34 (-0.49, -0.22)	-0.57 (-0.84, -0.33)
SAR + MTX	-0.50 (-0.68, -0.32)	-0.34 (-0.47, -0.22)	-0.49 (-0.63, -0.34)
SSZ + HCQ + MTX	-0.42 (-0.65, -0.23)	-0.30 (-0.44, -0.18)	-0.49 (-0.75, -0.21)
TCZ	-0.56 (-0.71, -0.44)	-0.36 (-0.49, -0.26)	-0.48 (-0.62, -0.34)
TCZ + MTX	-0.54 (-0.66, -0.42)	-0.36 (-0.47, -0.25)	-0.47 (-0.57, -0.37)
TOF + MTX	-0.47 (-0.61, -0.33)	-0.32 (-0.44, -0.23)	-0.64 (-0.81, -0.48)
TOF	-0.40 (-0.57, -0.25)	-0.29 (-0.41, -0.19)	-0.55 (-0.73, -0.35)
UPA + MTX	-0.45 (-0.65, -0.28)	-0.31 (-0.45, -0.20)	-0.58 (-0.82, -0.36)

Notes: **H1**, **H2**, and **H3** are the Treatment → ACR → HAQ, Treatment → ACR → EULAR → HAQ, and Treatment → HAQ pathways respectively. 95% credible intervals are in parentheses. Estimates are based on 6-month simulations of 1,000 patients and 1,000 parameters sets for each therapy. ΔHAQ denotes a change in the HAQ score at 6 months from baseline; a negative value indicates a reduction in the HAQ score. Mean ΔHAQ is calculated for each parameter set by averaging across the 1,000 patients. cDMARDs = conventional disease-modifying antirheumatic drugs; MTX = methotrexate; ABT IV = abatacept intravenous; ABT SC = abatacept subcutaneous; ADA = adalimumab; ADA BWWD = adalimumab-bwwd (biosimilar Samsung Bioepis); ANA = anakinra; BCT = baricitinib; CZP = certolizumab pegol; ETN = etanercept; ETN SZZS = etanercept-szzs (biosimilar Sandoz); ETN YKRO = etanercept-ykro (biosimilar Samsung Bioepis); GOL = golimumab; HCQ = hydroxychloroquine sulfate; IFX = infliximab; IFX QBTX = infliximab-qbtx (biosimilar Pfizer); RTX = rituximab; SAR = sarilumab; SSZ = sulfazalazine; TCZ = tocilizumab; TOF = tofacitinib; UPA = upadacitinib; ACR = American College of Rheumatology.

Estimates for **H1** and **H3** are generally similar but treatment response is considerably smaller when using **H2**. This suggests that the additional mapping between ACR response and EULAR response attenuates treatment response. Given these varying estimates of the change in HAQ during the initial treatment phase and the impact of HAQ on other important outcomes within the model including utility and health care costs, the choice of **H1-H3** (and in particular **H2** vs. **H1/H3**)

appears to have important consequences for value assessment.

8.4 HAQ progression in the absence of tDMARD treatment

The natural course of HAQ progression in the absence of tDMARDs develops over time according to an estimated natural course for patients remaining on cDMARDs or following discontinuation of the last tDMARD of the sequence (i.e., on NBT). The natural course of HAQ can either be assumed to change at a constant linear rate or be modeled using a LCGM that accounts for non-linear progression and heterogeneity across patients.

8.4.1 Constant linear rate of progression

The rate of progression in the linear case is based on the observational study by Wolfe and Michaud (2010). They assessed the development of HAQ over time at six month intervals for up to 11 years among 3,829 RA patients who switched from non-biologic treatment to biologic treatment and participated in the National Data Bank for Rheumatic Diseases (NDB) longitudinal study of RA outcomes. The annual HAQ progression rate prior to biologic therapy was 0.031 (95% confidence interval (95%CI): 0.026 to 0.036) and is assumed to reflect the course of progression of HAQ in the absence of tDMARDs.

Based on the same data, Michaud et al. (2011) reported overall and age-specific specific HAQ progression rates. The differences between the overall and age specific rates are as follows: <40: -0.020 (95%CI: -0.0223 to -0.0177); 40-64: -0.008 (95%CI: -0.0101 to -0.0059); $\geq$ 65 0.017 (95%CI: 0.0136 to 0.0204). These estimates are applied to the overall progression rate of 0.031 to obtain age specific HAQ progression rates (see Section D.1).

Table 9: Annual linear progression of HAQ in the absence of tDMARDs beyond 6 months

		95% CI		
	Estimate	Lower	Upper	Reference
Overall progression rate				
MTX or non-biologic treatment	0.031	0.026	0.036	Wolfe and Michaud (2010)
Change in overall progression rate by age				
<40	-0.020	-0.028	-0.012	Michaud et al. (2011)
40-64	-0.008	-0.010	-0.006	Michaud et al. (2011)
65+	0.017	0.013	0.021	Michaud et al. (2011)

Notes: 95% confidence intervals are calculated using a normal distribution. Confidence intervals for changes in HAQ progression rates by age assume no covariance between the overall progression rate and the age-specific rates reported by Michaud et al. (2011).

8.4.2 Latent class growth model

We also model the rate of HAQ progression in the absence of tDMARDs using a mixture model approach that has increasingly been used to model HAQ progression over time (Stevenson et al. 2016; Norton et al. 2013, 2014). These models suggest that different subgroups have distinct HAQ trajectories and that the rate of worsening of HAQ progression decreases over time. We use the LCGM estimated by Norton et al. (2014) and since we aim to model trajectories for cDMARDs and NBTs we chose the specification based on data from the Early Rheumatoid Arthritis Cohort Study (ERAS) cohort, which has a high percentage of patients receiving methotrexate and a very small percentage receiving biologics. Complete details of the LCGM are provided in Section D.2.

The Norton et al. (2014) LCGM determined that there are four classes of patients and thus four distinct HAQ trajectories. The probability of class membership depends on 7 variables: age, gender, DAS28, disease duration, rheumatoid factor, the ACR 1987 criteria for RA, and a measure of socioeconomic status. Age, gender, and the DAS28 are relevant to the way the population is defined within our model (see Section 5) and are therefore important determinants of the HAQ trajectory. Other variables (disease duration, rheumatoid factor, ACR criteria, and socioeconomic status) are not defined within our population. We consequently set disease duration (8.2 months), rheumatoid factor (0.73), and the socioeconomic status variable (0.49) equal to their mean values with the ERAS cohort. The ACR criteria was set to 1.

HAQ trajectories (in levels) by class are shown Figure 5. The dotted lines plot observed mean values. There are clear distinguishable classes as both the level of the HAQ score and its slope vary between groups. Norton et al. (2014) refer to the groups as "low", "moderate", "high", and "severe" groups, in order from the lowest to highest HAQ scores. The observed trends for the low, medium, and high groups follow a J-shaped pattern with a sharp drop following treatment initiation and an upward slope thereafter, while the severe group experiences persistently high HAQ scores. Since our model separates the initial treatment phase from the maintenance phase, we are only concerned with HAQ progression following the initial drop. As in Stevenson et al. (2016), we consequently only predict values from year 2 onward. The fitted values are the solid upward sloping lines in the plot.

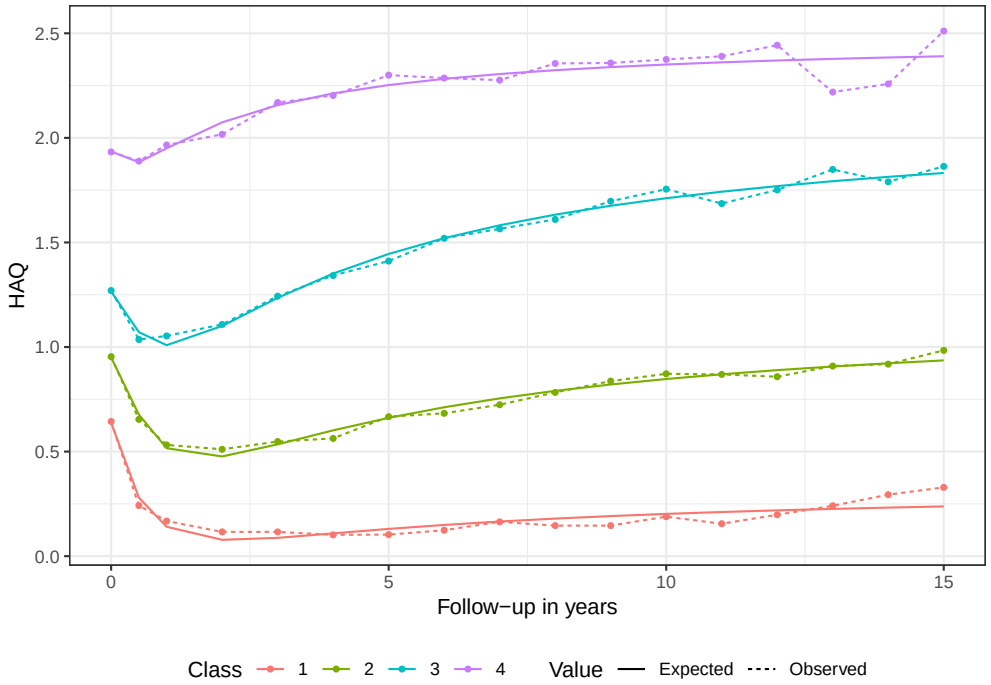


Figure 5: Observed and predicted HAQ trajectories in the ERAS dataset from the latent class growth model

Notes: The first three data points corresponds to years 0, 0.5, and 1, respectively; all other data points are spaced 1 year apart.

An important question for modeling disease progression in RA is how the rate of progression within each class in the LCGM compares to a constant linear trajectory. We examine this question in

Figure 6, which compares yearly rates of changes in HAQ using the LCGM and with constant annual rates of change (0.031 per year) based on the Wolfe and Michaud (2010) analysis. The LCGM was simulated over 30 years and differences between year t and year $t-1$ were used to assess changes in HAQ score from one year to the next.

In the moderate, high, and severe groups the rate of HAQ progression is higher initially in the LCGM than in the Wolfe and Michaud (2010) analysis; however, the LCGM modeled rate of HAQ progression declines over time and eventually begins to approach zero. In the low group, HAQ increases at a rate less than 0.031 per year and the rate of increase declines over time.

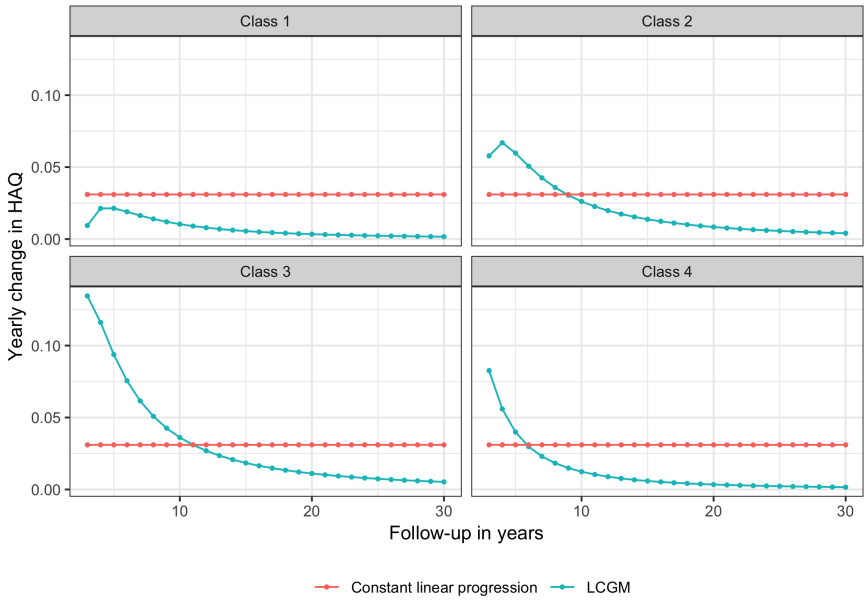


Figure 6: A comparison of predicted yearly changes in HAQ between a latent class growth model and constant linear progression from year 2 onwards

8.5 HAQ trajectory with tDMARD maintenance treatment

Based on the NDB longitudinal study, Wolfe and Michaud (2010) estimated the overall annual HAQ progression rate among RA patients who had switched to biologic treatment at -0.001 (95CI: -0.004 to 0.002). In a separate analysis, also based on NDB data, Michaud et al. (2011) reported annual HAQ progression rates by treatment adjusted for baseline HAQ score, age, sex, education, smoking, BMI, comorbidity, and RA onset. The average HAQ rate among patients on a biologic was -0.001 as well, which instills confidence that the reported HAQ progression rates for different biologics as reported by Michaud et al. (2011) can be directly compared with the overall annual HAQ progression rate of 0.031 reported by Wolfe and Michaud (2010). Accordingly, biologic specific HAQ progression rates by Michaud et al. (2011) are used in the model. For tDMARD treatments evaluated in the model for which no HAQ progression rate was reported by Michaud et al. (2011), the overall biologic rate of -0.001 is used.

8.6 Duration of maintenance treatment

Time to treatment discontinuation in the maintenance phase depends on the pathway (**S1-S6**) used to model treatment switching. If **S1** is selected, a single treatment discontinuation curve based on

an analysis from the CORRONA database is used for all patients. In **S2-S5**, time to treatment discontinuation is stratified by the level of disease activity, and in **S6** treatment duration depends on EULAR response.

8.6.1 Treatment duration in the US

We based our estimates of treatment duration during the maintenance phase for patients in the US with analyses of the CORRONA database (Strand et al. 2013). The analysis sample consisted of 6,209 patients age 18 or older treated between 2002 and 2011 receiving either TNF inhibitors or other bDMARDs. The mean age was 57.6 years, 43% of patients were biologic naive, the mean CDAI was 16, and just over 26% of patients had high disease activity (CDAI $\geq$ 22).

7 parametric survival models (exponential, Weibull, Gompertz, gamma, log-logistic, lognormal, and generalized gamma) were estimated on individual patient data reconstructed from a Kaplan-Meier curve from the CORRONA analysis using the algorithm developed in Guyot et al. (2012). We compared fit using the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC). The generalized gamma had the lowest AIC and BIC, so we consider it to be the preferred model. A plot of of the generalized gamma distribution against the Kaplan-Meier curve is shown in Figure 7. As can be seen in the plot, the shape of the survival curve estimated using a generalized gamma distribution tracks the Kaplan-Meier curve closely.

Table 10: AIC and BIC for parametric models of treatment duration from the CORRONA database

Distribution	AIC	BIC
Exponential	33,240	33,246
Weibull	33,182	33,196
Gompertz	32,963	32,977
Gamma	33,222	33,236
Log-logistic	32,848	32,861
Lognormal	32,650	32,663
Generalized gamma	32,507	32,527

We considered estimating separate time to discontinuation curves for each treatment, but did not for a number of the reasons cited in Stevenson et al. (2016). The majority of the literature focuses on anti-TNFs (e.g., infliximab, etanercept, and adalimumab) (e.g. Gomez-Reino and Carmona 2006; Yazici et al. 2009; Pan et al. 2009), which makes it difficult to estimate discontinuation curves for the other treatments. Furthermore, studies comparing rates of discontinuation across treatments tend to be observational because clinical trials are of short duration and do not reflect real-world patient populations. However, although observational studies provide accurate predictions on time to discontinuation, it is difficult to avoid bias from confounding when estimating differences across treatments because patients are not randomized into treatment and control groups (Souto et al. 2015) .

We also lack data on treatment duration for patients on cDMARDs. Following Stevenson et al. (2016), we assume that, conditional on continuing treatment at 6 months, treatment duration for tDMARDs is applicable to treatment duration for cDMARDs. This is, in turn, based on the assumption that cDMARDs are not likely to be more toxic than biologics used in combination with cDMARDs.

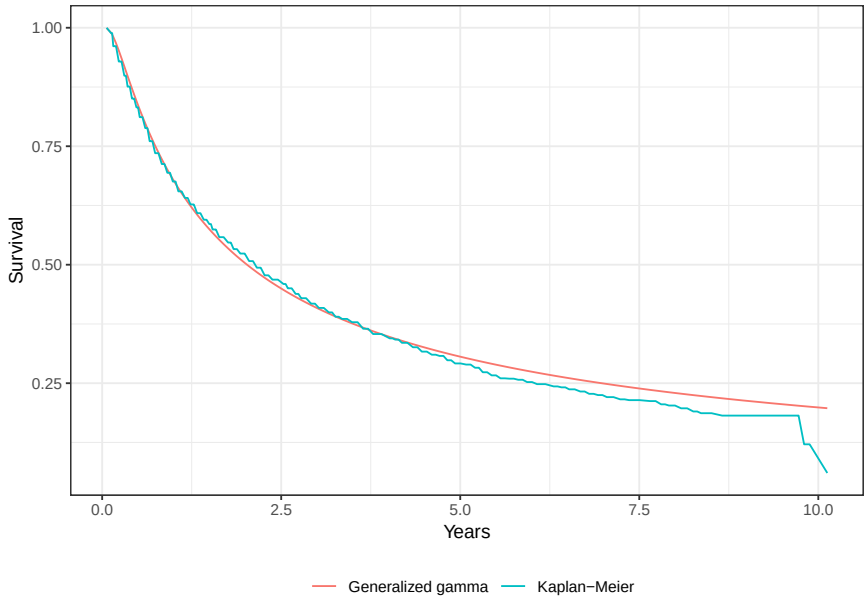


Figure 7: Generalized gamma and Kaplan-Meier time to treatment discontinuation curves using reconstructed individual patient data from the CORRONA database

8.6.2 Treatment duration by disease activity level

When **S2-S5** are selected, treatment duration is stratified by the level of disease activity. Since patients in the CORRONA database tended to have moderate disease activity (mean CDAI = 16), we use the CORRONA survival curve to model treatment duration for patients with moderate disease activity. We adjust this curve for patients in remission or low disease activity using the odds ratios reported in Zhang et al. (2011), which imply that patients in remission or with low disease activity have .52 times the odds of stopping treatment as patients with moderate disease activity. In particular, we adjust the probability of treatment failure at each point in time using the methodology described in Section A.1. As with the analysis described in Section 8.6.1, we then fit 7 parametric survival models to individual patient data reconstructed from the adjusted survival curve using the Guyot et al. (2012) algorithm. Generalized gamma time to treatment discontinuation curves stratified by disease activity level are shown in Figure 8. Survival curves for patients with severe disease activity are not displayed because patients with severe disease activity are assumed to switch treatments after the first 6 months of treatment.

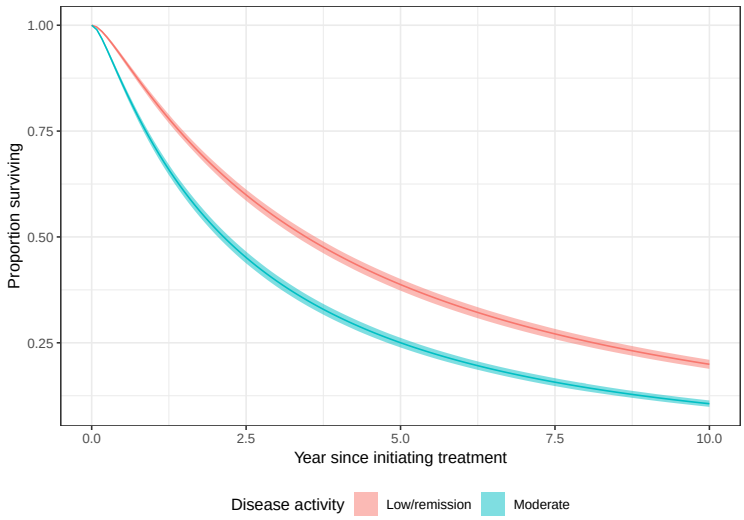


Figure 8: Generalized gamma time to treatment discontinuation curves by disease activity level

Notes: The shaded region denotes the simulation based 95% confidence interval (Mandel 2013).

8.6.3 Treatment duration by EULAR response

In **S6**, we stratify time to treatment discontinuation by EULAR response based on analyses of the British Society for Rheumatology Biologics Registers (BSRBR) database (Stevenson et al. 2016). We again fit 7 parametric survival models using reconstructed individual patient data. The survival curves reported in Stevenson et al. (2016) were used to create the patient data. The AIC and BIC of each model by EULAR response category are shown in Table 11.

Table 11: AIC and BIC for parametric models of treatment duration by EULAR response

	Moderate EULAR response		Good EULAR response	
Distribution	AIC	BIC	AIC	BIC
Exponential	38,840	38,847	15,126	15,132
Weibull	38,478	38,492	15,090	15,101
Gompertz	38,099	38,112	15,066	15,077
Gamma	38,587	38,600	15,098	15,110
Log-logistic	38,142	38,155	15,062	15,073
Lognormal	37,988	38,001	15,047	15,059
Generalized gamma	37,869	37,889	15,048	15,065

One concern is that the BSRBR is representative of the UK but not the US. As a result, we also estimate "adjusted" survival models appropriate for US based analyses. The adjustment is made in six steps using the analyses from the CORRONA database described in Section 8.6.1.

1. Calculate a hazard function based on a survival curve from an analysis of the CORRONA database. In particular, reconstruct individual patient data from the survival curve (Guyot

et al. 2012) and fit a spline-based survival model. Then use the spline-based model to estimate the hazard function $h(t)_{corrona}$.

2. Calculate a hazard function based on the BSRBR. To do so, first calculate hazard functions for both moderate and good EULAR responders using the same method described in step 1. Then calculate an overall hazard function with the proportion of moderate and good responders in the BSRBR analysis. Given that the number of moderate responders is $5,492$ and the number of good responders is $2,417$ the overall hazard function is $h(t)_{bsrbr} = \frac{5,492}{7,909} h(t)_{bsrbr,moderate} + \frac{2,417}{7,909} h(t)_{bbsrbr,good}$.

3. At each point in time, calculate the ratio of the CORRONA and BSRBR hazard functions: $HR(t) = h(t)_{corrona}/h(t)_{bbsrbr}$.

4. Apply the hazard ratio in step 3 to the BSRBR hazard functions for each EULAR response category. That is $h(t)_{bsrbr,moderate,adj} = h(t)_{bsrbr,moderate} \cdot HR(t)$ and $h(t)_{bsrbr,good,adj} = h(t)_{bsrbr,good} \cdot HR(t)$.

5. Generate survival curves using the hazard functions from step 4. Specifically, given a general hazard function $h(t)$, calculate the cumulative hazard function, $H(t) = \int_{z=0}^{t} h(z)dz$, convert this to a survival function using $S(t) = exp(-H(t))$, and reconstruct individual patient data using the survival curve.

6. Fit parametric survival models to the individual patient data generated in step 5.

Both adjusted and unadjusted survival curves by EULAR response fit using a generalized gamma distribution are shown in Figure 9. AIC and BIC for the parametric models fit in step 6 to the adjusted individual patient data are shown in Table 12.

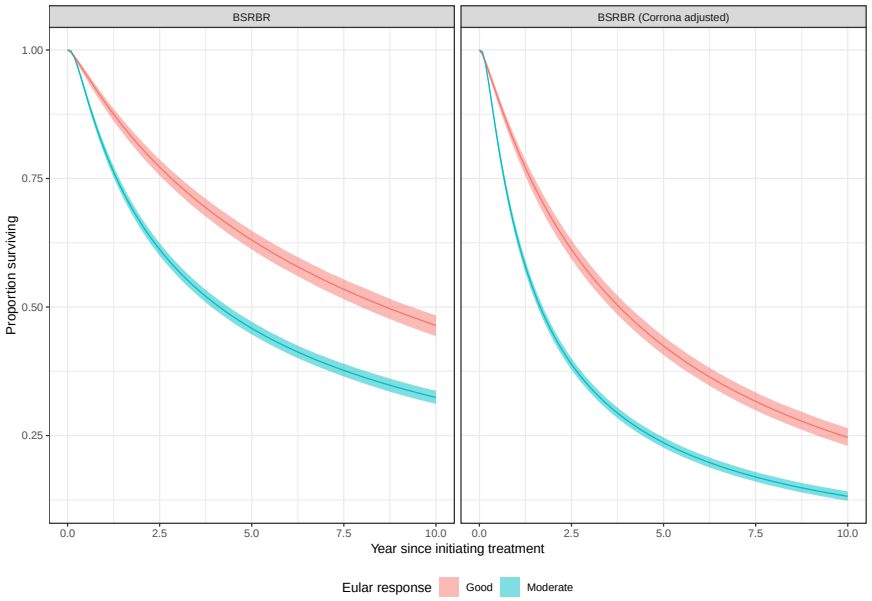


Figure 9: Generalized gamma survival curve of treatment duration using reconstructed individual patient data based on analyses from Stevenson et al. (2016) by EULAR response category

Notes: The shaded region denotes the simulation based 95% confidence interval (Mandel 2013).

Table 12: AIC and BIC for CORRONA adjusted parametric models of treatment duration by EULAR response

	Moderate EULAR response		Good EULAR response	
Distribution	AIC	BIC	AIC	BIC
Exponential	42,304	42,310	18,098	18,103
Weibull	41,946	41,959	18,051	18,062
Gompertz	41,569	41,582	18,039	18,050
Gamma	42,098	42,111	18,063	18,074
Log-logistic	41,406	41,419	18,037	18,049
Lognormal	41,235	41,248	18,004	18,016
Generalized gamma	41,110	41,129	18,000	18,017

8.7 Rebound post treatment

Since no data exists on the size of the HAQ rebound post treatment, we vary its size as a proportion of the initial 6-month HAQ decline. 1 is used as an upper bound, which implies that the HAQ rebound is equal to the improvement experienced at the end of the initial 6-month period with that treatment. 0.7 is currently used as a lower bound.

8.8 Serious infections

Based on the NMA by Singh et al. (2011) and in accordance with Stevenson et al. (2016), we assume a rate of 0.035 (95% CI: 0.027 to 0.046) infections per person-year with all tDMARDs and a rate of 0.026 (no CI reported) infections per person-year with cDMARDs. The rate of infection is assumed to be equal across tDMARDs because the published results for specific tDMARDs are estimated with very little precision. The standard error on the infection rate for tDMARDs is assumed to be the same as the standard error for cDMARDs since no standard error was reported for tDMARDs in Singh et al. (2011).

A patient in the IPS has a serious infection if the simulated time to serious infection occurs before the simulated time of treatment discontinuation. Table 13 shows the probability of this occurring when treatment duration is modeled using a generalized gamma distribution. The probability of a serious infection is relatively rare as only 3.82% of patients using cDMARDs and 8.55% of patients using tDMARDs have serious infections. However, differences between cDMARDs and tDMARDs are not insignificant as the probability of a serious infection is almost 5 percentage points higher with tDMARDs than with cDMARDs.

An important question related to the sensitivity of cost-effectiveness to the model specification is whether the probability of serious infections depends on the distribution used to model time to treatment discontinuation. We consequently simulated time to treatment discontinuation using each of the 7 possible probability distributions. We used the pathway **S1** to model treatment switching, so survival is based on the discussion in Section 8.6.1. Results from the simulation are reported in Table 14. There are very small differences across distributions, suggesting that the treatment duration distribution has almost no impact on the probability of serious infections.

Table 13: Probability of serious infection

	Probability		
		95% CI	
	Mean	Lower	Upper
cDMARDs or NBT	0.0382	0.0250	0.0533
tDMARDs	0.0856	0.0620	0.1089

Notes: Probabilities are estimated by simulating 1,000 patients and 1,000 parameter sets. Treatment duration is simulated using a generalized gamma distribution.

Table 14: Probability of serious infection with cDMARDs by distribution used to model treatment duration

Distribution	Mean probability
Exponential	0.0367
Weibull	0.0374
Gompertz	0.0381
Gamma	0.0382
Log-logistic	0.0391
Lognormal	0.0380
Generalized gamma	0.0382

Notes: Probabilities are estimated by simulating 1,000 patients and 1,000 parameter sets.

8.9 Utility

Two algorithms can be used to map HAQ to an EQ-5D utility score. Each is used to simulate utility for every patient in the model to obtain a distribution of utility over time. Our preferred algorithm is the mixture model developed by Hernández-Alava et al. (2013), which is described in detail in Section F.1. The second algorithm uses the logistic regression equation reported in Wailoo et al. (2006). Regression coefficients are reported in Section F.2.

Figure 10 compares results from the two algorithms. Mean utility scores from the Hernández-Alava et al. (2013) mixture model lie above those from the Wailoo et al. (2006) equation for all values of HAQ. Moreover, the slope of utility curve produced from the mixture model is steeper (although less so for the commonly observed HAQ scores between 1 and 1.5), implying that changes in HAQ from the mixture model predict larger changes in utility. Given that the mixture models have been shown to predict utility more accurately (Hernández-Alava et al. 2012, 2013, 2014), this suggests that standard models underestimate the quality-adjusted life-year benefits, and hence, the cost-effectiveness of treatments.

The utility score depends on serious infections in addition to HAQ. In particular, disutility due to serious infections is assumed to be 0.156 for the duration of the month of infection based on prior studies (Stevenson et al. 2016; Oppong et al. 2013). However, given the weak evidence for this estimate, the disutility of an infection is allowed to vary by 20% in either direction.

Finally, in the R package, we also allow users to incorporate treatment attributes unrelated to safety

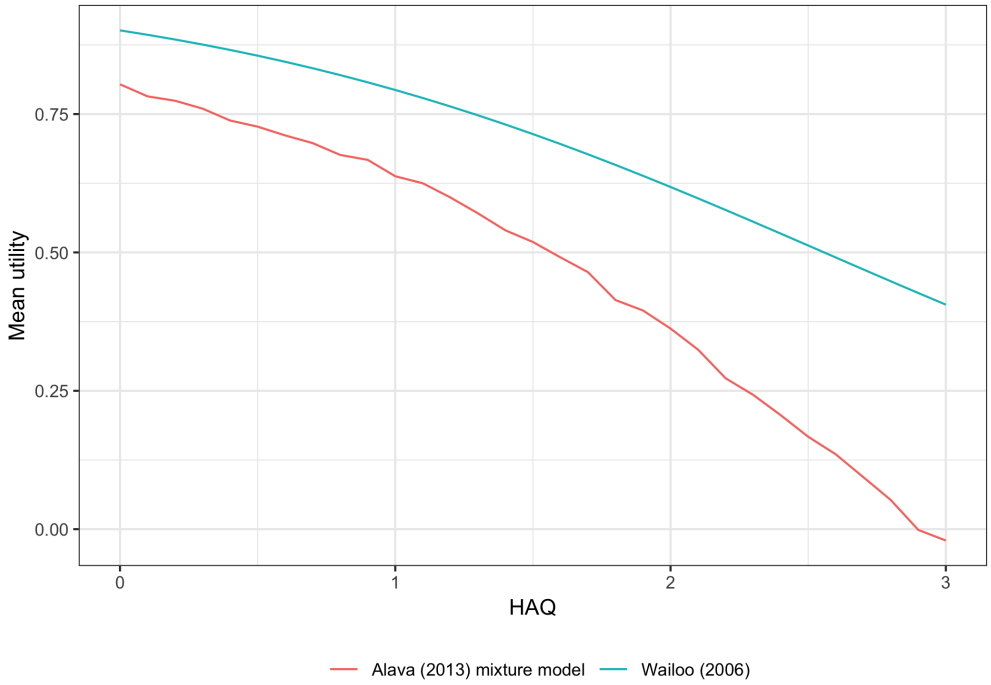


Figure 10: Simulated mean utility by current HAQ

or efficacy that might impact utility. In particular, users can specify a vector of variables and a vector of corresponding coefficients. Each coefficient is the impact of the corresponding variable on utility in a given 6 month period. By default, we include variables related to mode of administration (infusion, injection, oral) and years since FDA approval; however, since we have no evidence on the impact of each variable on utility, the coefficients are set to 0 in our default settings.

8.10 Mortality

The probability of death is simulated as a function of age/sex specific mortality from U.S. lifetables (Arias 2015), baseline HAQ, and changes in HAQ from baseline. Wolfe et al. (2003) estimate an odds ratio for the effect of HAQ on mortality of 2.22, which is applied to the absolute mortality rates of the general population (HAQ score of 0). To capture the effect of treatment on mortality, we assume that, for every 0.25-unit increase in HAQ score, subsequent 6-month mortality increases according to the hazard ratios reported in Michaud et al. (2012). Parameter estimates are shown in Table 15.

Figure 11 plots survival curves by gender for 1,000 patients with a baseline age of 55. Survival was simulated by setting the log odds ratios and log hazard ratios from Table 15 equal to their point estimates. Three scenarios are considerd. In scenario one, patients do not have RA (i.e., HAQ score of 0). In the second scenario, patients have baseline HAQ score of 1 but it does not increase over time. In the third scenario, patients still have a baseline HAQ score of 1, but it increases by 0.03 per year. The third scenario, therefore, utilizes the relationship between changes and HAQ and mortality from Michaud et al. (2012) while the second scenario does not.

Mean survival for females without RA was 82.5 years and declined to 77.0 for females with a constant baseline HAQ of 1 and to 72.4 when HAQ increased by 0.03 per year. Mean survival for males in the first, second, and third scenario were 79.4, 73.2, and 70.1 years respectively. Overall,

Table 15: Mortality parameters

		95% CI		
	Estimate	Lower	Upper	Reference
Impact of baseline HAQ on mortality				
Log odds of mortality	0.798	0.582	1.012	Wolfe et al. (2003)
Impact of 0.25-unit change in HAQ from baseline on mortality				
Log hazard ratio 0-6 months	0.113	0.077	0.157	Michaud et al. (2012)
Log hazard ratio >6-12 months	0.148	0.104	0.191	Michaud et al. (2012)
Log hazard ratio >12-24 months	0.148	0.095	0.191	Michaud et al. (2012)
Log hazard ratio >24-36 months	0.191	0.131	0.247	Michaud et al. (2012)
Log hazard ratio >36 months	0.174	0.104	0.239	Michaud et al. (2012)

Notes: 95% confidence intervals are calculated using normal distributions on the log odds and log hazard ratio scales.

the figure suggests that RA increases mortality and that larger increases in HAQ over time increase mortality by even more.

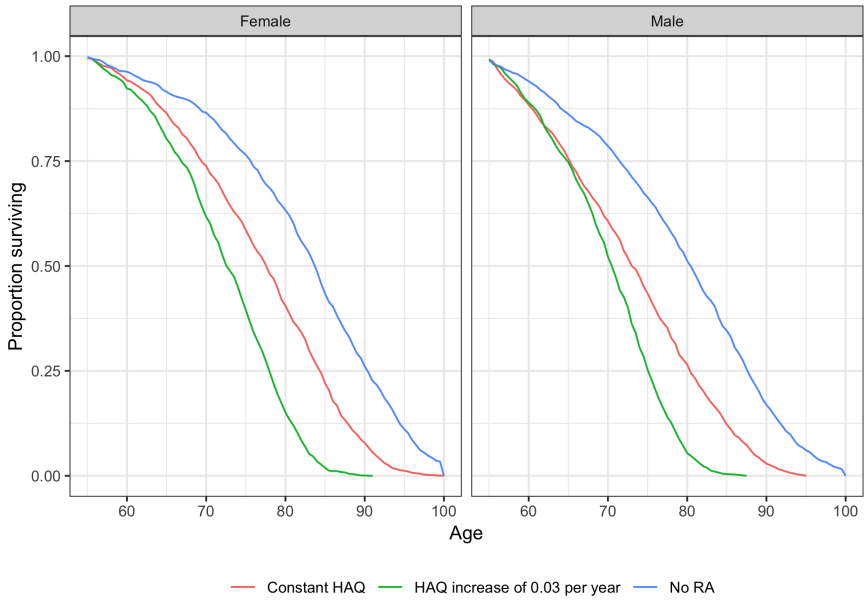


Figure 11: Simulated survival curve for a patient age 55

Notes: Baseline HAQ is 1 for the "Constant HAQ" and "HAQ increase of 0.03 per year" scenarios; baseline HAQ is 0 for the "No RA" scenario.

8.11 Costs

An overview of drug acquisition and administration costs is presented in Table 16. Costs are a function of dose and frequency of administration, strength and dosage form, price, and infusion costs. Since infliximab dosing depend on patient weight, the costs for infliximab reported in the table average over a patient population that is 21% male. The prices in the table are based on the wholesale acquisition cost (WAC) and do not include discounts or rebates so they may be higher than actual drug costs. In the simulation, a unique discount can be used for each drug; currently the discount is assumed to range from 20% to 30%. The methodology used to calculate drug acquisition and administration costs is described in more detail in Appendix G.

Table 16: Drug acquisition and administration cost

Drug	Dose and frequency of administration	Strength and dosage form	Number of doses first 6 months	Number of doses per year beyond the first 6 months	Price per unit	Infusion cost	Cost for the first 6 months	Cost per year beyond the first 6 months
Abatacept IV	750 mg IV at weeks 0, 2, 4 then Q4W	250mg vial	8	13	1,167.33	164	29,327	47,657
Abatacept SC	125 mg SC QW with IV loading dose	125mg/ml syringe	26	52	1,094.72	164	32,726	65,453
Adalimumab	40 mg EOW	40 mg/0.8 mL syringe or pen injector	13	26	2,587.05	0	33,631	67,263
Adalimumab-bwwd (biosimilar Samsung Bioepis)	40 mg EOW	40 mg/0.8 mL syringe or pen injector	13	26	2,069.64	0	26,905	53,810
Anakinra	100 mg daily	100 mg syringe	182	364	147.07	0	26,766	53,532
Baricitinib	2 mg daily	2 mg tablet	182	364	71.23	0	12,963	25,927
Certolizumab pegol	400 mg at weeks 0, 2, 4 then 200 mg Q2W	400 mg kit or syringe kit (200 mg 2)	16	26	4,327.43	0	69,238	112,513
Etanercept	50 mg QW	50 mg/0.98 mL syringe or pen injector	26	52	1,293.52	0	33,631	67,262
Etanercept-szzs (biosimilar Sandoz)	50 mg QW	50 mg/0.98 mL syringe or pen injector	26	52	1,034.81	0	26,905	53,810
Etanercept-ykro (biosimilar Samsung Bioepis)	50 mg QW	50 mg/0.98 mL syringe or pen injector	26	52	1,034.81	0	26,905	53,810
Golimumab	50 mg QM	50 mg/0.5 mL syringe or pen injector	6	12	4,809.02	0	28,854	57,708
Hydroxychloroquine sulfate	400 mg daily	200 mg tablet	182	364	3.88	0	1,413	2,826
Infliximab	3 mg/kg at 0, 2, and 6 weeks, 3mg/kg Q8W, 6 mg/kg Q6W after 6 months	100 mg vial	5	8.67	1,167.82	164	18,337	54,132
Infliximab-qbtx (biosimilar Pfizer)	3 mg/kg at 0, 2, and 6 weeks, 3mg/kg Q8W, 6 mg/kg Q6W after 6 months	100 mg vial	5	8.67	946.28	164	961	1,913
Methotrexate	15mg QW	15 mg injection	26	52	5.81	0	151	302
Rituximab	1000 mg at weeks 0, 2; then Q24 W	500 mg/50ml vial	4	4.33	4,697.60	164	38,236	41,423
Sarilumab	200 mg EOW	200mg/1.14 mL syringe	13	26	1,457.57	0	18,948	37,896
Sulfazalazine	1-2 g daily	500 mg tablet	182	364	0.24	0	176	352
Tocilizumab	162 mg SC EOW	162 mg/0.9 mL syringe	13	26	1,014.26	0	13,185	26,370
Tofaticinib	5 mg BID	5 mg tablet	364	728	74.68	0	27,182	54,364
Upadacitinib	15 mg daily	15 mg tablet	182	364	163.89	0	29,827	59,655

Notes: Costs in the table do not include rebates or discounts, but rebates and discounts are used in the simulation. Cost for infliximab are calculated by assuming that 21% of patients are male and that the weight of men and women are 89 kg and 75 kg respectively. Tocilizumab is dosed weekly if weight is greater than 100 kg; costs for tocilizumab reported in the table are for patients weighing less than 100 kg. IV = intravenous; SC = subcutaneous.

Parameters associated with resource use are show in in Table 17. Costs related to physician visits, chest X-rays, tuberculosis tests, and outpatient follow-up are based on Claxton et al. (2016). The cost per hospital day and the relationship between the HAQ score and the annual number of hospital days are from Carlson et al. (2015). Cost of any serious infection are assumed to be equal to the cost of pneumonia hospitalization at $5,873, based on Medicare reimbursement rates. Wolfe et al. (2005) provide an estimate of annual income loss in relation to HAQ scores: $4,372 (95% CI: 2,078 to 6,607; 2002 dollars) change per unit HAQ change, which are inflated to 2019 dollars.

Table 17: Resource use parameters

		95% CI		
	Estimate	Lower	Upper	Reference
Days in hospital per year				
HAQ: 0-<0.5	0.260	0.000	1.725	Carlson et al. (2015)
HAQ: 0.5-<1	0.130	0.000	1.409	Carlson et al. (2015)
HAQ: 1-<1.5	0.510	0.015	1.850	Carlson et al. (2015)
HAQ: 1.5-<2	0.720	0.092	1.979	Carlson et al. (2015)
HAQ: 2-<2.5	1.860	1.013	2.960	Carlson et al. (2015)
HAQ: >2.5	4.160	3.238	5.196	Carlson et al. (2015)
Cost per day in hospital	1,347	974	1,779	Carlson et al. (2015)
Cost per day in hospital	1,347	974	1,779	Carlson et al. (2015)
Cost per day in hospital	1,347	974	1,779	Carlson et al. (2015)
Cost per day in hospital	1,347	974	1,779	Carlson et al. (2015)
Cost per day in hospital	1,347	974	1,779	Carlson et al. (2015)
Cost per day in hospital	1,347	974	1,779	Carlson et al. (2015)
General management cost				
Chest x-ray	117	104	130	Claxton et al. (2016)
X-ray visit	57	48	65	Claxton et al. (2016)
Outpatient follow-up	201	171	231	Claxton et al. (2016)
Mantoux tuberculin skin test	32	32	32	Claxton et al. (2016)
Productivity loss				
Linear regression coefficient - HAQ	6,218	2,997	9,439	Wolfe et al. (2005)

Notes: 95% confidence intervals for hospital days per year by HAQ score and hospital cost per day are calculated by using the methods of moments to generate the parameters of the gamma distribution given a mean and standard error. The 95% confidence intervals for general management costs are based on normal distributions as assumed in Claxton et al. (2016). 95% confidence interval for productivity loss are calculated using a normal distribution and inflated to 2016 dollars.

8.12 Insurance value

In the IVI-RA Model interface, users have complete control over the probability of illness parameter and the marginal rate of substitution between the sick and the well states. In the IVI-RA Value Tool, we set the probability of obtaining RA in the next year equal to 0.000633, based on the annual incidence rate reported for individuals age 55 to 64 in Myasoedova et al. (2010). Furthermore, we set the the marginal rate of substitution between the sick and well states equal to 1.5 given that positive demand for health insurance suggests that it is positive, but note that there is considerable uncertainty around this estimate and that more research is required.

9 Simulation and uncertainty analysis

9.1 Individual patient simulation

The IPS is a discrete-time simulation that simulates individual patients one at a time. Model cycles, denoted by t, were chosen to be 6-months long to be consistent with most RCT and real-world data evidence. Algorithm 1 describes the main components of the IPS for a single patient and a single treatment. The full simulation cycles through each treatment in a treatment sequence and through each simulated patient.

Algorithm 1 Main components of the individual patient simulation

1. **First 6 months** ($t = 0$)
   (a) Simulate treatment switching using **S1-S6**, time to serious infection T_{si}, and death (Appendix E).
      i. **If S1-S6** leads to a treatment switch or if the sampled time to serious infection occurs during cycle 0 (i.e., $T_{si} = 0$), **then** stop treatment. It is assumed that HAQ does not change.
      Else, continue treatment. Simulate change in HAQ using **H1-H3** and time to treatment discontinuation T.
      ii. **If** patient died, **then** move to next patient.
2. **Maintenance phase** (for $t > 0$ and $t \leq T$)
   (a) Simulate death and change in HAQ.
   (b) **If** patient died, **then** move to next patient.
   (c) **If** $t = T$, **then** switch treatment. Treatment switch caused by a serious infection if time to serious infection occurred during or before cycle T (i.e., $T_{si} \leq T$).

9.2 Parameter uncertainty

Parameter uncertainty is quantified using PSA, which propagates uncertainty in the model input parameters throughout the model by randomly sampling the input parameters from their joint probability distribution (Baio and Dawid 2015; Claxton et al. 2005). Probability distributions are determined according to the distributional properties of the statistical estimates, which, in turn, depend on the statistical techniques used and the distributions of the underlying data. We use normal distributions for sample means, gamma distributions for right-skewed data (e.g., hospital costs), and Dirichlet distributions for multinomial data. The multivariate normal distribution is used for regression parameters estimated using frequentist techniques, provided that the variance-covariance from the statistical analysis is available. For parameters estimated using a Bayesian NMA, we fit multivariate normal distributions to the posterior distribution of the parameters generated from the Markov-Chain Monte-Carlo (MCMC) algorithm using sample means and the sample covariance matrix. When we lack evidence on a parameter, we typically assume a uniform distribution with lower and upper limits that reflect the degree of uncertainty in the parameter. The PSA parameter distributions are summarized in Table 18.

Table 18: Probabilistic sensitivity analysis parameter distributions

Parameter(s)	Distribution
Rebound factor	Uniform
NMA parameters - ACR response	Multivariate normal
NMA parameters - DAS28	Multivariate normal
NMA parameters - HAQ	Multivariate normal
Drug acquisition and administration cost	Fixed
Survival model parameters for treatment duration during maintenance phase	Multivariate normal
US lifetable mortality rates	Fixed
Mortality probability odds ratio - baseline HAQ	Normal
Mortality probability hazard ratio - change in HAQ from baseline	Normal
ACR response to EULAR response mapping	Dirichlet
ACR response to SDAI mapping	Uniform
ACR response to CDAI mapping	Uniform
ACR response to HAQ mapping	Normal
EULAR response to HAQ mapping	Normal
Linear HAQ progression - by therapy	Normal
Linear HAQ progression - by age	Normal
Latent class growth model for HAQ progression	Normal
Utility model - Hernández-Alava et al. (2013) mixture model	Multivariate normal
Utility model - Wailoo et al. (2006)	Normal
Hospital costs - hospital days by HAQ	Gamma
Hospital costs - hospital costs per day	Gamma
General management cost	Gamma
Serious infection - survival parameters	Normal
Serious infection - cost per infection	Uniform
Serious infection - utility loss	Uniform

9.3 Structural uncertainty

We consider structural uncertainty due to two factors:

- The relationship between health states within the model.
- The statistical model used to estimate parameters.

Table 19 summarizes the competing model structures, which are conditional on the perspective of the decision maker. In total, there are 12 x 2 x 8 x 2 = 384 possible model structures. The choice of model structure for the initial treatment phase (**H1-H3** and **S1-S6**) depends on the preferred measures of disease activity included in the model as well as whether statistical relationships should be modeled directly or indirectly. Likewise, model structures related to HAQ progression, treatment duration, and converting HAQ to utility all reflect uncertainty in the appropriate statistical model.

Table 19: Competing model structures

Component of model structure	Possible combinations
Initial effect of treatment on HAQ (**H1-H3**) and switching (**S1-S6**)	12
HAQ trajectory	2
Cause and probability distribution used to model treatment discontinuation	8
Utility algorithm	2

9.4 Implementation

We begin by describing the simulation procedure conditional on model structure, which uses PSA to capture uncertainty within but not between models. The procedure proceeds in two steps: first, model parameters are sampled from their joint probability distribution (Section 9.2), and second, for each parameter set, model outcomes are simulated one at a time for individual patients in the specified population (Section 5).

Analysts who wish to expand the analysis to capture uncertainty between models can follow the approach described in Bojke et al. (2009). In particular, for each randomly sampled parameter set, each model structure (or a subset of plausible model structures) can be simulated. The distribution of simulated outcomes across parameters and models will then reflect uncertainty both within and between models.

It's important to note that simulation output for an individual patient captures differences in outcomes across patients due to random variation (often referred to as first order uncertainty). This information might be useful to patients since it is needed to predict the distribution of their future outcomes conditional on their characteristics, but less useful to a decision maker concerned with making treatment decisions for a population or subset of a population. Analysts wishing to use the model for CEA or MCDA should therefore estimate mean outcomes by averaging over the simulated patients for each parameter set and model structure. The number of simulated patients should be sufficiently large so that mean outcomes are stable across model runs (i.e., so that first order uncertainty is eliminated).

Although CEA and MCDA is concerned with mean outcomes, that does not imply that it does not account for heterogeneity. Instead, since outcomes depend on the characteristics of each patient,

model averages are a function of the population analyzed. Subgroup analyses can be used to examine differences in cost-effectiveness across subgroups by simulating patients with certain shared characteristics.

Parameter and structural uncertainty imply decision uncertainty, or the degree to which decisions are made based on imperfect knowledge. Indeed, in CEA, with the aim to maximize health outcomes for a given budget, the optimal decision with current information is to choose the policy that maximizes the expected NMB; however, due to uncertainty, the incorrect policy may be considered the most cost-effective. To characterize uncertainty within a CEA framework, standard summary measures including 95% credible intervals for NMBs and other model outcomes, cost-effectiveness planes, CEACs, the CEAF, and the EVPI can be calculated from the simulated output. Since the EVPI is computationally costly, it can be approximated using meta-modeling techniques (Jalal et al. 2013, 2015; Heath et al. 2016).

10 Validation

We aim to validate the model using the five types of validation described by Eddy et al. (2012). Currently, we are able to use the first three validation types. First, we have checked the model for face validity by ensuring that simulated outcomes are consistent with current science and evidence. Second, we performed unit tests to verify that the individual units of code that are used to simulate the model return expected results. Third, we compared simulated results for key outcomes such as mortality, HAQ over time, and time to treatment discontinuation with real-world data and our underlying parameter values. In particular, we ran the model online under various scenarios using our R Shiny web application and checked the simulated outcomes.

In the future, we plan to use both external validation and predictive validation to help fine tune our model. External validation will be performed by comparing outcomes simulated using our model to real-world outcomes and predictive validity will involve using our model to forecast future events and comparing our forecasted outcomes to the observed outcomes.

11 Limitations and areas for improvement

The IVI-RA model is an open-source model that is part of the OSVP process and therefore designed to be updated and improved over time. We believe that there are number of potential areas for improvement.

- **Adverse events other than serious infections:** The current model does not account for side effects other than serious infections even though these are important to patients and can result in treatment switching.
- **Adverse events that vary across biologics:** The model allows the serious infection rate to differ between cDMARDs and tDMARDs but assumes that the infection rate is equal among tDMARDs. Future model versions might want to reconsider the evidence underlying this assumption.
- **Time to treatment discontinuation:** Our time to treatment discontinuation curves are based on scanned data and combine information from multiple sources. Direct analyses of databases like the CORRONA database or the National Data Bank for Rheumatic Diseases (NDB) could generate more accurate estimates of treatment duration as well the effect of treatment response or disease activity level on discontinuation rates.

- **Patient preferences:** In the current model, patient utility is a function of the HAQ score and varies according to age, gender, and unobserved patient-specific factors. In other words, utility depends on treatment (through the effect of treatment on HAQ) and the characteristics of the patient. Future iterations of the model should consider other ways that treatment influences utility and that utility varies across patients. For example, disease activity level or the number of previous therapies might help predict utility conditional on HAQ. Furthermore, surveys could be used to estimate the effect of treatment attributes such as route of administration or frequency of administration on utility. Finally, since unobserved patient-specific factors are very important predictors of utility, the model could be run for specific classes of patients within the mixture model (e.g., subgroups where HAQ has the largest effect on utility), although it might be difficult to identify these patient subgroups in a real-world setting.

- **Treatment effect modifiers:** There is currently little evidence (that we are aware of) suggesting that treatment effects vary across patients. When there is sufficient evidence in the literature related to treatment response heterogeneity, we will allow treatment response at 6 months to depend on the characteristics of the patient.

- **Treatment effects after treatment failure:** There are two main limitations in the model related to reductions in treatment response after failing a biologic; first, there are not enough RCTs to reliably estimate tDMARD-specific treatment effects for tDMARD experienced patients using a NMA, and second, treatment response likely does not only depend on whether a patient is tDMARD naive or experienced, but on the number of previous failures as well. Our current approach is to assume that treatment response is reduced for tDMARD experience patients based on evidence from Carlson et al. (2015). One possible extension is to use a Bayesian NMA approach in which the Carlson et al. (2015) results are used to generate priors for the tDMARD experienced group. As new RCTs become available, the posterior distributions from the Bayesian analysis would move further from the prior and closer to estimates from the trials. The estimates could be further improved by combing NMA results with real-world data and modeling reductions in treatment response as a flexible function of the number of failed biologics.

- **A LCGM for the progression of tDMARDs over time:** The LCGM can be used to model HAQ progression for patients using cDMARDs or on NBT; however, we only have estimates of constant linear progression of HAQ for patients on biologics. Future studies that use non-linear mixture models to model the long-term progression of disease for patients using tDMARDs are needed.

- **Long-term trends in disease activity:** The current model uses results from RCTs to model changes in disease activity during the first 6 months of treatment. But there is, to our knowledge, no evidence on the progression of disease activity over time. New studies are needed to model both the long-term impact of treatment on disease activity and the correlation between changes in disease activity and changes in HAQ.

- **The patient population:** Our population characteristics are based on summary data reported in the published literature. As a result, the sampled patient populations within the model do not account for correlations across all of the variables. Distributions estimated from patient databases like the CORRONA database or the NDB would yield more realistic patient populations.

- **Estimating the rebound effect:** One of the most important predictors of cost-effectiveness is the degree to which the HAQ score increases following treatment failure. Most models currently assume that the HAQ score increases by the same amount as the initial 6 month decline in the HAQ score, but there is little evidence to support this. Studies that attempt to quantify the rebound effect are critical.
- **Insurance value:** Additional research is needed on insurance value and its use in value assessment. First, the framework presented here assumes that there is a single probability of illness and that treatment benefits and costs can be attributed to that sick state. However, in practice, the probability of illness depends on age and illness (e.g., RA) worsens over time. Moreover, in RA, treatment benefits and costs depend on disease severity. Future research should consider insurance value in a dynamic context, in which the value to a healthy individual today depends on the probability of all future health states. Second, new research is needed on the marginal rate of substitution between the sick and well states. Lakdawalla et al. (2017) suggest a few promising approaches.

Appendices

A Rates, probabilities, and standard errors

A.1 Using odds ratios to adjust probabilities

Let p_1 be a baseline probability, β be a vector of log odds ratios, and x be a vector of regressors. We apply the log odds ratios to p_1 to generate a new probability p_2 with the logistic equation,

$$p_2 = \frac{1}{1 + \exp\left[-\left(\text{logit}(p_1) + x^T\beta\right)\right]}, \tag{A1}$$

where,

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) \tag{A2}$$

A.2 Converting rates and probabilities

Given a *constant* rate r during a given time period, we estimate the probability of an event occurring before time t using the exponential distribution,

$$p(\tau < t|r) = 1 - e^{-rt}. \tag{A3}$$

Given a probability p, the rate parameter is recovered by applying the log transformation,

$$r = \frac{-\ln(1-p)}{t}. \tag{A4}$$

A.3 Calculating standard errors from confidence intervals

Journal articles often report confidence intervals rather than standard errors. However, given that regression coefficients are asymptotically normally distributed, standard errors can be calculated from a confidence interval using the normal distribution. In particular, given a coefficient estimate β (e.g., a log hazard ratio, log odds ratio, or linear regression coefficient) and an upper bound u and lower bound l of a two-sided 95% confidence interval, we calculate the standard error as,

$$SE(\beta) = \frac{u-l}{2 \cdot \Phi^{-1}(0.975)}, \tag{A5}$$

where $\Phi^{-1}(p)$ is the quantile function of the normal distribution.

B Heterogeneous populations

When generating heterogeneous patient populations, we sample binary variables from binomial distributions, continuous uncorrelated variables from normal distributions, and continuous correlated variables from multivariate normal distributions. Truncated distributions are used when variables are restricted to lie within certain intervals.

In particular, the proportion of the female population is drawn from a binomial distribution while age, disease duration and the number of previous DMARDs are drawn from truncated normal distributions. Each sampled value of the number of previous DMARDs is rounded to the nearest integer. Baseline HAQ and three disease activity measures (DAS28, SDAI, and CDAI) are drawn from truncated multivariate normal distributions. The covariance matrix is calculated using the correlations reported in Aletaha et al. (2005) (Figure A1).

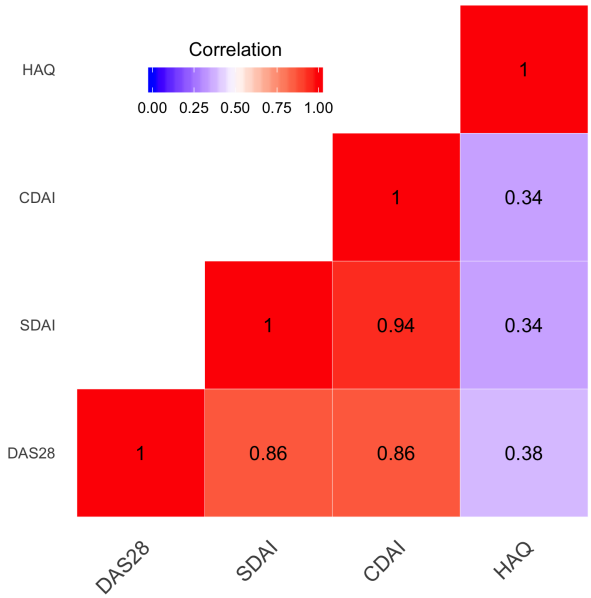


Figure A1: Correlations between disease activity measures and HAQ

We used the correlations from the routine cohort (during visit 1) rather than correlations in the inception cohort (at baseline) since the correlation between HAQ and the disease activity measures were more similar to those from the Leflunomide database (Smolen et al. 2003). That said, correlations between the three disease activity measures were nearly identical in each cohort. The one exception was that the correlation between SDAI and CDAI of 1 in the routine cohort seemed unreasonably high so we used the value of 0.94 from the inception cohort.

We used this sampling procedure to simulate 1,000 patients. Summary statistics from a simulated patient cohort of size 1,000 are shown in Table A1.

Table A1: Summary of characteristics for 1,000 simulated patients

		95 CI%	
	Mean	Lower	Upper
Age	54.95	29.83	77.97
Male	0.24	0.00	1.00
Weight (kg)	78.30	75.00	89.00
Previous DMARDs	3.42	0.00	7.00
DAS28	6.00	3.64	8.13
SDAI	42.95	18.67	66.95
CDAI	41.02	16.86	64.19
HAQ	1.50	0.23	2.67

C Mapping ACR response to changes in disease activity

Let DA denote disease activity, n_1 the number of patients with ACR 20 to <50 response, n_2 the number of patients with ACR 50 to <70 response, n_3 the number of patients with ACR ≥70

response, and N the number of patients with an ACR response greater than or equal to 20%. Mean changes in SDAI, CDAI, and DAS28 by overlapping ACR response categories are converted to mean changes by mutually exclusive ACR response categories as follows:

- **ACR 70**: Mean changes by ACR ≥70 were reported directly in Aletaha and Smolen (2005).
- **ACR 50 to <70**: Mean change in disease activity given ACR 50 to <70 response is calculated by solving for $\mathbb{E}[DA|50 \leq ACR < 70]$:

$$\mathbb{E}[DA|ACR \geq 50] = \frac{n_2}{N} \cdot \mathbb{E}[DA|50 \leq ACR < 70] + \frac{n_3}{N} \cdot \mathbb{E}[DA|ACR \geq 70]. \tag{A6}$$

- Mean change in disease activity given ACR 20 to <50 response is calculated by solving for $\mathbb{E}[DA|20 \leq ACR < 50]$

$$\mathbb{E}[DA|ACR \geq 20] = \frac{n_1}{N} \cdot \mathbb{E}[DA|20 \leq ACR < 50] + \frac{n_2 + n_3}{N} \cdot \mathbb{E}[DA|ACR \geq 50]. \tag{A7}$$

D HAQ progression

D.1 Effect of age on linear HAQ progression

Michaud et al. (2011) report an overall rate of linear HAQ progression and rates for three age groups (<40, 40-64, ≥ 65). Let β be the overall rate of progression and β_a be the rate of progression for age group a. To estimate the effect of age on the progression rate, we calculated the difference between the overall progression rate and the age specific rate, $\delta_a = \beta - \beta_a$. We estimated the standard error of this quantity assuming no covariance between β and β_a,

$$SE(\delta_a) = \sqrt{SE(\beta)^2 + SE(\beta_a)^2}. \tag{A8}$$

D.2 HAQ trajectory with a latent class growth model

Norton et al. (2014) model HAQ progression using a LCGM. The probability that individual i is a member of class c at time t is modeled using a multinomial logistic regression,

$$P(C_{it} = c) = \frac{\exp(w_{it}^T \delta_c)}{\sum_{s=1}^{4} \exp(w_{it}^T \delta_s)}, \tag{A9}$$

where δ_s is the vector of regression coefficients associated with class s and w_{it} is the corresponding vector of regressors. The variables included in w_{it} are age, gender, baseline DAS28, symptom duration, rheumatoid factor, ACR criteria, and socioeconomic status. Regression coefficients for classes 2-4 relative to class 1 are shown in Table A2. Older age and female gender are especially important predictors of membership in higher risk classes; a worse DAS28 score, rheumatoid factor

Table A2: Determinants of class membership in the ERAS cohort

		95% CI	
	Coefficient	Lower	Upper
Class 2: moderate			
Intercept	-3.496	-4.715	-2.277
Age at onset	0.025	0.011	0.039
Female gender	0.841	0.457	1.225
Disease duration (months)	0.304	0.147	0.461
DAS28 score	0.032	0.001	0.063
Rheumatoid factor positive	0.214	-0.251	0.679
ACR criteria for RA	0.278	-0.163	0.719
Socioeconomic status	0.993	0.276	1.710
Class 3: high			
Intercept	-6.686	-7.980	-5.392
Age at onset	0.037	0.023	0.051
Female gender	1.694	1.275	2.113
Disease duration (months)	0.573	0.424	0.722
DAS28 score	0.046	0.013	0.079
Rheumatoid factor positive	0.315	-0.175	0.805
ACR criteria for RA	0.413	-0.050	0.876
Socioeconomic status	1.119	0.449	1.789
Class 4: severe			
Intercept	-12.055	-14.215	-9.895
Age at onset	0.082	0.060	0.104
Female gender	1.976	1.449	2.503
Disease duration (months)	0.800	0.631	0.969
DAS28 score	0.042	0.001	0.083
Rheumatoid factor positive	0.298	-0.270	0.866
ACR criteria for RA	0.939	0.320	1.558
Socioeconomic status	1.429	0.682	2.176

Notes: Class 1, or the "low" group, is the reference category.

positivity, fulfillment of the 1987 ACR criteria, lower socioeconomic status, and longer disease duration are also predictors of membership in classes with worse HAQ progression.

The HAQ trajectory for a given class can be written as,

$$y^*_{itc} = \beta_{0c} + \beta_{1c}x_t + \beta_{2c}x_t^2 + \beta_{3c}x_t^3 + \epsilon_{it} \tag{A10}$$

$$y_{itc} = \begin{cases} 0 & \text{if } y^*_{itc} < 0 \\ y^*_{itc} & \text{if } 0 \leq y^*_{itc} \leq 3 \\ 3 & \text{if } y^*_{itc} > 3, \end{cases} \tag{A11}$$

where y_{itc} is the HAQ score, x_t is a variable that is a function of time, the β_{jc} are polynomial regression coefficients for members of class c, and ϵ_{it} is an error term.

Sam Norton generously provided us with statistical estimates of the 4 class LCGM used in Norton et al. (2014) from MPlus. Like Stevenson et al. (2016), we noted that the coefficient estimates the MPlus resulted in large fluctuations in the predicted HAQ scores, likely because three decimal places was not precise enough for the cubic term in Equation A10. We consequently used the coefficient estimates to predict the probability of class membership—which are less likely to be influenced by the number of reported decimal places—but estimated Equation A10 using the observed HAQ values reported in Figure 2 in Norton et al. (2014). However, since standard errors were artificially high using grouped data, we standard errors in Equation A10 were based on those reported in the original paper. Moreover, since we are only interested in the HAQ trajectory following the HAQ decline during the initial treatment phase, we limited our analysis to HAQ values from year 2 and onwards. Using the post year 2 data, we estimated Equation A10 using separate linear regressions with cubic polynomials for each class (Table A3). Like Norton et al. (2014), we set x_t equal to a reciprocal transformation of time,

$$x_t = 1 - \frac{1}{t+1} \tag{A12}$$

In the simulation model, we simulate the HAQ score at 6 months as a function of the baseline HAQ score and the change in HAQ during the initial treatment phase. Since the Norton et al. (2014) model is not conditional on the HAQ score in the previous period, we use it to predict changes in HAQ rather than the level of the HAQ score. More precisely, for a patient in a given class, we model the change in HAQ as,

$$\begin{aligned} \Delta y^*_{itc} &= y^*_{i,t,c} - y^*_{i,t-1,c} \\ &= \beta_{1c}(x_t - x_{t-1}) + \beta_{2c}(x_t^2 - x_{t-1}^2) + \beta_{3c}(x_t^3 - x_{t-1}^3) + (\epsilon_{i,t} - \epsilon_{i,t-1}). \end{aligned} \tag{A13}$$

Since Equation A10 was estimated on aggregated data, we did not have reliable estimates of ϵ_{it}. We consequently set $\epsilon_{i,t} - \epsilon_{i,t-1}$ equal to 0, which implies that we are generating a mean response rather than a predicted response. In other words, we are not simulating the random variation associated with each individual, but are still accurately simulating mean outcomes across populations or subpopulations.

Table A3: LCGM HAQ trajectory coefficients

	Coefficient	Standard error
Class 1: low		
Intercept	0.638	0.058
Linear	-1.009	0.074
Quadratic	-0.649	0.027
Cubic	1.355	0.003
Class 2: moderate		
Intercept	0.950	0.058
Linear	-0.109	0.020
Quadratic	-3.368	0.002
Cubic	3.699	0.064
Class 3: high		
Intercept	1.265	0.064
Linear	-0.132	0.056
Quadratic	-2.531	0.021
Cubic	3.538	0.002
Class 4: severe		
Intercept	1.935	0.063
Linear	-0.540	0.073
Quadratic	1.196	0.027
Cubic	-0.109	0.003

Notes: Class 1, or the "low" group, is the reference category.

E Simulating mortality

Death is simulated for each patient during each model cycle based on age, gender, baseline HAQ, and change in HAQ from baseline. A 0/1 death indicator is randomly drawn using the following procedure:

1. Find q_{xg}, the probability that a patient of gender g and age x will die before age $x+1$, from lifetables.

2. As described in Section A.1, adjust q_{gx} using the effect of a change in baseline HAQ on the odds of mortality, OR,

$$p_m = \frac{1}{1+\exp\left[-(\text{logit}(q_x)+\log(OR)\cdot HAQ)\right]}. \tag{A14}$$

3. Following Section A.2, convert the mortality probability, p_m, into a mortality rate, r_m.

$$r_m = -log(1-p_m). \tag{A15}$$

4. Adjust the mortality rate, r_m, using the estimated log hazard ratio of mortality, HR, of a

change in HAQ from baseline, Δ HAQ.

$$r_m = r_m \cdot \exp[\log(HR) \cdot \Delta HAQ] \tag{A16}$$

5. Following Section A.2, convert the mortality rate into a probability given a 6-month cycle length,

$$p_m = 1 - \exp[-r_m * (6/12)]. \tag{A17}$$

6. Randomly draw a 0/1 death indicator, d, given the probability of death, p_m,

$$d \sim \text{Bin}(1, p_m). \tag{A18}$$

F Simulate utility

F.1 Mixture model

The mixture model estimated by Hernández-Alava et al. (2013) simulates utility in two stages. In the first stage, we sampled pain for a given individual in a particular model cycle based on the HAQ score. In the second stage, we simulated utility as a function of HAQ, pain and age/sex.

F.1.1 Simulating pain

To simulate pain from HAQ, we used the summary statistics for pain and HAQ reported in Sarzi-Puttini et al. (2002). Pain was measured with the visual analog scale (VAS) with mean $\mu_{pain} = 61.65$ and standard deviation $\sigma_{pain} = 19.10$, while HAQ was reported to have mean $\mu_{haq} = 1.39$ and standard deviation $\sigma_{haq} = 0.59$.

We then estimated the correlation between pain and HAQ by digitally scanning the curve depicting the (linear) relationship between pain and HAQ (Figure 114) shown in Stevenson et al. (2016). Using the scanned data, we regressed pain on HAQ using simple ordinary least squares (OLS). The correlation between pain and HAQ, estimated as $\rho = 0.52$, was calculated by rearranging the OLS estimate for the slope, β, of the regression model,

$$\rho = \beta \cdot \frac{\sigma_{haq}}{\sigma_{pain}}. \tag{A19}$$

Pain was simulated using these parameters by assuming that pain was normally distributed conditional on HAQ,

$$pain|haq = h \sim N\left(\mu_{pain} + \rho\frac{\sigma_{pain}}{\sigma_{haq}}(h - \mu_{haq}), \sigma^2_{pain}(1 - \rho^2)\right). \tag{A20}$$

However, since the VAS is constrained to lie between 0 and 100, pain was drawn from a truncated normal distribution with a lower limit of 0 and an upper limit of 100.

F.1.2 Simulating utility

After simulating pain, we simulated utility with a mixture model. Within each class c, the HAQ score for patient i in period t was modeled as,

$$y_{it|C_{it}} = \begin{cases} 1 & \text{if } y^*_{it|C_{it}} > 0.883 \\ y^*_{it|C_{it}} & \text{otherwise} \end{cases} \tag{A21}$$

$$y^*_{it|C_{it}} = \alpha_{ic} + x_{it}^T \beta_c + \epsilon_{it} \tag{A22}$$

$$\alpha_{ic} = \gamma_c + z_i^T \kappa + \mu_i, \tag{A23}$$

where ϵ_{it} is a random error term and β_c is a vector of regression coefficients corresponding to the vector of variables x_{it}. α_{ic} is a random intercept for individual i and class c that is predicted by a class-specific intercept, γ_c, a vector of individual-specific variables z_i, a coefficient vector κ, and an error term, μ_i. Variables included in x_{it} are HAQ, HAQ^2, $Pain/100$, $Age/10$, and $Age/100$; z_i contains a single indicator variable, $Male$, equal to 1 if the patient is male and 0 if female.

The probability of class membership was modeled using a multinomial logit model,

$$P(C_{it} = c) = \frac{\exp(w_{it}^T \delta_c)}{\sum_{s=1}^{4} \exp(w_{it}^T \delta_s)}, \tag{A24}$$

where there are four possible classes and δ_c is a vector of coefficients corresponding to the vector of variables, w_{it} (which includes an intercept). Variables included in w_{it} other than the intercept are HAQ, $Pain/100$, and $Pain/100^2$.

We sampled from the mixture model as follows.

1. For each individual i, sample the error term, $\mu_i \sim N(0, \sigma^2_\mu)$.
2. For each individual i and time-period t:
   (a) Sample class membership conditional on w_{it}; that is, sample $C_{it} \sim \text{Cat}(p_1, p_2, p_3, p_4)$ where p_c is the probability of being in class c.
   (b) Predict the intercept α_{ic}.
   (c) Sample the error term, $\epsilon_{it} \sim N(0, \sigma^2_\epsilon)$.
   (d) Predict the HAQ score, y_{it}.

F.2 Logistic regression model

Wailoo et al. (2006) use a logistic regression equation to predict utility as a function of patient demographics, disease history, and current disease status. The regression coefficients from the model are shown in Table A4 and used to predict utility with the inverse logit function. Specifically, if the vector of coefficients is denoted by β and the corresponding vector of explanatory variables is denoted by the vector x, then predicted utility is given by $1/(1 + \exp(-x^T\beta))$.

Table A4: Logistic regression coefficient from Wailoo utility algorithm

	Estimate	Standard error
Intercept	2.0734	0.0263
Age	0.0058	0.0004
Disease duration	0.0023	0.0004
Baseline HAQ	-0.2004	0.0101
Male	-0.2914	0.0118
Number of previous DMARDs	0.0249	0.0028
Current HAQ	-0.8647	0.0103

Notes: Coefficients are from the logistic regression reported in Wailoo et al. (2006).

G Drug acquisition and administration costs

Drug acquisition and administration costs are calculated separately during the initial treatment phase and the maintenance phase since dosing typically differs. Costs are separated into acquisition costs and infusion costs. Infusion costs are calculated by multiplying the number of doses in a 6 month period by the cost of an infusion and acquisition costs are calculated as,

$$cost = \left\lceil \frac{dose_{amt}}{strength_{amt}} \right\rceil \cdot doses_{num} \cdot price, \tag{A25}$$

where $\lceil \cdot \rceil$ is the ceiling function and implies that products cannot be reused after opening, $dose_{amt}$ is the recommended dose of the drug, $strength_{amt}$ is the strength of the drug, $doses_{num}$ is the number of doses in a 6 month period, and $price$ is the price per unit of the treatment after discounts and rebates. For example, as shown in Table 16, both the strength and the dose of adalimumab are 50 mg, so costs (before discounts and rebates) for the initial 6 month period are calculated by multiplying the number of doses (13) by the WAC ($2,587.05).

When dosing depends on weight, costs are calculated separately for each patient in the simulation. In particular, costs are calculated as,

$$cost = \lceil weight \cdot dose_{amt}/strength_{amt} \rceil \cdot doses_{num} \cdot price, \tag{A26}$$

where $weight$ is patient weight, $dose_{amt}$ is the dose per weight, and $strength_{amt}$, $price$, and $doses_{num}$ are defined in the same way as in the non-weight based scenario. To illustrate, the acquisition cost (before discounts and rebates) for infliximab after the first 6 months is calculated by multiplying each patient's weight by the dose (6 mg/kg) and dividing by the size of a vial (100 mg), and then multiplying by the number of doses (8.67) and the price per unit ($1,167.82).

H Annualized costs and benefits

Letting t index time (in years), total costs ($\hat{cost}$) and QALYs ($\hat{qalys}$) for each patient simulated over a time horizon T are given by,

$$\hat{cost} = \sum_{t=1}^{T} c_t \beta_c^t \tag{A27}$$

$$\hat{qalys} = \sum_{t=1}^{T} q_t \beta_q^t, \tag{A28}$$

where costs and QALYs at time t, c_t and q_t, are discounted using the discount factors β_c and β_q, respectively. The discount factor is a function of the annual discount rate (typically assumed to be 0.03); that is, $\beta_s = 1/(1 + r_s)$ where r_s is the discount rate for $s = c, q$. The time horizon T is set to equal the maximum number of years that a patient could survive within the model—currently the maximum age that a patient could survive to is 100 given the default lifetables used within the model, so T is equal to 100 minus a patient's age at the start of the simulation.

Annualized QALY gains and costs, which are used to estimate the annual insurance value of treatment, can therefore be calculated using the formula for a geometric series and assuming a constant flow rate each model cycle (that is, by setting $c_t = c$ and $q_t = q$ in each time period). In particular, annualized costs c and QALYs q, can be derived by solving the following equations for c and q,

$$\hat{cost} = c\frac{1 - \beta_c^T}{1 - \beta_c} \tag{A29}$$

$$\hat{qalys} = q\frac{1 - \beta_q^T}{1 - \beta_q}. \tag{A30}$$

I Network Meta-Analysis

Treatment effects with tDMARDs relative to cDMARDs for moderate to severe RA patients who failed treatment with a conventional DMARD were estimated based on currently available evidence as reported in the literature. Relevant randomized controlled trials were identified with a systematic literature review and treatment effects were estimated by means of a network meta-analysis.

I.1 Systematic literature review to identify relevant studies

I.1.1 Eligibility criteria

Study eligibility criteria were defined in terms of the population, interventions, comparisons, outcomes, and study design (PICOS).

Table A5: Study eligibility criteria

Criteria	Description
Population	Adult (>18) patients with moderate to severe active rheumatoid arthritis who failed treatment with a conventional DMARD and were either tDMARD naive or experienced.
Interventions	Approved dosing regimens (or equivalent) of the following tDMARDs as monotherapy or in combination with a cDMARD: adalimumab, certolizumab pegol, etanercept, golimumab, infliximab, rituximab, abatacept, tocilizumab, sarilumab, anakinra, tofacitinib, baricitinib, upadacitinib, biosimilars; triple therapy (methotrexate + sulfasalazine + hydroxychloroquine)
Comparators	cDMARDs; placebo; any of the interventions of interest; any other intervention (or dosing regimen) that facilitates an indirect comparison between the interventions of interest
Outcomes	At least one of the following outcomes at 6 months of follow-up: ACR 20/50/70, HAQ-DI, DAS28
Study design	Randomized clinical trials
Other	Only full text reported in Enlgish were included. Studies only available as conference abstracts or presentations were excluded.

I.1.2 Literature search

Relevant studies were identified by searching the following databases: Medline, EMBASE, and Cochrane Central Register of Controlled Trials. The searches were executed May 2019 with the following predefined search strategies and corresponding results.

I.1.2.1 Medline

Table A6: Medline literature search strategy

Order	Search terms	Results
1	exp rheumatoid arthritis/	108,325
2	exp anakinra/	4,930
3	(Urine Interleukin 1 Inhibitor or Antril or Kineret or Anakinra).ti,ab.	1,538
4	exp abatacept/	2,724
5	(abatacept or Belatacept or Nulojix or Orencia or BMS-188667).ti,ab.	1,728
6	exp adalimumab/	4,699
7	(adalimumab or Humira or D2E7 Antibody).ti,ab.	5,935
8	exp certolizumab pegol/	518
9	(certolizumab pegol or certolizumab or Cimzia or Cimzias or CDP-870).ti,ab.	943
10	exp etanercept/	5,517
11	(etanercept or Enbrel or TNF Receptor Type II-IgG Fusion Protein or Recombinant Human Dimeric TNF Receptor Type II-IgG Fusion Protein).ti,ab.	6,417
12	golimumab.ti,ab.	903
13	(golimumab or Simponi or CNTO-148).ti,ab.	906
14	exp infliximab/	9,521
15	(infliximab or Remicade or Infliximab-dyyb or Inflectra or CenTNF or TA-650).ti,ab.	11,088
16	exp rituximab/	12,594
17	(rituximab or Mabthera or Rituxan or Rituximab CD20 Antibody or IDEC-102 or IDEC-C2B8 or IDEC-C2B8-anti-CD20).ti,ab.	18,312
18	tocilizumab.ti,ab.	2,330
19	(tocilizumab or atlizumab or Actemra or ACTPen).ti,ab.	2,350

Continued on next page

Table A6: Medline literature search strategy

Order	Search terms	Results
20	sarilumab.mp.	78
21	(sarilumab or iguratimod or Careram or Kolbet or T-614 or Kevzara or REGN-88 or SAR-153191).ti,ab.	161
22	tofacitinib.ti,ab.	858
23	(tasocitinib or tofacitinib citrate or Xeljanz or Cp-690,550 or CP-690550 or Jaquinus).ti,ab.	140
24	baricitinib.mp.	180
25	(LY309104 or INCB028050 or Olumiant).ti,ab.	10
26	upadacitinib.mp.	31
27	ABT-494.ti,ab.	8
28	exp methotrexate/	36,577
29	(amethopterin or mexate or methotrexate or Otrexup or VIBEX MTX or DepoMethotrexate or Jylamvo).ti,ab.	39,062
30	exp sulfasalazine/	4,011
31	(sulfasalazine or salicylazosulfapyridine or salazosulfapyridine or Pyralin or azulfadine or asulfidine or Colo-Pleon or Colo Pleon or Pleon or Ulcol or sulfasalazin or Ucine or salazopyrin).ti,ab.	3,643
32	exp hydroxychloroquine/	2,762
33	(hydroxychloroquine or oxychlorochin or oxychloroquine or hydroxychlorochin or plaquenil).ti,ab.	3,528
34	(28 or 29) and (30 or 31) and (32 or 33)	356
35	triple therapy.ti,ab.	5,454
36	or/2-29,34-35	105,887
37	Randomized Controlled Trials as Topic/	123,625
38	randomized controlled trial/	482,117
39	Random Allocation/	98,976
40	Double Blind Method/	151,262
41	Single Blind Method/	26,752
42	clinical trial/	516,228
43	clinical trial, phase i.pt.	18,921
44	clinical trial, phase ii.pt.	30,565
45	clinical trial, phase iii.pt.	15,033
46	clinical trial, phase iv.pt.	1,707
47	controlled clinical trial.pt.	93,070
48	randomized controlled trial.pt.	482,117
49	multicenter study.pt.	250,224
50	clinical trial.pt.	516,228
51	exp Clinical Trials as topic/	325,643
52	or/37-51	1,293,078
53	(clinical adj trial$).tw.	332,293
54	((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw.	163,549
55	PLACEBOS/	34,343
56	placebo$.tw.	204,048
57	randomly allocated.tw.	26,182
58	(allocated adj2 random$).tw.	29,328
59	or/53-58	588,391
60	52 or 59	1,534,440
61	case report.tw.	276,026
62	letter/	1,026,586
63	historical article/	351,546
64	or/61-63	1,639,252
65	60 not 64	1,499,796
66	1 and 36 and 65	3,747

I.1.2.2 Embase

Table A7: Embase literature search strategy

Order	Search terms	Results
1	exp rheumatoid arthritis/	187,909
2	*Rheumatoid Arthritis/dt, dm, co, th, dr	34,018
3	1 or 2	187,909
4	exp anakinra/	1,822
5	(Urine Interleukin 1 Inhibitor or Antril or Kineret or Anakinra).ti,ab.	3,292
6	exp adalimumab/	29,758
7	(adalimumab or Humira or D2E7 Antibody).ti,ab.	15,909
8	exp certolizumab pegol/	5,753
9	(certolizumab pegol or Cimzia or Cimzias or CDP-870).ti,ab.	1,811
10	exp etanercept/	29,146
11	(etanercept or Enbrel or TNF Receptor Type II-IgG Fusion Protein or Recombinant Human Dimeric TNF Receptor Type II-IgG Fusion Protein).ti,ab.	13,565
12	exp golimumab/	5,898
13	(golimumab or Simponi or CNTO-148).ti,ab.	3,191
14	exp infliximab/	46,385
15	(infliximab or Remicade or Infliximab-dyyb or Inflectra or CenTNF or TA-650).ti,ab.	23,567
16	exp rituximab/	71,192
17	(rituximab or Mabthera or Rituxan or Rituximab CD20 Antibody or IDEC-102 or IDEC-C2B8 or IDEC-C2B8-anti-CD20).ti,ab.	40,561
18	exp abatacept/	8,278
19	(abatacept or Belatacept or Nulojix or Orencia or BMS-188667).ti,ab.	4,866
20	exp tocilizumab/	9,614
21	(tocilizumab or atlizumab or Actemra or ACTPen).ti,ab.	6,161
22	exp sarilumab/	374
23	exp iguratimod/	221
24	(sarilumab or iguratimod or Careram or Kolbet or T-614 or Kevzara or REGN-88 or SAR-153191).ti,ab.	438
25	exp tofacitinib/	3,251
26	(tasocitinib or tofacitinib citrate or Xeljanz or Cp-690,550 or CP-690550 or Jaquinus).ti,ab.	293
27	baricitinib.mp.	763
28	(LY309104 or INCB028050 or Olumiant).ti,ab.	25
29	upadacitinib.mp.	192
30	(ABT-494).ti,ab.	43
31	exp methotrexate/	167,405
32	(amethopterin or mexate or methotrexate or Otrexup or VIBEX MTX or DepoMethotrexate or Jylamvo).ti,ab.	62,417
33	exp sulfasalazine/	24,168
34	(sulfasalazine or salicylazosulfapyridine or salazosulfapyridine or Pyralin or azulfadine or asulfidine or Colo-Pleon or Colo Pleon or Pleon or Ulcol or sulfasalazin or Ucine or salazopyrin).ti,ab.	5,854
35	exp hydroxychloroquine/	21,806
36	(hydroxychloroquine or oxychlorochin or oxychloroquine or hydroxychlorochin or plaquenil).ti,ab.	7,062
37	(31 or 32) and (33 or 34) and (35 or 36)	4,507
38	(triple therapy).ti,ab.	9,763
39	or/4-30, 37, 38	166,313
40	Clinical trial/	957,596
41	randomized controlled trial/	548,967
42	controlled clinical trial/	462,492
43	multicenter study/	215,446

Continued on next page

Table A7: EMBASE literature search strategy

Order	Search terms	Results
44	Phase 3 clinical trial/	39,689
45	Phase 4 clinical trial/	3,404
46	exp RANDOMIZATION/	82,570
47	single blind procedure/	35,014
48	double blind procedure/	160,287
49	crossover procedure/	59,104
50	PLACEBO/	333,751
51	randomi?ed controlled trial$.tw.	201,791
52	rct.tw.	32,266
53	(random$ adj2 allocat$).tw.	39,742
54	single blind$.tw.	22,863
55	double blind$.tw.	197,482
56	((treble or triple) adj blind$).tw.	956
57	placebo$.tw.	288,822
58	Prospective study/	518,790
59	or/40-58	2,131,018
60	case study/	61,210
61	case report.tw.	381,208
62	Abstract report/ or letter/	1,099,876
63	Conference proceeding.pt.	0
64	Conference abstract.pt.	3,404,073
65	Editorial.pt.	600,353
66	Letter.pt.	1,063,962
67	Note.pt.	751,642
68	or/60-67	6,241,081
69	59 not 68	1,609,889
70	3 and 39 and 69	6,981

I.1.2.3 Cochrane Central Register of Controlled Trials

Table A8: Cochrane Central Register of Controlled Trials literature search strategy

Order	Search terms	Results
1	exp Arthritis, Rheumatoid/	5,326
2	exp Interleukin 1 Receptor Antagonist Protein/	253
3	(Urine Interleukin 1 Inhibitor or Antril or Kineret or Anakinra).ti,ab.	289
4	adalimumab.ti,ab.	2,529
5	(Humira or D2E7 Antibody).ti,ab.	322
6	certolizumab pegol.ti,ab.	446
7	(Cimzia or Cimzias or CDP-870).ti,ab.	30
8	etanercept.ti,ab.	1,886
9	(Enbrel or TNF Receptor Type II-IgG Fusion Protein or Recombinant Human Dimeric TNF Receptor Type II-IgG Fusion Protein).ti,ab.	211
10	(Golimumab or CNTO-148).ti,ab.	602
11	Simponi.ti,ab.	26
12	infliximab.ti,ab.	2,129
13	Remicade.ti,ab.	204
14	rituximab.ti,ab.	4,112
15	(Mabthera or Rituxan or Rituximab CD20 Antibody or IDEC-102 or IDEC-C2B8 or IDEC-C2B8-anti-CD20).ti,ab.	372

Continued on next page

Table A8: Cochrane Central Register of Controlled Trials literature search strategy

Order	Search terms	Results
16	abatacept.ti,ab.	610
17	(Belatacept or Nulojix or Orenciaor BMS-188667).ti,ab.	247
18	tocilizumab.ti,ab.	875
19	(atlizumab or Actemra or ACTPen).ti,ab.	44
20	Sarilumab.mp.	167
21	iguratimod.ti,ab.	41
22	(sarilumab or Careram or Kolbet or T-614 or Kevzara or REGN-88 or SAR-153191).ti,ab.	178
23	tofacitinib.ti,ab.	517
24	(tasocitinib or tofacitinib citrate or Xeljanz or Cp-690,550 or CP-690550 or Jaquinus).ti,ab.	123
25	baricitinib.mp.	279
26	(LY309104 or INCB028050 or Olumiant).ti,ab.	28
27	upadacitinib.mp.	97
28	ABT-494.ti,ab.	82
29	methotrexate.mp.	10,721
30	(amethopterin or mexate or methotrexate or Otrexup or VIBEX MTX or DepoMethotrexate or Jylamvo).ti,ab.	8,915
31	sulfasalazine.mp.	962
32	(sulfasalazine or salicylazosulfapyridine or salazosulfapyridine or Pyralin or azulfadine or asulfidine or Colo-Pleon or Colo Pleon or Pleon or Ulcol or sulfasalazin or Ucine or salazopyrin).ti,ab.	841
33	hydroxychloroquine.mp.	912
34	(hydroxychloroquine or oxychlorochin or oxychloroquine or hydroxychlorochin or plaquenil).ti,ab.	751
35	(29 or 30) and (31 or 32) and (33 or 34)	192
36	triple therapy.ti,ab.	2,896
37	or/2-28,35-36	15,492
38	1 and 37	1,090

I.1.3 Study selection

Two investigators working independently scanned all abstracts identified in the literature search. The same two investigators independently reviewed relevant abstracts in full-text. Discrepancies occurring between the studies selected by the two investigators were resolved by involving a third investigator and reaching consensus.

I.1.4 Data extraction

Two investigators working independently extracted relevant data on study characteristics, interventions, patient characteristics, and outcomes from the final list of selected eligible studies. Discrepancies observed between the data extracted by the two investigators were resolved by involving a third investigator and reaching consensus.

I.2 Analyses

In order to perform a network meta-analysis where the risk of biased relative treatment effect estimates is limited we need to have 1) a single evidence network where each randomized controlled trial has at least one intervention in common with another trial; and 2) no differences in study

designs and the distribution of patient characteristics that act as relative treatment effect-modifiers across the studies in the network (ref). For both the tDMARD naive population and the tDMARD experienced population a connected evidence network could be defined, however for the latter population the treatment history across studies was considered too different to obtain valid results from a network meta-analysis, and the limited number of studies would not allow adjusting for these differences with statistical techniques. Hence, network meta-analyses were only performed for the tDMARD naive population. Studies that reported results for a mixed population regarding prior tDMARD use were excluded from the analyses; only studies with results reported for a tDMARD naive population were included.

Treatment effects at 6 months with each of the interventions in the network relative to cDMARDs were estimated in terms of ACR response, change from baseline in HAQ, and change from baseline in DAS28 with Bayesian random effects network meta-analysis models as presented by Dias et al. 2013. To avoid influencing the observed results by prior belief, uninformative prior distributions were used for the estimated treatment effect and between-study heterogeneity parameters. Posterior distributions for the model parameters are derived with the Markov chain Monte Carlo method using the JAGS software package (http://mcmcjags.sourceforge.net/). The studies included in each analysis for the tDMARD naive population were considered sufficiently similar regarding study design and distribution of patient characteristics that may act as relative treatment effect-modifiers. Accordingly, no meta-regression was performed to adjust for between-trial differences and the obtained estimated for each of the interventions were considered reflective of this target population of interest.

I.2.1 ACR response at 6 months

The four mutually exclusive ACR response categories were estimated from the overlapping ACR categories using a ordered probit model appropriate for ordered categorical data (Dias et al. 2013). The model assumes that there is an underlying continuous variable (ACR20/50/70) categorized by specifying different cutoffs corresponding to the point at which an individual moves from one category to the next in each trial. The advantage of this approach over an analysis that considers ACR categories separately is that all possible outcomes are analyzed simultaneously based on the same randomized controlled trials, allowing for consistent estimates by category.

More specifically, let r_{jkl} be the number of patients in trial j for treatment k in the mutually exclusive category $l = 1, 2, 3, 4$. The model can be written as,

$$r_{jkl} \sim \text{Multinomial}(p_{jk1}, p_{jk2}, p_{jk3}, p_{jk4}, n_{jk}), \tag{A31}$$

where p_{jkl} is the probability that a patient from trial j and treatment k is in category l and there are n_{jk} patients in trial j and treatment k. Probabilities are modeled using a probit function,

$$\Phi^{-1}(p_{jkl}) = \begin{cases} u_{jb} + z_{jl} & \text{if } k = b \\ u_{jb} + z_{jl} + \delta_{jbk} & \text{if } k \succ b, \end{cases} \tag{A32}$$

where u_j is a trial specific intercept, z_{jl} is a cutpoint for trial j and category l, and δ_{jbk} is the effect of treatment k relative to treatment b. The cutpoint for category c, z_{jc}, is modeled as random,

$$z_{jc} \sim N(v_c, \sigma_z^2). \tag{A33}$$

The study-specific relative treatment effects are also drawn from a common population distribution with mean d_{bk} and variance τ^2,

$$\delta_{jbk} \sim N(d_{bk}, \tau^2), \tag{A34}$$

where $d_{bk} = d_{Ak} - d_{Ab}$. To generate treatment responses, we estimate the response for treatment A by averaging μ_{jA} across trials containing treatment A. In particular, letting $S_A = \{\mu_{1A}, \ldots \mu_{|S_A|A}\}$ be the set of all trials containing treatment A, we estimate,

$$A = \frac{1}{|S_A|} \sum_{\mu_A \in S_A} \mu_A. \tag{A35}$$

We calculate the probability of ACR $< 20\%$ improvement, ACR $< 50\%$ improvement, and ACR $< 70\%$ improvement with treatment k as,

$$P(ACR_k < 70) = \phi(A + z_3 + d_{kA}) \tag{A36}$$
$$P(ACR_k < 50) = \phi(A + z_2 + d_{kA}) \tag{A37}$$
$$P(ACR_k < 20) = \phi(A + d_{kA}). \tag{A38}$$

The probabilities of overlapping ACR response (i.e., ACR 20/50/70) are then,

$$P(ACR_k > 70) = \gamma \cdot (1 - P(ACR_k < 70)) \tag{A39}$$
$$P(ACR_k > 50) = \gamma \cdot (1 - P(ACR_k < 50)) \tag{A40}$$
$$P(ACR_k > 20) = \gamma \cdot (1 - P(ACR_k < 20)), \tag{A41}$$

where γ is the reduction in treatment response at a given line of therapy. $\gamma = 1$ is a patient is bDMARD naive and on average, equal to .84 after failing a biologic. The mutually exclusive ACR response categories are easily derived from the overlapping categories.

To avoid influencing the observed results by prior belief, uninformative prior distributions were used for the estimated model parameters. Posterior distributions for the model parameters are derived with the Markov chain Monte Carlo method.

I.2.2 Change in HAQ and DAS28 at 6 months

The models of changes in HAQ and DAS28 at 6 months use a normal likelihood (since the sample mean is approximately normally distributed by the central limit theorem if the sample size is reasonably large) and an identity link.

More specifically, let y_{jk} be the outcome of interest in trial j and treatment k, and consider the model,

$$y_{jk} \sim N(\theta_{jk}, \sigma^2_{jk}), \tag{A42}$$

where,

$$\theta_{jk} = \begin{cases} \mu_{jb} & \text{if } k = b \\ \mu_{jb} + \delta_{jbk} & \text{if } k \succ b. \end{cases} \tag{A43}$$

δ_{jbk} is modeled using a random effect with $d_{AA} = 0$,

$$\delta_{jbk} \sim N(d_{bk}, \sigma^2), \tag{A44}$$

where $d_{bk} = d_{Ak} - d_{Ab}$. As with the ACR response model, we allow treatment response to depend on patient characteristics by modeling d_{bk} as a function of covariates for each individual patient i,

$$d_{bk} = x_i^T \beta_{bk}, \tag{A45}$$

In the simulation, we allow for treatment effect modifiers by modeling d_{bk} as a function of covariates for each individual patient i,

$$d_{bk} = x_i^T \beta_{bk}. \tag{A46}$$

The absolute treatment effect is estimated by calculating A as in Equation A35. The absolute treatment effect for treatment k is then,

$$\gamma(A + d_{kA}), \tag{A47}$$

where γ is defined as in Equation A39. Uninformative priors were used to derive the posterior distributions.

I.3 Evidence base

I.3.1 Study identification and selection

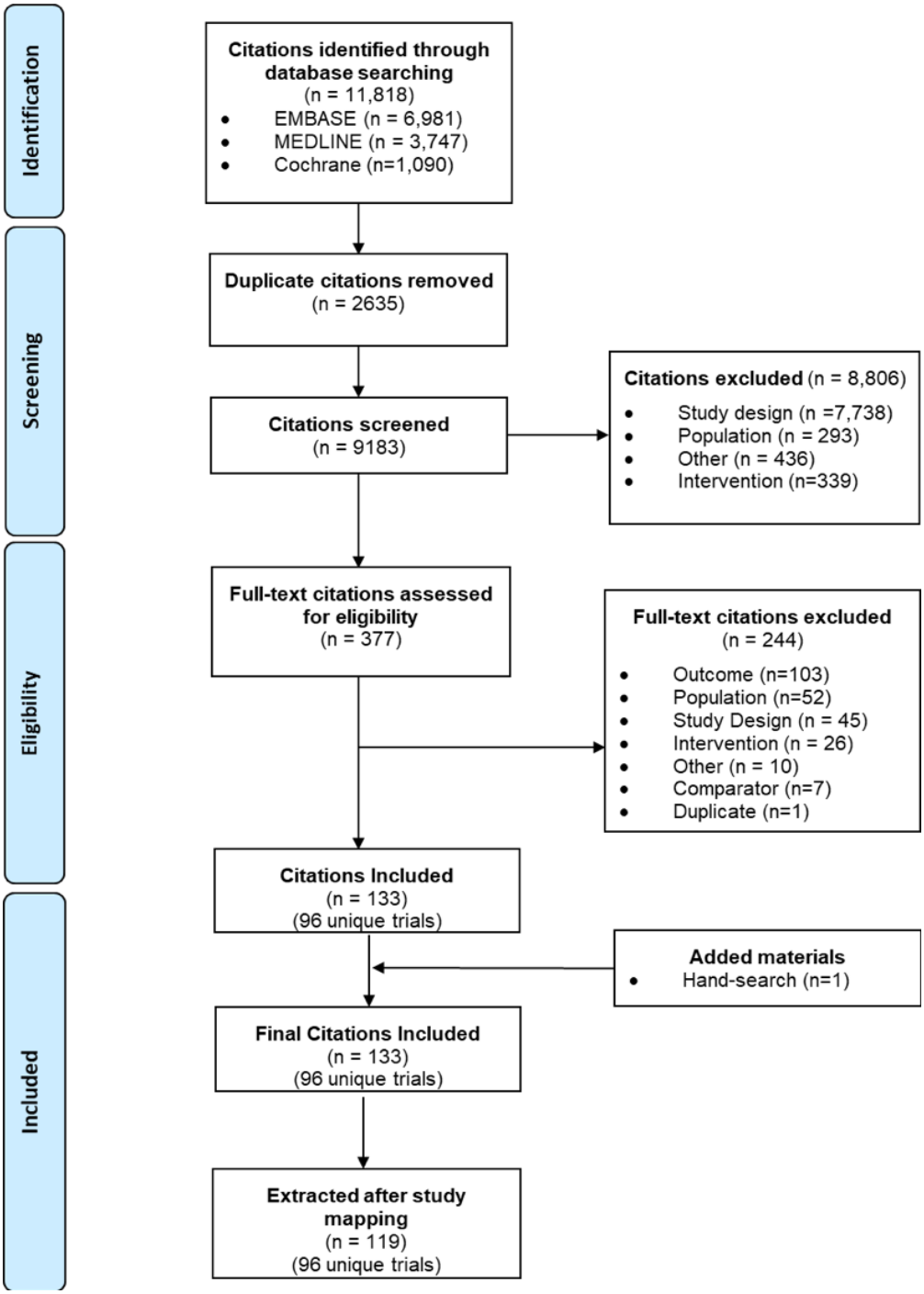


Figure A2: Summary of the study identification and selection process

I.3.2 Included studies

Table A9: Studies meeting the eligibility criteria for inclusion in the evidence base

Trial ID	Author and Year	Title	Journal
ACQUIRE	Genovese, 2011	Subcutaneous abatacept versus intravenous abatacept: A phase iiib noninferiority study in patients with an inadequate response to MTX	Arthritis and Rheumatism
ACT-RAY	Dougados, 2013	Adding tocilizumab or switching to tocilizumab monotherapy in methotrexate inadequate responders: 24-week symptomatic and structural results of a 2-year randomised controlled strategy trial in rheumatoid arthritis (act-ray)	Annals of the Rheumatic Diseases
	Dougados, 2014	Clinical, radiographic and immunogenic effects after 1 year of tocilizumab-based treatment strategies in rheumatoid arthritis: The act-ray study	Annals of the Rheumatic Diseases
ACT-STAR	Weinblatt, 2013	Tocilizumab as monotherapy or in combination with nonbiologic disease-modifying antirheumatic drugs: Twenty-four-week results of an open-label, clinical practice study	Arthritis Care and Research
ADACTA	Gabay, 2013	Tocilizumab monotherapy versus adalimumab monotherapy for treatment of rheumatoid arthritis (adacta): A randomised, double-blind, controlled phase 4 trial	Lancet
AIM	Kremer, 2006	Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: A randomized trial	Annals of Internal Medicine
AMPLE	Weinblatt, 2013	Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: Findings of a phase iiib, multinational, prospective, randomized study	Arthritis and Rheumatism
ARMADA	Weinblatt, 2003	Adalimumab, a fully human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: The armada trial	Arthritis and Rheumatism
ASCERTAIN	Emery, 2018	Safety and tolerability of subcutaneous sarilumab and intravenous tocilizumab in patients with rheumatoid arthritis	Rheumatology
ATTAIN	Emery, 2005	Abatacept has beneficial effects in rheumatoid arthritis patients with an inadequate response to anti-tnfalpha therapy	Clinical and Experimental Rheumatology
	Genovese, 2005	Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition	New England Journal of Medicine
	Westhovens, 2006	Improved health-related quality of life for rheumatoid arthritis patients treated with abatacept who have inadequate response to anti-tnf therapy in a double-blind, placebo-controlled, multicentre randomized clinical trial	Rheumatology
ATTAIN;AIM	Wells, 2009	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	Annals of the Rheumatic Diseases
ATTEST	Schiff, 2008	Efficacy and safety of abatacept or infliximab vs placebo in attest: A phase iii, multi-centre, randomised, double-blind, placebo-controlled study in patients with rheumatoid arthritis and an inadequate response to methotrexate	Annals of the Rheumatic Diseases
ATTRACT	Lipsky, 2000	Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study group	New England Journal of Medicine
	Maini, 2004	Sustained improvement over two years in physical function, structural damage, and signs and symptoms among patients with rheumatoid arthritis treated with and methotrexate	Arthritis and Rheumatism

Continued on next page

Trial ID	Author and Year	Title	Journal
	Maini, 1999	Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: A randomised phase iii trial. Attract study group	Lancet
Bao 2011	Bao et al, 2011	Secondary failure to treatment with recombinant human il-1 receptor antagonist in chinese patients with rheumatoid arthritis	Clinical Rheumatology
BREVACTA	Kivitz, 2014	Subcutaneous tocilizumab versus placebo in combination with disease-modifying antirheumatic drugs in patients with rheumatoid arthritis	Arthritis Care and Research
CHANGE	Miyasaka, 2008	Clinical investigation in highly disease-affected rheumatoid arthritis patients in japan with adalimumab applying standard and general evaluation: The change study	Modern Rheumatology
Choy 2012	Choy, 2012	Certolizumab pegol plus mtx administered every 4 weeks is effective in patients with ra who are partial responders to mtx	Rheumatology
Cohen 2002	Cohen et al, 2002	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 and Rheumatism
	Cohen et al, 2003	Interleukin 1 receptor antagonist anakinra improves functional status in patients with rheumatoid arthritis	Journal of Rheumatology
Cohen 2004	Cohen et al, 2004	A multicentre, double blind, randomised, placebo controlled trial of anakinra (kineret), a recombinant interleukin 1 receptor antagonist, in patients with rheumatoid arthritis treated with background methotrexate	Annals of the Rheumatic Diseases
Cohen 2017	Cohen, 2017	Efficacy and safety of the biosimilar abp 501 compared with adalimumab in patients with moderate to severe rheumatoid arthritis: A randomised, double-blind, phase iii equivalence study	Annals of the Rheumatic Diseases
Cohen 2018	Cohen, 2018	A randomized controlled trial comparing pf-06438179/gp1111 (an infliximab biosimilar) and infliximab reference product for treatment of moderate to severe active rheumatoid arthritis despite methotrexate therapy	Arthritis Research and Therapy
Combe 2006	Combe, 2009	Efficacy, safety and patient-reported outcomes of combination etanercept and sulfasalazine versus etanercept alone in patients with rheumatoid arthritis: A double-blind randomised 2-year study	Annals of the Rheumatic Diseases
DANCER	Emery, 2006	The efficacy and safety of rituximab in patients with active rheumatoid arthritis despite methotrexate treatment: Results of a phase iib randomized, double-blind, placebo-controlled, dose-ranging trial	Arthritis and Rheumatism
	Mease, 2008	Improved health-related quality of life for patients with active rheumatoid arthritis receiving rituximab: Results of the dose-ranging assessment: International clinical evaluation of rituximab in rheumatoid arthritis (dancer) trial	Journal of rheumatology
De Filippis 2006	De Filippis, 2006	Improving outcomes in tumour necrosis factor a treatment: Comparison of the efficacy of the tumour necrosis factor a blocking agents etanercept and infliximab in patients with active rheumatoid arthritis	Panminerva Medica
DE019	Keystone, 2004	Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: A randomized, placebo-controlled, 52-week trial	Arthritis and Rheumatism
Edwards 2004	Edwards, 2004	Efficacy of b-cell-targeted therapy with rituximab in patients with rheumatoid arthritis	New England Journal of Medicine

Continued on next page

Trial ID	Author and Year	Title	Journal
Elmedany 2019	Elmedany, 2019	Efficacy and safety profile of intravenous tocilizumab versus intravenous abatacept in treating female saudi arabian patients with active moderate-to-severe rheumatoid arthritis	Clinical Rheumatology
Emery 2017	Emery, 2017	A phase iii randomised, double-blind, parallel-group study comparing sb4 with etanercept reference product in patients with active rheumatoid arthritis despite methotrexate therapy	Annals of the Rheumatic Diseases
EQUIRA	Matucci-Cerinic, 2018	Efficacy, safety and immunogenicity of gp2015, an etanercept biosimilar, compared with the reference etanercept in patients with moderate-to-severe rheumatoid arthritis: 24-week results from the comparative phase iii, randomised, double-blind equira study	Rheumatic and Musculoskeletal Diseases
ETN Study 309	Combe, 2006	Etanercept and sulfasalazine, alone and combined, in patients with active rheumatoid arthritis despite receiving sulfasalazine: A double-blind comparison	Annals of the Rheumatic Diseases
FAST4WARD	Fleischmann, 2009	Efficacy and safety of certolizumab pegol monotherapy every 4 weeks in patients with rheumatoid arthritis failing previous disease-modifying antirheumatic therapy: The fast4ward study	Annals of the Rheumatic Diseases
Fleischmann 2012	Fleischmann, 2012	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 and Rheumatism
	Wallenstein, 2016	Effects of the oral janus kinase inhibitor tofacitinib on patient-reported outcomes in patients with active rheumatoid arthritis: Results of two phase 2 randomised controlled trials	Clinical and Experimental Rheumatology
Fleischmann 2018	Fleischmann, 2018	A comparative clinical study of pf-06410293, a candidate adalimumab biosimilar, and adalimumab reference product (humira) in the treatment of active rheumatoid arthritis	Arthritis Research and Therapy
GO AFTER	Smolen, 2015	Golimumab in patients with active rheumatoid arthritis after treatment with tumor necrosis factor alpha inhibitors: Findings with up to five years of treatment in the multicenter, randomized, double-blind, placebo-controlled, phase 3 go-after study	Arthritis Research and Therapy
GO FURTHER	Bingham, 2014	The effect of intravenous golimumab on health-related quality of life in rheumatoid arthritis: 24-week results of the phase iii go-further trial	Journal of Rheumatology
GO-AFTER	Smolen, 2009	Golimumab in patients with active rheumatoid arthritis after treatment with tumour necrosis factor alpha inhibitors (go-after study): A multicentre, randomised, double-blind, placebo-controlled, phase iii trial	Lancet
	Smolen, 2013	Insights into the efficacy of golimumab plus methotrexate in patients with active rheumatoid arthritis who discontinued prior anti-tumour necrosis factor therapy: Post-hoc analyses from the go-after study	Annals of the Rheumatic Diseases
GO-FORTH	Tanaka, 2012	Golimumab in combination with methotrexate in japanese patients with active rheumatoid arthritis: Results of the go-forth study	Annals of the Rheumatic Diseases
	Tanaka, 2016	Clinical efficacy, radiographic progression, and safety through 156 weeks of therapy with subcutaneous golimumab in combination with methotrexate in japanese patients with active rheumatoid arthritis despite prior methotrexate therapy: Final results of the randomized go-forth trial	Modern rheumatology
GO-FORWARD	Genovese, 2012	Effect of golimumab on patient-reported outcomes in rheumatoid arthritis: Results from the go-forward study	Journal of rheumatology

Continued on next page

Trial ID	Author and Year	Title	Journal
	Keystone, 2009	Golimumab, a human antibody to tumour necrosis factor alpha given by monthly subcutaneous injections, in active rheumatoid arthritis despite methotrexate therapy: The go-forward study	Annals of the Rheumatic Diseases
	Keystone, 2010	Golimumab in patients with active rheumatoid arthritis despite methotrexate therapy: 52-week results of the go-forward study	Annals of the Rheumatic Diseases
GO-FURTHER	Weinblatt, 2013	Intravenous golimumab is effective in patients with active rheumatoid arthritis despite methotrexate therapy with responses as early as week 2: Results of the phase 3, randomised, multicentre, double-blind, placebo-controlled go-further trial	Annals of the Rheumatic Diseases
GO-LIVE	Kremer, 2010	Golimumab, a new human anti-tumor necrosis factor alpha antibody, administered intravenously in patients with active rheumatoid arthritis: Forty-eight-week efficacy and safety results of a phase iii randomized, double-blind, placebo-controlled study	Arthritis and Rheumatism
	Weinblatt, 2013	Radiographic benefit and maintenance of clinical benefit with intravenous golimumab therapy in patients with active rheumatoid arthritis despite methotrexate therapy: Results up to 1 year of the phase 3, randomised, multicentre, double blind, placebo controlled go-further trial	Annals of the Rheumatic Diseases
GO-SAVE	Huffstutter, 2017	Clinical response to golimumab in rheumatoid arthritis patients who were receiving etanercept or adalimumab: Results of a multicenter active treatment study	Current Medical Research and Opinion
HERA	Bae, 2017	A phase iii, multicentre, randomised, double-blind, active-controlled, parallel-group trial comparing safety and efficacy of hd203, with innovator etanercept, in combination with methotrexate, in patients with rheumatoid arthritis: The hera study	Annals of the Rheumatic Diseases
HIKARI	Yamamoto, 2014	Efficacy and safety of certolizumab pegol without methotrexate co-administration in japanese patients with active rheumatoid arthritis: The hikari randomized, placebo-controlled trial	Modern rheumatology
Iwahashi 2014	Iwahashi, 2014	Efficacy, safety, pharmacokinetics and immunogenicity of abatacept administered subcutaneously or intravenously in japanese patients with rheumatoid arthritis and inadequate response to methotrexate: A phase ii/iii, randomized study	Modern rheumatology
Janmshidi 2017	Jamshidi, 2017	A phase iii, randomized, two-armed, double-blind, parallel, active controlled, and non-inferiority clinical trial to compare efficacy and safety of biosimilar adalimumab (cinnora(r)) to the reference product (humira(r)) in patients with active rheumatoid arthritis	Arthritis Research and Therapy
JESMR	Kameda, 2010	Etanercept (etn) with methotrexate (mtx) is better than etn monotherapy in patients with active rheumatoid arthritis despite mtx therapy: A randomized trial	Modern rheumatology
J-RAPID	Yamamoto, 2014	Efficacy and safety of certolizumab pegol plus methotrexate in japanese rheumatoid arthritis patients with an inadequate response to methotrexate: The j-rapid randomized, placebo-controlled trial	Modern rheumatology
Kim 2007	Kim, 2007	A randomized, double-blind, placebo-controlled, phase iii study of the human anti-tumor necrosis factor antibody adalimumab administered as subcutaneous injections in korean rheumatoid arthritis patients treated with methotrexate	APLAR Journal of Rheumatology
Kremer 2003	Kremer, 2003	Treatment of rheumatoid arthritis by selective inhibition of t-cell activation with fusion protein ctla4ig	New England Journal of Medicine

Continued on next page

Trial ID	Author and Year	Title	Journal
Kremer 2005	Kremer, 2005	Treatment of rheumatoid arthritis with the selective costimulation modulator abatacept: Twelve-month results of a phase iib, double-blind, randomized, placebo-controlled trial	Arthritis and Rheumatism
Kremer 2012	Kremer, 2012	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 and Rheumatism
LARA	Machado, 2014	Open-label observation of addition of etanercept versus a conventional disease-modifying antirheumatic drug in subjects with active rheumatoid arthritis despite methotrexate therapy in the latin american region	Journal of Clinical Rheumatology
Li 2016	Li, 2016	Efficacy and safety results from a phase 3, randomized, placebo-controlled trial of subcutaneous golimumab in chinese patients with active rheumatoid arthritis despite methotrexate therapy	International Journal of Rheumatic Diseases
LITHE	Kremer, 2011	Tocilizumab inhibits structural joint damage in rheumatoid arthritis patients with inadequate responses to methotrexate: Results from the double-blind treatment phase of a randomized placebo-controlled trial of tocilizumab safety and prevention of structural joint damage at one year	Arthritis and Rheumatism
Matsubara 2018	Matsubara, 2018	Abatacept in combination with methotrexate in japanese biologic-naive patients with active rheumatoid arthritis: A randomised placebocontrolled phase iv study	Rheumatic and Musculoskeletal Diseases
Matsuno 2018a	Matsuno, 2018	Phase iii, multicentre, double-blind, randomised, parallel-group study to evaluate the similarities between lbec0101 and etanercept reference product in terms of efficacy and safety in patients with active rheumatoid arthritis inadequately responding to methotrexate	Annals of the Rheumatic Diseases
MOBILITY	Genovese, 2015	Sarilumab plus methotrexate in patients with active rheumatoid arthritis and inadequate response to methotrexate: Results of a phase iii study	Arthritis and Rheumatology
	Strand, 2016	Sarilumab plus methotrexate improves patient-reported outcomes in patients with active rheumatoid arthritis and inadequate responses to methotrexate: Results of a phase iii trial	Arthritis Research and Therapy
MONARCH	Burmester, 2017	Efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy for the treatment of patients with active rheumatoid arthritis (monarch): A randomised, double-blind, parallel-group phase iii trial	Annals of the Rheumatic Diseases
	Strand, 2018	Patient-reported outcomes from a randomized phase iii trial of sarilumab monotherapy versus adalimumab monotherapy in patients with rheumatoid arthritis	Arthritis Research and Therapy
Moreland 1999	Mathias, 2000	Health-related quality of life and functional status of patients with rheumatoid arthritis randomly assigned to receive etanercept or placebo	Clinical Therapeutics
	Moreland, 1999	Etanercept therapy in rheumatoid arthritis. A randomized, controlled trial	Annals of Internal Medicine
MUSASHI	Ogata, 2014	Phase iii study of the efficacy and safety of subcutaneous versus intravenous tocilizumab monotherapy in patients with rheumatoid arthritis	Arthritis Care and Research
Niu 2011	Niu et al, 2011	Regulatory immune responses induced by il-1 receptor antagonist in rheumatoid arthritis	Molecular Immunology
OPTION	Smolen, 2008	Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (option study): A double-blind, placebo-controlled, randomised trial	Lancet
ORAL SCAN	Fleischmann, 2017	Efficacy of tofacitinib in patients with rheumatoid arthritis stratified by background methotrexate dose group	Clinical Rheumatology

Continued on next page

Trial ID	Author and Year	Title	Journal
	van der Heijde, 2013	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 and Rheumatism
ORAL-SOLO	Strand, 2015	Effects of tofacitinib monotherapy on patient-reported outcomes in a randomized phase 3 study of patients with active rheumatoid arthritis and inadequate responses to dmards	Arthritis Research and Therapy
ORAL-STANDARD	Strand, 2016	Tofacitinib or adalimumab versus placebo: Patient-reported outcomes from a phase 3 study of active rheumatoid arthritis	Rheumatology
	van Vollenhoven, 2012	Tofacitinib or adalimumab versus placebo in rheumatoid arthritis	New England Journal of Medicine
ORAL-STEP	Burmester, 2013	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
ORAL-STRATEGY	Fleischmann, 2017	Efficacy and safety of tofacitinib monotherapy, tofacitinib with methotrexate, and adalimumab with methotrexate in patients with rheumatoid arthritis (oral strategy): A phase 3b/4, double-blind, head-to-head, randomised controlled trial	Lancet
ORAL-SYNC	Kremer, 2013	Tofacitinib in combination with nonbiologic disease-modifying antirheumatic drugs in patients with active rheumatoid arthritis: A randomized trial	Annals of Internal Medicine
	Li, 2018	Efficacy and safety of tofacitinib in chinese patients with rheumatoid arthritis	Chinese Medical Journal
	Strand, 2017	Tofacitinib in combination with conventional disease-modifying antirheumatic drugs in patients with active rheumatoid arthritis: Patient-reported outcomes from a phase iii randomized controlled trial	Arthritis Care and Research
RA-BEACON	Genovese, 2016	Baricitinib in patients with refractory rheumatoid arthritis	New England Journal of Medicine
	Genovese, 2018	Response to baricitinib based on prior biologic use in patients with refractory rheumatoid arthritis	Rheumatology
	Smolen, 2017	Patient-reported outcomes from a randomised phase iii study of baricitinib in patients with rheumatoid arthritis and an inadequate response to biological agents (ra-beacon)	Annals of the Rheumatic Diseases
RA-BEAM	Keystone, 2017	Patient-reported outcomes from a phase 3 study of baricitinib versus placebo or adalimumab in rheumatoid arthritis: Secondary analyses from the ra-beam study	Annals of the Rheumatic Diseases
	Taylor, 2017	Baricitinib versus placebo or adalimumab in rheumatoid arthritis	New England Journal of Medicine
RA-BEAM;RA-BUILD;RA-BEACON	Tanaka, 2018	Efficacy and safety of baricitinib in japanese patients with rheumatoid arthritis: Subgroup analyses of four multinational phase 3 randomized trials	Modern rheumatology
RA-BUILD-A; RA-BUILD-Ba	Dougados, 2017	Baricitinib in patients with inadequate response or intolerance to conventional synthetic dmards: Results from the ra-build study	Annals of the Rheumatic Diseases
	Emery, 2017	Patient-reported outcomes from a phase iii study of baricitinib in patients with conventional synthetic dmard-refractory rheumatoid arthritis	Rheumatic and Musculoskeletal Diseases
RACAT	O'Dell, 2013	Therapies for active rheumatoid arthritis after methotrexate failure	New England Journal of Medicine

Continued on next page

Trial ID	Author and Year	Title	Journal
RADIATE	Emery, 2008	Il-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: Results from a 24-week multicentre randomised placebo-controlled trial	Annals of the Rheumatic Diseases
	Strand, 2012	Improvements in health-related quality of life after treatment with tocilizumab in patients with rheumatoid arthritis refractory to tumour necrosis factor inhibitors: Results from the 24-week randomized controlled radiate study	Rheumatology
RAPID-1	Keystone, 2008	Certolizumab pegol plus methotrexate is significantly more effective than placebo plus methotrexate in active rheumatoid arthritis: Findings of a fifty-two-week, phase iii, multicenter, randomized, double-blind, placebo-controlled, parallel-group study	Arthritis and Rheumatism
RAPID-2	Smolen, 2009	Efficacy and safety of certolizumab pegol plus methotrexate in active rheumatoid arthritis: The rapid 2 study. A randomised controlled trial	Annals of the Rheumatic Diseases
	Strand, 2011	Certolizumab pegol plus methotrexate provides broad relief from the burden of rheumatoid arthritis: Analysis of patient-reported outcomes from the rapid 2 trial	Annals of the Rheumatic Diseases
RAPID-C	Bi, 2019	Efficacy and safety of certolizumab pegol in combination with methotrexate in methotrexate-inadequate responder chinese patients with active rheumatoid arthritis: 24-week results from a randomised, double-blind, placebo-controlled phase 3 study	Clinical and Experimental Rheumatology
RA-SCORE	Peterfy, 2016	Mri assessment of suppression of structural damage in patients with rheumatoid arthritis receiving rituximab: Results from the randomised, placebo-controlled, double-blind ra-score study	Annals of the Rheumatic Diseases
RED SEA	Jobanputra, 2012	A randomised efficacy and discontinuation study of etanercept versus adalimumab (red sea) for rheumatoid arthritis: A pragmatic, unblinded, non-inferiority study of first tnf inhibitor use: Outcomes over 2 years	BMJ Open
REFLEX	Cohen, 2006	Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: Results of a multicenter, randomized, double-blind, placebo-controlled, phase iii trial evaluating primary efficacy and safety at twenty-four weeks	Arthritis and Rheumatism
	Keystone, 2008	Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid arthritis refractory to anti-tumor necrosis factor therapy	Arthritis and Rheumatism
ROSE	Yazici, 2012	Efficacy of tocilizumab in patients with moderate to severe active rheumatoid arthritis and a previous inadequate response to disease-modifying antirheumatic drugs: The rose study	Annals of the Rheumatic Diseases
SAMURAI	Nishimoto, 2007	Study of active controlled monotherapy used for rheumatoid arthritis, an il-6 inhibitor (samurai): Evidence of clinical and radiographic benefit from an x ray reader-blinded randomised controlled trial of tocilizumab	Annals of the Rheumatic Diseases
SA-RA-KAKEHASI	Tanaka, 2019	Sarilumab plus methotrexate in patients with active rheumatoid arthritis and inadequate response to methotrexate: Results of a randomized, placebo-controlled phase iii trial in japan	Arthritis Research and Therapy
SATORI	Nishimoto, 2009	Study of active controlled tocilizumab monotherapy for rheumatoid arthritis patients with an inadequate response to methotrexate (satori): Significant reduction in disease activity and serum vascular endothelial growth factor by il-6 receptor inhibition therapy	Modern rheumatology

Continued on next page

Trial ID	Author and Year	Title	Journal
SELECT-BEYOND	Genovese, 2018	Safety and efficacy of upadacitinib in patients with active rheumatoid arthritis refractory to biologic disease-modifying anti-rheumatic drugs (select-beyond): A double-blind, randomised controlled phase 3 trial	Lancet
SELECT-NEXT	Burmester, 2018	Safety and efficacy of upadacitinib in patients with rheumatoid arthritis and inadequate response to conventional synthetic disease-modifying anti-rheumatic drugs (select-next): A randomised, double-blind, placebo-controlled phase 3 trial	Lancet
SERENE	Emery, 2010	Efficacy and safety of different doses and retreatment of rituximab: A randomised, placebo-controlled trial in patients who are biological naive with active rheumatoid arthritis and an inadequate response to methotrexate (study evaluating rituximab's efficacy in mtx inadequate responders (serene))	Annals of the Rheumatic Diseases
	Khan, 2011	Rituximab after methotrexate failure in rheumatoid arthritis: Evaluation of the serene trial	Expert Opinion on Biological Therapy
STAR	Furst, 2003	Adalimumab, a fully human anti tumor necrosis factor-alpha monoclonal antibody, and concomitant standard antirheumatic therapy for the treatment of rheumatoid arthritis: Results of star (safety trial of adalimumab in rheumatoid arthritis)	Journal of rheumatology
START	Westhovens, 2006	The safety of infliximab, combined with background treatments, among patients with rheumatoid arthritis and various comorbidities: A large, randomized, placebo-controlled trial	Arthritis and Rheumatism
Strand 2006	Strand, 2006	Sustained benefit in rheumatoid arthritis following one course of rituximab: Improvements in physical function over 2 years	Rheumatology
SUMMACTA	Burmester, 2014	A randomised, double-blind, parallel-group study of the safety and efficacy of subcutaneous tocilizumab versus intravenous tocilizumab in combination with traditional disease-modifying antirheumatic drugs in patients with moderate to severe rheumatoid arthritis (summacta study)	Annals of the Rheumatic Diseases
SURPRISE	Kaneko, 2016	Comparison of adding tocilizumab to methotrexate with switching to tocilizumab in patients with rheumatoid arthritis with inadequate response to methotrexate: 52-week results from a prospective, randomised, controlled study (surprise study)	Annals of the Rheumatic Diseases
SWITCH	Brown, 2018	Alternative tumour necrosis factor inhibitors (tnfi) or abatacept or rituximab following failure of initial tnfi in rheumatoid arthritis: The switch rct	NIHR Health Technology Assessment
Takeuchi 2013	Takeuchi, 2013	Phase ii dose-response study of abatacept in japanese patients with active rheumatoid arthritis with an inadequate response to methotrexate	Modern rheumatology
TAME	Greenwald, 2011	Evaluation of the safety of rituximab in combination with a tumor necrosis factor inhibitor and methotrexate in patients with active rheumatoid arthritis: Results from a randomized controlled trial	Arthritis and Rheumatism
Tanaka 2012	Tanaka, 2012	A study on the selection of dmards for the combination therapy with adalimumab	The Kobe Journal of Medical Sciences
TARGET	Fleischmann, 2017	Sarilumab and nonbiologic disease-modifying antirheumatic drugs in patients with active rheumatoid arthritis and inadequate response or intolerance to tumor necrosis factor inhibitors	Arthritis and Rheumatology
TEMPO	Klareskog, 2004	Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: Double-blind randomised controlled trial	Lancet
	van der Heijde, 2005	Comparison of different definitions to classify remission and sustained remission: 1 year tempo results	Annals of the Rheumatic Diseases

Continued on next page

Trial ID	Author and Year	Title	Journal
	van der Heijde, 2006	Patient reported outcomes in a trial of combination therapy with etanercept and methotrexate for rheumatoid arthritis: The tempo trial	Annals of the Rheumatic Diseases
	van der Heijde, 2006	Comparison of etanercept and methotrexate, alone and combined, in the treatment of rheumatoid arthritis: Two-year clinical and radiographic results from the tempo study, a double-blind, randomized trial	Arthritis and Rheumatism
TOWARD	Genovese, 2008	Interleukin-6 receptor inhibition with tocilizumab reduces disease activity in rheumatoid arthritis with inadequate response to disease-modifying antirheumatic drugs: The tocilizumab in combination with traditional disease-modifying antirheumatic drug therapy study	Arthritis and Rheumatism
van de Putte 2004	van de Putte, 2004	Efficacy and safety of adalimumab as monotherapy in patients with rheumatoid arthritis for whom previous disease modifying antirheumatic drug treatment has failed	Annals of the Rheumatic Diseases
VOLTAIRE-RA	Cohen, 2018	Similar efficacy, safety and immunogenicity of adalimumab biosimilar bi 695501 and humira reference product in patients with moderately to severely active rheumatoid arthritis: Results from the phase iii randomised voltaire-ra equivalence study	Annals of the Rheumatic Diseases
Weinblatt 1999	Weinblatt, 1999	A trial of etanercept, a recombinant tumor necrosis factor receptor: Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate	New England Journal of Medicine
Weinblatt 2018	Weinblatt, 2018	Phase iii randomized study of sb5, an adalimumab biosimilar, versus reference adalimumab in patients with moderate-to-severe rheumatoid arthritis	Arthritis and Rheumatology

Table A10: Inclusion and exclusion criteria of the individual studies

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
ACQUIRE	–	CRP levels of >=0.8 mg/dL	>=10 SJC and >=12 TJCj	MTX for >= 3 month (.=15 mg/week)	inadequate response to 3 months of MTX therapy (15 mg/week)	prior exposure to rituximab
ACT-RAY	–	–	–	MTX dose >=12 weeks, with a stable dose of at least 15 mg/week for 6 weeks or longer before starting study treatment.	inadequate response to MTX	–
ACT-STAR	>=6 months	No requirement	>=4 SJC and >=4 TJCg	history of bDMARDs or cDMARD use	inadequate response to bDMARDs or cDMARDs	–
ADACTA	>=6 months	No requirement	–	MTX (current or past use); MTX intolerant patients were permitted	inadequate response to MTX, be unable to tolerate MTX, or be inappropriate candidates for continued MTX treatment in the judgment of the investigator	prior exposure to bDMARDs
AIM	>=1 year	CRP of >=10.0mg/l	>=10 SJC and >=12 TJCj	MTX (>=15 mg/wk) for >=3 months (28 day stable dose prior to entry)	inadequate response to MTX (>=15mg/week)	–
AMPLE	<=5 years	–	–	prior history of MTX use	inadequate response to MTX	bDMARDS
ARMADA	–	–	>=9 TJC and >=6 SJCj	MTX >=6 months or longer (28 day stable dose prior to entry)	inadequate response to MTX	anti-CD4 therapy or TNF antagonists
ASCERTAIN	>=3 months	hs-CRP of >=4mg/l	>=4 TJC and >=4 SJCg	>=1 TNF; continuous tx with >=1 cDMARD (except for simultaneously use of LEF and MTX) for >=12 consecutive weeks prior to screening and on a stable dose for >=6 weeks	>=1 TNF or patients intolerant of >=1 TNF	–
ATTAIN	>=1 year	CRP of >=1mg/dL	>=10 SJC and >=12 TJCj	anti-TNF alpha therapy (etanercept, infliximab, or both); oral DMARD for >=3 months (stable dose 28 days prior to study entry)	inadequate response to antiâTNF therapy with etanercept, infliximab, or both at the approved dose after >=3 months	–
ATTEST	>=1 year	CRP of >=1 mg/dL	>=10 SJC and >=12 TJCj	MTX (>=15 mg/week) for >=3 months prior to randomisation (stable for at least 28 days)	inadequate response to MTX	any prior abatacept or anti-TNF therapy

Continued on next page

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
ATTRACT	–	ESR of >28mm/h and CRP of >2mg/dL	>=6 SJC and >=6 TJCj	MTX for >=3 months with no break in treatment of more than 2 weeks during this period (stable dose >=12.5mg/week 4 weeks prior to screening)	inadequate response to MTX	cDMARDs other than MTX
Bao 2011	–	CRP of >= 15mg/l or ESR of >=28mm/ha	>=6 SJC or >=6 TJCj	MTX (7.5-25mg per week) for >=12 weeks	active disease despite MTX	–
BREVACTA	>=6 months	CRP of >=10 mg/l and/or ESR of >=28 mm/h	>=6 SJC and >=8 tenderj	>=1 DMARD (stable dose >=8 weeks prior to baseline); up to 20% of population could have additional taken >=1 anti-TNF agents	inadequate response to >=1 DMARDs (in up to 20% of patients, could include inadequete response to >=1 anti-TNF agent)	–
CHANGE	–	CRP of >=2mg/dL	>=10 SJC and >=12 TJCj (excluding distal interphalangeal joints)	>=1 DMARD	inadequate response to >=1 DMARD	any TNF antagonist or an alkylating agent
Choy 2012	>=6 months	ESR >=28 mm/h (or CRP >10 mg/la	>=9 SJC and >=9 TJCj	MTX >=6 months (stable dose 10-25mg/week for >=8 weeks prior to treatment); 10-15 mg/week was deemed acceptable in cases where a dosage reduction had been necessary because of toxicity	partial response to MTX	prior treatment with any TNF-a inhibitor
Cohen 2002	>6 months but <12 years	CRP >1.5mg/db	>=6 SJCbj	MTX (15-25mg/week) for >=6 consecutive months	inadequate response to MTX	–
Cohen 2004	>=6 months	CRP of >=15 mg/l or ESR >= 28 mm/h	>=6 SJC and >=9 TJCj	MTX (stable dose 10â25 mg/week) for >=24 weeks	inadequate response to MTX	prior treatment with an IL1Ra.
Cohen 2017	>=3 months	ESR of >=28 mm/hour or CRP of >=1.0mg/dL	>=6 SJC and >=6 TJCj	MTX for >=12 consecutive weeks (stable oral dose of 7.5â25 mg/week for â¥8 weeks before to tx)	inadequate response to MTX	>=2 or more biologic therapies for RA; Previous receipt of HUMIRAÂ® (adalimumab) or a biosimilar of adalimumab
Cohen 2018	>=4 months	hs-CRP of >= 10 mg/L	>=6 SJC and >=6 TJCj	MTX >=12 weeks (10-25mg/week) and oral folic/folinic acid (>=5mg/week) for >=21 days prior (patients intolerant to 10â25 mg/wk could enroll with an MTX dose as low as 7.5 mg/wk)	inadequate response to MTX	infliximab or lymphocyte-depleting therapies (e.g., rituximab, alemtuzumab)

Continued on next page

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
DANCER	>=6 months	ESR >=28 mm/h and CRP >=1.5mg/dL	–	MTX >= 12 weeks (stable dose 10-25mg/week prior to randomization)	1-5 DMARDs, manifesting as a lack or loss of response to treatment	–
De Filippis 2006	>2 years	ESR >22mg/h, CRP >1.9 mg/dc	>5 SJC and >10 TJCcj	DMARDs for >6 months, included a stable dose of MTX in the 3 months prior to study entry	inadequate response to cDMARDs	–
DE019	–	CRP >1mg/dL	>=9 TJC and >=6 SJCj	MTX >=3 months (stable dose of 12.5â25 mg/week [or 10 mg/week in patients intolerant to MTX] for >=4 weeks)	inadequate response or intolerance to MTX	prior use of anti-CD4 antibody therapy or TNF antagonists
Edwards 2004	–	CRP of >= 15 mg/L or ESR of >=28 mm per houra	>=8 SJC and >=8 TJCj	MTX at >=10 mg/week	inadequate response to MTX	–
Elmedany 2019	–	–	–	at least 1 TNF	failed to achieve remission on at least 1 TNF	–
Emery 2017	>=6 months and <15 years	ESR >=28mm/h or CRP >=1.0 mg/dL despite MTX for 6 months	>=6 SJC and >=6 TJCj	MTX >=6 months (stable dose of 10-25mg/week for >=4 weeks prior to screening)	inadequate response to MTX	previous bDMARD use
EQUIRA	>=6 months	CRP >5 mg/L or ESR â¥28 mm/h	–	MTX 10â25 mg/week	inadequate response to MTX 10â25 mg/week following dose escalation according to local standards	any previous exposure to ETN; previous use of >2 biologics (allowed only if the therapy was efficient and not failing and was withdrawn because of other reasons that were not due to efficacy failure or safety issues)
ETN Study 309	<20 years	ESR >=28mm or CRP >=20mg/L	>=6 SJC and >=10 PJCj	ssz (2-3 g daily) for >=4 months before screening without signs of toxicity	inadequate response to cDMARD	etanercept or other TNF antagonists
FAST4WARD	>=6 months	ESR >28 mm/h or CRP of >=.10 mg/La	>=9 TJC and >=9 SJCj	>=1 DMARD	inadequate response of intolerance to >=1 DMARD	prior treatment with TNFa inhibitors
Fleischmann 2012	>=6 months	ESR ULN or CRP >= 7 mg/L	>=6 TJC and >=6 SJCg	>=1 DMARD	failure on >=1 DMARD due to lack of efficacy or toxicity	prior history of TNF failure
Fleischmann 2018	>=4 months	CRP >=8mg/L	>=6 SJC and >=6 TJCj	MTX for >=12 weeks with stable dose for >=4 weeks	inadequate response to MTX	no more than 2 biologic agents
GO-AFTER	>=3 months	–	>=4 SJC and >=4 TJCj	>1 dose of etanercept, adalimumab, or infliximab	inadequate response to TNF and MTX	–

Continued on next page

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
GO-FORTH	>=3 months	CRP >1.5 mg/dL or ESR of >28 mm/h	>=4 SJC and >=4 SJC g	MTX for >=3 months (stable dose of 15 mg-25 mg/week during 4 weeks prior to screening)	inadequate response to MTX	–
GO-FORWARD	>=3 months	CRP of >= 1.5 mg/dL or ESR >= 28 mm/h	>=4 SJC and >=4 TJCg	MTX for >=3 months (stable dose of 15 mg-25 mg/week during 4 weeks prior to screening)	inadequate response to MTX	TNF inhibitors or rituximab
GO-FURTHER	>=3 months	CRP >=1.0 mg/dL	>=6 SJC and >=6 TJCg	MTX for >=3 months (stable dose of 15 mg-25 mg/week during 4 weeks prior to screening)	inadequate response to MTX	–
GO-LIVE	–	CRP of >= 1.5 mg/dL or ESR of >= 28 mm/h	>= 4 SJC and >= 4 TJCj	tolerated MTX (15mg/week) >=3 months (stable dose 15-25mg for 4 weeks prior to screening)	inadequate response to MTX	any prior receipt of rituximab, abatacept, or natalizumab
GO-SAVE	–	–	>=6 SJC and >=6 TJCg	MTX h at a stable dose (7.5â25 mg/week) for 4 weeks and maintained unless MTX toxicity occurred	inadequate response to etanercept+MTX or adalimumab + MTX	biologics for RA other than adalimumab and etanercept; concomintant DMARDs other than MTX, ssz, or hcq
HERA	–	ESR of >=28 mm/h or CRP of >=1.0mg/dL	>=6 SJC and >=6 TJCj	MTX >=6 months prior to screening	inadequate response to MTX >=6 months prior to screening	–
HIKARI	>=6 months	ESR of >=28 mm/hour or CRP of >=2.0 mg/dL	>=6 SJC and >=6 TJCj	>=1 prior DMARD (including MTX)	inadequate response of intolerance to >=1 DMARD	2 or more TNF inhibitors and/or who had failed more than 1 TNF alpha inhibitor
Iwahashi 2014	–	CRP of >= 0.8 mg/dL	>=10 SJC and >= 12 TJCj	MTX >=3 months (stable dose 6-8mg/week prior to randomization)	inadequate response to MTX	any bDMARD; abatacept; exposure to any biologic not currently approved in japan
Jamshidi 2017	>=6 months	CRP of >20 mg/L	–	>=1 cDMARD for >=12 months	inadequate response to >=1 cDMARD for >=12 months	bDMARDS including any TNF inhibitor
JESMR	–	CRP of > 2 mg/dL or ESR of >= 28 mm/h	>=6 SJC and >=6 TJCj	MTX (>=6mg/week) for >=3 months (stable dose at least 4 weeks prior to study entry)	inadequate response to MTX	bDMARDS
J-RAPID	6 months-15 years	ESR of >=30 mm/hour or CRP of >=1.5 mg/dL	>=9 TJC and >=9 SJCj	MTX (6-8mg/week) >=2 months	inadequate response to MTX	2 or more TNF inhibitors and/or who had failed more than 1 TNF alpha inhibitor
Kim 2007	–	–	>=6 SJC and >=9 TJCj	MTX (10-30 mg weekly) for >=6 months; previous reception of >=1 DMARD other than MTX	inadequate response to 2-4 DMARDS	–
Kremer 2003	–	CRP of >=1 mg/dL	>=10 SJC and >= 12 TJCj	MTX (10-30 mg weekly) for >= 6 months (stable dose 28 days prior to enrollment)	inadequate response to MTX	–

Continued on next page

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
Kremer 2005	–	CRP >1mg/dL	>=10 SJC and >=12 TJCj	MTX (10-30mg/week) for at least 6 months, stable dose for 28 days prior to enrollment	inadequate response to MTX	–
Kremer 2012	>=6 months	ESR ULN or CRP >=7 mg/L	>=6 TJC and >=6 PJCg	MTX continuously for 4 months	inadequate response to MTX	–
LARA	>=3 months	ESR of >= 28 mm/h	>=6 SJC and >=8 TJCj	previous history of MTX use	inadequate response to MTX	Previous treatment with ETN or other bDMARDS
Li 2016	>=6 months	CRP >= 15 mg/L or ESR >= 28 mm/h	>= 4 SJC and >=4 TJCg	MTX (stable dose 7.5â20 mg/week) â¥ 4 weeks before study agent initiation	inadequate response to MTX	bDMARD
LITHE	>=6 months	CRP >=1mg/dL	>=10 SJC, >=12 TJCj	MTX (10-30 mg weekly) for >=6 months (stable dose 28 days prior to enrollment)	inadequate response to a stable dose of MTX	prior treatment failure with a TNF agent
Matsubara 2018	<5 years	CRP >=2.0mg/dL or ESR >=28 mm/h	>=6 SJC and >=6 TJCj	MTX (>=6mg/week) for >=3 months	inadequate response to MTX	prior exposure to bDMARDs
Matsuno 2018a	–	ESR of >=28mm/h	>=6 SJC and >=6 TJCj	MTX (<=16 mg/week with less than 2-week drug withdrawal) >=12 weeks prior to screening (stable dose of >= 6mg/week during 4 weeks prior to the screening)	inadequate response to MTX	–
MOBILITY	–	hs-CRP of > 6 mg/L	>=8TJC and >=6 SJCg	MTX >= 12 weeks (stable dose for at least 6 weeks prior to screening visit)	inadequate response to MTX	history of nonresponse to bDMARDS
MONARCH	–	CRP of >=8 mg/L or ESR of >28mm/h	>=6 SJC or >=8 TJC g	MTX dose (10â25 mg/week or 6â25 mg/week for patients within Asia-Pacific region) for â¥12 weeks OR intolerant of or considered inappropriate candidates for continued treatment with MTX	inadequate response, intolerance, or inappropriate candidacy for continued MTX treatment	prior exposure to bDMARDs, including IL6 receptor agonists or JAK inhibitors
Moreland 1999	–	ESR>= 28 mm/h or CRP >= 20 mg/La	–	history of use of 1-4 DMARDs	inadequate response to 1-4 DMARDs	–
MUSASHI	>=6 months	>=30mm/h and CRP >=1.0 mg/dL	>=8 TJC and >=6 SJCj	history of cDMARD use	inadequate response to any synthetic DMARD	–
Niu 2011	–	ESR of >=28 mm/h, or a CRP of >=2.0 mg/dL	>=4 SJC and >=6 TJCj	MTX (stable dose 7.5â15 mg per week)	inadequate response to MTX	–

Continued on next page

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
OPTION	>=6 months	CRP of >=18 mg/K or ESR of >=28 mm/h	>=6 SJC and >=8 TJCj	MTX for >=12 weeks prior to study start (stable dose 10-25 mg/week for 8+ weeks)	inadequate response to MTX	–
ORAL-SCAN	–	ESR of >28 mm/hour or CRP of >7mg/L	>=6 TJC/PJC and >=6 SJCg	MTX (15â25 mg weekly) for 6 weeks (stable doses 15 mg were allowed only if there were safety issues at higher doses). Prior use of biologic or nonbiologic DMARDs was permitted	inadequate response to MTX	–
ORAL-STANDARD	–	ESR of >=28 mm/h or CRP a >7mg/L	>=6 TJC/PCJ or >=6 SJCg	MTX (7.5-25 mg weekly)	inadequate response to MTX	bDMARDs; adalimumab; lack of response to prior anti-TNF biologic
ORAL-STEP	–	ESR of > 28 mm/h or CRP of > 66Â·67 nmol/L (7 mg/L)	>=6 TJC/PJC and >=6 SJCg	MTX (stable dose 7.5-25mg/week) for >=6 months (continuous for >=4 months)	inadequate response or intolerance to one or more approved TNFi	–
ORAL-STRATEGY	–	CRP >= 3mg/L	>=4 TJC/PJC and >=4 SJCi	MTX at a stable dose of >=15-25 mg; patients who had responded inadequately or had an adverse event secondary to treatment with a biological DMARD could be included but had to have discontinued the biological DMARD for a minimum period of time before randomisation	inadequate response to MTX	previous treatment with adalimumab or tofacitinib
ORAL-SYNC	–	ESR >=28mm/h or CRP >66.7nmol/L	>=4 TJC/PJCand >=4 SJCg	>=1 cDMARD or bDMARD; Patients receiving background MTX (25 mg/wk) required at least 4 months of therapy therapy with stable dosing 6 weeks before receiving the study drug.	inadequate response to >=1 cDMARD or bDMARD (stably dosed)	–
RA-BEACON	–	CRP >=3 mg/L	>=6 TJC and >=6 SJCg	>=1 TNF inhibitors; patients who had received other biologics DMARDs could also participate (bDMARDs must have been discontinued at least 4 weeks prior to randomization (>=6 months for rituximab)	inadequate response >=1 TNF inhibitor	–
RA-BEAM	–	CRP >=6 mg per litter	>=6 SJC and >=6 TJCj	MTX >=12 weeks (stable dose 15-25 mg/week >=8 weeks prior to entry)	inadequate response to MTX	bDMARDs

Continued on next page

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
RA-BUILD-A; RA-BUILD-Bh	–	hsCRP of >=10 mg/L or ESR of >=28 mm/h	>=6 TJC and >=6 SJCg	>=1 cDMARD for >=12 weeks prior to study entry (stable dose 8 weeks prior to study entry) or intolerance to >=1 cDMARD (For participants not receiving a cDMARD at the time of entry, the investigator will document in the participant's history that the participant had failed, was unable to tolerate, or had a contraindication to treatment with a cDMARD)	inadequate response or intolerance to >=1 cDMARD	bDMARDs at any time
RACAT	–	–	–	MTX (stable dose 15-25mg/week) for >=12 weeks	inadequate response to MTX	–
RADIATE	–	ESR of >28 mm/h or CRP of >1.0 mg/dL	>=6 SJC and >=8 TJCj	MTX >=12 weeks (stable dose >=8 weeks); prior TNF use	inadequate response to current anti-rheumatic therapies, including MTX; inadequate response or intolerance to treatment with 1 or more anti-TNF therapies within 1 year of entering study;	–
RAPID-1	>=6 months and <15 years	ESR of >=30 mm/h and CRP of >15mg/L	>=9 TJC and >=9 SJCj	MTX for >=6 months (stable dose >=10mg/week for >=2 months prior)	inadequate response to MTX	–
RAPID-2	>=6 months	–	–	MTX for >=6 months (stable dose >=10mg/week for >=2 week month before baseline)	inadequately response to MTX	–
RAPID-C	>=6 months	ESR of >=30 mm/hour and CRP >15 mg/L	>=6 TJC and >=6 SJCj	MTX for at least 3 months prior to the baseline visit, with a stable dose of â¥10 mg/wk for >=2 weeks prior to baseline	inadequate response to MTX	TNF failure
RA-SCORE	>=3 months and =<10 years	–	–	MTX (12.5-25m/week) for >=12 weeks (stable dose 4 weeks prior); 7.5 mg/week or 10 mg/week were permitted only in cases of documented intolerance to higher doses.	inadequate response to MTX	bDMARDS or with a B cell modulating or cell depleting therapy.
RED SEA	–	–	–	>=2 DMARDS	cDMARDs	any TNF inhibitor

Continued on next page

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
REFLEX	>=6 months	CRP of >=1.5 mg/dL or ESR >=28mm/h	>=8 SJ >=8 TJCg	MTX (10-25mg/week) for >=12 weeks prior to screening (last 4 weeks stable dose); prior use or intolerance to >=1 TNF inhibitor (infliximab, adalimumab, or etanercept)	inadequate response to previous or current treatment with the anti-TNF agents infliximab, adalimumab, or etanercept, or were intolerant to at least 1 administration of these agents.	–
ROSE	>=6 months	CRP of >=95.25 nmol/l and ESR of >=28mm/h	>=6 SJC or >=6 TJCh	history of use of >=1 cDMARD	inadequate clinical response >=1cDMARD as determined by the investigator	unsuccessful treatment with an anti-TNF agent; previous treatment with tocilizumab
SAMURAI	>=6 months	ESR of >=30 mm/h and CRP of >=20 mg/k	>=6 TJC and >=6 SJCj	–	inadequate response to at least >=DMARD or immunosuppresent	–
SA-RA-KAKEHASI	>=3 months	CRP >= 0.6 mg/ dl	>=8 TJC and >=6g	MTX >=12 weeks (stable dose 6â16 mg/week â¥ 6 weeks prior to screening)	inadequate response to MTX	prior TNF of bDMARD failure
SATORI	>=6 months	ESRof >=30mm/h and CRP of >=1.0 mg/dL	>=6 TJC and >=6 SJCh	MTX (>=8mg/week) for >=8 weeks	inadequate response to MTX	any DMARD or immunosuppressant other than MTX
SELECT-BEYOND	>=3 months	hsCRP >= 3 mg/L	>=6 SJC and >=6 TJCj	>=1 bDMARD	inadequate response or intolerance to >=1 bDMARD	prior exposure to JAK inhibitor
SELECT-NEXT	>=3 months	hsCRP of >=3 mg/L	>=6 SJC >=6 TJCg	prior cDMARD exposure; the protocol allowed the enrollment of up to 20% of patients with exposure to no more than 1 bDMARD	inadequate response to at least one of the following cDMARDs: MTX ,sulfasalazine, or leflunomide	inadequate response to bDMARD; any previous exposure to a JAK inhibitor
SERENE	>=6 months	CRP of >= 0.6 mg/dL (6 mg/L) or ESR of >=28 mm/h	>=8 SJC and >=8 TJCg	MTX (10-25 mg/week) for >=12 weeks	inadequate response to at least one of the following cDMARDs: MTX ,sulfasalazine, or leflunomide	bDMARDs
STAR	>=3 months	–	>=6 SJC and >=9 TJCj	–	–	previous exposure to anti-CD4 therapy or biologic DMARDs (e.g., TNF antagonists, interleukin-1 receptor antagonists)
START	–	–	>=6 SJC and >=6 TJCj	MTX for >= 3 months prior to randomization (stable dose at least 4 weeks prior)	inadequate response to MTX	–
Strand 2006	–	CRP of 1.5 mg/dL and/or ESR of 30mm/ha	>=8 TJC and >=8 SJCj	MTX (>=10 mg/week) for >=16 weeks	inadequate response to 1-5 DMARDS	–

Continued on next page

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
SUMMACTA	>=6 months	CRP >=10mg/L and/or ESR >=28mm/h	>=4 SJC and >=4 TJCg	permitted DMARDs at a stable dose for >=8 weeks prior to baseline; biologic agents had to be discontinued prior to study entry	inadequate response to current DMARDs	previous treatment with tocilizumab, alkylating agents or cell depleting therapies
SURPRISE	–	–	–	MTX (stable dose >=6 mg/week) for >=8 weeks before enrollment	inadequate response to current DMARDs	prior exposure to biologics
SWITCH	>=24 weeks	–	–	>=2 cDMARDs and 1 TNFI	>=2 cDMARDs including MTX (failure per NICE/BSR guidelines); and persistent RA despite having been treated with a current initial TNFI agent for at least 12 weeks	>1 TNFI or other bDMARD
Takeuchi 2013a	–	CRP >=1.0 mg/df	>=10 SJC or >=12 TJCg	MTX for >=12 weeks (6-8 mg QW)	active disease despite MTX therapy	–
TAME	>=6 months	No requirement for CRP or ESR	>=5 SJC and >=5 TJCj	MTX at least 12 weeks immediately prior to randomization	MTX	–
Tanaka 2012	–	–	–	>=cDMARD or bDMARD	inadequate repsonse to >=1 cDMARD or bDMARD	–
TARGET	>=6 months	CRP of >=8 mg/L	>=6 SJC and >=8 TJCg	>=1 TNF inhibitor for >=3 months; Continuous treatment (â¥12 weeks before randomization) with 1 or a combination of cDMARDs and on stable dose(s) for â¥6 weeks before screening	inadequate response to >=1 anti-TNF inhibitor and/or intolerance to â¥1 anti-TNF inhibitor resulting in or requiring their discontinuation	prior treatment with any cell-depleting agents including, but not limited to, rituximab without a normal lymphocyte and CD 19+ lymphocyte count; prior treatment with antiâIL-6 or IL-6 receptor antagonist therapies, including, but not limited to, tocilizumab or sarilumab
TEMPO	>=6 months	ESR of >= 28 mm/h or greater or CRP of >=20 mg/a	>=10 SJC and >=12 PJCj	>=1 DMARD other than MTX; Individuals previously treated with MTX MTX could be enrolled provided they had not had clinically important toxic effects or lack of response, at the discretion of the investigator, and had not been treated with MTX within 6 months of enrolment	inadequate response to >=1 DMARD	etanercept or other TNF agonists
TOWARD	>=6 months	CRP of >= 1 mg/dL or an ESR of >= 28 mm/h	>=6 SJC and >=8 TJCj	(MTX, chloroquine, hydroxychloroquine, parenteral gold, sulfasalazine, azathioprine, and leflunomide) for 8 weeks prior to study entry.	inadequate response to current anti-rheumatic therapies, including 1 or more traditional DMARDs;	any cell-depleting therpy; TNF inhibitor failure

Continued on next page

Trial	Disease duration	Acute-phase reactant	Swollen and tender joint count	Prior treatment requirement	Prior treatment failure requirement	Exclusion criteria prior treatment history
van de Putte 2004	–	-	>=12 TJC and >=10 SJCg	>=1 prior DMARDs; Patients taking traditional DMARDs at the time of recruitment were required to undergo a 4 week washout period before the initial injection of the study drug.	inadequate response to >=1 previous DMARD	–
VOLTAIRE-RA	>=6 months	ESR of >28 mm/hour or a CRP of >1.0 mg/dL	>=6 SJC and >=6 TJCg	MTX (15-25 mg/week) for >=12 weeks prior to day 1, stable dose for >=4 weeks prior (10mg per week permitted if intolerance to higher dose)	inadequate response to MTX	>=2 bDMARDS; adalimumab or adalimumab biosimilar
Weinblatt 1999	–	–	>=6 SJC and >=6 TJCg	stable dose of 15 to 25 mg per week for the last four weeks (weekly doses as low as 10 mg were acceptable for patients who could not tolerate higher doses)	inadequate response to MTX	–
Weinblatt 2018	6 months-15 years	ESR of >=28 mm/hour or CRP of >=1.0 mg/dL	>=6 SJC and >=6 TJCg	MTX for â¥6 months (stable dosage of MTX (10â 25 mg/week) for â¥4 weeks)	inadequate response to MTX	previous exposure to bDMARDS

Notes: a. study required either of the two acute phase reactant criteria specified above, or alternatively morning stiffness lasting 45 minutes or longer in lieu of fulfillment of CRP/ESR;
b. study required at least 2 of the following: >=9 tender joints or painful joints, morning stiffness >=45 min, or CRP >1.5 mg/dL;
c. study required at least 3 of the 4 following features: ESR >22mg/h, CRP >1.9mg/dL, morning stiffness >45 min, >5 swollen joints and >10 tender joints;
d. patients were required to meet at least two of the following criteria at baseline: 1) CRP >1.5 mg/dL or ESR of 28mm/h, 2) morning stiffness lasting >=30 minutes, radiographic evidence of bone erosion, or 4) anti-cyclic circullinated peptide antibody;
e. in addition to meeting either the CRP or ESR requirements, patients were required to have the presence of IgG anti-cyclic citrullinated peptide antibodies or rheumatoid factor (RF);
f. study entry required one or more of the following: >=10 swollen joints (66-joint count), >=12 tender joints (68-joint count), or CRP >=1.0mg/dL;
g. out of 66 swollen joints and 68 tender joints evaluated;
h. out of 46 swollen joints and 49 tender joints evaluated;
i. out of 28 swollen joints and 28 tender joints evaluated;
j. number of joints evaluated not specified; h. The study design of RA-BUILD permitted but did not require concomitant cDMARD background therapy (which was not based on random assignment, but at the discretion of the investigator). Subgroup data stratified by background cDMARD were therefore used within the analysis, and the corresponding results were treated as two separate trials (RA-BUILD-A and RA-BUILD-B).

I.3.3 Subset of studies that provide evidence for estimation of treatment effects among tDMARD naive population

Table A11: Criteria for selection of subset of studies that provide evidence for estimation of treatment effects among the tDMARD naive population

Criteria for selection	Criteria for exclusion	Comments
Trials that permitted up to 20% tDMARD experienced patients in their population as determined by either demographic information, study inclusion criteria, or both.	Trials with only one arm meeting the specified cutoff for prior tDMARD exposure were ineligible. For example, if 19% of Arm X's patients were previously exposed to tDMARDs, and 23% of Arm Y's patients were previously exposed, this trial was deemed ineligible for the network meta-analysis.	Trials that prohibited specific tDMARD treatments, specific tDMARD drug classes, or both were considered to be tDMARD naÃ¯ve (e.g. Participants were excluded if they received prior TNF inhibitor treatment), unless it was explicitly stated that at least some participants had been previously exposed to other tDMARD agents. This assumption does not include trials that only excluded prior treatment with one or more of the drugs being investigated in the trial. Trials were included in the network if the publication specified up to 20% exposure to any tDMARD agent, a specific tDMARD drug or drug class (such as TNF inhibitors), or both within the study protocol or demographics (e.g. up to 20% of participants could have received prior TNFi treatment or 15% of arm A and 17% of arm B received prior TNFi treatment).

I.3.3.1 Study characteristics

Table A12: Study characteristics, tDMARD naive population

Trial	Region	Multicenter	Masking	Treatment	Availability of ACR 20/50/70 at 6 months f-up	Availability of DAS28 at 6 months f-up	Availability of HAQ-DI at 6 months f-up
ACQUIRE	multinational	Yes	double-blind	ABT (125mg SC) + cDMARD ABT (10mg/kg IV) + cDMARD	Y	Y	Y
ACT-RAY	multinational	Yes	double-blind	TOC (8mg/kg IV) + cDMARD TOC (8mg/kg IV)	Y	Y	Y
ADACTA	multinational	Yes	double-blind	TOC (8mg/kg IV) ADA (40mg SC)	Y	Y	Y
AIM	multinational	Yes	double-blind	ABT (10mg/kg IV) + cDMARD cDMARD	Y	Y	Y
AMPLE	multinational	Yes	single-blind	ABT (125mg SC) + cDMARD ADA (40mg SC) + cDMARD	Y	Y	
ARMADA	USA, Canada	Yes	double-blind	ADA (20mg SC) + cDMARD ADA (80mg SC) + cDMARD ADA (40mg SC) + cDMARD cDMARD	Y		Y
ATTEST	multinational	Yes	double-blind	ABT (10mg/kg IV) + cDMARD IFX (3mg/kg IV Q8WEEK) + cDMARD cDMARD	Y	Y	Y
ATTRACT	multinational	Yes	double-blind	IFX (3mg/kg IV Q4WEEK) + cDMARD IFX (10mg/kg IV Q8WEEK) + cDMARD IFX (10mg/kg IV Q4WEEK) + cDMARD cDMARD IFX (3mg/kg IV Q8WEEK) + cDMARD			
Bao 2011	China	No	double-blind	ANA (80mg SC) + cDMARD cDMARD	Y	Y	
CHANGE	Japan	Yes	double-blind	ADA (20mg SC Q2WEEK) ADA (40mg SC) ADA (80mg SC) Placebo	Y		Y
Choy 2012	multinational	Yes	double-blind	CTZ (400mg SC) + cDMARD cDMARD	Y		Y
Cohen 2002	multinational	Yes	double-blind	cDMARD ANA (0.04mg/kg SC) + cDMARD ANA (0.1mg/kg SC) + cDMARD ANA (0.4mg/kg SC) + cDMARD ANA (1mg/kg SC) + cDMARD ANA (2mg/kg SC) + cDMARD	Y		Y
Cohen 2004	multinational	Yes	double-blind	ANA (100mg SC) + cDMARD cDMARD	Y		Y
Cohen 2018	multinational	Yes	double-blind	IFX-Pfizer (3mg/kg IV) + cDMARD IFX (3mg/kg IV Q8WEEK) + cDMARD			Y

Continued on next page

Trial	Region	Multicenter	Masking	Treatment	Availability of ACR 20/50/70 at 6 months f-up	Availability of DAS28 at 6 months f-up	Availability of HAQ-DI at 6 months f-up
De Filippis 2006	Italy	No	NR	ETN (50mg SC) + cDMARD IFX (3mg/kg IV Q8WEEK) + cDMARD			
DE019	USA, Canada	Yes	double-blind	ADA (20mg SC) + cDMARD ADA (40mg SC) + cDMARD cDMARD	Y		Y
Edwards 2004	multinational	Yes	double-blind	cDMARD RTX (1000mg IV) + cDMARD RTX (1000mg IV)	Y	Y	
Emery 2017	multinational	Yes	double-blind	ETN (50mg SC) + cDMARD ETN-SB4 (50mg SC) + cDMARD	Y	Y	
EQUIRA	multinational	Yes	double-blind	ETN-GP2015 (50mg SC) + cDMARD ETN (50mg SC) + cDMARD	Y	Y	Y
ETN Study 309	multinational	Yes	double-blind	ETN (50mg SC) cDMARD ETN (50mg SC) + cDMARD ETN (50mg SC) + cDMARD	Y	Y	
FAST4WARD	multinational	Yes	double-blind	CTZ (400mg SC) Placebo	Y		Y
Fleischmann 2012	multinational	Yes	double-blind	Placebo TOF (1mg PO) TOF (3mg PO) TOF (5mg PO) TOF (10mg PO) TOF (15mg PO) ADA (40mg SC)	Y	Y	Y
Fleischmann 2018	multinational	Yes	double-blind	ADA-Pfizer (40mg SC) + cDMARD ADA (40mg SC) + cDMARD	Y	Y	Y
GO-FORTH	Japan	Yes	double-blind	GOL (100mg SC) + cDMARD GOL (50mg SC) + cDMARD cDMARD	Y	Y	Y
GO-FORWARD	multinational	Yes	double-blind	GOL (100mg SC) GOL (100mg SC) + cDMARD cDMARD GOL (50mg SC) + cDMARD	Y		Y
GO-FURTHER	multinational	Yes	double-blind	cDMARD GOL (2mg/kg IV) + cDMARD	Y	Y	Y
GO-LIVE	multinational	Yes	double-blind	GOL (4mg/kg IV) GOL (4mg/kg IV) + cDMARD cDMARD GOL (2mg/kg IV) GOL (2mg/kg IV) + cDMARD	Y		
HIKARI	Japan	–	double-blind	cDMARD	Y	Y	Y

Continued on next page

Trial	Region	Multicenter	Masking	Treatment	Availability of ACR 20/50/70 at 6 months f-up	Availability of DAS28 at 6 months f-up	Availability of HAQ-DI at 6 months f-up
				CTZ (200mg SC) + cDMARD			
Iwahashi 2014	Japan	Yes	double-blind	ABT (125mg SC) + cDMARD	Y	Y	Y
				ABT (10mg/kg IV) + cDMARD			
Jamshidi 2017	Iran	Yes	double-blind	ADA-Cinnora (40mg SC) + cDMARD	Y	Y	
				ADA (40mg SC) + cDMARD			
JESMR	Japan	No	open-label	ETN (50mg SC)	Y		
				ETN (50mg SC) + cDMARD			
J-RAPID	Japan	–	double-blind	cDMARD	Y	Y	Y
				CTZ (100mg SC) + cDMARD			
				CTZ (200mg SC) + cDMARD			
				CTZ (400mg SC) + cDMARD			
Kim 2007	Korea	–	double-blind	cDMARD	Y		Y
				ADA (40mg SC)			
Kremer 2003	multinational	Yes	double-blind	ABT (2mg/kg IV) + cDMARD	Y		
				cDMARD			
				ABT (10mg/kg IV) + cDMARD			
Kremer 2012	multinational	Yes	double-blind	cDMARD	Y		
				TOF (1mg PO) + cDMARD			
				TOF (3mg PO) + cDMARD			
				TOF (5mg PO) + cDMARD			
				TOF (10mg PO) + cDMARD			
				TOF (15mg PO) + cDMARD			
				TOF (20mg PO) + cDMARD			
LARA	multinational	Yes	open-label	ETN (50mg SC) + cDMARD	Y	Y	Y
				cDMARD			
Li 2016	China	Yes	double-blind	cDMARD	Y		Y
				GOL (50mg SC) + cDMARD			
LITHE	multinational	Yes	double-blind	TOC (8mg/kg IV) + cDMARD	Y	Y	Y
				TOC (4mg/kg IV) + cDMARD			
				cDMARD			
Matsubara 2018	Japan	Yes	double-blind	ABT (10mg/kg IV) + cDMARD	Y	Y	
				cDMARD			
MOBILITY	multinational	Yes	double-blind	cDMARD	Y		Y
				SAR (200mg SC) + cDMARD			
				SAR (150mg SC) + cDMARD			
MONARCH	multinational	Yes	double-blind	ADA (40mg SC)	Y	Y	Y
				SAR (200mg SC)			
Moreland 1999	North America	Yes	double-blind	ETN (10mg SC)	Y		Y
				ETN (50mg SC)			
				Placebo			
Niu 2011	multinational	Yes	double-blind	cDMARD	Y		
				ANA (80mg SC) + cDMARD			

Continued on next page

Trial	Region	Multicenter	Masking	Treatment	Availability of ACR 20/50/70 at 6 months f-up	Availability of DAS28 at 6 months f-up	Availability of HAQ-DI at 6 months f-up
OPTION	multinational	Yes	double-blind	TOC (4mg/kg IV) + cDMARD TOC (8mg/kg IV) + cDMARD cDMARD	Y	Y	Y
ORAL-SCAN	multinational	Yes	double-blind	TOF (10mg PO) + cDMARD TOF (5mg PO) + cDMARD cDMARD			Y
ORAL-STANDARD	multinational	Yes	double-blind	TOF (10mg PO) + cDMARD TOF (5mg PO) + cDMARD ADA (40mg SC) + cDMARD cDMARD	Y		Y
ORAL-STRATEGY	multinational	Yes	double-blind	TOF (5mg PO) TOF (5mg PO) + cDMARD ADA (40mg SC) + cDMARD	Y	Y	Y
ORAL-SYNC	multinational	Yes	double-blind	TOF (5mg PO) + cDMARD cDMARD TOF (10mg PO) + cDMARD			
RA-BEAM	multinational	Yes	double-blind	BCT (4mg PO) + cDMARD cDMARD cDMARD ADA (40mg SC) + cDMARD	Y	Y	Y
RA-BUILD-Aa	multinational	Yes	double-blind	cDMARD BCT (2mg PO) + cDMARD BCT (4mg PO) + cDMARD			
RA-BUILD-Ba	multinational	Yes	double-blind	Placebo BCT (2mg PO) BCT (4mg PO)			
RACAT	USA, Canada	Yes	double-blind	SSZ + HCQ + MTX SSZ + HCQ + MTX ETN (50mg SC) + cDMARD	Y	Y	Y
RAPID-1	multinational	Yes	double-blind	CTZ (200mg SC) + cDMARD CTZ (400mg SC) + cDMARD cDMARD	Y		
RAPID-2	multinational	Yes	double-blind	CTZ (200mg SC) + cDMARD CTZ (400mg SC) + cDMARD cDMARD	Y	Y	Y
RA-SCORE	multinational	Yes	double-blind	cDMARD RTX (1000mg IV) + cDMARD RTX (500mg IV) + cDMARD	Y	Y	Y
RED SEA	England	–	double-blind	ADA (40mg SC) ETN (50mg SC)			
SATORI	Japan	–	double-blind	TOC (8mg/kg IV) cDMARD	Y		Y

Continued on next page

Trial	Region	Multicenter	Masking	Treatment	Availability of ACR 20/50/70 at 6 months f-up	Availability of DAS28 at 6 months f-up	Availability of HAQ-DI at 6 months f-up
SELECT-NEXT	multinational	Yes	double-blind	cDMARD UPA (15mg PO) + cDMARD UPA (30mg PO) + cDMARD		Y	Y
SERENE	multinational	Yes	double-blind	RTX (1000mg IV) + cDMARD cDMARD RTX (500mg IV) + cDMARD	Y	Y	
STAR	USA, Canada	Yes	double-blind	ADA (40mg SC) + cDMARD cDMARD	Y		
START	Belgium	No	double-blind	IFX (10mg/kg IV) + cDMARD IFX (3mg/kg IV Q8WEEK) + cDMARD cDMARD	Y		
SURPRISE	Japan	Yes	double-blind	TOC (8mg/kg IV) TOC (8mg/kg IV) + cDMARD	Y	Y	Y
Takeuchi 2013a	Japan	Yes	double-blind	ABT (10mg/kg IV) + cDMARD ABT (2mg/kg IV) + cDMARD cDMARD	Y	Y	Y
TEMPO	multinational	Yes	double-blind	ETN (50mg SC) cDMARD ETN (50mg SC) + cDMARD ETN (50mg SC) + cDMARD	Y		Y
TOWARD	multinational	Yes	double-blind	TOC (8mg/kg IV) + cDMARD cDMARD	Y	Y	Y
van de Putte 2004	multinational	Yes	double-blind	ADA (40mg SC) Placebo ADA (20mg SC Q2WEEK) ADA (20mg SC QWEEK) ADA (40mg SC QWEEK)	Y	Y	Y
Weinblatt 1999	USA	Yes	double-blind	ETN (50mg SC) cDMARD	Y		
Weinblatt 2018	Poland, Lithuania	Yes	double-blind	ADA-SB5 (40mg SC) + cDMARD ADA (40mg SC) + cDMARD	Y	Y	

I.3.3.2 Patient characteristics

Table A13: Patient characteristics, tDMARD naive population

Trial	Intervention	N	Age (mean,(SD))	Male (n,(%))	Caucasian (n,(%))	Asian (n,(%))	TJC (mean,(SD))	SJC (mean,(SD))	DAS28 CRP (mean,(SD))	DAS28 ESR (mean,(SD))	HAQ-DI (mean,(SD))
ACQUIRE	ABT (125mg SC) + cDMARD	736	49.9 (13.2)	– (15.6)	– (74.7)	– (–)	30.1 (14.1)	20.4 (9.6)	6.2 (0.9)	– (–)	1.7 (0.7)
	ABT (10mg/kg IV) + cDMARD	721	50.1 (12.6)	– (19.6)	– (74.5)	– (–)	29.1 (13.3)	19.4 (8.6)	6.2 (0.8)	– (–)	1.7 (0.7)
ACT-RAY	TOC (8mg/kg IV) + cDMARD	279	53.0 (13.4)	50 (18.1)	– (–)	– (–)	25.8 (13.9)	14.4 (8.9)	– (–)	6.3 (1.0)	1.5 (0.7)
	TOC (8mg/kg IV)	277	53.6 (11.9)	59 (21.4)	– (–)	– (–)	26.6 (15.2)	15.3 (10.2)	– (–)	6.4 (1.0)	1.5 (0.6)
ADACTA	TOC (8mg/kg IV)	163	54.4 (13.0)	34 (21.0)	145 (89.0)	– (–)	15.9 (6.7)c	11.3 (5.3)c	– (–)	6.7 (0.9)	1.6 (0.6)
	ADA (40mg SC)	163	53.3 (12.4)	29 (18.0)	133 (82.0)	– (–)	16.5 (7.0)c	12.4 (5.4)c	– (–)	6.8 (0.9)	1.7 (0.6)
AIM	ABT (10mg/kg IV) + cDMARD	433	51.5 (12.9)	– (22.2)	– (87.5)	– (–)	31.0 (13.2)	21.4 (8.8)	– (–)	– (–)	1.7 (0.7)
	cDMARD	219	50.4 (12.4)	– (18.3)	– (88.1)	– (–)	32.3 (13.6)	22.1 (8.8)	– (–)	– (–)	1.7 (0.6)
AMPLE	ADA (40mg SC) + cDMARD	328	51.0 (12.8)	– (17.6)	– (78.0)	– (–)	26.3 (15.8)	15.9 (10.0)	5.5 (1.1)	– (–)	1.5 (0.7)
	ABT (125mg SC) + cDMARD	318	51.4 (12.6)	– (18.6)	– (80.8)	– (–)	25.4 (15.3)	15.8 (9.8)	5.5 (1.1)	– (–)	1.5 (0.7)
ARMADA	ADA (80mg SC) + cDMARD	73	55.5 (11.7)	– (24.7)	– (–)	– (–)	30.3 (15.7)	17.0 (8.2)	– (–)	– (–)	1.6 (0.7)
	ADA (20mg SC) + cDMARD	69	53.5 (12.4)	– (24.6)	– (–)	– (–)	28.5 (14.4)	17.6 (8.7)	– (–)	– (–)	1.5 (0.6)
	ADA (40mg SC) + cDMARD	67	57.2 (11.4)	– (25.4)	– (–)	– (–)	28.0 (12.7)	17.3 (8.6)	– (–)	– (–)	1.6 (0.6)
	cDMARD	62	56.0 (10.8)	– (17.7)	– (–)	– (–)	28.7 (15.2)	16.9 (9.5)	– (–)	– (–)	1.6 (0.6)
ATTEST	IFX (3mg/kg IV Q8WEEK) + cDMARD	165	49.1 (12.0)	– (17.6)	– (80.6)	– (–)	31.7 (14.5)	20.3 (8.0)	– (–)	6.8 (0.9)	1.7 (0.7)
	ABT (10mg/kg IV) + cDMARD	156	49.0 (12.5)	– (16.7)	– (80.8)	– (–)	31.6 (13.9)	21.3 (8.6)	– (–)	6.9 (1.0)	1.8 (0.6)
	cDMARD	110	49.4 (11.5)	– (12.7)	– (76.4)	– (–)	30.3 (11.7)	20.1 (7.0)	– (–)	6.8 (1.0)	1.8 (0.7)
ATTRACT	cDMARD	88	51.0a	– (20.0)	78 (89.0)	– (–)	24.0a	19.0a	– (–)	– (–)	– (–)
	IFX (10mg/kg IV Q8WEEK) + cDMARD	87	55.0a	– (23.0)	79 (91.0)	– (–)	30.0a	20.0a	– (–)	– (–)	– (–)
	IFX (3mg/kg IV Q4WEEK) + cDMARD	86	51.0a	– (23.0)	76 (88.0)	– (–)	31.0a	20.0a	– (–)	– (–)	– (–)
	IFX (3mg/kg IV Q8WEEK) + cDMARD	86	56.0a	– (19.0)	80 (93.0)	– (–)	32.0a	19.0a	– (–)	– (–)	– (–)
	IFX (10mg/kg IV Q4WEEK) + cDMARD	81	52.0a	– (27.0)	76 (94.0)	– (–)	35.0a	23.0a	– (–)	– (–)	– (–)

Continued on next page

Trial	Intervention	N	Age (mean,(SD))	Male (n,(%))	Caucasian (n,(%))	Asian (n,(%))	TJC (mean,(SD))	SJC (mean,(SD))	DAS28 CRP (mean,(SD))	DAS28 ESR (mean,(SD))	HAQ-DI (mean,(SD))
Bao 2011	ANA (80mg SC) + cDMARD	42	45.0 (10.0)	9 (21.4)	– (–)	– (–)	11.4 (6.5)	7.8 (4.5)	– (–)	– (–)	.6 (0.7)
	cDMARD	12	45.0 (11.0)	2 (16.7)	– (–)	– (–)	10.4 (7.1)	6.1 (4.0)	– (–)	– (–)	.7 (0.6)
CHANGE	ADA (40mg SC)	91	56.9 (10.3)	19 (20.9)	– (–)	– (–)	24.4 (10.7)	19.1 (7.3)	– (–)	– (–)	1.6 (0.7)
	Placebo	87	53.4 (12.8)	20 (23.0)	– (–)	– (–)	23.7 (8.8)	19.3 (7.0)	– (–)	– (–)	1.4 (0.8)
	ADA (20mg SC Q2WEEK)	87	54.8 (12.5)	18 (20.7)	– (–)	– (–)	24.6 (11.1)	19.2 (8.4)	– (–)	– (–)	1.6 (0.8)
	ADA (80mg SC)	87	54.3 (10.9)	15 (17.2)	– (–)	– (–)	24.9 (10.7)	20.8 (7.9)	– (–)	– (–)	1.8 (0.7)
Choy 2012	CTZ (400mg SC) + cDMARD	126	53.0 (12.3)	35 (27.8)	– (–)	– (–)	29.0 (11.6)	22.8 (9.4)	6.2 (1.0)	– (–)	– (–)
	cDMARD	121	55.6 (11.7)	41 (33.9)	– (–)	– (–)	31.0 (12.9)	22.2 (9.6)	6.3 (1.0)	– (–)	– (–)
Cohen 2002	cDMARD	74	53.0 (–)	– (14.9)	67 (90.5)	– (–)	28.1 (13.9)	18.4 (9.8)	– (–)	– (–)	1.4 (0.6)
	ANA (0.04mg/kg SC) + cDMARD	63	52.6 (–)	– (22.2)	56 (88.9)	– (–)	23.9 (11.4)	18.8 (8.7)	– (–)	– (–)	1.4 (0.6)
	ANA (0.1mg/kg SC) + cDMARD	74	53.0 (–)	– (20.3)	67 (90.5)	– (–)	25.9 (14.8)	18.3 (9.2)	– (–)	– (–)	1.5 (0.7)
	ANA (0.4mg/kg SC) + cDMARD	77	52.8 (–)	– (23.4)	64 (83.1)	– (–)	27.1 (13.0)	19.1 (9.2)	– (–)	– (–)	1.5 (0.6)
	ANA (1mg/kg SC) + cDMARD	59	49.0 (–)	– (15.3)	51 (86.4)	– (–)	22.0 (12.9)	17.6 (8.8)	– (–)	– (–)	1.3 (0.6)
	ANA (2mg/kg SC) + cDMARD	72	54.1 (–)	– (37.5)	66 (91.7)	– (–)	24.6 (12.8)	17.4 (8.1)	– (–)	– (–)	1.3 (0.6)
Cohen 2004	ANA (100mg SC) + cDMARD	250	56.0 (–)	– (21.0)	– (86.0)	– (–)	26.8 (15.7)	20.1 (11.7)	– (–)	– (–)	1.4 (0.6)
	cDMARD	251	57.0 (–)	– (25.0)	– (87.0)	– (–)	24.5 (13.1)	20.0 (10.2)	– (–)	– (–)	1.3 (0.6)
Cohen 2018	IFX (3mg/kg IV Q8WEEK) + cDMARD	326	52.8 (12.9)	62 (19.0)	247 (76.0)	45 (13.8)	25.7 (12.9)	16.3 (8.7)	6.0 (0.9)	– (–)	1.6 (0.7)
	IFX-Pfizer (3mg/kg IV) + cDMARD	324	52.8 (13.3)	66 (20.4)	257 (79.0)	46 (14.2)	24.7 (13.9)	16.1 (9.4)	6.0 (1.0)	– (–)	1.6 (0.6)
De Filippis 2006	ETN (50mg SC) + cDMARD	16	44.7 (14.2)	– (–)	– (–)	– (–)	22.4 (8.1)	16.9 (7.3)	– (–)	– (–)	1.9 (0.7)
	IFX (3mg/kg IV Q8WEEK) + cDMARD	16	46.8 (10.9)	– (–)	– (–)	– (–)	20.9 (10.0)	14.7 (5.0)	– (–)	– (–)	1.7 (0.7)
DE019	ADA (20mg SC) + cDMARD	212	57.3 (10.5)	52 (24.5)	– (85.4)	– (–)	27.9 (13.6)	19.6 (9.9)	– (–)	– (–)	1.4 (0.6)
	ADA (40mg SC) + cDMARD	207	56.1 (13.5)	49 (23.7)	– (83.6)	– (–)	27.3 (12.7)	19.3 (9.8)	– (–)	– (–)	1.5 (0.6)
	cDMARD	200	56.1 (12.0)	54 (27.0)	– (83.0)	– (–)	28.1 (13.8)	19.0 (9.5)	– (–)	– (–)	1.5 (0.6)
Edwards 2004	RTX (1000mg IV)	40	54.0 (10.0)	– (27.0)	– (–)	– (–)	34.0 (15.0)	21.0 (11.0)	– (–)	6.8 (1.0)	– (–)
	RTX (1000mg IV) + cDMARD	40	54.0 (12.0)	– (25.0)	– (–)	– (–)	32.0 (16.0)	23.0 (13.0)	– (–)	6.8 (0.9)	– (–)
	cDMARD	40	54.0 (11.0)	– (20.0)	– (–)	– (–)	32.0 (13.0)	19.0 (10.0)	– (–)	6.9 (0.8)	– (–)
Emery 2017	ETN-SB4 (50mg SC) + cDMARD	299	52.1 (11.7)	50 (16.7)	– (93.3)	– (3.7)	23.5 (11.9)	15.4 (7.5)	– (–)	6.5 (0.9)	1.5 (0.6)

Continued on next page

Trial	Intervention	N	Age (mean,(SD))	Male (n,(%))	Caucasian (n,(%))	Asian (n,(%))	TJC (mean,(SD))	SJC (mean,(SD))	DAS28 CRP (mean,(SD))	DAS28 ESR (mean,(SD))	HAQ-DI (mean,(SD))
	ETN (50mg SC) + cDMARD	297	51.6 (11.6)	44 (14.8)	– (91.9)	– (4.4)	23.6 (12.6)	15.0 (7.3)	– (–)	6.5 (0.9)	1.5 (0.6)
EQUIRA	ETN (50mg SC) + cDMARD	190	53.2 (12.7)	40 (21.1)	185 (97.0)	3 (1.6)	14.8 (5.8)c	11.1 (5.4)c	5.6 (0.8)	– (–)	1.4 (0.6)
	ETN-GP2015 (50mg SC) + cDMARD	186	55.2 (11.2)	28 (15.1)	180 (97.0)	(.0)	14.2 (6.2)c	10.5 (5.3)c	5.4 (0.9)	– (–)	1.5 (0.6)
ETN Study 309	ETN (50mg SC)	103	51.3 (13.5)	22 (21.4)	– (–)	– (–)	29.7 (14.7)	19.1 (10.1)	– (–)	5.1 (1.1)	1.7 (0.6)
	ETN (50mg SC) + cDMARD	101	50.6 (12.3)	20 (19.8)	– (–)	– (–)	31.3 (14.1)	19.4 (10.4)	– (–)	5.2 (1.2)	1.6 (0.6)
	cDMARD	50	53.3 (12.8)	9 (18.0)	– (–)	– (–)	31.3 (14.0)	18.7 (11.1)	– (–)	5.0 (1.1)	1.6 (0.5)
FAST4WARD	CTZ (400mg SC)	111	52.7 (12.7)	24 (21.6)	– (–)	– (–)	29.6 (13.7)	21.2 (10.1)	– (–)	6.3 (1.1)	1.4 (0.6)
	Placebo	109	54.9 (11.6)	12 (11.0)	– (–)	– (–)	28.3 (12.5)	19.9 (9.3)	– (–)	6.3 (0.9)	1.6 (0.7)
Fleischmann 2012	Placebo	59	53.0 (13.7)	– (11.0)	43 (72.9)	6 (10.2)	25.9 (–)	16.9 (–)	5.6 (–)	6.6 (–)	1.5 (–)
	TOF (1mg PO)	54	55.0 (13.3)	– (14.8)	44 (81.5)	5 (9.3)	27.0 (–)	16.7 (–)	5.5 (–)	6.5 (–)	1.6 (–)
	TOF (3mg PO)	51	53.0 (12.2)	– (13.7)	38 (74.5)	5 (9.8)	24.6 (–)	15.9 (–)	5.4 (–)	6.4 (–)	1.5 (–)
	TOF (5mg PO)	49	54.0 (13.5)	– (12.2)	36 (73.5)	6 (12.2)	27.1 (–)	17.4 (–)	5.6 (–)	6.6 (–)	1.4 (–)
	TOF (10mg PO)	61	52.0 (10.9)	– (13.1)	44 (72.1)	5 (8.2)	25.7 (–)	16.3 (–)	5.5 (–)	6.5 (–)	1.5 (–)
	TOF (15mg PO)	57	53.0 (13.0)	– (12.3)	46 (80.7)	4 (7.0)	25.9 (–)	16.9 (–)	5.5 (–)	6.5 (–)	1.6 (–)
	ADA (40mg SC)	53	54.0 (11.9)	– (15.1)	43 (81.1)	4 (7.5)	24.1 (–)	14.9 (–)	5.4 (–)	6.3 (–)	1.4 (–)
Fleischmann 2018	ADA (40mg SC) + cDMARD	300	53.5 (12.9)	71 (23.7)	256 (85.0)	17 (5.7)	26.7 (14.8)	17.0 (9.8)	6.1 (0.9)	– (–)	1.7 (0.6)
	ADA-Pfizer (40mg SC) + cDMARD	297	51.5 (13.6)	56 (18.9)	261 (88.0)	16 (5.4)	24.3 (12.3)	15.4 (7.8)	5.9 (0.9)	– (–)	1.5 (0.6)
GO-FORTH	cDMARD	90	51.1 (11.6)	15 (17.0)	– (–)	– (100.0)	13.2 (7.8)	11.4 (6.6)	– (–)	5.6 (1.0)	1.0 (0.7)
	GOL (100mg SC) + cDMARD	90	50.0 (12.2)	9 (10.3)	– (–)	– (100.0)	12.9 (7.6)	11.5 (6.6)	– (–)	5.5 (1.0)	.9 (0.6)
	GOL (50mg SC) + cDMARD	89	50.4 (9.9)	13 (15.1)	– (–)	– (100.0)	13.1 (8.4)	11.8 (6.7)	– (–)	5.5 (1.2)	1.0 (0.6)
GO-FORWARD	GOL (100mg SC)	133	51.0a	28 (21.1)	– (–)	– (–)	22.0a	11.0a	4.8a	6.0a	1.4 (–)
	cDMARD	133	52.0a	24 (18.0)	– (–)	– (–)	21.0a	12.0a	4.9a	6.1a	1.3 (–)
	GOL (100mg SC) + cDMARD	89	50.0a	17 (19.1)	– (–)	– (–)	23.0a	12.0a	4.9a	5.9a	1.4 (–)
	GOL (50mg SC) + cDMARD	89	52.0a	17 (19.1)	– (–)	– (–)	26.0a	13.0a	5.1a	6.1a	1.4 (–)
GO-FURTHER	GOL (2mg/kg IV) + cDMARD	395	51.9 (12.6)	69 (17.5)	– (–)	– (–)	26.4 (13.9)	15.0 (8.2)	6.0 (0.8)	– (–)	1.6 (0.6)
	cDMARD	197	51.4 (11.3)	40 (20.3)	– (–)	– (–)	25.9 (14.1)	14.8 (8.5)	5.9 (0.9)	– (–)	1.6 (0.7)
GO-LIVE	GOL (4mg/kg IV)	129	48.4 (–)	24 (18.6)	86 (67.0)	13 (10.1)	26.5 (24.0)	15.2 (14.0)	– (–)	– (–)	1.5 (–)
	GOL (2mg/kg IV) + cDMARD	129	49.7 (–)	30 (23.3)	88 (68.0)	10 (7.8)	26.8 (23.0)	15.5 (13.0)	– (–)	– (–)	1.5 (–)
	cDMARD	129	50.2 (–)	26 (20.2)	92 (71.0)	11 (8.5)	28.2 (23.0)	16.1 (13.0)	– (–)	– (–)	1.5 (–)

Continued on next page

Trial	Intervention	N	Age (mean,(SD))	Male (n,(%))	Caucasian (n,(%))	Asian (n,(%))	TJC (mean,(SD))	SJC (mean,(SD))	DAS28 CRP (mean,(SD))	DAS28 ESR (mean,(SD))	HAQ-DI (mean,(SD))
	GOL (4mg/kg IV) + cDMARD	128	49.6 (–)	25 (19.5)	88 (69.0)	10 (7.8)	27.1 (23.0)	15.3 (14.0)	– (–)	– (–)	1.5 (–)
	GOL (2mg/kg IV)	128	49.9 (–)	21 (16.4)	93 (73.0)	9 (7.0)	28.1 (24.0)	15.7 (13.0)	– (–)	– (–)	1.6 (–)
HIKARI	CTZ (200mg SC) + cDMARD	116	56.0 (10.2)	33 (28.4)	– (–)	– (100.0)	16.2 (9.6)	13.8 (7.5)	– (–)	6.1 (0.9)	1.1 (0.7)
	cDMARD	114	55.4 (9.8)	26 (22.8)	– (–)	– (100.0)	17.6 (10.3)	15.5 (8.6)	– (–)	6.3 (1.0)	1.2 (0.7)
Iwahashi 2014	ABT (125mg SC) + cDMARD	59	56.1 (12.3)	21 (35.6)	– (–)	– (–)	20.9 (9.3)	16.4 (7.0)	5.6 (0.8)	– (–)	1.3 (0.7)
	ABT (10mg/kg IV) + cDMARD	59	55.2 (13.6)	11 (18.6)	– (–)	– (–)	22.3 (9.9)	17.6 (7.2)	6.0 (0.9)	– (–)	1.3 (0.6)
Jamshidi 2017	ADA-Cinnora (40mg SC) + cDMARD	69	48.3 (12.7)	10 (14.7)	– (–)	– (–)	9.5 (8.2)	10.0 (7.4)	– (–)	5.5 (1.2)	– (–)
	ADA (40mg SC) + cDMARD	69	47.6 (11.5)	8 (11.8)	– (–)	– (–)	9.7 (8.0)	9.5 (7.0)	– (–)	5.5 (1.3)	– (–)
JESMR	ETN (50mg SC) + cDMARD	77	56.5 (11.1)	15 (20.0)	– (–)	– (–)	14.9 (8.0)	12.6 (6.5)	– (–)	– (–)	– (–)
	ETN (50mg SC)	74	58.1 (12.6)	9 (12.7)	– (–)	– (–)	15.0 (9.4)	12.5 (6.1)	– (–)	– (–)	– (–)
J-RAPID	CTZ (400mg SC) + cDMARD	85	55.4 (10.3)	16 (18.8)	– (–)	– (100.0)	20.5 (10.2)	16.6 (7.4)	– (–)	6.3 (0.8)	1.1 (0.6)
	CTZ (200mg SC) + cDMARD	82	50.6 (11.4)	13 (15.9)	– (–)	– (100.0)	19.0 (9.0)	16.6 (8.4)	– (–)	6.2 (0.8)	1.1 (0.7)
	cDMARD	77	51.9 (11.1)	11 (14.3)	– (–)	– (100.0)	19.6 (10.4)	17.4 (10.0)	– (–)	6.5 (0.9)	1.2 (0.7)
	CTZ (100mg SC) + cDMARD	72	54.3 (10.6)	14 (19.4)	– (–)	– (100.0)	21.2 (13.3)	18.4 (10.7)	– (–)	6.3 (0.9)	1.2 (0.7)
Kim 2007	ADA (40mg SC)	65	48.5 (10.2)	3 (4.6)	– (–)	– (–)	19.2 (9.2)	12.2 (5.6)	– (–)	– (–)	– (–)
	cDMARD	63	49.8 (10.5)	9 (14.3)	– (–)	– (–)	20.3 (8.6)	12.8 (5.8)	– (–)	– (–)	– (–)
Kremer 2003	cDMARD	119	54.7 (–)	53 (–)	– (87.0)	– (–)	29.2 (13.0)	21.8 (8.8)	– (–)	– (–)	– (–)
	ABT (10mg/kg IV) + cDMARD	115	55.8 (–)	40 (–)	– (87.0)	– (–)	30.8 (12.2)	21.3 (8.4)	– (–)	– (–)	– (–)
	ABT (2mg/kg IV) + cDMARD	105	54.4 (–)	42 (–)	– (87.0)	– (–)	28.2 (12.0)	20.2 (8.9)	– (–)	– (–)	– (–)
Kremer 2012	cDMARD	69	53.0 (13.4)	– (18.8)	58 (84.1)	(–)	21.6 (–)	15.7 (–)	– (–)	6.1 (–)	1.2 (–)
	TOF (1mg PO) + cDMARD	70	52.0 (11.6)	– (18.6)	61 (87.1)	(–)	23.6 (–)	16.5 (–)	– (–)	6.4 (–)	1.6 (–)
	TOF (3mg PO) + cDMARD	68	51.0 (14.9)	– (23.5)	54 (79.4)	1 (1.5)	22.8 (–)	15.7 (–)	– (–)	6.1 (–)	1.4 (–)
	TOF (5mg PO) + cDMARD	71	52.0 (12.8)	– (19.7)	63 (88.7)	(–)	21.5 (–)	14.1 (–)	– (–)	6.1 (–)	1.4 (–)
	TOF (10mg PO) + cDMARD	74	56.0 (10.4)	– (25.7)	64 (86.5)	(–)	24.8 (–)	14.7 (–)	– (–)	6.4 (–)	1.3 (–)
	TOF (15mg PO) + cDMARD	75	54.0 (11.1)	– (12.0)	65 (86.7)	(–)	23.7 (–)	15.3 (–)	– (–)	6.2 (–)	1.4 (–)
	TOF (20mg PO) + cDMARD	80	54.0 (10.8)	– (12.2)	72 (90.0)	(–)	23.1 (–)	15.2 (–)	– (–)	6.3 (–)	1.5 (–)

Continued on next page

Trial	Intervention	N	Age (mean,(SD))	Male (n,(%))	Caucasian (n,(%))	Asian (n,(%))	TJC (mean,(SD))	SJC (mean,(SD))	DAS28 CRP (mean,(SD))	DAS28 ESR (mean,(SD))	HAQ-DI (mean,(SD))
LARA	ETN (50mg SC) + cDMARD	281	48.4 (12.0)	33 (11.7)	134 (48.0)	– (–)	25.1 (11.9)	18.2 (8.4)	– (–)	6.6 (0.7)	1.6 (0.7)
	cDMARD	142	48.6 (11.3)	14 (9.9)	65 (46.0)	– (–)	26.2 (12.3)	19.3 (10.1)	– (–)	6.7 (0.7)	1.6 (0.7)
Li 2016	cDMARD	132	46.7 (12.2)	28 (21.2)	– (–)	– (–)	22.5 (14.8)	11.8 (7.4)	5.5 (1.1)	– (–)	1.2 (0.7)
	GOL (50mg SC) + cDMARD	132	47.7 (11.5)	22 (16.7)	– (–)	– (–)	22.9 (15.4)	10.7 (7.0)	5.4 (1.1)	– (–)	1.3 (0.7)
LITHE	TOC (4mg/kg IV) + cDMARD	401	51.4 (12.6)	– (16.0)	– (–)	– (–)	27.9 (14.2)	17.0 (9.8)	– (–)	6.5 (0.9)	1.5 (0.6)
	TOC (8mg/kg IV) + cDMARD	401	53.4 (11.7)	– (18.0)	– (–)	– (–)	29.3 (15.2)	17.3 (9.5)	– (–)	6.6 (1.0)	1.5 (0.6)
	cDMARD	394	51.3 (12.4)	– (17.0)	– (–)	– (–)	27.9 (14.8)	16.6 (9.2)	– (–)	6.5 (1.0)	1.5 (0.6)
Matsubara 2018	ABT (10mg/kg IV) + cDMARD	203	56.6 (12.5)	38 (18.7)	– (–)	– (–)	13.8 (8.9)	13.0 (8.0)	4.9 (1.0)	– (–)	1.0 (0.7)
	cDMARD	202	54.8 (12.1)	27 (13.4)	– (–)	– (–)	13.9 (8.3)	12.3 (6.8)	4.7 (1.1)	– (–)	.9 (0.6)
MOBILITY	SAR (150mg SC) + cDMARD	400	50.1 (11.9)	– (20.0)	– (86.3)	– (8.3)	27.2 (14.2)	16.6 (9.0)	6.0 (0.9)	– (–)	1.6 (0.6)
	SAR (200mg SC) + cDMARD	399	50.8 (11.8)	– (15.0)	– (86.0)	– (8.3)	26.5 (14.5)	16.8 (9.7)	6.0 (0.9)	– (–)	1.7 (0.6)
	cDMARD	398	50.9 (11.2)	– (19.0)	– (86.2)	– (8.0)	26.8 (13.7)	16.7 (9.3)	5.9 (0.9)	– (–)	1.6 (0.7)
MONARCH	ADA (40mg SC)	185	53.6 (11.9)	35 (18.9)	164 (88.6)	– (–)	26.7 (13.6)	17.5 (10.3)	6.0 (0.9)	6.8 (0.8)	1.6 (0.6)
	SAR (200mg SC)	184	50.9 (12.6)	27 (14.7)	171 (92.9)	– (–)	28.0 (13.2)	18.6 (10.7)	6.0 (0.9)	6.8 (0.8)	1.6 (0.6)
Moreland 1999	Placebo	80	51.0 (–)	– (24.0)	– (89.0)	– (–)	35.0 (–)d	25.0 (–)d	– (–)	– (–)	1.7 (–)
	ETN (10mg SC)	76	53.0 (–)	– (16.0)	– (96.0)	– (–)	34.0 (–)d	25.0 (–)d	– (–)	– (–)	1.7 (–)
	ETN (50mg SC)	78	53.0 (–)	– (26.0)	– (94.0)	– (–)	33.0 (–)d	25.0 (–)d	– (–)	– (–)	1.6 (–)
Niu 2011	cDMARD	12	45.3 (–)	– (16.7)	– (–)	– (–)	12.3 (5.8)	10.3 (4.6)	– (–)	– (–)	.7 (0.3)
	ANA (80mg SC) + cDMARD	38	46.1 (–)	– (18.4)	– (–)	– (–)	11.7 (5.4)	11.8 (6.5)	– (–)	– (–)	.7 (0.4)
OPTION	TOC (4mg/kg IV) + cDMARD	214	51.4 (12.8)	38 (18.0)	– (–)	– (–)	33.2 (15.6)	20.0 (10.9)	– (–)	6.8 (0.9)	1.6 (0.6)
	TOC (8mg/kg IV) + cDMARD	205	50.8 (11.8)	30 (15.0)	– (–)	– (–)	31.9 (15.5)	19.5 (11.3)	– (–)	6.8 (0.9)	1.6 (0.6)
	cDMARD	204	50.6 (12.1)	45 (22.0)	– (–)	– (–)	32.8 (16.1)	20.7 (11.7)	– (–)	6.8 (0.9)	1.5 (0.6)
ORAL-SCAN	TOF (5mg PO) + cDMARD	321	53.7 (11.6)	52 (16.2)	– (47.4)	– (–)	24.1 (–)	14.1 (–)	5.2 (–)	6.3 (–)	1.4 (0.7)
	TOF (10mg PO) + cDMARD	316	52.0 (11.4)	33 (13.6)	– (45.6)	– (–)	23.0 (–)	14.4 (–)	5.2 (–)	6.3 (–)	1.4 (0.7)
	cDMARD	156	– (–)	– (–)	– (–)	– (–)	22.9 (–)	14.2 (–)	– (–)	– (–)	1.3 (0.7)
ORAL-STANDARD	TOF (5mg PO) + cDMARD	204	53.0 (11.9)	30 (14.7)	– (–)	– (–)	28.5 (–)	16.7 (–)	5.4 (0.9)	– (–)	1.5 (–)
	ADA (40mg SC) + cDMARD	204	52.5 (11.7)	42 (20.6)	– (–)	– (–)	26.7 (–)	16.4 (–)	5.3 (0.9)	– (–)	1.5 (–)
	TOF (10mg PO) + cDMARD	201	52.9 (11.8)	33 (16.4)	– (–)	– (–)	26.1 (–)	15.8 (–)	5.4 (0.8)	– (–)	1.5 (–)
	cDMARD	108	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
ORAL-STRATEGY	ADA (40mg SC) + cDMARD	386	50.7 (13.4)	66 (17.0)	293 (76.0)	40 (11.0)	15.2 (6.7)c	11.0 (5.4)c	5.7 (1.0)	6.5 (1.0)	1.6 (0.6)

Continued on next page

Trial	Intervention	N	Age (mean,(SD))	Male (n,(%))	Caucasian (n,(%))	Asian (n,(%))	TJC (mean,(SD))	SJC (mean,(SD))	DAS28 CRP (mean,(SD))	DAS28 ESR (mean,(SD))	HAQ-DI (mean,(SD))
	TOF (5mg PO)	386	49.7 (12.2)	65 (17.0)	296 (77.0)	41 (11.0)	15.4 (6.5)c	11.2 (5.6)c	5.7 (0.9)	6.5 (0.9)	1.6 (0.6)
	TOF (5mg PO) + cDMARD	376	50.0 (13.4)	65 (17.0)	286 (76.0)	38 (10.0)	15.6 (6.5)c	11.8 (5.7)c	5.8 (0.9)	6.6 (0.9)	1.6 (0.6)
ORAL-SYNC	TOF (10mg PO) + cDMARD	318	51.9 (11.8)	– (18.9)	– (54.7)	– (–)	26.6 (16.1)	14.4 (9.7)	– (–)	6.4 (1.1)	1.4 (0.7)
	TOF (5mg PO) + cDMARD	318	52.7 (11.7)	– (16.2)	– (54.9)	– (–)	25.0 (15.3)	14.5 (10.3)	– (–)	6.3 (1.0)	1.4 (0.7)
	cDMARD	158	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	1.4 (0.7)
RA-BEAM	cDMARD	488	53.0 (2.0)	106 (21.7)	– (–)	– (–)	23.0 (14.0)	16.0 (9.0)	5.7 (1.0)	6.4 (–)	1.6 (0.7)
	BCT (4mg PO) + cDMARD	487	54.0 (2.0)	112 (23.0)	– (–)	– (–)	23.0 (13.0)	15.0 (8.0)	5.8 (0.9)	6.5 (–)	1.6 (0.7)
	ADA (40mg SC) + cDMARD	330	53.0 (12.0)	79 (23.9)	– (–)	– (–)	23.0 (14.0)	15.0 (9.0)	5.8 (0.9)	6.4 (–)	1.6 (0.7)
RA-BUILD-Af	BCT (2mg PO) + cDMARD	229	52.0 (12.0)	45 (20.0)	– (–)	– (–)	24.0 (14.0)	14.0 (9.0)	5.6 (1.0)	6.3 (1.0)	1.5 (0.6)
	cDMARD	228	51.0 (13.0)	39 (17.0)	– (–)	– (–)	24.0 (15.0)	13.0 (7.0)	5.5 (0.9)	6.2 (1.0)	1.5 (0.6)
	BCT (4mg PO) + cDMARD	227	52.0 (12.0)	40 (18.0)	– (–)	– (–)	24.0 (14.0)	14.0 (7.0)	5.6 (0.9)	6.2 (0.9)	1.6 (0.6)
RA-BUILD-Bf	BCT (2mg PO) + cDMARD	229	52.0 (12.0)	45 (20.0)	– (–)	– (–)	24.0 (14.0)	14.0 (9.0)	5.6 (1.0)	6.3 (1.0)	1.5 (0.6)
	cDMARD	228	51.0 (13.0)	39 (17.0)	– (–)	– (–)	24.0 (15.0)	13.0 (7.0)	5.5 (0.9)	6.2 (1.0)	1.5 (0.6)
	BCT (4mg PO) + cDMARD	227	52.0 (12.0)	40 (18.0)	– (–)	– (–)	24.0 (14.0)	14.0 (7.0)	5.6 (0.9)	6.2 (0.9)	1.6 (0.6)
RACAT	SSZ + HCQ + MTX	178	57.8 (13.0)	101 (56.7)	– (90.4)	– (–)	13.4 (6.6)	11.1 (5.3)	– (–)	5.8 (0.9)	1.4 (0.8)
	ETN (50mg SC) + cDMARD	175	56.0 (13.2)	90 (51.4)	– (83.4)	– (–)	13.3 (6.4)	11.3 (5.2)	– (–)	5.9 (0.9)	1.5 (0.8)
RAPID-1	CTZ (200mg SC) + cDMARD	393	51.4 (11.6)	– (17.6)	– (–)	– (–)	30.8 (12.4)	21.7 (9.9)	– (–)	6.9a	– (–)
	CTZ (400mg SC) + cDMARD	390	52.4 (11.7)	– (16.4)	– (–)	– (–)	31.1 (13.3)	21.5 (9.8)	– (–)	6.9a	– (–)
	cDMARD	199	52.2 (11.2)	– (16.1)	– (–)	– (–)	29.8 (13.0)	21.2 (9.7)	– (–)	7.0a	– (–)
RAPID-2	CTZ (200mg SC) + cDMARD	246	52.2 (11.1)	40 (16.3)	– (–)	– (–)	30.1 (14.5)	20.5 (9.6)	– (–)	6.9 (0.8)	1.6 (0.6)
	CTZ (400mg SC) + cDMARD	246	51.9 (11.8)	54 (22.0)	– (–)	– (–)	30.0 (13.9)	21.0 (10.2)	– (–)	6.8 (0.8)	1.6 (0.6)
	cDMARD	127	51.5 (11.8)	20 (15.7)	– (–)	– (–)	30.4 (13.4)	21.9 (9.7)	– (–)	6.8 (0.9)	1.6 (0.6)
RA-SCORE	cDMARD	63	50.3 (11.9)	15 (23.8)	– (–)	– (–)	14.9 (6.7)	11.4 (6.1)	5.6 (1.1)	6.3 (1.1)	1.5 (0.8)
	RTX (500mg IV) + cDMARD	62	48.7 (11.1)	17 (27.4)	– (–)	– (–)	15.2 (7.5)	12.5 (7.1)	5.6 (1.1)	6.3 (1.2)	1.4 (0.7)
	RTX (1000mg IV) + cDMARD	60	50.7 (11.7)	10 (16.7)	– (–)	– (–)	14.0 (6.9)	10.9 (5.9)	5.3 (1.0)	6.0 (1.1)	1.3 (0.7)
RED SEA	ADA (40mg SC)	60	55.0 (12.5)	15 (–)	– (–)	– (–)	– (–)	– (–)	5.6 (0.9)	– (–)	– (–)

Continued on next page

Trial	Intervention	N	Age (mean,(SD))	Male (n,(%))	Caucasian (n,(%))	Asian (n,(%))	TJC (mean,(SD))	SJC (mean,(SD))	DAS28 CRP (mean,(SD))	DAS28 ESR (mean,(SD))	HAQ-DI (mean,(SD))
	ETN (50mg SC)	60	53.2 (13.4)	18 (–)	– (–)	– (–)	– (–)	– (–)	5.8 (0.9)	– (–)	– (–)
SATORI	cDMARD	66	50.8 (12.2)	16 (25.0)	– (–)	– (100.0)	14.2 (8.6)e	12.7 (7.5)e	6.2 (0.9)	– (–)	– (–)
	TOC (8mg/kg IV)	61	52.6 (10.6)	6 (9.8)	– (–)	– (100.0)	13.8 (7.5)e	12.4 (5.9)e	6.1 (0.9)	– (–)	– (–)
SELECT-NEXT	UPA (15mg PO) + cDMARD	221	55.3 (11.5)	39 (18.0)	– (–)	– (–)	25.2 (13.8)	16.0 (10.0)	5.7 (1.0)	– (–)	1.5 (0.6)
	cDMARD	221	56.0 (12.2)	55 (25.0)	– (–)	– (–)	24.7 (15.0)	15.4 (9.2)	5.6 (0.8)	– (–)	1.4 (0.6)
	UPA (30mg PO) + cDMARD	219	55.8 (11.3)	47 (21.0)	– (–)	– (–)	26.2 (14.3)	16.2 (10.6)	5.7 (0.9)	– (–)	1.5 (0.6)
SERENE	RTX (1000mg IV) + cDMARD	172	51.3 (12.6)	32 (18.8)	– (80.6)	– (–)	28.7 (15.0)	19.5 (10.3)	5.9 (1.0)	6.5 (1.1)	– (–)
	cDMARD	172	52.2 (12.4)	25 (14.5)	– (82.6)	– (–)	30.2 (15.9)	20.9 (11.3)	6.0 (1.0)	6.5 (1.0)	– (–)
	RTX (500mg IV) + cDMARD	168	51.9 (12.9)	34 (20.4)	– (80.2)	– (–)	27.1 (14.1)	18.6 (9.6)	5.8 (0.9)	6.4 (1.0)	– (–)
STAR	ADA (40mg SC) + cDMARD	318	55.0 (12.8)	65 (20.4)	– (89.0)	– (–)	27.3 (13.0)	20.9 (11.0)	– (–)	– (–)	1.4 (0.6)
	cDMARD	318	55.8 (12.4)	66 (20.8)	– (85.8)	– (–)	27.6 (13.8)	21.3 (11.2)	– (–)	– (–)	1.4 (0.6)
START	cDMARD	363	52.0a	61 (16.8)	– (–)	– (–)	22.0a	15.0a	– (–)	– (–)	– (–)
	IFX (10mg/kg IV) + cDMARD	361	52.0a	80 (22.2)	– (–)	– (–)	22.0a	15.0a	– (–)	– (–)	– (–)
	IFX (3mg/kg IV Q8WEEK) + cDMARD	360	53.0a	72 (20.0)	– (–)	– (–)	22.0a	15.0a	– (–)	– (–)	– (–)
SURPRISE	TOC (8mg/kg IV) + cDMARD	115	55.8 (11.7)	15 (13.0)	– (–)	– (100.0)	9.6 (7.5)	7.6 (5.3)	– (–)	5.1 (1.1)	1.0 (0.7)
	TOC (8mg/kg IV)	111	56.3 (2.7)	15 (13.5)	– (–)	– (100.0)	10.1 (9.0)	9.9 (7.6)	– (–)	5.3 (1.2)	1.0 (0.7)
Takeuchi 2013a	ABT (10mg/kg IV) + cDMARD	61	53.4 (11.3)	– (19.6)	– (–)	– (–)	21.8 (9.3)	16.6 (6.7)	6.0 (0.7)	– (–)	1.3 (0.6)
	ABT (2mg/kg IV) + cDMARD	67	52.5 (11.1)	– (14.9)	– (–)	– (–)	21.0 (8.2)	17.6 (6.5)	5.8 (0.7)	– (–)	1.2 (0.7)
	cDMARD	66	53.4 (12.0)	– (21.3)	– (–)	– (–)	21.6 (8.2)	17.5 (6.1)	6.0 (0.7)	– (–)	1.5 (0.7)
TEMPO	ETN (50mg SC) + cDMARD	231	52.5 (12.4)	60 (26.0)	– (–)	– (–)	34.2 (14.8)	22.1 (11.3)	5.5 (1.2)	– (–)	– (–)
	cDMARD	228	53.0 (12.8)	48 (21.0)	– (–)	– (–)	33.1 (13.4)	22.6 (10.7)	5.5 (1.2)	– (–)	– (–)
	ETN (50mg SC)	223	53.2 (13.8)	52 (23.0)	– (–)	– (–)	35.0 (14.5)	23.0 (10.7)	5.7 (1.1)	– (–)	– (–)
TOWARD	TOC (8mg/kg IV) + cDMARD	805	53.0 (13.0)	– (19.0)	– (72.0)	– (9.0)	30.1 (16.0)	19.7 (11.6)	– (–)	6.7 (1.0)	1.5 (0.6)
	cDMARD	415	54.0 (13.0)	– (16.0)	– (72.0)	– (10.0)	29.1 (14.8)	18.7 (10.8)	– (–)	6.6 (1.0)	1.5 (0.6)
van de Putte 2004	ADA (40mg SC)	113	52.7 (13.3)	23 (20.4)	– (–)	– (–)	33.7 (15.9)	20.5 (10.6)	– (–)	7.1 (0.9)	1.8 (0.6)
	ADA (20mg SC QWEEK)	112	54.4 (11.8)	31 (27.3)	– (–)	– (–)	35.3 (14.9)	19.8 (9.7)	– (–)	7.1 (0.9)	1.9 (0.6)
	Placebo	110	53.5 (13.2)	25 (22.7)	– (–)	– (–)	35.5 (14.2)	19.8 (9.3)	– (–)	7.1 (0.9)	1.9 (0.6)

Continued on next page

Trial	Intervention	N	Age (mean,(SD))	Male (n,(%))	Caucasian (n,(%))	Asian (n,(%))	TJC (mean,(SD))	SJC (mean,(SD))	DAS28 CRP (mean,(SD))	DAS28 ESR (mean,(SD))	HAQ-DI (mean,(SD))
	ADA (20mg SC Q2WEEK)	106	53.1 (12.2)	22 (20.8)	– (–)	– (–)	33.9 (14.4)	19.6 (8.7)	– (–)	7.1 (0.9)	1.9 (0.6)
	ADA (40mg SC QWEEK)	103	51.8 (11.8)	22 (21.4)	– (–)	– (–)	33.8 (14.0)	19.3 (8.8)	– (–)	7.0 (0.8)	1.8 (0.6)
Weinblatt 1999	cDMARD	30	53.0 (–)	27 (–)	– (83.0)	– (–)	28.0 (–)	17.0 (–)	– (–)	– (–)	1.5 (–)
	ETN (50mg SC)	59	48.0 (–)	10 (–)	– (76.0)	– (–)	28.0 (–)	20.0 (–)	– (–)	– (–)	1.5 (–)
Weinblatt 2018	ADA (40mg SC) + cDMARD	273	52.5 (11.9)	49 (17.9)	269 (99.0)	4 (1.5)	24.1 (10.8)	15.5 (7.5)	– (–)	6.5 (0.7)	1.4 (0.6)
	ADA-SB5 (40mg SC) + cDMARD	271	49.8 (12.7)	54 (19.9)	271 (100.0)	(–)	23.9 (11.7)	15.8 (8.0)	– (–)	6.5 (0.7)	1.3 (0.6)

Notes: a. median reported in lieu of mean
b. evaluated out of 68 tender joints and 66 swollen joints respectively, unless other specified
c. 28 joints evaluates
d. 71 tender joints and 68 swollen joints evaluated
e. 49 tender joints and 56 swollen joints evaluated
f. the study design of RA-BUILD permitted but did not require concomitant cDMARD background therapy (which was not based on random assignment, but the discretion of the investigator). Subgroup data stratified by background cDMARD therapy was therefore used within the analysis, and the corresponding reults were treated as two separate trials (RA-BUILD-A and RA-BUILD-B). Baseline demographic data depicted here reflect that of the overall population in lieu of subgroup specific data, which were not unavailable

I.3.3.3 Evidence network

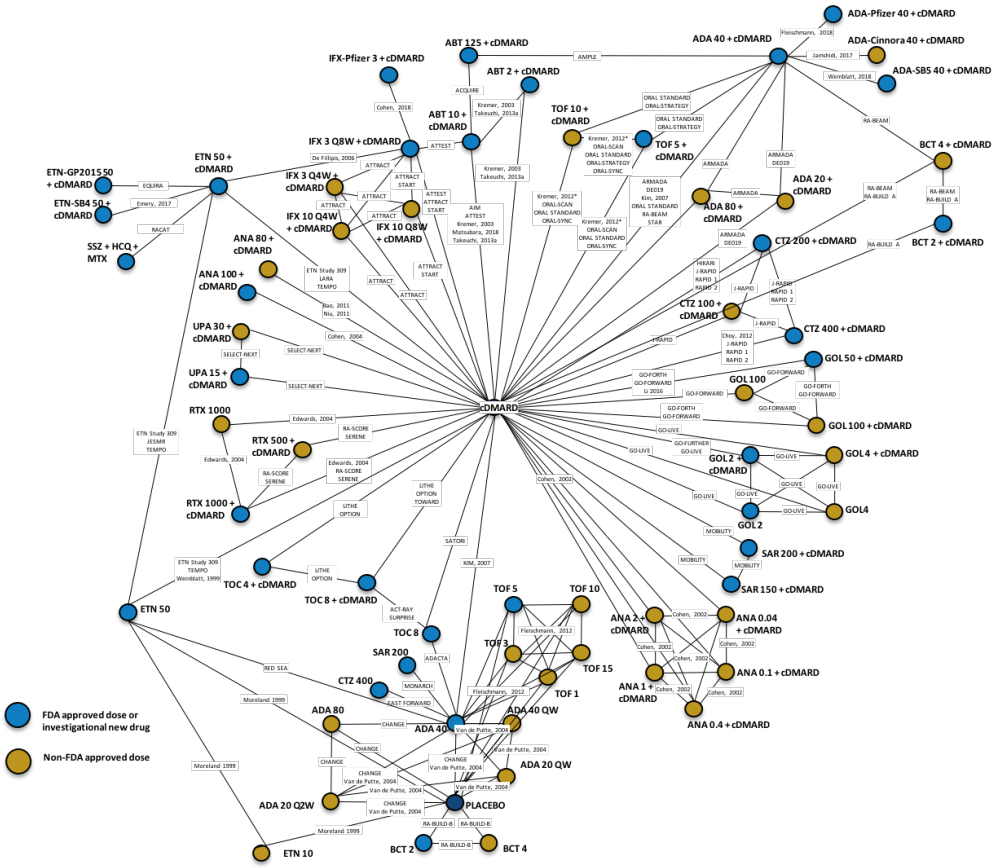


Figure A3: Evidence network, tDMARD naive population

I.3.3.4 Study specific 6-month data used for estimation of treatment effects, tDMARD naive population

Table A14: Study specific data

Trial ID	Treatment	ACR 20 (%)	ACR 50 (%)	ACR 70 (%)	ΔDAS28 (SE)	ΔHAQ-DI (SE)
ACQUIRE	ABT (10mg/kg IV) + cDMARD	535/721 (74.3%)	350/721 (48.6%)	174/721 (24.2%)	-2.57 (0.05)	-0.7 (0.02)
ACQUIRE	ABT (125mg SC) + cDMARD	550/736 (74.8%)	369/736 (50.2%)	189/736 (25.8%)	-2.55 (0.05)	-0.69 (0.02)
ACT-RAY	TOC (8mg/kg IV)	194/276 (70.3%)	110/276 (40.2%)	70/276 (25.4%)	-3.21 (0.08)	-0.55 (0.03)

Continued on next page

Table A14: Study specific data

Trial ID	Treatment	ACR 20 (%)	ACR 50 (%)	ACR 70 (%)	ΔDAS28 (SE)	ΔHAQ (SE)
ACT-RAY	TOC (8mg/kg IV) + cDMARD	198/277 (71.5%)	126/277 (45.5%)	67/277 (24.5%)	-3.43 (0.08)	-0.56 (0.04)
ADACTA	ADA (40mg SC)	80/162 (49.4%)	45/162 (27.8%)	29/162 (17.9%)	-1.8 (0.11)	-0.5 (0.09)
ADACTA	TOC (8mg/kg IV)	106/163 (65%)	77/163 (47.2%)	53/163 (32.5%)	-3.3 (0.11)	-0.7 (0.09)
AIM	ABT (10mg/kg IV) + cDMARD	287/424 (67.9%)	169/424 (39.9%)	83/424 (19.8%)	-2.38 (0.06)	-0.62 (0.03)
AIM	cDMARD	84/214 (39.7%)	35/214 (16.8%)	13/214 (6.5%)	-1.29 (0.09)	-0.52 (0.05)
AMPLE	ABT (125mg SC) + cDMARD	210/318 (66.2%)	128/318 (40.5%)	67/318 (21.2%)	-2.06 (0.08)	
AMPLE	ADA (40mg SC) + cDMARD	217/328 (66.2%)	132/328 (40.5%)	76/328 (23.2%)	-2.12 (0.07)	
ARMADA	cDMARD	9/62 (14.5%)	5/62 (8.1%)	3/62 (4.8%)		-0.27 (0.07)
ARMADA	ADA (20mg SC) + cDMARD	33/69 (47.8%)	22/69 (31.9%)	7/69 (10.1%)		-0.54 (0.07)
ARMADA	ADA (40mg SC) + cDMARD	45/67 (67.2%)	37/67 (55.2%)	18/67 (26.9%)		-0.62 (0.08)
ARMADA	ADA (80mg SC) + cDMARD	48/73 (65.8%)	31/73 (42.5%)	14/73 (19.2%)		-0.59 (0.06)
ATTEST	cDMARD	45/110 (41.8%)	22/110 (20%)	10/110 (9.1%)	-1.48 (0.15)	-0.31 (0.06)
ATTEST	ABT (10mg/kg IV) + cDMARD	104/156 (66.7%)	63/156 (40.4%)	31/156 (20.5%)	-2.53 (0.12)	-0.69 (0.05)
ATTEST	IFX (3mg/kg IV Q8WEEK) + cDMARD	98/165 (59.4%)	61/165 (37%)	39/165 (24.2%)	-2.25 (0.12)	-0.61 (0.05)
ATTRACT	cDMARD	20/88 (22.9%)	NA/88 (NA%)	NA/88 (NA%)		
ATTRACT	IFX (10mg/kg IV Q4WEEK) + cDMARD	39/87 (44.9%)	NA/87 (NA%)	NA/87 (NA%)		
ATTRACT	IFX (10mg/kg IV Q8WEEK) + cDMARD	43/81 (53.8%)	NA/81 (NA%)	NA/81 (NA%)		
ATTRACT	IFX (3mg/kg IV Q4WEEK) + cDMARD	49/86 (57.5%)	NA/86 (NA%)	NA/86 (NA%)		
ATTRACT	IFX (3mg/kg IV Q8WEEK) + cDMARD	46/86 (53.7%)	NA/86 (NA%)	NA/86 (NA%)		
Bao 2011	cDMARD	2/12 (17%)	0/12 (0%)	0/12 (0%)	-1.28 (0.23)	
Bao 2011	ANA (80mg SC) + cDMARD	27/42 (64%)	15/42 (38%)	7/42 (17%)	-1.69 (0.3)	

Continued on next page

Table A14: Study specific data

Trial ID	Treatment	ACR 20 (%)	ACR 50 (%)	ACR 70 (%)	ΔDAS28 (SE)	ΔHAQ (SE)
CHANGE	ADA (20mg SC Q2WEEK)	25/87 (28.7%)	14/87 (16.1%)	9/87 (10.3%)		-0.2 (0.05)
CHANGE	ADA (40mg SC)	40/91 (44%)	22/91 (24.2%)	11/91 (12.1%)		-0.2 (0.06)
CHANGE	ADA (80mg SC)	44/87 (50.6%)	28/87 (32.2%)	13/87 (14.9%)		-0.4 (0.06)
CHANGE	Placebo	12/87 (13.8%)	5/87 (5.7%)	1/87 (1.1%)		0.1 (0.06)
Choy 2012	cDMARD	27/119 (22.9%)	7/119 (5.9%)	2/119 (1.7%)		-0.09 (0.1)
Choy 2012	CTZ (400mg SC) + cDMARD	56/124 (45.9%)	22/124 (18%)	0/124 (0%)		-0.32 (0.1)
Cohen 2002	cDMARD	11/48 (23%)	1/48 (4%)	0/48 (0%)		-0.15 (0.18)
Cohen 2002	ANA (0.04mg/kg SC) + cDMARD	11/63 (19%)	8/63 (13%)	3/63 (5%)		
Cohen 2002	ANA (0.1mg/kg SC) + cDMARD	13/46 (30%)	9/46 (20%)	3/46 (7%)		
Cohen 2002	ANA (0.4mg/kg SC) + cDMARD	19/55 (36%)	6/55 (11%)	1/55 (2%)		
Cohen 2002	ANA (1mg/kg SC) + cDMARD	24/59 (42%)	14/59 (24%)	5/59 (10%)		-0.37 (0.17)
Cohen 2002	ANA (2mg/kg SC) + cDMARD	16/46 (35%)	7/46 (17%)	3/46 (7%)		-0.51 (0.2)
Cohen 2004	cDMARD	55/251 (22%)	20/251 (8%)	5/251 (2%)		-0.18 (0.03)
Cohen 2004	ANA (100mg SC) + cDMARD	95/250 (38%)	42/250 (17%)	15/250 (6%)		-0.29 (0.03)
Cohen 2018	IFX-Pfizer (3mg/kg IV) + cDMARD					-0.595 (0.06)
Cohen 2018	IFX (3mg/kg IV Q8WEEK) + cDMARD					-0.571 (0.06)
De Filippis 2006	ETN (50mg SC) + cDMARD	9/15 (60%)	3/15 (25.8%)	NA/15 (NA%)		
De Filippis 2006	IFX (3mg/kg IV Q8WEEK) + cDMARD	9/15 (60%)	4/15 (32.4%)	NA/15 (NA%)		
DE019	cDMARD	59/200 (29.5%)	19/200 (9.5%)	5/200 (2.5%)		-0.24 (0.04)
DE019	ADA (20mg SC) + cDMARD	129/212 (60.8%)	87/212 (41%)	37/212 (17.5%)		-0.6 (0.04)
DE019	ADA (40mg SC) + cDMARD	131/207 (63.3%)	81/207 (39.1%)	43/207 (20.8%)		-0.56 (0.04)
Edwards 2004	cDMARD	15/40 (38%)	5/40 (13%)	2/40 (5%)	-1.3 (0.19)	
Edwards 2004	RTX (1000mg IV)	26/40 (65%)	13/40 (33%)	6/40 (15%)	-2.2 (0.22)	

Continued on next page

Table A14: Study specific data

Trial ID	Treatment	ACR 20 (%)	ACR 50 (%)	ACR 70 (%)	ΔDAS28 (SE)	ΔHAQ (SE)
Edwards 2004	RTX (1000mg IV) + cDMARD	29/40 (73%)	17/40 (43%)	9/40 (23%)	-2.6 (0.21)	
Emery 2017	ETN-SB4 (50mg SC) + cDMARD	220/299 (73.8%)	128/299 (43%)	69/299 (23.2%)	-2.6 (0.08)	
Emery 2017	ETN (50mg SC) + cDMARD	213/297 (71.7%)	116/297 (39.1%)	59/297 (19.9%)	-2.5 (0.08)	
EQUIRA	ETN-GP2015 (50mg SC) + cDMARD	147/168 (88%)	107/168 (64.1%)	56/168 (33.5%)	-2.78 (0.1)	-0.57 (0.09)
EQUIRA	ETN (50mg SC) + cDMARD	143/155 (92.9%)	110/155 (71%)	66/155 (42.6%)	-2.78 (0.11)	-0.67 (0.09)
ETN Study 309	cDMARD	14/50 (28%)	7/50 (14%)	1/50 (2%)	-0.8 (0.19)	
ETN Study 309	ETN (50mg SC)	76/103 (73.8%)	47/103 (46.6%)	22/103 (21.4%)	-2.38 (0.13)	
ETN Study 309	ETN (50mg SC) + cDMARD	74/100 (74%)	52/100 (52%)	25/100 (25%)	-2.48 (0.14)	
FAST4WARD	CTZ (400mg SC)	50/111 (45.5%)	25/111 (22.7%)	6/111 (5.5%)		-0.36 (0.11)
FAST4WARD	Placebo	10/109 (9.3%)	4/109 (3.7%)	0/109 (0%)		0.13 (0.11)
Fleischmann 2012	Placebo	14/59 (25.4%)	6/59 (10.2%)	4/59 (6.8%)	-1.43 (0.18)	-0.37 (0.16)
Fleischmann 2012	TOF (10mg PO)	39/61 (65.5%)	27/61 (44.3%)	22/61 (37.7%)	-2.85 (0.17)	-0.72 (0.15)
Fleischmann 2012	TOF (15mg PO)	38/57 (66.7%)	31/57 (54.4%)	18/57 (33.3%)	-2.83 (0.18)	-0.82 (0.15)
Fleischmann 2012	TOF (1mg PO)	13/54 (24.1%)	3/54 (7.4%)	3/54 (5.6%)	-1.04 (0.18)	
Fleischmann 2012	TOF (3mg PO)	19/51 (37.3%)	14/51 (27.5%)	6/51 (13.7%)	-2.02 (0.19)	
Fleischmann 2012	TOF (5mg PO)	24/49 (51%)	17/49 (34.7%)	9/49 (20.4%)	-2.35 (0.19)	
Fleischmann 2018	ADA-Pfizer (40mg SC) + cDMARD	251/297 (84.6%)	180/297 (60.9%)	89/297 (30.3%)	-2.77 (0.08)	-0.654 (0.06)
Fleischmann 2018	ADA (40mg SC) + cDMARD	236/300 (78.7%)	167/300 (55.7%)	94/300 (31.4%)	-2.85 (0.08)	-0.674 (0.06)
GO-FORTH	cDMARD	29/88 (33%)	13/88 (14.8%)	5/88 (5.7%)	-0.6 (0.15)	-0.03 (0.06)
GO-FORTH	GOL (100mg SC) + cDMARD	65/87 (74.7%)	42/87 (48.3%)	19/87 (21.8%)	-2.04 (0.12)	-0.45 (0.05)
GO-FORTH	GOL (50mg SC) + cDMARD	61/86 (70.9%)	36/86 (41.9%)	23/86 (26.7%)	-2.05 (0.13)	-0.33 (0.05)
GO-FORWARD	cDMARD	37/133 (27.8%)	18/133 (13.5%)	7/133 (5.3%)		-0.13 (0.05)
GO-FORWARD	GOL (100mg SC)	47/133 (35.3%)	26/133 (19.5%)	15/133 (11.3%)		-0.24 (0.06)
GO-FORWARD	GOL (100mg SC) + cDMARD	53/89 (59.6%)	29/89 (32.6%)	13/89 (14.6%)		-0.45 (0.06)
GO-FORWARD	GOL (50mg SC) + cDMARD	53/89 (59.6%)	33/89 (37.1%)	18/89 (20.2%)		-0.47 (0.06)
GO-FURTHER	cDMARD	62/197 (31.6%)	26/197 (13.2%)	8/197 (4.1%)	-0.8 (0.1)	-0.21 (0.04)

Continued on next page

Table A14: Study specific data

Trial ID	Treatment	ACR 20 (%)	ACR 50 (%)	ACR 70 (%)	ΔDAS28 (SE)	ΔHAQ (SE)
GO-FURTHER	GOL (2mg/kg IV) + cDMARD	255/395 (64.6%)	138/395 (34.9%)	69/395 (17.7%)	-2 (0.07)	-0.53 (0.03)
GO-LIVE	cDMARD	32/129 (24.8%)	12/129 (9.3%)	4/129 (3.1%)		
GO-LIVE	GOL (2/4mg/kg IV)	67/257 (26.1%)	26/257 (10.1%)	12/257 (4.7%)		
GO-LIVE	GOL (2/4mg/kg IV) + cDMARD	112/257 (43.6%)	56/257 (21.8%)	18/257 (7%)		
GO-LIVE	GOL (2mg/kg IV)	29/128 (22.7%)	11/128 (8.6%)	4/128 (3.1%)		
GO-LIVE	GOL (2mg/kg IV) + cDMARD	48/129 (37.2%)	24/129 (18.6%)	8/129 (6.2%)		
GO-LIVE	GOL (4mg/kg IV)	38/129 (29.5%)	15/129 (11.6%)	8/129 (6.2%)		
GO-LIVE	GOL (4mg/kg IV) + cDMARD	64/128 (50%)	32/128 (25%)	10/128 (7.8%)		
HIKARI	cDMARD	12/114 (11.4%)	6/114 (6.1%)	1/114 (0.9%)	-0.21 (0.12)	0.12 (0.05)
HIKARI	CTZ (200mg SC) + cDMARD	74/116 (63.8%)	54/116 (46.6%)	30/116 (25.9%)	-2.06 (0.12)	-0.48 (0.05)
Iwahashi 2014	ABT (10mg/kg IV) + cDMARD	49/59 (83.1%)	36/59 (62.7%)	17/59 (30.5%)	-2.75 (0.18)	-0.61 (0.15)
Iwahashi 2014	ABT (125mg SC) + cDMARD	53/59 (91.5%)	38/59 (66.1%)	22/59 (37.3%)	-2.97 (0.18)	-0.62 (0.15)
J-RAPID	cDMARD	19/77 (24.7%)	13/77 (16.9%)	1/77 (1.3%)	-0.63 (0.15)	-0.18 (0.06)
J-RAPID	CTZ (100mg SC) + cDMARD	43/72 (61.1%)	31/72 (44.4%)	19/72 (26.4%)	-2.11 (0.16)	-0.43 (0.06)
J-RAPID	CTZ (200mg SC) + cDMARD	60/82 (73.2%)	45/82 (54.9%)	24/82 (29.3%)	-2.46 (0.15)	-0.55 (0.05)
J-RAPID	CTZ (400mg SC) + cDMARD	61/85 (71.8%)	45/85 (54.1%)	26/85 (30.6%)	-2.69 (0.14)	-0.57 (0.05)
Jamshidi 2017	ADA-Cinnora (40mg SC) + cDMARD	62/68 (92%)	52/68 (77%)	31/68 (47%)	-2.93 (0.16)	
Jamshidi 2017	ADA (40mg SC) + cDMARD	60/68 (89%)	51/68 (75%)	36/68 (53%)	-2.92 (0.16)	
JESMR	ETN (50mg SC)	44/69 (63.8%)	32/69 (47.8%)	18/69 (26.1%)		
JESMR	ETN (50mg SC) + cDMARD	65/73 (90.4%)	47/73 (64.4%)	28/73 (38.4%)		
Kim 2007	cDMARD	23/63 (36.5%)	9/63 (14.3%)	4/63 (7.9%)		-0.2 (0.06)
Kim 2007	ADA (40mg SC)	40/65 (61.5%)	28/65 (43.1%)	13/65 (21.5%)		-0.5 (0.07)

Continued on next page

Table A14: Study specific data

Trial ID	Treatment	ACR 20 (%)	ACR 50 (%)	ACR 70 (%)	ΔDAS28 (SE)	ΔHAQ (SE)
Kremer 2003	cDMARD	42/119 (35.3%)	14/119 (11.8%)	2/119 (1.7%)		
Kremer 2003	ABT (10mg/kg IV) + cDMARD	69/115 (60%)	41/115 (36.5%)	18/115 (16.5%)		
Kremer 2003	ABT (2mg/kg IV) + cDMARD	43/105 (41.9%)	24/105 (22.9%)	11/105 (10.5%)		
Kremer 2012	cDMARD	23/69 (34.4%)	14/69 (21.2%)	7/69 (10.3%)		
Kremer 2012	TOF (10mg PO) + cDMARD	40/74 (55%)	24/74 (33.1%)	15/74 (21.1%)		
Kremer 2012	TOF (15mg PO) + cDMARD	43/75 (58.6%)	31/75 (42%)	26/75 (34.7%)		
Kremer 2012	TOF (1mg PO) + cDMARD	28/70 (41.4%)	20/70 (29.6%)	16/70 (23.4%)		
Kremer 2012	TOF (20mg PO) + cDMARD	42/80 (52.7%)	30/80 (38.2%)	18/80 (23.3%)		
Kremer 2012	TOF (3mg PO) + cDMARD	35/68 (52.8%)	17/68 (26.2%)	15/68 (22.6%)		
Kremer 2012	TOF (5mg PO) + cDMARD	33/71 (47.6%)	22/71 (32.2%)	16/71 (23.5%)		
LARA	cDMARD	71/142 (50%)	33/142 (23.2%)	16/142 (11.3%)	-1.7 (0.12)	-0.5 (0.1)
LARA	ETN (50mg SC) + cDMARD	232/279 (83.2%)	173/279 (62%)	97/279 (34.8%)	-3.2 (0.09)	-0.9 (0.1)
Li 2016	cDMARD	21/132 (15.9%)	9/132 (6.8%)	2/132 (1.5%)		0.15 (0.06)
Li 2016	GOL (50mg SC) + cDMARD	56/132 (42.4%)	25/132 (18.9%)	8/132 (6.1%)		-0.26 (0.05)
LITHE	cDMARD	106/393 (27%)	38/393 (9.7%)	7/393 (2%)	-1.49 (0.09)	-0.32 (0.03)
LITHE	TOC (4mg/kg IV) + cDMARD	201/399 (50.6%)	100/399 (25.1%)	43/399 (11%)	-2.45 (0.08)	-0.45 (0.03)
LITHE	TOC (8mg/kg IV) + cDMARD	224/398 (56.3%)	128/398 (32.2%)	50/398 (12.6%)	-3.28 (0.08)	-0.51 (0.03)
Matsubara 2018	cDMARD	47/202 (23.3%)	27/202 (13.4%)	10/202 (5%)	-0.48 (0.1)	
Matsubara 2018	ABT (10mg/kg IV) + cDMARD	145/203 (71.9%)	110/203 (54.2%)	68/203 (33.5%)	-2.26 (0.09)	
MOBILITY	cDMARD	133/398 (33.4%)	67/398 (17%)	27/398 (7%)		-0.32 (0.03)
MOBILITY	SAR (150mg SC) + cDMARD	232/400 (58%)	148/400 (37%)	80/400 (20%)		-0.56 (0.03)
MOBILITY	SAR (200mg SC) + cDMARD	265/399 (66.4%)	183/399 (46%)	99/399 (25%)		-0.57 (0.03)
MONARCH	ADA (40mg SC)	108/185 (58.4%)	55/185 (29.7%)	22/185 (11.9%)	-1.97 (0.09)	-0.43 (0.05)
MONARCH	SAR (200mg SC)	132/184 (71.7%)	84/184 (45.7%)	43/184 (23.4%)	-2.86 (0.09)	-0.61 (0.05)
Moreland 1999	ETN (10mg SC)	38/76 (51%)	18/76 (24%)	6/76 (9%)		-0.696 (0.13)

Continued on next page

Table A14: Study specific data

Trial ID	Treatment	ACR 20 (%)	ACR 50 (%)	ACR 70 (%)	ΔDAS28 (SE)	ΔHAQ (SE)
Moreland 1999	ETN (50mg SC)	46/78 (59%)	31/78 (40%)	11/78 (15%)		-0.651 (0.13)
Moreland 1999	Placebo	8/80 (11%)	4/80 (5%)	0/80 (1%)		-0.132 (0.12)
Niu 2011	cDMARD	6/12 (54.5%)	1/12 (9.1%)	1/12 (9.1%)		
Niu 2011	ANA (80mg SC) + cDMARD	25/38 (66.7%)	17/38 (47.2%)	13/38 (36.1%)		
OPTION	cDMARD	54/204 (26%)	22/204 (11%)	4/204 (2%)	-1.52 (0.12)	-0.34 (0.11)
OPTION	TOC (4mg/kg IV) + cDMARD	102/213 (48%)	67/213 (31%)	26/213 (12%)	-2.74 (0.11)	-0.52 (0.1)
OPTION	TOC (8mg/kg IV) + cDMARD	120/205 (59%)	90/205 (44%)	45/205 (22%)	-3.4 (0.1)	-0.55 (0.09)
ORAL-SCAN	cDMARD					-0.25 (0.2)
ORAL-SCAN	TOF (10mg PO) + cDMARD					-0.62 (0.11)
ORAL-SCAN	TOF (5mg PO) + cDMARD					-0.56 (0.13)
ORAL-STANDARD	cDMARD	30/106 (28.3%)	12/106 (12.26%)	2/106 (1.89%)		
ORAL-STANDARD	ADA (40mg SC) + cDMARD	94/199 (47.2%)	55/199 (27.64%)	18/199 (9.05%)		-0.52 (0.08)
ORAL-STANDARD	TOF (10mg PO) + cDMARD	103/196 (52.6%)	67/196 (34.69%)	41/196 (21.04%)		-0.61 (0.08)
ORAL-STANDARD	TOF (5mg PO) + cDMARD	101/196 (51.5%)	71/196 (36.73%)	39/196 (19.9%)		-0.58 (0.08)
ORAL-STRATEGY	ADA (40mg SC) + cDMARD	274/386 (71%)	169/386 (44%)	80/386 (21%)	-2.51 (0.07)	-0.5 (0.03)
ORAL-STRATEGY	TOF (5mg PO)	249/384 (65%)	147/384 (38%)	70/384 (18%)	-2.11 (0.07)	-0.5 (0.03)
ORAL-STRATEGY	TOF (5mg PO) + cDMARD	275/376 (73%)	173/376 (46%)	94/376 (25%)	-2.31 (0.07)	-0.6 (0.03)
RA-BEAM	cDMARD	179/488 (37%)	94/488 (19%)	39/488 (8%)	-1.13 (0.06)	-0.35 (0.56)
RA-BEAM	ADA (40mg SC) + cDMARD	219/330 (66%)	150/330 (45%)	72/330 (22%)	-2.27 (0.08)	-0.63 (0.61)
RA-BEAM	BCT (4mg PO) + cDMARD	360/487 (74%)	246/487 (51%)	145/487 (30%)	-2.53 (0.06)	-0.75 (0.65)
RA-BUILD-A	cDMARD	48/109 (44%)	22/109 (20%)	NA/109 (NA%)		
RA-BUILD-A	BCT (2mg PO) + cDMARD	72/111 (65%)	47/111 (42%)	NA/111 (NA%)		
RA-BUILD-A	BCT (4mg PO) + cDMARD	76/114 (67%)	48/114 (42%)	NA/114 (NA%)		
RA-BUILD-B	BCT (2mg PO)	9/18 (50%)	7/18 (39%)	NA/18 (NA%)		
RA-BUILD-B	BCT (4mg PO)	7/13 (54%)	5/13 (38%)	NA/13 (NA%)		
RA-BUILD-B	Placebo	2/17 (12%)	2/17 (12%)	NA/17 (NA%)		
RA-SCORE	cDMARD	18/63 (28.6%)	6/63 (11.1%)	1/63 (1.6%)	-0.85 (0.17)	-0.19 (0.14)

Continued on next page

Table A14: Study specific data

Trial ID	Treatment	ACR 20 (%)	ACR 50 (%)	ACR 70 (%)	ΔDAS28 (SE)	ΔHAQ (SE)
RA-SCORE	RTX (1000mg IV) + cDMARD	31/60 (51.7%)	16/60 (26.7%)	4/60 (8.3%)	-1.64 (0.17)	-0.44 (0.14)
RA-SCORE	RTX (500mg IV) + cDMARD	31/62 (51.6%)	15/62 (24.2%)	7/62 (11.3%)	-1.69 (0.17)	-0.425 (0.14)
RACAT	ETN (50mg SC) + cDMARD	90/163 (55.2%)	58/163 (35.6%)	26/163 (16%)	-2.06 (0.11)	-0.51 (0.07)
RACAT	SSZ + HCQ + MTX	89/159 (56%)	41/159 (25.8%)	8/159 (5%)	-1.79 (0.1)	-0.44 (0.06)
RAPID-1	cDMARD	27/199 (13.6%)	15/199 (7.6%)	5/199 (3%)		
RAPID-1	CTZ (200mg SC) + cDMARD	228/393 (58.8%)	145/393 (37.1%)	84/393 (21.4%)		
RAPID-1	CTZ (400mg SC) + cDMARD	236/390 (60.8%)	155/390 (39.9%)	80/390 (20.6%)		
RAPID-2	cDMARD	11/127 (8.7%)	3/127 (3.1%)	1/127 (0.8%)	-0.5 (0.09)	-0.14 (0.04)
RAPID-2	CTZ (200mg SC) + cDMARD	140/246 (57.3%)	79/246 (32.5%)	39/246 (15.9%)	-2.27 (0.09)	-0.5 (0.03)
RAPID-2	CTZ (400mg SC) + cDMARD	141/246 (57.6%)	81/246 (33.1%)	26/246 (10.6%)	-2.46 (0.08)	-0.5 (0.03)
SATORI	cDMARD	16/64 (25%)	6/64 (10.9%)	4/64 (6.3%)		-0.434 (0.14)
SATORI	TOC (8mg/kg IV)	48/61 (80.3%)	30/61 (49.2%)	17/61 (29.5%)		-0.621 (0.14)
SELECT-NEXT	cDMARD	79/221 (36%)	33/221 (15%)	13/221 (6%)	-1.02 (0.09)	-0.26 (0.08)
SELECT-NEXT	UPA (15mg PO) + cDMARD	141/221 (64%)	83/221 (38%)	46/221 (21%)	-2.25 (0.09)	-0.61 (0.08)
SELECT-NEXT	UPA (30mg PO) + cDMARD	145/219 (66%)	94/219 (43%)	59/219 (27%)	-2.38 (0.09)	-0.55 (0.08)
SERENE	cDMARD	40/172 (23.3%)	15/172 (9.3%)	8/172 (5.2%)	-0.75 (0.1)	
SERENE	RTX (1000mg IV) + cDMARD	86/170 (50.6%)	44/170 (25.9%)	17/170 (10%)	-1.69 (0.1)	
SERENE	RTX (500mg IV) + cDMARD	91/167 (54.5%)	43/167 (26.3%)	15/167 (9%)	-1.76 (0.1)	
STAR	cDMARD	110/318 (34.9%)	35/318 (11.3%)	11/318 (3.5%)		
STAR	ADA (40mg SC) + cDMARD	167/318 (52.8%)	91/318 (28.9%)	47/318 (14.8%)		
START	cDMARD	87/361 (25.5%)	33/361 (9.7%)	16/361 (4.7%)		
START	IFX (10mg/kg IV) + cDMARD	205/361 (61%)	119/361 (35.4%)	54/361 (16.1%)		
START	IFX (3mg/kg IV Q8WEEK) + cDMARD	199/360 (58%)	110/360 (32.1%)	48/360 (14%)		
SURPRISE	TOC (8mg/kg IV)	77/111 (69.4%)	60/111 (54.1%)	37/111 (34.2%)	-2.7 (0.14)	-0.4 (0.06)

Continued on next page

Table A14: Study specific data

Trial ID	Treatment	ACR 20 (%)	ACR 50 (%)	ACR 70 (%)	ΔDAS28 (SE)	ΔHAQ (SE)
SURPRISE	TOC (8mg/kg IV) + cDMARD	86/115 (74.8%)	63/115 (54.8%)	37/115 (33%)	-2.9 (0.12)	-0.4 (0.05)
Takeuchi 2013a	ABT (10mg/kg IV) + cDMARD	47/61 (77%)	28/61 (45.9%)	13/61 (21.3%)	-2.5 (0.17)	-0.53 (0.14)
Takeuchi 2013a	ABT (2mg/kg IV) + cDMARD	42/67 (62.7%)	25/67 (37.3%)	11/67 (16.4%)	-1.8 (0.17)	-0.34 (0.14)
Takeuchi 2013a	cDMARD	14/66 (21.2%)	4/66 (6.1%)	0/66 (0%)	-0.7 (0.17)	-0.1 (0.14)
TEMPO	cDMARD	167/228 (73.5%)	91/228 (40.1%)	34/228 (15.2%)		-0.631 (0.07)
TEMPO	ETN (50mg SC)	182/223 (82%)	130/223 (58.5%)	79/223 (35.5%)		-0.688 (0.07)
TEMPO	ETN (50mg SC) + cDMARD	163/231 (70.6%)	92/231 (40.1%)	38/231 (16.6%)		-0.893 (0.07)
TOWARD	cDMARD	101/413 (24.5%)	37/413 (9%)	11/413 (2.9%)	-1.16 (0.07)	-0.2 (0.05)
TOWARD	TOC (8mg/kg IV) + cDMARD	488/803 (60.8%)	301/803 (37.6%)	164/803 (20.5%)	-3.17 (0.05)	-0.5 (0.04)
van de Putte 2004	ADA (20mg SC Q2WEEK)	34/106 (32.5%)	16/106 (15.8%)	8/106 (8.31%)	-1.3 (0.16)	-0.29 (0.06)
van de Putte 2004	ADA (20mg SC QWEEK)	41/112 (37%)	22/112 (20.2%)	10/112 (9.73%)	-1.6 (0.16)	-0.39 (0.06)
van de Putte 2004	ADA (40mg SC QWEEK)	56/103 (55%)	34/103 (33.8%)	18/103 (18.2%)	-2 (0.16)	-0.49 (0.05)
van de Putte 2004	ADA (40mg SC)	48/113 (43.2%)	25/113 (22.9%)	13/113 (12.3%)	-1.7 (0.15)	-0.38 (0.06)
van de Putte 2004	Placebo	21/110 (19.8%)	9/110 (8.89%)	1/110 (1.8%)	-0.7 (0.12)	-0.07 (0.05)
Weinblatt 1999	cDMARD	8/30 (27%)	0/30 (3%)	0/30 (0%)		
Weinblatt 1999	ETN (50mg SC)	41/59 (71%)	23/59 (39%)	8/59 (15%)		
Weinblatt 2018	ADA-SB5 (40mg SC) + cDMARD	183/269 (68%)	98/269 (36.4%)	47/269 (17.5%)	-2.74 (0.08)	
Weinblatt 2018	ADA (40mg SC) + cDMARD	184/273 (67.4%)	100/273 (36.6%)	50/273 (18.3%)	-2.68 (0.08)	

Note: ΔDAS28 and ΔHAQ denote differences between the end of the trial and baseline.

I.4 Comparing the IVI network meta-analysis to the NICE network meta-analysis

To help ensure that differences in cost-effectiveness estimates from our model relative to others are not driven by the NMA results, we compared our NMA estimates to estimates reported by NICE in Stevenson et al. (2016). We focus on ACR response, since the NICE report and other models use treatment pathways similar to **H1** and **H2** and rarely use DAS28 to inform treatment duration. As shown in Table A15, our results are similar and the NICE point estimates are generally within the 95% credible intervals surrounding our point estimates.

Table A15: A comparison of NICE and IVI estimates of ACR response probabilities

	IVI			NICE		
	ACR20	ACR50	ACR70	ACR20	ACR50	ACR70
cDMARDs	0.291 (0.277, 0.306)	0.120 (0.111, 0.130)	0.040 (0.036, 0.044)	0.298	0.123	0.042
ABT IV + MTX	0.636 (0.546, 0.720)	0.394 (0.306, 0.485)	0.199 (0.139, 0.269)	0.573	0.328	0.156
ABT SC + MTX	0.632 (0.486, 0.760)	0.392 (0.258, 0.537)	0.200 (0.109, 0.311)	0.638	0.391	0.199
ADA + MTX	0.588 (0.495, 0.669)	0.346 (0.263, 0.426)	0.166 (0.113, 0.222)	0.615	0.368	0.183
ADA	0.501 (0.334, 0.645)	0.271 (0.145, 0.399)	0.120 (0.052, 0.202)	0.499	0.264	0.115
ADA BWWD + MTX	0.585 (0.369, 0.791)	0.352 (0.169, 0.574)	0.175 (0.063, 0.347)	-	-	-
ANA + MTX	0.460 (0.243, 0.683)	0.243 (0.092, 0.440)	0.105 (0.028, 0.234)	-	-	-
BCT	0.599 (0.172, 0.924)	0.389 (0.059, 0.794)	0.218 (0.016, 0.590)	-	-	-
BCT + MTX	0.554 (0.345, 0.760)	0.321 (0.154, 0.535)	0.153 (0.055, 0.308)	-	-	-
CZP	0.581 (0.286, 0.832)	0.355 (0.116, 0.634)	0.181 (0.038, 0.406)	-	-	-
CZP + MTX	0.737 (0.639, 0.821)	0.507 (0.394, 0.616)	0.289 (0.198, 0.390)	0.564	0.319	0.150
ETN	0.598 (0.493, 0.706)	0.356 (0.257, 0.469)	0.173 (0.109, 0.256)	0.645	0.398	0.205
ETN + MTX	0.584 (0.466, 0.690)	0.343 (0.240, 0.453)	0.165 (0.100, 0.242)	0.713	0.472	0.263
ETN SZZS + MTX	0.499 (0.263, 0.742)	0.276 (0.104, 0.511)	0.126 (0.033, 0.294)	-	-	-
ETN YKRO + MTX	0.612 (0.379, 0.820)	0.378 (0.176, 0.618)	0.194 (0.065, 0.390)	-	-	-
GOL + MTX	0.615 (0.482, 0.744)	0.375 (0.252, 0.513)	0.187 (0.106, 0.292)	0.642	0.395	0.202
IFX + MTX	0.585 (0.481, 0.701)	0.344 (0.253, 0.460)	0.165 (0.107, 0.253)	0.595	0.348	0.169
IFX QBTX + MTX	-	-	-	-	-	-
Placebo	0.183 (0.088, 0.299)	0.065 (0.024, 0.125)	0.019 (0.005, 0.042)	0.175	0.059	0.016
RTX	0.486 (0.276, 0.713)	0.264 (0.113, 0.477)	0.118 (0.036, 0.261)	-	-	-
RTX + MTX	0.560 (0.422, 0.704)	0.323 (0.205, 0.466)	0.152 (0.080, 0.252)	0.573	0.328	0.156
SAR	0.645 (0.373, 0.851)	0.415 (0.175, 0.664)	0.223 (0.064, 0.440)	-	-	-
SAR + MTX	0.617 (0.423, 0.801)	0.381 (0.206, 0.591)	0.195 (0.080, 0.364)	-	-	-
SSZ + HCQ + MTX	0.519 (0.279, 0.752)	0.294 (0.112, 0.524)	0.138 (0.037, 0.299)	0.503	0.266	0.117
TCZ	0.685 (0.554, 0.798)	0.447 (0.313, 0.584)	0.241 (0.142, 0.358)	0.717	0.477	0.266
TCZ + MTX	0.667 (0.562, 0.761)	0.427 (0.321, 0.535)	0.224 (0.148, 0.313)	0.706	0.464	0.256
TOF + MTX	0.586 (0.453, 0.704)	0.346 (0.229, 0.466)	0.167 (0.093, 0.253)	-	-	-
TOF	0.498 (0.332, 0.684)	0.271 (0.144, 0.441)	0.121 (0.050, 0.235)	-	-	-
UPA + MTX	0.569 (0.369, 0.764)	0.335 (0.168, 0.540)	0.162 (0.062, 0.313)	-	-	-

Notes: ACR20/50/70 categories are the probability of at least a 20/50/70% improvement. 95% credible intervals are in parentheses. IVI estimates are based on 6-month simulations of 1,000 patients and 1,000 parameters sets for each therapy. NICE estimates are from Table 37 in Stevenson et al. (2017). cDMARDs = conventional disease-modifying antirheumatic drugs; MTX = methotrexate; ABT IV = abatacept intravenous; ABT SC = abatacept subcutaneous; ADA = adalimumab; ADA BWWD = adalimumab-bwwd (biosimilar Samsung Bioepis); ANA = anakinra; BCT = baricitinib; CZP = certolizumab pegol; ETN = etanercept; ETN SZZS = etanercept-szzs (biosimilar Sandoz); ETN YKRO = etanercept-ykro (biosimilar Samsung Bioepis); GOL = golimumab; HCQ = hydroxychloroquine sulfate; IFX = infliximab; IFX QBTX = infliximab-qbtx (biosimilar Pfizer); RTX = rituximab; SAR = sarilumab; SSZ = sulfazalazine; TCZ = tocilizumab; TOF = tofacitinib; UPA = upadacitinib; ACR = American College of Rheumatology.

I.5 Excluded publications after full-text screening

Table A16: Publications not meeting the systematic review eligibility criteria; excluded from the evidence base

Author and Year	Title	Journal	Reason	Subreason
Aalbers, 2015	Intra-articular etanercept treatment in inflammatory arthritis: A randomized double-blind placebo-controlled proof of mechanism clinical trial validating tnf as a potential therapeutic target for local treatment	Joint, bone, spine	Outcomes	No outcomes of interest at 24 weeks
Abe, 2006	A multicenter, double-blind, randomized, placebo controlled trial of infliximab combined with low dose methotrexate in japanese patients with rheumatoid arthritis	Journal of rheumatology	Outcomes	No outcomes of interest at 24 weeks
Abu-Zaid, 2018	The effectiveness of etanercept and adalimumab on anemia of chronic disease and serum hepcidin in patients with rheumatoid arthritis, a comparative study	Egyptian Rheumatologist	Study design	Non-interventional
Allaart, 2007	Treatment of recent-onset rheumatoid arthritis: Lessons from the best study	Journal of rheumatology.	Population	cDMARD nave
Alten, 2018	Abatacept used in combination with non-methotrexate disease-modifying antirheumatic drugs: A descriptive analysis of data from interventional trials and the real-world setting	Arthritis Research	Therapy	Outcomes
No outcomes of interest at 24 weeks				
Alten, 2019	Randomised, double-blind, phase iii study comparing the infliximab biosimilar, pf-06438179/gp1111, with reference infliximab: Efficacy, safety and immunogenicity from week 30 to week 54	Open Heart	Outcomes	No outcomes of interest at 24 weeks
Alten, 2017	Janus kinase inhibitor baricitinib for rheumatoid arthritis : Randomized, double-blind, placebo- and active-controlled, phase 3 study evaluating the efficacy and safety of baricitinib in patients with moderately to severely active rheumatoid arthritis who have had an inadequate response to methotrexate therapy (ra-beam)	Der internist	Other	Language
Anonymous, 2003	Adalimumab (humira) for rheumatoid arthritis	Medical Letter on Drugs	Therapeutics	Study design
Review				
Anonymous, 2010	Tocilizumab (actemra) for rheumatoid arthritis	Medical Letter on Drugs	Therapeutics	Study design
Review				
Anonymous, 2017	Sarilumab (kevzara) for rheumatoid arthritis	Medical Letter on Drugs and Therapeutics	Study design	Review
Anonymous, 2018	Baricitinib (olumiant) for rheumatoid arthritis	Medical Letter on Drugs	Therapeutics	Study design
Review				
Antoni, 1999	Combination therapy of the chimeric monoclonal anti-tumor necrosis factor alpha antibody (infliximab) with methotrexate in patients with rheumatoid arthritis	Clinical and experimental rheumatology	Study design	Review

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Apsangikar, 2018	A prospective, randomized, double-blind, comparative clinical study of efficacy and safety of a biosimilar adalimumab with innovator product in patients with active rheumatoid arthritis on a stable dose of methotrexate	Indian Journal of Rheumatology	Intervention	Not of interest
Bae, 2013	Improved health outcomes with etanercept versus usual dmard therapy in an asian population with established rheumatoid arthritis	BMC musculoskeletal disorders	Outcomes	No outcomes of interest at 24 weeks
Bankhurst, 1999	Etanercept and methotrexate combination therapy	Clinical and experimental rheumatology	Study design	Not of interest
Bao, 2016	Good response to infliximab in rheumatoid arthritis following failure of interleukin-1 receptor antagonist	International journal of rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Bao , 2016	Good response to infliximab in rheumatoid arthritis following failure of interleukin-1 receptor antagonist	International Journal of Rheumatic Diseases	Outcomes	No outcomes of interest at 24 weeks
Bathon, 2000	A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis	New England journal of medicine	Population	cDMARD nave
Bay-Jensen, 2014	Serological biomarkers of joint tissue turnover predict tocilizumab response at baseline	Journal of Clinical Rheumatology	Outcomes	No outcomes of interest
Bazzichi, 2019	Subcutaneous tocilizumab alone or with a csdmard in rheumatoid arthritis patients: Subanalysis of italian data from a multicenter phase iiib/iv trial	Clinical Rheumatology	Study design	Non-comparative post-hoc analysis
Beals, 2017	Magnetic resonance imaging of the hand and wrist in a randomized, double-blind, multicenter, placebo-controlled trial of infliximab for rheumatoid arthritis: Comparison of dynamic contrast enhanced assessments with semi-quantitative scoring	Plos one	Outcomes	No outcomes of interest at 24 weeks
Bingham, 2015	Maintenance of clinical and radiographic benefit with intravenous golimumab therapy in patients with active rheumatoid arthritis despite methotrexate therapy: Week-112 efficacy and safety results of the open-label long-term extension of a phase iii, double-blind, randomized, placebo-controlled trial	Arthritis Care and Research	Outcomes	No outcomes of interest at 24 weeks
Bingham, 2015	Humoral immune response to vaccines in patients with rheumatoid arthritis treated with tocilizumab: Results of a randomised controlled trial (visara)	Annals of the Rheumatic Diseases	Outcomes	No outcomes of interest at 24 weeks
Bobbio-Pallavicini, 2007	High iga rheumatoid factor levels are associated with poor clinical response to tumour necrosis factor alpha inhibitors in rheumatoid arthritis	Annals of the Rheumatic Diseases	Outcomes	No outcomes of interest
Boers, 2001	Demonstration of response in rheumatoid arthritis patients who are nonresponders according to the american college of rheumatology 20Boyle, 2015	The jak inhibitor tofacitinib suppresses synovial jak1-stat signalling in rheumatoid arthritis	Annals of the rheumatic diseases	Outcomes
No outcomes of interest at 24 weeks				

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Breedveld, 2005	Association between baseline radiographic damage and improvement in physical function after treatment of patients with rheumatoid arthritis	Annals of the rheumatic diseases	Outcomes	No outcomes of interest
Bresnihan, 1999	Treatment of rheumatoid arthritis with interleukin 1 receptor antagonist	–	Outcomes	No outcomes of interest at 24 weeks
Bresnihan, 2002	Effects of anakinra on clinical and radiological outcomes in rheumatoid arthritis	–	Study design	Review
Bresnihan , 1998	Treatment of rheumatoid arthritis with recombinant human interleukin-1 receptor antagonist	Arthritis	Rheumatism	Population
cDMARD nave				
Buch, 2019	Can switching to abatacept therapy in patients with rheumatoid arthritis on background methotrexate reverse tnf-inhibitor-induced antinuclear autoantibody/ double-stranded DNA autoantibody conversion? An analysis of the ample and attest trials	Clinical and Experimental Rheumatology	Outcomes	No outcomes of interest
Burmester, 2016	Efficacy and safety of subcutaneous tocilizumab versus intravenous tocilizumab in combination with traditional dmards in patients with ra at week 97 (summacta)	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Burmester, 2017	Tocilizumab combination therapy or monotherapy or methotrexate monotherapy in methotrexate-naive patients with early rheumatoid arthritis: 2-year clinical and radiographic results from the randomised, placebo-controlled function trial	Annals of the rheumatic diseases	Population	cDMARD nave
Bykerk, 2012	Tocilizumab in patients with active rheumatoid arthritis and inadequate responses to dmards and/or tnf inhibitors: A large, open-label study close to clinical practice	Annals of the Rheumatic Diseases	Study design	Non-interventional
Calguneri, 1999	Combination therapy versus monotherapy for the treatment of patients with rheumatoid arthritis	Clinical and experimental rheumatology	Population	cDMARD nave
Campion , 1996	Dose-range and dose-frequency study of recombinant human interleukin-1 receptor antagonist in patients with rheumatoid arthritis. The il-1ra arthritis study group	Arthritis	Rheumatism	Population
cDMARD nave				
Carubbi, 2016	Safety and efficacy of intra-articular anti-tumor necrosis factor alpha agents compared to corticosteroids in a treat-to-target strategy in patients with inflammatory arthritis and monoarthritis flare	International Journal of Immunopathology and Pharmacology	Comparator	Corticosteroids
Cella, 2005	Validation of the functional assessment of chronic illness therapy fatigue scale relative to other instrumentation in patients with rheumatoid arthritis	Journal of Rheumatology	Study design	Non-comparative post-hoc
Charles, 2000	Assessment of antibodies to double-stranded DNA induced in rheumatoid arthritis patients following treatment with infliximab, a monoclonal antibody to tumor necrosis factor alpha: Findings in open-label and randomized placebo-controlled trials	Arthritis	Rheumatism	Outcomes

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
No outcomes of interest				
Charles-Schoeman, 2017	Improvement of high-density lipoprotein function in patients with early rheumatoid arthritis treated with methotrexate monotherapy or combination therapies in a randomized controlled trial	Arthritis	rheumatology	Population
cDMARD nave				
Chen, 2009	Randomized, double-blind, placebo-controlled, comparative study of human anti-tnf antibody adalimumab in combination with methotrexate and methotrexate alone in taiwanese patients with active rheumatoid arthritis	Journal of the Formosan Medical Association	Outcomes	No outcomes of interest at 24 weeks
Chen, 2006	The effect of etanercept on anti-cyclic citrullinated peptide antibodies and rheumatoid factor in patients with rheumatoid arthritis	Annals of the rheumatic diseases	Outcomes	No outcomes of interest
Choe, 2017	A randomised, double-blind, phase iii study comparing sb2, an infliximab biosimilar, to the infliximab reference product remicade in patients with moderate to severe rheumatoid arthritis despite methotrexate therapy	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Choy, 2002	Efficacy of a novel pegylated humanized anti-tnf fragment (cdp870) in patients with rheumatoid arthritis: A phase ii double-blinded, randomized, dose-escalating trial	Rheumatology	Outcomes	No outcomes of interest at 24 weeks
Choy, 2008	Factorial randomised controlled trial of glucocorticoids and combination disease modifying drugs in early rheumatoid arthritis	Annals of the rheumatic diseases	Intervention	Ciclosporin
Cohen, 2004	Secondary addition of methotrexate to partial responders to etanercept alone is effective in severe rheumatoid arthritis	Annals of the Rheumatic Diseases	Outcomes	No outcomes of interest
Cohen, 2016	A phase i pharmacokinetics trial comparing pf-05280586 (a potential biosimilar) and rituximab in patients with active rheumatoid arthritis	British journal of clinical pharmacology	Outcomes	No outcomes of interest
Collison, 2018	Selective inhibition of jak1 shows promise for ra	Nature Reviews Rheumatology	Study design	Review
Combe, 2014	Efficacy and safety of golimumab as add-on therapy to disease-modifying antirheumatic drugs: Results of the go-more study	Annals of the rheumatic diseases	Population	Low disease activity population
Conaghan, 2013	Impact of intravenous abatacept on synovitis, osteitis and structural damage in patients with rheumatoid arthritis and an inadequate response to methotrexate: The asset randomised controlled trial	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Coombs, 2010	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	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Cuomo, 2006	A comparison between the simplified disease activity index (sdai) and the disease activity score (das28) as measure of response to treatment in patients undergoing different therapeutic regimens	Reumatismo	Other	Language

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
de Jong, 2013	Induction therapy with a combination of dmards is better than methotrexate monotherapy: First results of the treach trial	Annals of the rheumatic diseases	Population	cDMARD nave
De Stefano, 2010	Comparison of combination therapies in the treatment of rheumatoid arthritis: Leflunomide-anti-tnf-alpha versus methotrexate-anti-tnf-alpha	Clinical rheumatology	Intervention	Unspecified treatment
Dehoratius, 2018	Satisfaction with subcutaneous golimumab and its autoinjector among rheumatoid arthritis patients with inadequate response to adalimumab or etanercept	The Patient: Patient-Centered Outcomes Research	Outcomes	No outcomes of interest
den Broeder, 2002	A single dose, placebo controlled study of the fully human anti-tumor necrosis factor-alpha antibody adalimumab (d2e7) in patients with rheumatoid arthritis	Journal of rheumatology	Outcomes	No outcomes of interest
Deodhar, 2016	The effect of etanercept on traditional metabolic risk factors for cardiovascular disease in patients with rheumatoid arthritis	Clinical rheumatology	Outcomes	No outcomes of interest
Detert, 2016	Effects of treatment with etanercept versus methotrexate on sleep quality, fatigue and selected immune parameters in patients with active rheumatoid arthritis	Clinical	Experimental Rheumatology	Study design
Non-randomized				
Dischereit, 2013	Infliximab improves bone metabolism and bone mineral density in rheumatoid arthritis and ankylosing spondylitis: A prospective 2-year study	Clinical rheumatology	Study design	Non-randomized
Domanska, 2017	Comparative usability study for a certolizumab pegol autoinjection device in patients with rheumatoid arthritis	Expert opinion on drug delivery	Outcomes	No outcomes of interest
Duan, 2015	Efficacy and safety evaluation of a combination of iguratimod and methotrexate therapy for active rheumatoid arthritis patients: A randomized controlled trial	Clinical rheumatology	Population	cDMARD nave
Durez, 2007	Treatment of early rheumatoid arthritis: A randomized magnetic resonance imaging study comparing the effects of methotrexate alone, methotrexate in combination with infliximab, and methotrexate in combination with intravenous pulse methylprednisolone	Arthritis and rheumatism	Population	cDMARD nave
Egeth, 2017	Patient and healthcare professionals preference for brenzys vs. Enbrel autoinjector for rheumatoid arthritis: A randomized crossover simulated-use study	Advances in therapy	Outcomes	No outcomes of interest
Emery, 2015	Evaluating drug-free remission with abatacept in early rheumatoid arthritis: Results from the phase 3b, multicentre, randomised, active-controlled avert study of 24 months, with a 12-month, double-blind treatment period	Annals of the rheumatic diseases	Population	cDMARD nave
Emery, 2010	Impact of t-cell costimulation modulation in patients with undifferentiated inflammatory arthritis or very early rheumatoid arthritis: A clinical and imaging study of abatacept (the adjust trial)	Annals of the Rheumatic Diseases	Population	Undifferentiated arthritis
Emery, 2006	Treatment of rheumatoid arthritis patients with abatacept and methotrexate significantly improved health-related quality of life	Journal of rheumatology	Outcomes	No outcomes of interest

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Emery, 2017	52-week results of the phase 3 randomized study comparing sb4 with reference etanercept in patients with active rheumatoid arthritis	Rheumatology	Outcomes	No outcomes of interest at 24 weeks
Engvall, 2010	Infliximab therapy increases body fat mass in early rheumatoid arthritis independently of changes in disease activity and levels of leptin and adiponectin: A randomised study over 21 months	Arthritis research	therapy	Population
cDMARD nave				
Eriksson, 2013	Biological vs. Conventional combination treatment and work loss in early rheumatoid arthritis: A randomized trial	JAMA internal medicine	Population	cDMARD nave
Eriksson, 2016	Infliximab versus conventional combination treatment and seven-year work loss in early rheumatoid arthritis: Results of a randomized swedish trial	Arthritis Care and Research	Population	cDMARD nave
Fernandez-Nebro, 2005	Treatment of rheumatic inflammatory disease in 25 patients with secondary amyloidosis using tumor necrosis factor alpha antagonists	American Journal of Medicine	Study design	Non-randomized
Ferraccioli, 2002	Rescue of combination therapy failures using infliximab, while maintaining the combination or monotherapy with methotrexate: Results of an open trial	Rheumatology	Outcomes	No outcomes of interest at 24 weeks
Fleischmann, 2012	Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis	New England journal of medicine	Comparator	No active comparator at 24 weeks (placebo crossover only)
Fleischmann , 2017	Efficacy and safety of tofacitinib monotherapy, tofacitinib with methotrexate, and adalimumab with methotrexate in patients with rheumatoid arthritis (oral strategy): A phase 3b/4, double-blind, head-to-head, randomised controlled trial	Lancet (london, england)	Duplicate publication	–
Fleischmann , 2003	Anakinra, a recombinant human interleukin-1 receptor antagonist (r-methuil-1ra), in patients with rheumatoid arthritis: A large, international, multicenter, placebo-controlled trial	Arthritis and Rheumatism	Population	cDMARD nave
Furst, 2007	Open-label, pilot protocol of patients with rheumatoid arthritis who switch to infliximab after an incomplete response to etanercept: The opposite study	Annals of the rheumatic diseases	Study design	Pilot study
Furst, 2015	Two dosing regimens of certolizumab pegol in patients with active rheumatoid arthritis	Arthritis Care and Research	Intervention	Discontinuation/withdrawal study
Galarraga, 2009	Etanercept improves inflammation-associated arterial stiffness in rheumatoid arthritis	Rheumatology	Outcomes	No outcomes of interest
Gao, 2010	Therapeutic effect of infliximab on moderate and severe active rheumatoid arthritis	Nan fang yi ke da xue xue bao [Journal of Southern Medical University]	Other	Language
Gashi, 2014	Treatment of rheumatoid arthritis with biologic dmards (rituximab and etanercept)	Medical archives	Outcomes	No outcomes of interest
Genovese, 2002	Etanercept versus methotrexate in patients with early rheumatoid arthritis: Two-year radiographic and clinical outcomes	Arthritis and rheumatism	Population	cDMARD nave

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Genovese, 2017	Peficitinib, a jak inhibitor, in combination with limited conventional synthetic disease-modifying antirheumatic drugs in the treatment of moderate-to-severe rheumatoid arthritis	Arthritis	rheumatology	Outcomes
No outcomes of interest at 24 weeks				
Genovese, 2011	Atacicept in patients with rheumatoid arthritis and an inadequate response to tumor necrosis factor antagonist therapy: Results of a phase ii, randomized, placebo-controlled, dose-finding trial	Arthritis and Rheumatism	Intervention	Atacicept
Genovese, 2016	Efficacy and safety of abt-494, a selective jak-1 inhibitor, in a phase iib study in patients with rheumatoid arthritis and an inadequate response to methotrexate	Arthritis	rheumatology	Outcomes
No outcomes of interest at 24 weeks				
Genovese , 2004	Combination therapy with etanercept and anakinra in the treatment of patients with rheumatoid arthritis who have been treated unsuccessfully with methotrexate	Arthritis and Rheumatism	Comparator	Etanacerpt monotherapy vs. etanercept + anakinra
Gerards, 2003	Cyclosporin a monotherapy versus cyclosporin a and methotrexate combination therapy in patients with early rheumatoid arthritis: A double blind randomised placebo controlled trial	Annals of the Rheumatic Diseases	Intervention	Cyclosporin
Gerlag, 2010	Preclinical and clinical investigation of a ccr5 antagonist, azd5672, in patients with rheumatoid arthritis receiving methotrexate	Arthritis and rheumatism	Outcomes	No outcomes of interest at 24 weeks
Gerlag, 2019	Effects of b-cell directed therapy on the preclinical stage of rheumatoid arthritis: The prairi study	Annals of the Rheumatic Diseases	Population	Preclinical RA
Goekoop-Ruiterman, 2005	Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the best study): A randomized, controlled trial	Arthritis and rheumatism	Population	cDMARD nave
Gomez-Garcia, 2013	Reduced numbers of circulating cd28-negative cd4+ cells in patients with rheumatoid arthritis chronically treated with abatacept	International journal of rheumatic diseases	Outcomes	No outcomes of interest
Gonzalez-Juanatey, 2006	Effect of anti-tumor necrosis factor alpha therapy on the progression of subclinical atherosclerosis in severe rheumatoid arthritis	Arthritis	Rheumatism	Study design
Non-randomized				
Gottenberg, 2016	Nontnf-targeted biologic vs a second anti-tnf drug to treat rheumatoid arthritis in patients with insufficient response to a first anti-tnf drug: A randomized clinical trial	Journal of the American Medical Association	Intervention	Unspecified treatment
Guler-Yuksel, 2008	Changes in bone mineral density in patients with recent onset, active rheumatoid arthritis	Annals of the rheumatic diseases	Population	cDMARD nave

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Hara, 2007	Efficacy and safety of iguratimod compared with placebo and salazosulfapyridine in active rheumatoid arthritis: A controlled, multicenter, double-blind, parallel-group study	Modern rheumatology	Population	cDMARD nave
Haraoui, 2011	Safety and effectiveness of rituximab in patients with rheumatoid arthritis following an inadequate response to 1 prior tumor necrosis factor inhibitor: The reset trial	Journal of Rheumatology	Study design	Single-arm
Haugeberg, 2009	Bone loss in patients with active early rheumatoid arthritis: Infliximab and methotrexate compared with methotrexate treatment alone. Explorative analysis from a 12-month randomised, double-blind, placebo-controlled study	Annals of the rheumatic diseases	Population	cDMARD nave
Hazlewood, 2012	Abatacept use after failure of multiple biologic agents in patients with severe rheumatoid arthritis	Journal of Clinical Rheumatology	Study design	Observational
Heath, 2010	Selective depletion of b lymphocytes with rituximab preserves b-cell function	Nature Reviews Endocrinology	Study design	Review
Heimans, 2013	Health-related quality of life and functional ability in patients with early arthritis during remission steered treatment: Results of the improved study	Arthritis Research and Therapy	Population	Undifferentiated/early RA
Huang, 2009	Adalimumab plus methotrexate for the treatment of rheumatoid arthritis: A multi-center randomized, double-blind, placebo-controlled clinical study	Zhonghua nei ke za zhi [Chinese journal of internal medicine]	Other	Language
Huizinga, 2014	Sarilumab, a fully human monoclonal antibody against il-6r[alpha] in patients with rheumatoid arthritis and an inadequate response to methotrexate: Efficacy and safety results from the randomised saril-ra-mobility part a trial	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Huizinga, 2014	Sarilumab, a fully human monoclonal antibody against il-6ralpha in patients with rheumatoid arthritis and an inadequate response to methotrexate: Efficacy and safety results from the randomised saril-ra-mobility part a trial	Annals of the Rheumatic Diseases	Outcomes	No outcomes of interest at 24 weeks
Jani, 2016	A prospective, randomized, double-blind, multicentre, parallel-group, active controlled study to compare efficacy and safety of biosimilar adalimumab (exemptia; zrc-3197) and adalimumab (humira) in patients with rheumatoid arthritis	International journal of rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Jiang , 2000	A multicenter, double-blind, dose-ranging, randomized, placebo-controlled study of recombinant human interleukin-1 receptor antagonist in patients with rheumatoid arthritis: Radiologic progression and correlation of genant and larsen scores	Arthritis	Rheumatism	Outcomes
No outcomes of interest				
Jones, 2010	Comparison of tocilizumab monotherapy versus methotrexate monotherapy in patients with moderate to severe rheumatoid arthritis: The ambition study	Annals of the rheumatic diseases	Population	cDMARD nave
Kaeley, 2018	Similar improvements in patient-reported outcomes among rheumatoid arthritis patients treated with two different doses of methotrexate in combination with adalimumab: Results from the musica trial	Rheumatology and Therapy	Intervention	Dose randomization

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Kaine, 2012	Evaluation of abatacept administered subcutaneously in adults with active rheumatoid arthritis: Impact of withdrawal and reintroduction on immunogenicity, efficacy and safety (phase iiib allow study)	Annals of the rheumatic diseases	Intervention	Discontinuation/withdrawal study
Kastanek, 2002	Using anakinra for adult rheumatoid arthritis	–	Study design	Review
Kastbom, 2007	Fcgamma receptor type iiia genotype and response to tumor necrosis factor alpha-blocking agents in patients with rheumatoid arthritis	Arthritis Rheumatism	Study design	Non-randomized
Kavanaugh, 2008	Assessment of rituximab's immunomodulatory synovial effects (arise trial). 1: Clinical and synovial biomarker results	Annals of the Rheumatic Diseases	Study design	Single-arm trial
Kay, 2008	Golimumab in patients with active rheumatoid arthritis despite treatment with methotrexate: A randomized, double-blind, placebo-controlled, dose-ranging study	Arthritis and rheumatism	Outcomes	No outcomes of interest at 24 weeks
Keystone, 2003	Role of adalimumab in the treatment of early rheumatoid arthritis	Clinical and experimental rheumatology	Population	subanalysis of early RA
Keystone, 2015	Two-year radiographic and clinical outcomes from the canadian methotrexate and etanercept outcome study in patients with rheumatoid arthritis	Rheumatology	Outcomes	No outcomes of interest at 24 weeks
Keystone, 2016	Two-year radiographic and clinical outcomes from the canadian methotrexate and etanercept outcome study in patients with rheumatoid arthritis	Rheumatology	Intervention	Dose tapering
Keystone, 2004	Once-weekly administration of 50 mg etanercept in patients with active rheumatoid arthritis: Results of a multicenter, randomized, double-blind, placebo-controlled trial	Arthritis and rheumatism	Outcomes	No outcomes of interest at 24 weeks
Keystone, 2015	Safety and efficacy of baricitinib at 24 weeks in patients with rheumatoid arthritis who have had an inadequate response to methotrexate	Annals of the rheumatic diseases	Comparator	No active comparator at 24 weeks (placebo crossover only)
Keystone, 2009	Rituximab inhibits structural joint damage in patients with rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitor therapies	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Kim, –	Randomized comparison of etanercept with usual therapy in an asian population with active rheumatoid arthritis: The appeal trial	International Journal of Rheumatic Diseases.	Outcomes	No outcomes of interest at 24 weeks
Kim, 2012	Randomized comparison of etanercept with usual therapy in an asian population with active rheumatoid arthritis: The appeal trial	International journal of rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Kim, 2013	A clinical trial and extension study of infliximab in korean patients with active rheumatoid arthritis despite methotrexate treatment	Journal of korean medical science	Study design	No outcomes of interest at 24 weeks
Kirkham, 2014	Effects of golimumab, an anti-tumour necrosis factor-alpha human monoclonal antibody, on lipids and markers of inflammation	Annals of the Rheumatic Diseases	Outcomes	No outcomes of interest at 24 weeks

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Kivitz, 2018	Usability and patient preference phase 3 study of the sarilumab pen in patients with active moderate-to-severe rheumatoid arthritis	Rheumatology and Therapy	Intervention	Dose randomization
Kivitz, 2006	Clinical assessment of pain, tolerability, and preference of an autoinjection pen versus a prefilled syringe for patient self-administration of the fully human, monoclonal antibody adalimumab: The touch trial	Clinical Therapeutics	Study design	Single-arm trial
Kivitz, 2018	Two-year efficacy and safety of subcutaneous tocilizumab in combination with disease-modifying antirheumatic drugs including escalation to weekly dosing in rheumatoid arthritis	Journal of Rheumatology	Outcomes	No outcomes of interest at 24 weeks
Kosinski, 2002	Health-related quality of life in early rheumatoid arthritis: Impact of disease and treatment response	American journal of managed care	Population	cDMARD nave
Kremer, 2009	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 and rheumatism	Outcomes	No outcomes of interest at 24 weeks
Kremer, 2004	Benefit/risk of leflunomide in rheumatoid arthritis	Clinical and Experimental Rheumatology	Study design	Review
Kremer, 2016	A phase iib study of abt-494, a selective jak-1 inhibitor, in patients with rheumatoid arthritis and an inadequate response to anti-tumor necrosis factor therapy	Arthritis	rheumatology	Outcomes
No outcomes of interest at 24 weeks				
Lan, 2004	A comparative study of etanercept plus methotrexate and methotrexate alone in taiwanese patients with active rheumatoid arthritis: A 12-week, double-blind, randomized, placebo-controlled study	Journal of the Formosan Medical Association	Study design	No outcomes of interest at 24 weeks
Langer, 2003	Kineret: Efficacy and safety in daily clinical practice: An interim analysis of the kineret response assessment initiative (kreative) protocol	International Journal of Clinical Pharmacology Research	Study design	Observational
Lazzerini, 2008	Arrhythmic risk during acute infusion of infliximab: A prospective, single-blind, placebo-controlled, crossover study in patients with chronic arthritis	Journal of Rheumatology	Outcomes	No outcomes of interest
Lindegaard, 2016	Doubling the single-dose infusion rate of tocilizumab in rheumatoid arthritis is safe and efficacious	Scandinavian journal of rheumatology	Intervention	Dose-randomization
Lisbona, 2008	Etanercept reduces synovitis as measured by magnetic resonance imaging in patients with active rheumatoid arthritis after only 6 weeks	Journal of rheumatology	Outcomes	No outcomes of interest
Lu, 2009	Multicenter, randomized, double-blind, controlled trial of treatment of active rheumatoid arthritis with t-614 compared with methotrexate	Arthritis and rheumatism	Population	cDMARD nave
Lu, 2008	Safety and efficacy of t-614 in the treatment of patients with active rheumatoid arthritis: A double blind, randomized, placebo-controlled and multicenter trial	Chinese medical journal	Population	cDMARD nave

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
MacIsaac, 2014	Pre-treatment whole blood gene expression is associated with 14-week response assessed by dynamic contrast enhanced magnetic resonance imaging in infliximab-treated rheumatoid arthritis patients	Plos one	Outcomes	No outcomes of interest at 24 weeks
Maini, 1998	Therapeutic efficacy of multiple intravenous infusions of anti-tumor necrosis factor alpha monoclonal antibody combined with low-dose weekly methotrexate in rheumatoid arthritis	Arthritis and rheumatism	Outcomes	No outcomes of interest
Maini, 2006	Double-blind randomized controlled clinical trial of the interleukin-6 receptor antagonist, tocilizumab, in european patients with rheumatoid arthritis who had an incomplete response to methotrexate	Arthritis and rheumatism	Outcomes	No outcomes of interest at 24 weeks
Malottki, 2011	Adalimumab, etanercept, infiximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: A systematic review and economic evaluation	Health Technology Assessment	Study design	Systematic literature review
Manders, 2015	Cost-effectiveness of abatacept, rituximab, and tnfi treatment after previous failure with tnfi treatment in rheumatoid arthritis: A pragmatic multi-centre randomised trial	Arthritis research	therapy	Intervention
Unspecified treatment				
Mandl, 2012	Metrologic properties of ultrasound versus clinical evaluation of synovitis in rheumatoid arthritis: Results of a multicenter, randomized study	Arthritis and rheumatism	Outcomes	No outcomes of interest
Matsuno, 2018	A randomized double-blind parallel-group phase iii study to compare the efficacy and safety of ni-071 and infliximab reference product in japanese patients with active rheumatoid arthritis refractory to methotrexate	Modern Rheumatology.	Outcomes	No outcomes of interest at 24 weeks
Mazurov, 2014	The quality of life in patients with rheumatoid arthritis treated with rituximab	Klinicheskaia meditsina	Other	Language
Mease, 2010	Efficacy and safety of retreatment in patients with rheumatoid arthritis with previous inadequate response to tumor necrosis factor inhibitors: Results from the sunrise trial	Journal of rheumatology	Intervention	Discontinuation/withdrawal study
Migliore, 2012	May etanercept and pth (1-34) association heal erosions in early rheumatoid arthritis? A pilot study	European review for medical and pharmacological sciences	Intervention	Teriparatide
Montecucco, 2005	In early rheumatoid arthritis the combination of methotrexate and infliximab over 2 years reduces the progression of radiological lesions more than methotrexate alone	Clinical Experimental Rheumatology	Outcomes	
No outcomes of interest				
Moreland, 2004	Adalimumab in rheumatoid arthritis	Current rheumatology reports	Other	Review
Moreland, 2004	Infliximab in rheumatoid arthritis	Current rheumatology reports	Other	Review
Moreland, 2006	Efficacy and safety of rituximab in rheumatoid arthritis patients refractory to methotrexate	Current Rheumatology Reports	Intervention	methylprednisone/prednisone combination therapies

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Moreland, 2006	Efficacy of costimulation blockade with abatacept in rheumatoid arthritis patients refractory to tumor necrosis factor-alpha inhibition	Current Rheumatology Reports	Other	Review
Moreland, 2012	A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: The treatment of early aggressive rheumatoid arthritis trial	Arthritis and rheumatism	Population	cDMARD nave
Mori, 2018	Tofacitinib therapy for rheumatoid arthritis: A direct comparison study between biologic-naive and experienced patients	Internal Medicine	Study design	Non-randomized
Muller-Ladner, 2012	Comparison of patient satisfaction with two different etanercept delivery systems. A randomised controlled study in patients with rheumatoid arthritis	Zeitschrift fur rheumatologie	Study design	Pooled analysis
Neva, 2000	Combination drug therapy retards the development of rheumatoid atlantoaxial subluxations	Arthritis and rheumatism	Outcomes	No outcomes of interest
O'Dell, 2002	Treatment of rheumatoid arthritis with methotrexate and hydroxychloroquine, methotrexate and sulfasalazine, or a combination of the three medications: Results of a two-year, randomized, double-blind, placebo-controlled trial	Arthritis and rheumatism	Population	cDMARD nave
Ostergaard, 2015	Mri assessment of early response to certolizumab pegol in rheumatoid arthritis: A randomised, double-blind, placebo-controlled phase iiib study applying mri at weeks 0, 1, 2, 4, 8 and 16	Annals of the rheumatic diseases	Comparator	No active comparator at 24 weeks (placebo crossover only)
Pandi Kumar, 2018	A prospective study on comparing the efficacy of combination therapy and monotherapy of dmards in patients with rheumatoid arthritis	Research Journal of Pharmacy and Technology	Outcomes	No outcomes of interest at 24 weeks
Pavelka, 2017	Maintenance of remission with combination etanercept-dmard therapy versus dmards alone in active rheumatoid arthritis: Results of an international treat-to-target study conducted in regions with limited biologic access	Rheumatology international	Population	Low disease activity population
Porter, 2016	Tumour necrosis factor inhibition versus rituximab for patients with rheumatoid arthritis who require biological treatment (orbit): An open-label, randomised controlled, non-inferiority, trial	Lancet (london, england)	Intervention	Unspecified treatment
Quinn, 2005	Very early treatment with infliximab in addition to methotrexate in early, poor-prognosis rheumatoid arthritis reduces magnetic resonance imaging evidence of synovitis and damage, with sustained benefit after infliximab withdrawal: Results from a twelve-month randomized, double-blind, placebo-controlled trial	Arthritis and rheumatism	Population	cDMARD nave
Radstake, 2009	Formation of antibodies against infliximab and adalimumab strongly correlates with functional drug levels and clinical responses in rheumatoid arthritis	Annals of the Rheumatic Diseases	Study design	Non-randomized
Raffeiner, 2013	Adopting low-dose etanercept strategy in the long-term management of rheumatoid arthritis patients	Clinical Drug Investigation	Study design	Review
Ramos-Remus, 2008	The option trial: Inhibition of the interleukin-6 receptor with tocilizumab in patients with rheumatoid arthritis	Future Rheumatology	Other	Review

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Rau, 2004	Rapid alleviation of signs and symptoms of rheumatoid arthritis with intravenous or subcutaneous administration of adalimumab in combination with methotrexate	Scandinavian journal of rheumatology	Study design	No outcomes of interest at 24 weeks
Rexhepi, 2018	Evaluation of the efficacy of combined therapy of methotrexate and etanercept versus methotrexate as a mono-therapy	Open Access Macedonian Journal of Medical Sciences	Outcomes	No outcomes of interest at 24 weeks
Rezaei, 2013	Evaluation of hand bone loss by digital x-ray radiogrammetry as a complement to clinical and radiographic assessment in early rheumatoid arthritis: Results from the swefot trial	BMC Musculoskeletal Disorders	Population	cDMARD nave
Roshique, 2015	Efficacy and safety of a biosimilar rituximab in biologic naive patients with active rheumatoid arthritis	Clinical Rheumatology	Study design	Non-randomized
Roux, 2011	Etanercept compared to intraarticular corticosteroid injection in rheumatoid arthritis: Double-blind, randomized pilot study	Journal of rheumatology	Comparator	Betamethasone
Rubbert-Roth, 2010	Efficacy and safety of various repeat treatment dosing regimens of rituximab in patients with active rheumatoid arthritis: Results of a phase iii randomized study (mirror)	Rheumatology	Intervention	Dose-randomization
Russell, 2007	Abatacept improves both the physical and mental health of patients with rheumatoid arthritis who have inadequate response to methotrexate treatment	Annals of the rheumatic diseases	Outcomes	No outcomes of interest
Saleem, 2008	Does the use of tumour necrosis factor antagonist therapy in poor prognosis, undifferentiated arthritis prevent progression to rheumatoid arthritis?	Annals of the rheumatic diseases	Population	Undifferentiated arthritis
Salgado, 2013	The jak inhibitor tofacitinib for active rheumatoid arthritis: Results from phase iii trials	International Journal of Clinical Rheumatology	Study design	Review
Saunders, 2008	Triple therapy in early active rheumatoid arthritis: A randomized, single-blind, controlled trial comparing step-up and parallel treatment strategies	Arthritis and rheumatism	Population	cDMARD nave
Schiff, 2014	Rheumatoid arthritis secondary non-responders to tnf can attain an efficacious and safe response by switching to certolizumab pegol: A phase iv, randomised, multicentre, double-blind, 12-week study, followed by a 12-week open-label phase	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Schiff, 2014	Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: Two-year efficacy and safety findings from ample trial	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Scott, 2015	Tumour necrosis factor inhibitors versus combination intensive therapy with conventional disease modifying anti-rheumatic drugs in established rheumatoid arthritis: Tacit non-inferiority randomised controlled trial	BMJ	Outcomes	No outcomes of interest at 24 weeks
Sennels, 2008	Circulating levels of osteopontin, osteoprotegerin, total soluble receptor activator of nuclear factor-kappa b ligand, and high-sensitivity c-reactive protein in patients with active rheumatoid arthritis randomized to etanercept alone or in combination with methotrexate	Scandinavian Journal of Rheumatology	Study design	No outcomes of interest at 24 weeks

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Shi, 2013	The efficacy and safety of tocilizumab combined with disease-modifying anti-rheumatoid drugs in the treatment of active rheumatoid arthritis: A multi-center, randomized, double-blinded, placebo-controlled trial	Zhonghua nei ke za zhi [Chinese journal of internal medicine]	Other	Language
Smeets, 2003	Tumor necrosis factor alpha blockade reduces the synovial cell infiltrate early after initiation of treatment, but apparently not by induction of apoptosis in synovial tissue	Arthritis and rheumatism	Outcomes	No outcomes of interest at 24 weeks
Smolen, 2016	Head-to-head comparison of certolizumab pegol versus adalimumab in rheumatoid arthritis: 2-year efficacy and safety results from the randomised exxelerate study	Lancet (london, england)	Outcomes	No outcomes of interest at 24 weeks
Smolen, 2017	A randomised, double-blind trial to demonstrate bioequivalence of gp2013 and reference rituximab combined with methotrexate in patients with active rheumatoid arthritis	Annals of the rheumatic diseases	Outcomes	No outcomes of interest
Smolen, 2018	Safety, immunogenicity and efficacy after switching from reference infliximab to biosimilar sb2 compared with continuing reference infliximab and sb2 in patients with rheumatoid arthritis: Results of a randomised, double-blind, phase iii transition study	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Smolen, 2014	Adjustment of therapy in rheumatoid arthritis on the basis of achievement of stable low disease activity with adalimumab plus methotrexate or methotrexate alone: The randomised controlled optima trial	Lancet (london, england)	Population	cDMARD nave
Smolen, 2014	Certolizumab pegol in rheumatoid arthritis patients with low to moderate activity: The certain double-blind, randomised, placebo-controlled trial	Annals of the Rheumatic Diseases.	Population	Low to moderate disease activity
Smolen, 2015	Certolizumab pegol in rheumatoid arthritis patients with low to moderate activity: The certain double-blind, randomised, placebo-controlled trial	Annals of the rheumatic diseases	Population	cDMARD nave
Smolen, 2005	Evidence of radiographic benefit of treatment with infliximab plus methotrexate in rheumatoid arthritis patients who had no clinical improvement: A detailed subanalysis of data from the anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study	Arthritis and rheumatism	Outcomes	No outcomes of interest at 24 weeks
Smolen, 2013	Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (preserve): A randomised controlled trial	Lancet (london, england)	Intervention	Discontinuation/withdrawal study
Sonomoto, 2014	Effects of tofacitinib on lymphocytes in rheumatoid arthritis: Relation to efficacy and infectious adverse events	Rheumatology	Study design	Pooled analysis
Soubrier, 2009	Evaluation of two strategies (initial methotrexate monotherapy vs its combination with adalimumab) in management of early active rheumatoid arthritis: Data from the guepard trial	Rheumatology	Population	cDMARD nave
St Clair, 2002	The relationship of serum infliximab concentrations to clinical improvement in rheumatoid arthritis: Results from attract, a multicenter, randomized, double-blind, placebo-controlled trial	Arthritis and rheumatism	Outcomes	No outcomes of interest at 24 weeks

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Strand, 2015	Effects of tofacitinib monotherapy on patient-reported outcomes in a randomized phase 3 study of patients with active rheumatoid arthritis and inadequate responses to dmards	Arthritis Care and Research	Comparator	No active comparator at 24 weeks (placebo crossover only)
Strand, 2015	Tofacitinib with methotrexate in third-line treatment of patients with active rheumatoid arthritis: Patient-reported outcomes from a phase iii trial	Arthritis Care and Research	Outcomes	No outcomes of interest at 24 weeks
Strand, 2015	The impact of rheumatoid arthritis on work and predictors of overall work impairment from three therapeutic scenarios	International Journal of Clinical Rheumatology	Intervention	Discontinuation/withdrawal study
Strand, 2009	Rapid and sustained improvements in health-related quality of life, fatigue, and other patient-reported outcomes in rheumatoid arthritis patients treated with certolizumab pegol plus methotrexate over 1 year: Results from the rapid 1 randomized controlled trial	Arthritis research	therapy	Outcomes
No outcomes of interest at 24 weeks				
Stubenrauch, 2010	Subset analysis of patients experiencing clinical events of a potentially immunogenic nature in the pivotal clinical trials of tocilizumab for rheumatoid arthritis: Evaluation of an antidrug antibody elisa using clinical adverse event-driven immunogenicity testing	Clinical Therapeutics	Study design	Pooled analysis
Suh, 2019	Long-term efficacy and safety of biosimilar ct-p10 versus innovator rituximab in rheumatoid arthritis: 48-week results from a randomized phase iii trial	BioDrugs	Outcomes	No outcomes of interest at 24 weeks
Sun, 2016	Efficacy and safety of combined etanercept and iguratimod for active rheumatoid arthritis	Biomedical Research (India)	Population	cDMARD nave
Tak, 2011	Inhibition of joint damage and improved clinical outcomes with rituximab plus methotrexate in early active rheumatoid arthritis: The image trial	Annals of the rheumatic diseases	Population	cDMARD nave
Takeuchi, 2013	A phase 3 randomized, double-blind, multicenter comparative study evaluating the effect of etanercept versus methotrexate on radiographic outcomes, disease activity, and safety in japanese subjects with active rheumatoid arthritis	Modern rheumatology	Outcomes	No outcomes of interest at 24 weeks
Takeuchi, 2015	Evaluation of the pharmacokinetic equivalence and 54-week efficacy and safety of ct-p13 and innovator infliximab in japanese patients with rheumatoid arthritis	Modern rheumatology	Outcomes	No outcomes of interest at 24 weeks
Tam, 2012	Infliximab is associated with improvement in arterial stiffness in patients with early rheumatoid arthritis – a randomized trial	Journal of rheumatology	Population	cDMARD nave
Tanaka, 2016	Efficacy and safety of baricitinib in japanese patients with active rheumatoid arthritis receiving background methotrexate therapy: A 12-week, double-blind, randomized placebo-controlled study	Journal of rheumatology	Outcomes	No outcomes of interest at 24 weeks

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Tanaka, 2018	Efficacy and safety of baricitinib in japanese patients with active rheumatoid arthritis: A 52-week, randomized, single-blind, extension study	Modern rheumatology	Study design	Single-arm extension
Tanaka, 2011	Phase ii study of tofacitinib (cp-690,550) combined with methotrexate in patients with rheumatoid arthritis and an inadequate response to methotrexate	Arthritis Care and Research	Outcomes	No outcomes of interest at 24 weeks
Tanaka, 2019	Modified- versus immediate-release tofacitinib in japanese rheumatoid arthritis patients: A randomized, phase iii, non-inferiority study	Rheumatology	Intervention	Dose randomization
Tanaka, 2015	Efficacy and safety of tofacitinib as monotherapy in japanese patients with active rheumatoid arthritis: A 12-week, randomized, phase 2 study	Modern rheumatology	Outcomes	No outcomes of interest at 24 weeks
Taylor, 2004	Comparison of ultrasonographic assessment of synovitis and joint vascularity with radiographic evaluation in a randomized, placebo-controlled study of infliximab therapy in early rheumatoid arthritis	Arthritis and rheumatism	Outcomes	No outcomes of interest
Taylor, 2006	Ultrasonographic and radiographic results from a two-year controlled trial of immediate or one-year-delayed addition of infliximab to ongoing methotrexate therapy in patients with erosive early rheumatoid arthritis	Arthritis and rheumatism	Outcomes	No outcomes of interest at 24 weeks
Tony, 2019	Brief report: Safety and immunogenicity of rituximab biosimilar gp 2013 after switch from reference rituximab in patients with active rheumatoid arthritis	Arthritis Care and Research	Outcomes	No outcomes of interest
Van De Putte, 2003	Efficacy and safety of the fully human anti-tumour necrosis factor alpha monoclonal antibody adalimumab (d2e7) in dmard refractory patients with rheumatoid arthritis: A 12 week, phase ii study	Annals of the Rheumatic Diseases	Outcomes	No outcomes of interest at 24 weeks
van der Heijde, 2008	The safety and efficacy of adding etanercept to methotrexate or methotrexate to etanercept in moderately active rheumatoid arthritis patients previously treated with monotherapy	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
van der Heijde, 2019	Tofacitinib in combination with methotrexate in patients with rheumatoid arthritis: Clinical efficacy, radiographic, and safety outcomes from a twenty-four-month, phase iii study	Arthritis and Rheumatology.	Outcomes	No outcomes of interest at 24 weeks
van der Kooij, 2009	Patient-reported outcomes in a randomized trial comparing four different treatment strategies in recent-onset rheumatoid arthritis	Arthritis and rheumatism	Population	cDMARD nave
van der Kooij, 2009	Drug-free remission, functioning and radiographic damage after 4 years of response-driven treatment in patients with recent-onset rheumatoid arthritis	Annals of the rheumatic diseases	Population	cDMARD nave
van Jaarsveld, 2000	Aggressive treatment in early rheumatoid arthritis: A randomised controlled trial. On behalf of the rheumatic research foundation utrecht, the netherlands	Annals of the rheumatic diseases	Population	cDMARD nave
Van Riel, 2006	Efficacy and safely of combination etanercept and methotrexate versus etanercept alone in patients with rheumatoid arthritis with an inadequate response to methotrexate: The adore study	Annals of the Rheumatic Diseases	Outcomes	No outcomes of interest at 24 weeks

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Van Riel, 2008	Patient-reported health outcomes in a trial of etanercept monotherapy versus combination therapy with etanercept and methotrexate for rheumatoid arthritis: The adore trial	Annals of the Rheumatic Diseases	Study design	No outcomes of interest at 24 weeks
van Riel, 2006	Efficacy and safety of combination etanercept and methotrexate versus etanercept alone in patients with rheumatoid arthritis with an inadequate response to methotrexate: The adore study	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
van Vollenhoven, 2009	Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in patients with early rheumatoid arthritis (swefot trial): 1-year results of a randomised trial	Lancet (london, england)	Population	cDMARD nave
van Vollenhoven, 2011	Atacicept in patients with rheumatoid arthritis and an inadequate response to methotrexate: Results of a phase ii, randomized, placebo-controlled trial	Arthritis and rheumatism	Intervention	Atacicept
van Vollenhoven, 2016	Full dose, reduced dose or discontinuation of etanercept in rheumatoid arthritis	Annals of the rheumatic diseases	Population	Low disease activity population
van Vollenhoven, 2015	Safety and efficacy of atacicept in combination with rituximab for reducing the signs and symptoms of rheumatoid arthritis: A phase ii, randomized, double-blind, placebo-controlled pilot trial	Arthritis	rheumatology	Intervention
atacicept				
Weinblatt, 2006	Safety of the selective costimulation modulator abatacept in rheumatoid arthritis patients receiving background biologic and nonbiologic disease-modifying antirheumatic drugs: A one-year randomized, placebo-controlled study	Arthritis and rheumatism	Outcomes	No outcomes of interest at 24 weeks
Weinblatt, 2017	A phase iii study evaluating continuation, tapering, and withdrawal of certolizumab pegol after one year of therapy in patients with early rheumatoid arthritis	Arthritis	rheumatology	Population
cDMARD nave				
Weinblatt, 2018	Switching from reference adalimumab to sb5 (adalimumab biosimilar) in patients with rheumatoid arthritis: Fifty-two-week phase iii randomized study results	Arthritis and Rheumatology	Outcomes	No outcomes of interest at 24 weeks
Weinblatt, 2012	Efficacy and safety of certolizumab pegol in a broad population of patients with active rheumatoid arthritis: Results from the realistic phase iiib study	Arthritis	rheumatology	Outcomes
No outcomes of interest at 24 weeks				
Weinblatt, 2015	Twenty-eight-week results from the realistic phase iiib randomized trial: Efficacy, safety and predictability of response to certolizumab pegol in a diverse rheumatoid arthritis population	Arthritis research	therapy	Outcomes
No outcomes of interest at 24 weeks				

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Weinblatt, 2008	Efficacy and safety of etanercept 50 mg twice a week in patients with rheumatoid arthritis who had a suboptimal response to etanercept 50 mg once a week: Results of a multicenter, randomized, double-blind, active drug-controlled study	Arthritis and rheumatism	Intervention	Dose randomization
Weinblatt, 2007	Selective costimulation modulation using abatacept in patients with active rheumatoid arthritis while receiving etanercept: A randomised clinical trial	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Weisman, 2003	Efficacy, pharmacokinetic, and safety assessment of adalimumab, a fully human anti-tumor necrosis factor-alpha monoclonal antibody, in adults with rheumatoid arthritis receiving concomitant methotrexate: A pilot study	Clinical therapeutics	Intervention	Dose randomization
Weisman, 2007	A placebo-controlled, randomized, double-blinded study evaluating the safety of etanercept in patients with rheumatoid arthritis and concomitant comorbid diseases	Rheumatology	Population	Restricted to patients with comorbidities
Westhovens, 2006	A phase i study assessing the safety, clinical response, and pharmacokinetics of an experimental infliximab formulation for subcutaneous or intramuscular administration in patients with rheumatoid arthritis	Journal of rheumatology	Outcomes	No outcomes of interest at 24 weeks
Westhovens, 2015	Maintenance of remission following 2 years of standard treatment then dose reduction with abatacept in patients with early rheumatoid arthritis and poor prognosis	Annals of the rheumatic diseases	Population	Low disease activity population
Wijesinghe, 2017	Leflunomide is equally efficacious and safe compared to low dose rituximab in refractory rheumatoid arthritis given in combination with methotrexate: Results from a randomized double blind controlled clinical trial	BMC Musculoskeletal Disorders	Comparator	Leflunomide
Williams, 2016	Comparative assessment of clinical response in patients with rheumatoid arthritis between pf-05280586, a proposed rituximab biosimilar, and rituximab	British journal of clinical pharmacology	Study design	Modelling study
Wislowska, 2007	Preliminary evaluation in rheumatoid arthritis activity in patients treated with tnf-alpha blocker plus methotrexate versus methotrexate or leflunomide alone	Rheumatology International	Study design	Non-randomized
Xia, 2016	Iguratimod in combination with methotrexate in active rheumatoid arthritis : Therapeutic effects	Zeitschrift fur rheumatologie	Intervention	Iguaratimod
Yoo, 2013	A randomised, double-blind, parallel-group study to demonstrate equivalence in efficacy and safety of ct-p13 compared with innovator infliximab when coadministered with methotrexate in patients with active rheumatoid arthritis: The planetra study	Annals of the rheumatic diseases	Outcomes	No outcomes of interest at 24 weeks
Yoo, 2016	A phase iii randomized study to evaluate the efficacy and safety of ct-p13 compared with reference infliximab in patients with active rheumatoid arthritis: 54-week results from the planetra study	Arthritis research	therapy	Outcomes
No outcomes of interest at 24 weeks				

Continued on next page

Author and Year	Title	Journal	Reason	Subreason
Yoo, 2017	A multicentre randomised controlled trial to compare the pharmacokinetics, efficacy and safety of ct-p10 and innovator rituximab in patients with rheumatoid arthritis	Annals of the rheumatic diseases	Study design	Phase 1
Yoo, 2017	Efficacy, safety and pharmacokinetics of up to two courses of the rituximab biosimilar ct-p10 versus innovator rituximab in patients with rheumatoid arthritis: Results up to week 72 of a phase i randomized controlled trial	BioDrugs	Outcomes	No outcomes of interest at 24 weeks
Zhang, 2006	Infliximab versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: A preliminary study from china	APLAR Journal of Rheumatology	Outcomes	No outcomes of interest at 24 weeks
Zhang, 2013	Pharmacokinetics and pharmacodynamics of tocilizumab after subcutaneous administration in patients with rheumatoid arthritis	International journal of clinical pharmacology and therapeutics	Outcomes	No outcomes of interest at 24 weeks
Zhao, 2017	Analysis of efficacy and safety of treatment of active rheumatoid arthritis with iguratimod and methotrexate	Biomedical Research (India)	Population	cDMARD nave
Zhou, 2007	Pharmacokinetics and safety of golimumab, a fully human anti-tnf-alpha monoclonal antibody, in subjects with rheumatoid arthritis	Journal of clinical pharmacology	Outcomes	No outcomes of interest

References

Aletaha, D., Nell, V. P., Stamm, T., Uffmann, M., Pflugbeil, S., Machold, K., and Smolen, J. S. (2005). Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis research & therapy*, 7(4):R796.

Aletaha, D. and Smolen, J. (2005). The simplified disease activity index (sdai) and the clinical disease activity index (cdai): a review of their usefulness and validity in rheumatoid arthritis. *Clinical and experimental rheumatology*, 23(5):S100.

Anderson, J., Caplan, L., Yazdany, J., Robbins, M. L., Neogi, T., Michaud, K., Saag, K. G., O'dell, J. R., and Kazi, S. (2012). Rheumatoid arthritis disease activity measures: American college of rheumatology recommendations for use in clinical practice. *Arthritis care & research*, 64(5):640–647.

Arias, E. (2015). United states life tables, 2011. *National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System*, 64(11):1–62.

Athanasakis, K., Tarantilis, F., Tsalapati, K., Konstantopoulou, T., Vritzali, E., and Kyriopoulos, J. (2015). Cost-utility analysis of tocilizumab monotherapy in first line versus standard of care for the treatment of rheumatoid arthritis in greece. *Rheumatology international*, 35(9):1489–1495.

Baio, G. and Dawid, A. P. (2015). Probabilistic sensitivity analysis in health economics. *Statistical methods in medical research*, 24(6):615–634.

Barton, G. R., Briggs, A. H., and Fenwick, E. A. (2008). Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (ceac), the cost-effectiveness acceptability frontier (ceaf), and the expected value of perfection information (evpi). *Value in Health*, 11(5):886–897.

Basu, A. and Meltzer, D. (2007). Value of information on preference heterogeneity and individualized care. *Medical Decision Making*, 27(2):112–127.

Black, W. C. (1990). The ce plane: a graphic representation of cost-effectiveness. *Medical decision making*, 10(3):212–214.

Bojke, L., Claxton, K., Sculpher, M., and Palmer, S. (2009). Characterizing structural uncertainty in decision analytic models: a review and application of methods. *Value in Health*, 12(5):739–749.

Brennan, A., Bansback, N., Reynolds, A., and Conway, P. (2003). Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the uk. *Rheumatology*, 43(1):62–72.

Briggs, A. H., Claxton, K., and Sculpher, M. J. (2006). *Decision modelling for health economic evaluation*. Handbooks in Health Economic E.

Briggs, A. H. et al. (1999). A bayesian approach to stochastic cost-effectiveness analysis. *Health economics*, 8(3):257–261.

Carlson, J. J., Ogale, S., Dejonckheere, F., and Sullivan, S. D. (2015). Economic evaluation of tocilizumab monotherapy compared to adalimumab monotherapy in the treatment of severe active rheumatoid arthritis. *Value in Health*, 18(2):173–179.

Claxton, K., Sculpher, M., McCabe, C., Briggs, A., Akehurst, R., Buxton, M., Brazier, J., and O'Hagan, T. (2005). Probabilistic sensitivity analysis for nice technology assessment: not an optional extra. *Health economics*, 14(4):339–347.

Claxton, L., Jenks, M., Taylor, M., Wallenstein, G., Mendelsohn, A. M., Bourret, J. A., Singh, A., Moynagh, D., and Gerber, R. A. (2016). An economic evaluation of tofacitinib treatment in rheumatoid arthritis: Modeling the cost of treatment strategies in the united states. *Journal of managed care & specialty pharmacy*, 22(9):1088–1102.

Curtis, J. R., Jain, A., Askling, J., Bridges, S. L., Carmona, L., Dixon, W., Finckh, A., Hyrich, K., Greenberg, J. D., Kremer, J., et al. (2010). A comparison of patient characteristics and outcomes in selected european and us rheumatoid arthritis registries. In *Seminars in arthritis and rheumatism*, volume 40, pages 2–14. Elsevier.

Deighton, C., Hyrich, K., Ding, T., Ledingham, J., Lunt, M., Luqmani, R., Kiely, P., Bukhari, M., Abernethy, R., Ostor, A., et al. (2010). Bsr and bhpr rheumatoid arthritis guidelines on eligibility criteria for the first biological therapy. *Rheumatology*, 49(6):1197–1199.

Dias, S., Sutton, A. J., Ades, A., and Welton, N. J. (2013). Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Medical Decision Making*, 33(5):607–617.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., and Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*. Oxford university press.

Eddy, D. M., Hollingworth, W., Caro, J. J., Tsevat, J., McDonald, K. M., and Wong, J. B. (2012). Model transparency and validation: a report of the ispor-smdm modeling good research practices task force–7. *Medical Decision Making*, 32(5):733–743.

Espinoza, M. A., Manca, A., Claxton, K., and Sculpher, M. J. (2014). The value of heterogeneity for cost-effectiveness subgroup analysis: conceptual framework and application. *Medical Decision Making*, 34(8):951–964.

Fenwick, E., Claxton, K., and Sculpher, M. (2001). Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health economics*, 10(8):779–787.

Garber, A. M. and Phelps, C. E. (1997). Economic foundations of cost-effectiveness analysis. *Journal of health economics*, 16(1):1–31.

Garrison, L. P., Kamal-Bahl, S., and Towse, A. (2017). Toward a broader concept of value: identifying and defining elements for an expanded cost-effectiveness analysis. *Value in Health*, 20(2):213–216.

Gibson, L., Alava, M. H., and Wailoo, A. (2015). Progression of disease in people with rheumatoid arthritis treated with non biologic therapies. Technical report.

Gomez-Reino, J. J. and Carmona, L. (2006). Switching tnf antagonists in patients with chronic arthritis: an observational study of 488 patients over a four-year period. *Arthritis research & therapy*, 8(1):R29.

Guyot, P., Ades, A., Ouwens, M. J., and Welton, N. J. (2012). Enhanced secondary analysis of survival data: reconstructing the data from published kaplan-meier survival curves. *BMC medical research methodology*, 12(1):9.

Heath, A., Manolopoulou, I., and Baio, G. (2016). Estimating the expected value of partial perfect information in health economic evaluations using integrated nested laplace approximation. *Statistics in medicine*, 35(23):4264–4280.

Hernández-Alava, M., Wailoo, A., Wolfe, F., and Michaud, K. (2013). The relationship between eq-5d, haq and pain in patients with rheumatoid arthritis. *Rheumatology*, 52(5):944–950.

Hernández-Alava, M., Wailoo, A., Wolfe, F., and Michaud, K. (2014). A comparison of direct and indirect methods for the estimation of health utilities from clinical outcomes. *Medical Decision Making*, 34(7):919–930.

Hernández-Alava, M., Wailoo, A. J., and Ara, R. (2012). Tails from the peak district: adjusted limited dependent variable mixture models of eq-5d questionnaire health state utility values. *Value in Health*, 15(3):550–561.

Institute for Clinical and Economic Review (2017). Targeted immune modulators for rheumatoid arthritis: Effectiveness & value. Technical report.

Ioannidis, J. P. and Garber, A. M. (2011). Individualized cost-effectiveness analysis. *PLoS medicine*, 8(7):e1001058.

Jalal, H., Dowd, B., Sainfort, F., and Kuntz, K. M. (2013). Linear regression metamodeling as a tool to summarize and present simulation model results. *Medical Decision Making*, 33(7):880–890.

Jalal, H., Goldhaber-Fiebert, J. D., and Kuntz, K. M. (2015). Computing expected value of partial sample information from probabilistic sensitivity analysis using linear regression metamodeling. *Medical Decision Making*, 35(5):584–595.

Keeney, R. L. and Raiffa, H. (1993). *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press.

Lakdawalla, D., Malani, A., and Reif, J. (2015). The insurance value of medical innovation. Technical report, National Bureau of Economic Research.

Lakdawalla, D., Malani, A., and Reif, J. (2017). The insurance value of medical innovation. *Journal of public economics*, 145:94–102.

Lakdawalla, D. N., Romley, J. A., Sanchez, Y., Maclean, J. R., Penrod, J. R., and Philipson, T. (2012). How cancer patients value hope and the implications for cost-effectiveness assessments of high-cost cancer therapies. *Health Affairs*, 31(4):676–682.

Madan, J., Ades, T., Barton, P., Bojke, L., Choy, E., Helliwell, P., Jobanputra, P., Stein, K., Stevens, A., Tosh, J., et al. (2015). Consensus decision models for biologics in rheumatoid and psoriatic arthritis: recommendations of a multidisciplinary working party. *Rheumatology and Therapy*, 2(2):113–125.

Mandel, M. (2013). Simulation-based confidence intervals for functions with complicated derivatives. *The American Statistician*, 67(2):76–81.

Meltzer, D. O., Smith, P. C., et al. (2011). Theoretical issues relevant to the economic evaluation of health technologies. *Handbook of health economics*, 2:433–469.

Michaud, K., Vera-Llonch, M., and Oster, G. (2012). Mortality risk by functional status and health-related quality of life in patients with rheumatoid arthritis. *The Journal of rheumatology*, 39(1):54–59.

Michaud, K., Wallenstein, G., and Wolfe, F. (2011). Treatment and nontreatment predictors of health assessment questionnaire disability progression in rheumatoid arthritis: a longitudinal study of 18,485 patients. *Arthritis care & research*, 63(3):366–372.

Myasoedova, E., Crowson, C. S., Kremers, H. M., Therneau, T. M., and Gabriel, S. E. (2010). Is the incidence of rheumatoid arthritis rising?: results from olmsted county, minnesota, 1955–2007. *Arthritis & Rheumatology*, 62(6):1576–1582.

Norton, S., Fu, B., Scott, D. L., Deighton, C., Symmons, D. P., Wailoo, A. J., Tosh, J., Lunt, M., Davies, R., Young, A., et al. (2014). Health assessment questionnaire disability progression in early rheumatoid arthritis: systematic review and analysis of two inception cohorts. In *Seminars in arthritis and rheumatism*, volume 44, pages 131–144. Elsevier.

Norton, S., Sacker, A., Dixey, J., Done, J., Williams, P., and Young, A. (2013). Trajectories of functional limitation in early rheumatoid arthritis and their association with mortality. *Rheumatology*, page ket253.

Oppong, R., Kaambwa, B., Nuttall, J., Hood, K., Smith, R. D., and Coast, J. (2013). The impact of using different tariffs to value eq-5d health state descriptions: an example from a study of acute cough/lower respiratory tract infections in seven countries. *The European journal of health economics*, 14(2):197–209.

Pan, S. M. D., Dehler, S., Ciurea, A., Ziswiler, H.-R., Gabay, C., and Finckh, A. (2009). Comparison of drug retention rates and causes of drug discontinuation between anti–tumor necrosis factor agents in rheumatoid arthritis. *Arthritis Care & Research*, 61(5):560–568.

Prevoo, M., Van't Hof, M., Kuper, H., Van Leeuwen, M., Van De Putte, L., and Van Riel, P. (1995). Modified disease activity scores that include twenty-eight-joint counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis & Rheumatology*, 38(1):44–48.

Ramiro, S., Sepriano, A., Chatzidionysiou, K., Nam, J. L., Smolen, J. S., van der Heijde, D., Dougados, M., van Vollenhoven, R., Bijlsma, J. W., Burmester, G. R., et al. (2017). Safety of synthetic and biological dmards: a systematic literature review informing the 2016 update of the eular recommendations for management of rheumatoid arthritis. *Annals of the rheumatic diseases*, pages annrheumdis–2016.

Sanders, G. D., Neumann, P. J., Basu, A., Brock, D. W., Feeny, D., Krahn, M., Kuntz, K. M., Meltzer, D. O., Owens, D. K., Prosser, L. A., et al. (2016). Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. *Jama*, 316(10):1093–1103.

Sarzi-Puttini, P., Fiorini, T., Panni, B., Turiel, M., Cazzola, M., and Atzeni, F. (2002). Correlation of the score for subjective pain with physical disability, clinical and radiographic scores in recent onset rheumatoid arthritis. *BMC musculoskeletal disorders*, 3(1):18.

Singh, J. A., Saag, K. G., Bridges, S. L., Akl, E. A., Bannuru, R. R., Sullivan, M. C., Vaysbrot, E., McNaughton, C., Osani, M., Shmerling, R. H., et al. (2016). 2015 american college of

rheumatology guideline for the treatment of rheumatoid arthritis. *Arthritis & rheumatology*, 68(1):1–26.

Singh, J. A., Wells, G. A., Christensen, R., Tanjong Ghogomu, E., Maxwell, L. J., MacDonald, J. K., Filippini, G., Skoetz, N., Francis, D. K., Lopes, L. C., et al. (2011). Adverse effects of biologics: a network meta-analysis and cochrane overview. *The Cochrane Library*.

Smolen, J., Breedveld, F., Schiff, M., Kalden, J., Emery, P., Eberl, G., Van Riel, P., and Tugwell, P. (2003). A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology*, 42(2):244–257.

Souto, A., Maneiro, J. R., and Gómez-Reino, J. J. (2015). Rate of discontinuation and drug survival of biologic therapies in rheumatoid arthritis: a systematic review and meta-analysis of drug registries and health care databases. *Rheumatology*, 55(3):523–534.

Stephens, S., Botteman, M. F., Cifaldi, M. A., and van Hout, B. A. (2015). Modelling the cost-effectiveness of combination therapy for early, rapidly progressing rheumatoid arthritis by simulating the reversible and irreversible effects of the disease. *BMJ open*, 5(6):e006560.

Stevenson, M., Archer, R., Tosh, J., Simpson, E., Everson-Hock, E., Stevens, J., Hernandez-Alava, M., Paisley, S., Dickinson, K., Scott, D., et al. (2016). Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying antirheumatic drugs and after the failure of conventional disease-modifying antirheumatic drugs only: systematic review and economic evaluation. *Health Technology Assessment*, 20(35):1–610.

Stevenson, M. D., Wailoo, A. J., Tosh, J. C., Hernandez-Alava, M., Gibson, L. A., Stevens, J. W., Archer, R. J., Simpson, E. L., Hock, E. S., Young, A., et al. (2017). The cost-effectiveness of sequences of biological disease-modifying antirheumatic drug treatment in england for patients with rheumatoid arthritis who can tolerate methotrexate. *The Journal of Rheumatology*, pages jrheum–160941.

Strand, V., Williams, S., Miller, P., Saunders, K., Grant, S., and Kremer, J. (2013). Op0064 discontinuation of biologic therapy in rheumatoid arthritis (ra): Analysis from the consortium of rheumatology researchers of north america (corrona) database. *Annals of the Rheumatic Diseases*, 72(Suppl 3):A71–A72.

Thokala, P., Devlin, N., Marsh, K., Baltussen, R., Boysen, M., Kalo, Z., Longrenn, T., Mussen, F., Peacock, S., Watkins, J., et al. (2016). Multiple criteria decision analysis for health care decision making-an introduction: report 1 of the ispor mcda emerging good practices task force. *Value in health*, 19(1):1–13.

Tosh, J., Brennan, A., Wailoo, A., and Bansback, N. (2011). The sheffield rheumatoid arthritis health economic model. *Rheumatology*, 50(suppl 4):iv26–iv31.

Van Hout, B. A., Al, M. J., Gordon, G. S., and Rutten, F. F. (1994). Costs, effects and c/e-ratios alongside a clinical trial. *Health economics*, 3(5):309–319.

Wailoo, A., Brennan, A., Bansback, N., Nixon, R., Wolfe, F., and Michaud, K. (2006). Modeling the cost effectiveness of etanercept, adalimumab and anakinra compared to infliximab in the treatment of patients with rheumatoid arthritis in the medicare program. *Rockville, MD: Agency for Healthcare Research and Quality*.

Wailoo, A. J., Bansback, N., Brennan, A., Michaud, K., Nixon, R. M., and Wolfe, F. (2008). Biologic drugs for rheumatoid arthritis in the medicare program: a cost-effectiveness analysis. *Arthritis & Rheumatology*, 58(4):939–946.

Wolfe, F. and Michaud, K. (2010). The loss of health status in rheumatoid arthritis and the effect of biologic therapy: a longitudinal observational study. *Arthritis research & therapy*, 12(2):R35.

Wolfe, F., Michaud, K., Choi, H. K., and Williams, R. (2005). Household income and earnings losses among 6,396 persons with rheumatoid arthritis. *The Journal of rheumatology*, 32(10):1875–1883.

Wolfe, F., Michaud, K., Gefeller, O., and Choi, H. K. (2003). Predicting mortality in patients with rheumatoid arthritis. *Arthritis & Rheumatism*, 48(6):1530–1542.

Yazici, Y., Krasnokutsky, S., Barnes, J. P., Hines, P. L., Wang, J., and Rosenblatt, L. (2009). Changing patterns of tumor necrosis factor inhibitor use in 9074 patients with rheumatoid arthritis. *The Journal of rheumatology*, 36(5):907–913.

Zhang, J., Shan, Y., Reed, G., Kremer, J., Greenberg, J. D., Baumgartner, S., and Curtis, J. R. (2011). Thresholds in disease activity for switching biologics in rheumatoid arthritis patients: experience from a large us cohort. *Arthritis care & research*, 63(12):1672–1679.