Janus Kinase Inhibitors for Rheumatoid Arthritis: Effectiveness and Value

Modeling Analysis Plan

August 5, 2019

Institute for Clinical and Economic Review

Table of Contents

1. Approach

This analysis plan details our modeling approach and outcomes to be assessed for the economic evaluation of Janus kinase (JAK) inhibitors for highly active rheumatoid arthritis (RA). This economic evaluation serves as an update to the 2017 ICER RA review, but with a focus on JAK inhibitors. We chose to evaluate only the JAK inhibitors, and not drugs from other classes in this update, because of their increasing preference and use among clinicians in real-world practice, as informed by discussions with stakeholders. Refer to the Research Protocol for details on the systematic review of the clinical evidence on this topic.

The primary aim of this analysis will be to estimate the cost-effectiveness of JAK inhibitors for highly active RA patients using a decision analytic model. The model will compare three JAK inhibitors, upadacitinib (investigational drug under FDA review, AbbVie), baricitinib (Olumiant®, Eli Lilly and Company), and tofacitinib (Xeljanz®, Pfizer) to each other, as well as to conventional disease modifying antirheumatic drugs (cDMARD) and to adalimumab (Humira®, AbbVie). The base-case analysis will take a health care sector perspective (i.e., focus on direct medical care costs only), and a lifetime time horizon. Productivity losses will be considered in a scenario analysis. The model will be developed in heRo3, with some components of the model (e.g., survival distributions) being developed in RStudio.

2. Methods

2.1 Overview and Model Structure

We will develop a *de novo* decision analytic model for this evaluation, informed by key clinical trials and prior relevant economic models. The model will be based on the economic evaluation conducted for ICER's 2017 RA review, but with several changes made to accommodate recommendations that reflect current clinical practice in RA. The base-case analysis will take a health care sector perspective and thus focus on direct medical care costs only. Costs and outcomes will be discounted at 3% per year.

The primary model will focus on an intention-to-treat and treat-to-target analysis, with a hypothetical cohort of patients with severely active RA for whom prior treatment with cDMARDs has failed. Upon model entry, the hypothetical patient cohort will be initiated on a treatment, and treatment response will be assessed at three months. In the base-case analysis, a targeted immune modulator (TIM) will be added to a cDMARD, such as methotrexate. Pending data availability, we will model monotherapy with TIMs in a scenario analysis. Treatment switching will be based on disease activity as measured by disease activity score in 28 joints (DAS28) (Table 1), with those in remission and with low disease activity remaining on the same treatment as for the first three months, while those with moderate/high disease activity switch to subsequent line of therapy at the end of the three-month cycle.

Table 1. Disease Activity Based on DAS28 Categories

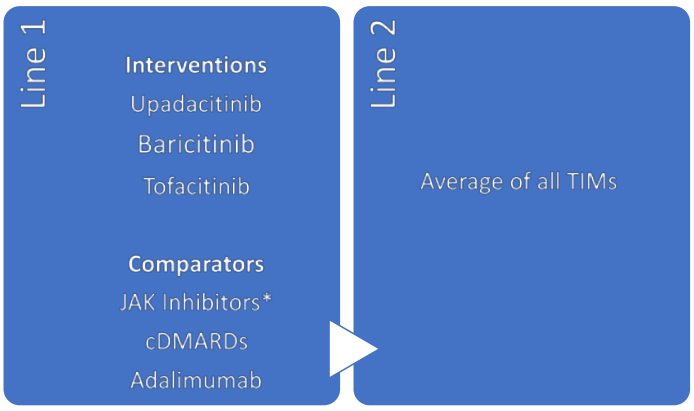
DAS28 Scores	Disease Activity
<2.6	Remission
2.6 to ≤3.2	Low disease activity (LDA)
>3.2 to ≤5.1	Moderate disease activity (MDA)
>5.1	High disease activity (HDA)

DAS28: disease activity score in 28 joints
Source: Canhão et al., 2018[1]

In a real-world clinical setting, it is not uncommon for patients to cycle through multiple therapies before finding a treatment option that they best respond to and tolerate. However, there is a lack of guidelines or published real-world evidence to standardize treatment sequencing for RA patients. In addition, the purpose of this analysis is to compare the selected treatments to each other, not to determine the cost-effectiveness of various treatment sequences. Thus, treatment switching will be to a second line that is an average of all TIMs across the different classes. After the first three months on treatment line one, those with MDA/HDA will switch to this second-line average of all TIMs, while those in remission/LDA will switch to this second line average of all TIMs over time for

reasons such as loss of efficacy, adverse events, patient and clinician preferences, and access restrictions. While this approach to modeling RA treatments does not reflect real-world practice, given that all treatment arms have a standardized treatment sequence beyond first line, it helps to compare the cost-effectiveness of the studied drugs rather than of different treatment sequences.

Figure 1. RA Treatment Sequence



cDMARD: conventional disease modifying antirheumatic drug, JAK: Janus kinase, TIM: targeted immune modulator
*JAK inhibitors will be compared to each other.

After initiating treatment with a TIM, the model will relate the DAS28-based response to the Health Assessment Questionnaire for RA Disability Index (HAQ) after three months of therapy. We found one other published model that related the DAS28 to HAQ score, but at six months of therapy, through a simulation model.[2] Other previously published models, including the previous ICER RA model, mapped the American College of Rheumatology (ACR) response or the European League Against Rheumatism (EULAR) to the HAQ after six months of therapy.[3-5] Our model will use a three-month cycle length, because we heard from clinicians that this more closely aligns with the time point that clinicians use in the recommended treat-to-target approach.[6] We found clinical trial data on the proportions of patients with different categories of disease activity based on the DAS28 at three months for all treatments included in line one but did not find a robust DAS28 to HAQ mapping algorithm at three or six months. Hence, we propose using a mapping algorithm from EULAR to HAQ (Table 2). The EULAR response is divided into three categories: "Good," "Moderate," and "None," and is based on the baseline DAS28 and the change in DAS28 from baseline at the time point measured. Here, we assumed remission as defined by DAS28 as equivalent to "Good" response, LDA as equivalent to "Moderate" response, and MDA and HDA as equivalent to "None" on the EULAR scale. While the HAQ to EULAR response mapping indicates HAQ change at six months, we assumed this to be the same at three months, which likely overestimates the benefit and biases the results in favor of the TIMs.

Table 2. Relationship Between EULAR and HAQ

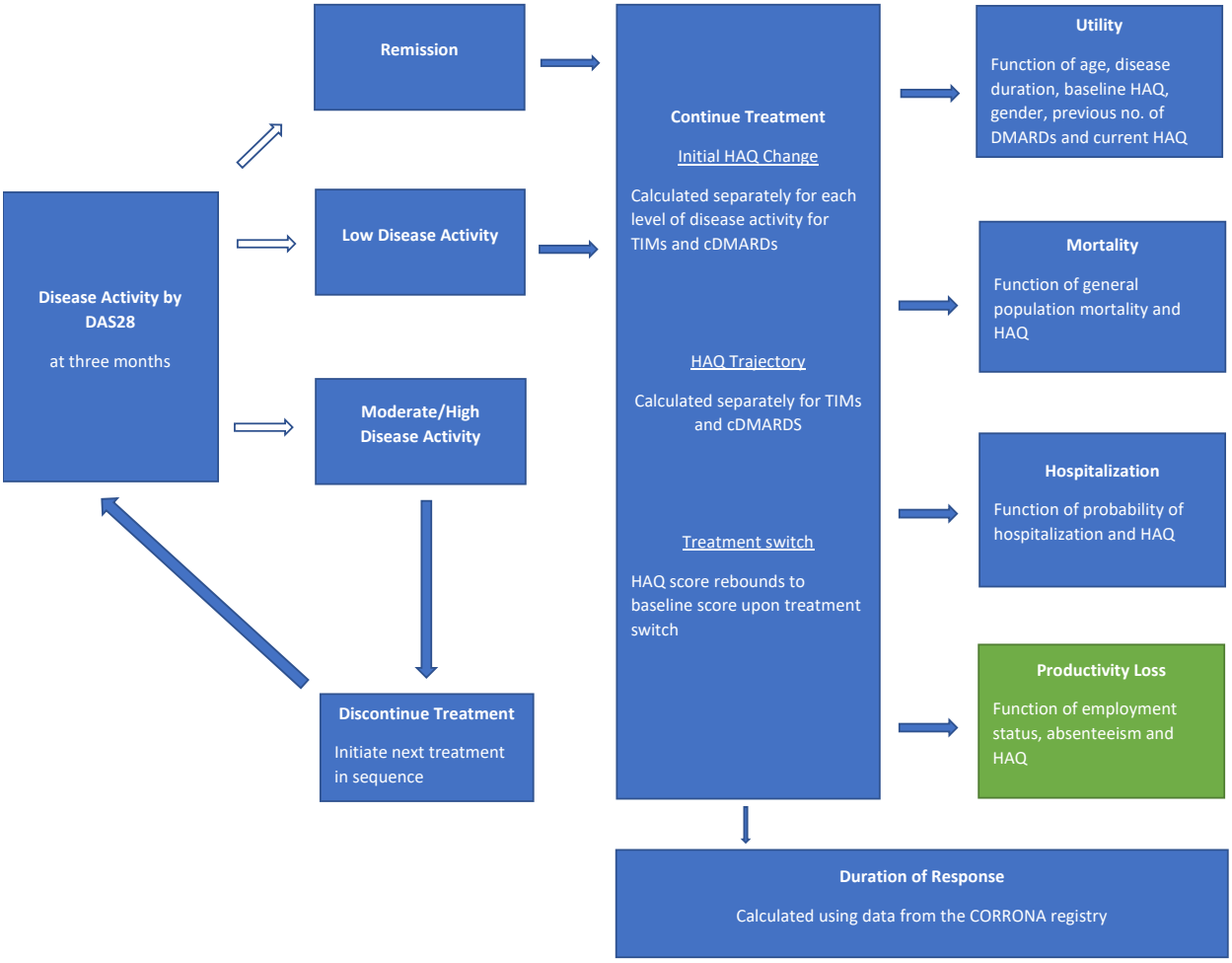
EULAR Response	Mean HAQ Change	Standard Error
Good	-0.672	0
Moderate	-0.317	0.048
None	0	0.112

HAQ: Health Assessment Questionnaire, EULAR: European League Against Rheumatism
HAQ change mapping to EULAR response categories was estimated from the British Society for Rheumatology Biologics Register (BSRBR) and has been used in the other published economic evaluations.

The HAQ score will then be linked to utility, mortality, hospitalizations, and productivity. Simulated utility scores and mortality will be used to calculate the quality-adjusted life years (QALYs) gained, with hospitalizations costs and productivity loss costs contributing to the health care sector perspective and societal perspective analyses, respectively (Figure 2). Long-term HAQ scores will be simulated until treatment discontinuation or death, with different estimates for long-term HAQ changes applied to those on TIMs and cDMARDs.

Patients remain in the model until they die. All patients can transition to death from all causes and from RA-related mortality.

Figure 2. Model Schematic



cDMARD: conventional disease modifying antirheumatic drug, DAS28: disease activity score in 28 joints, HAQ: Health Assessment Questionnaire, TIM: targeted immune modulator

Productivity losses will be measured in the modified societal perspective scenario analysis.

2.2 Key Model Choices and Assumptions

Our model includes several assumptions stated below.

Table 3. Modeling Assumptions

Assumption	Rationale
A treat-to-target approach is used, with treatment switching dependent on disease activity as measured by DAS28.	We use a treat-to-target approach to align with real-world clinical practice, using DAS28 to assess the likelihood of treatment switching.
A three-month cycle length will be adopted, rather than the commonly used six-month cycle length seen in several previously published RA economic models,[2-4] including the model developed for the 2017 ICER RA review.	The three-month cycle length more closely aligns with the average length of time clinicians wait before assessing the need for treatment switching using a treat-to-target approach.[6] Additionally, we have clinical trial data on the proportion of patients with different levels of disease activity as defined by the DAS28 at three months for all treatments included in line one.
We will adopt the EULAR to HAQ mapping algorithm for the different DAS28 disease activity categories, assuming remission to reflect a "Good" EULAR response, LDA to reflect "Moderate" EULAR response, and MDA and HDA to reflect "None" EULAR response.	All trials report DAS28 categories by remission, LDA, and an MDA/HDA combination, but we found no robust published evidence mapping DAS28 to HAQ.
No dose increase will be assumed for those in the LDA category as measured by the DAS28 at three months after initiation of a new TIM.	Clinical experts indicated that a dose increase for those with LDA but not in remission is patient-specific and not necessarily uniformly practiced for all drugs.
At three months after initiation of a new TIM, those with MDA or HDA as measured by the DAS28 will be assumed to switch treatment, with the second-line treatment efficacy an average of all RA treatments.	Clinical experts reported that they would most likely initiate a treatment switch to a new TIM if patients show MDA or HDA, irrespective of which disease activity measure is used.
Following line one therapy, patients switch to a second and final line of therapy comprising an average of all TIMs.	Although patients can switch through multiple lines of therapy, there is a lack of guidelines or published real-world evidence to standardize treatment sequencing for RA patients. In addition, the purpose of this analysis is to compare the selected treatments to each other, not to determine the cost-effectiveness of various treatment sequences. We assume that some patients will try and fail multiple regimens before achieving remission, but the proportions will be similar in those who fail a JAK inhibitor and those who initially fail another TIM. Thus, treatment switching, when necessary, will be to a second line that is an average of all TIMs.

Upon treatment discontinuation, HAQ rebounds to baseline HAQ; in a scenario analysis, the magnitude of rebound will be varied to be not more than the HAQ improvement in the first three months of treatment.	We are unaware of any robust data on the magnitude of HAQ rebound upon treatment discontinuation. We hence assume a rebound to baseline HAQ and vary this in a scenario analysis.
We assumed the same discontinuation rate among those with remission/LDA for all TIMs following the initial three months of therapy.	Prior evaluations attempting class-level economic evaluations have cited errors of confounding in observational studies reporting discontinuation rates. Additionally, because these therapies have been approved all not at once but over time, there is no good comparison of discontinuation rates between treatments.[4,5]
We assumed the discontinuation rate for cDMARDs to be the same as TIMs.	Prior evaluations have highlighted that cDMARDs have safety profiles similar to TIMs and have hence assumed the same discontinuation rates in both categories of drugs.[4,5]
The rate of serious infection is assumed to be the same for all TIMs.	Serious infection measured in the trials do not reflect long-term data, with real-world evidence potentially reflecting infection rates from a sequence rather than a single TIM, and with differences in patient baseline characteristics rendering them non-comparable among TIMs. Due to such inaccurate representations of serious infection caused by specific TIMs prior models have used the same rate of serious infection, an approach we believe is reasonable to follow.[2,4,5]
In the cDMARD arm, HAQ degradation over time will be assumed at 0.0269 per year for the first 15 years in the model, after which a constant rate of 15*0.0269 = 0.4035 will be applied.	Findings from the National Databank on Rheumatic Diseases show a degradation of HAQ over time among patients not on TIMs.[7] We will alter this HAQ progression using other data estimates from the NDB, as done in a scenario in a prior published economic evaluation.[2,8]
In the TIM arm, a long-term HAQ improvement of -0.001 annually will be assumed and will be standard for all TIMs assessed.	Data from the NDB estimated a long-term improvement in HAQ at -0.001 annually among patients on TIMs.[9]
Cost of treatment for those with MDA/HDA will be assumed the full length of the cycle (three months).	We found no data on specific time points within a three-month observation period where patients would lose response and will hence assume the cost of treatment for the length of the cycle.

2.3 Populations

The primary population of focus for the economic evaluation will include adults in the US with severely active RA with inadequate response to or intolerance to cDMARDs. The model will simulate a hypothetical homogeneous cohort of patients, with baseline characteristics of patients with highly active RA similar to US RA registries as summarized by Curtis and colleagues.[10] Other models, including the one in the 2017 ICER RA review, have adopted characteristics from this study by Curtis et al. (Table 4).[2,10] While the TIMs are indicated in populations with non-highly active disease as well, we chose to model patients with only severely active RA to reflect the population seen in the key clinical trials and in the study by Curtis et al.

Table 4. Baseline Population Characteristics

	Mean Value	Source
Age	55 years	Curtis et al., 2010[10]
Female (%)	79%	
Weight	75 kg (female) 89 kg (male)	
Baseline HAQ	1.5	
Baseline DAS28	6	

DAS28: disease activity score in 28 joints, HAQ: Health Assessment Questionnaire, kg: kilogram

2.4 Interventions

The list of interventions assessed will follow the scope set out for the clinical review, and was developed with input from stakeholders on which drugs to include. The full list of interventions is as follows:

- Upadacitinib ([investigational], AbbVie)
- Baricitinib (Olumiant®, Eli Lilly)
- Tofacitinib (Xeljanz®, Pfizer)

Comparators

Comparators include the following:

- JAK inhibitors (to each other)
- Methotrexate (cDMARD)
- Adalimumab (Humira®, AbbVie)

In the base-case analysis, cDMARDs will be added on to TIM therapy. In a scenario analysis, we will evaluate the cost-effectiveness of monotherapy with TIMs, pending data availability.

2.5 Input Parameters

Clinical Inputs

Treatment Response

Inputs on the proportion of patients with different levels of disease activity will be derived from a network meta-analysis (NMA) comprising the JAK inhibitors, cDMARDs, and adalimumab. For the subsequent line of therapy, an average of disease activity proportions based on the DAS28 for all TIMs will be determined (Table 5).

Table 5. Treatment Response at Three Months using DAS28

	Proportion of Patients Achieving Different Categories of Disease Activity by DAS28 at Three Months		
	<2.6 (Remission)	2.6 to ≤3.2 (LDA)	>3.2 (MDA and HDA)
JAK Inhibitors	NMA	NMA	NMA
All TIMs	TBD	TBD	TBD

DAS28: disease activity score in 28 joints, HDA: high disease activity, JAK: Janus kinase, LDA: low disease activity, MDA: moderate disease activity, NMA: network meta-analysis, TBD: to be determined, TIM: targeted immune modulator

Discontinuation

Among those treated with TIMs, the proportion of patients in MDA and HDA at three months after initiation of therapy will switch to subsequent line of therapy. These proportions will be estimated from an NMA. Patients in remission and with LDA at three months after treatment initiation will be assumed to continue on initial therapy. We will include estimates of treatment discontinuation due to other reasons such as loss of efficacy, serious adverse events including infections, and physician and patient preferences, based on data from an observational study of RA patients in the CORRONA registry.[11] The study included a sample of over 6,000 RA adult patients treated between 2002 and 2011 receiving TIMs, predominantly TNF-inhibitors. We will digitize the reported Kaplan-Meier (KM) curves and fit relevant parametric distributions to the curve based on Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC), and extrapolate the fitted curve over the modeled time horizon. Because the sampled population in the CORRONA registry comprised patients with MDA, we will adjust this curve to represent discontinuation among patients with remission/LDA using an odds ratio (OR) of 0.52 as reported by Zhang et al.[12] Following Stevenson et al. and the IVI RA modeling group, we assumed the same long-term discontinuation rate for all TIMs due to issues of bias and confounding found in observational studies for specific TIMs.[2,4,5]

For cDMARDs, again following the methods adopted by Stevenson et al. and the IVI RA modeling group, we assumed that those who were on cDMARD treatment for at least three months had the same treatment duration as those on TIMs.[2,4,5]

Mortality

Gender and age-specific mortality will be sourced from the Human Mortality Database's US-specific tables.[13] Prior evidence suggests that improved (lower) HAQ scores are associated with lower likelihood of death and that HAQ was the most significant predictor of mortality in RA patients.[14] The quantitative relationship between HAQ and mortality was assumed to be the same as that used in the 2017 ICER RA review and was based on a published US RA cost-effectiveness study.[3] The mortality equation to be used is:

US RA-severity specific mortality rate = All-cause mortality*1.33^{HAQ}

Adverse Events

We included only adverse events related to serious infection, aligning with approaches used in prior economic evaluations of RA treatments.[2,4,5] As stated in Table 3, we assumed that the rate of serious infection was uniform across TIMs, as published estimates on specific TIMs do not represent long-term data and are likely inaccurately estimated as mentioned in previously published literature. Estimates on serious infections were sourced from an NMA by Singh et al. (Table 6).[15] As in the 2017 ICER RA review and as used in prior models, we attributed a disutility of 0.156 for a one-month period following a serious infection, along with relevant costs of treating the infection.

Table 6. Adverse Events (Serious Infection)

Parameter	Value (95% CI)*	Source
TIM	0.035 (0.027 – 0.046)	Singh et al., 2011[15]
cDMARD	0.026 (NR)	

cDMARD: conventional disease modifying antirheumatic drug, CI: confidence interval, NR: not reported, TIM: targeted immune modulator

*Calculated as per person-year.

Health State Utilities

As in the 2017 ICER RA review, the relationship between HAQ and utility score was based on Wailoo and colleagues' publication.[16] The utility scores from Wailoo and colleagues were based on health state time-tradeoff evaluations made by a US general population sample using the EuroQol (EQ-5D) index, one of the most widely used instruments in health state valuation.[17] We compared the Wailoo et al. utility change from HAQ score moving from 1.0 to 1.5 to the utility change from a more advanced mathematical model.[18] Although the Wailoo et al. relationship produces a higher

utility within the HAQ range of 1.0 to 1.5, the change in utility for this HAQ range was approximately 0.1 and this change was deemed consistent with the other model. Uncertainty in the Wailoo et al. mapping will be evaluated in parameter sensitivity analyses. EQ-5D scores were calculated using this equation:

EQ-5D score = 1 – 1/(1+ exp(2.0734 + 0.0058*age + 0.0023*disease duration – 0.2004*baseline HAQ – 0.2914*male + 0.0249*previous DMARDs – 0.8647*current HAQ))

Additionally, a disutility (-0.156) was assigned for one month to individuals who experienced a serious infection.[4]

Drug Utilization

The inputs used to model drug utilization and associated costs are shown in Table 7.

Table 7. Treatment Regimen Recommended Dosage

Generic Name	Upadacitinib	Baricitinib	Tofacitinib	Methotrexate	Adalimumab
Brand Name	Investigational	Olumiant®	Xeljanz®	Generic	Humira®
Manufacturer	AbbVie	Eli Lilly	Pfizer	Multiple manufacturers	AbbVie
Route of Administration	Oral	Oral	Oral	Oral	Subcutaneous injection
Dosing	15 mg once daily	2 mg once daily	5 mg twice daily*	7.5 mg once weekly	40 mg every other week

mg: milligram

*Extended release version is dosed at 11 mg once daily.

Cost Inputs

Drug Costs

Drug costs will include the cost of acquisition. We have obtained net price data from SSR Health[19] that combined information on net US dollar sales with information on unit sales to derive net pricing at the unit level across all payer types. Data on the approved drugs of interest were current through the first quarter of 2019. We have estimated net prices for these drugs by comparing the four-quarter rolling averages (i.e., second quarter 2018 through first quarter 2019) of both net prices and wholesale acquisition cost (WAC) per unit to arrive at an average discount from WAC for each drug. We apply this derived discount to the latest WAC[20] of the TIMs of interest. We will derive a net price for each TIM using the WAC and discount from SSR and average the net price across all TIMs to estimate the second-line treatment cost.

Upadacitinib is currently under FDA review, and therefore has no published price. We will assume its price to be the average WAC of the other JAK inhibitors, discounted by 25%, the JAK inhibitor class-level discount derived from SSR Health.[19] Additionally, we will also calculate the price of upadacitinib required to reach cost-effectiveness thresholds ranging from $50,000 per QALY to $150,000 per QALY. All annual prices presented in Table 8 below assume 100% compliance.

For the cost of cDMARDs, we use the mean WAC of the multiple generic versions of methotrexate, aligning with the ICER Reference Case.

Table 8. Drug Costs

Drug	WAC per Unit	Discount from WAC	Net Price per Unit	Annual WAC	Annual Net Price
Upadacitinib – 15 mg Tab	--	25%*	--	$40,284†	$30,213†
Baricitinib (Olumiant®) – 2 mg Tab	$71.23	19%	$57.59	$26,017	$21,033
Tofacitinib (Xeljanz®) – 5 mg Tab	$74.68	34%	$49.50	$54,552	$36,159
Adalimumab (Humira®) – 40 mg/0.8 ml Sol	$2,587.05	34%	$1,696.21	$67,263	$44,102
Methotrexate Sodium (Generic) – 2.5 mg Tab	$2.55	--	$2.55	$398	$398

mg: milligram, WAC: wholesale acquisition cost

*Discount calculated as the average discount estimated for the other two JAK inhibitors.

†Assumed price, calculated as the average of the annual price of the other two JAK inhibitors.

Non-Drug Costs

Administration and Monitoring Costs

Oral treatments will be assumed to have no administration costs. Subcutaneous treatments will include costs for an annual office visit for training on self-administration and as necessary for subcutaneous administration. The administration costs for treatments administered intravenously include the cost for an intravenous infusion administered in a physician's office, calculated by multiplying the hourly infusion cost by the number of hours required for the infusion. Administration cost inputs that will be calculated for each drug are detailed in Table 9. All administration costs represent current 2019 US dollar values.

Table 9. Administration Costs

	Cost	Source
Subcutaneous Injection Administration (HCPCS Code: 96401)	$80.73	Physician's Fee Schedule, Centers for Medicare & Medicaid Services (CMS)[21]
Office Visit (HCPCS Code: 99213)	$75.32	
Intravenous Injection Administration – First Hour (HCPCS Code: 96413)	$143.08	
Intravenous Injection Administration – Each Additional Hour (HCPCS Code: 96415)	$30.99	

HCPCS: Healthcare Common Procedure Coding System

Monitoring Costs

Drug monitoring costs include office visits, tuberculosis tests, liver tests, and complete blood count tests, as appropriate for each medication. Table 10 details monitoring cost inputs. All monitoring costs have been inflated to 2018 US dollar values.

Table 10. Monitoring Costs

	Cost*	Source
Tuberculosis Test (HCPCS Code: 86480)	$84.83	Centers for Medicare & Medicaid Services (CMS)[22]
Liver Function Blood Test Panel (HCPCS Code: 80076)	$7.63	
Complete Blood Cell Count (HCPCS Code: 85025)	$10.65	
Chest X-Ray (HCPCS Code: 71020)	$8.19	

HCPCS: Healthcare Common Procedure Coding System

*Average Medicare Standardized Payment.

Non-Drug Health Care Costs

The cost of hospitalization will be based on the relationship between HAQ and hospitalization, an approach followed by previously published models.[2,3] As seen in Table 11, the number of hospitalization days increases with worsening (increasing) HAQ score. The cost of serious infection will be assumed to be the weighted average cost of treating pneumonia and cellulitis, two commonly occurring serious infections in RA patients (Table 11). This approach is based on that used in the 2017 ICER RA review. Both hospitalization and serious infection treatment costs have been inflated to 2018 US dollar values.

Table 11. Other Key Health Care Costs

HAQ Range	Hospitalization Days Per Year	Cost Per Day of Hospitalization	Source
HAQ: 0 to <0.5 HAQ: 0.5 to <1 HAQ: 1 to <1.5 HAQ: 1.5 to <2 HAQ: 2 to <2.5 HAQ: ≥2.5	0.260 0.130 0.510 0.720 1.860 4.160	$1,330 (95% CI: $931 to $1,729)	Carlson et al., 2015[3]
Cost of Serious Infection*	$9,013		Medicare Provider Utilization and Payment Data, 2016[23]

CI: confidence interval, HAQ: Health Assessment Questionnaire
*Weighted average of costs for pneumonia (2/3) and cellulitis (1/3).

2.6 Model Outcomes

Model outcomes will include life years (LYs) gained, QALYs gained, equal value of LY gained (evLYG), and total costs for each intervention over a lifetime horizon. All outcomes will be reported as discounted values, using a discount rate of 3% per annum. Additionally, we will also include cumulative time in remission for each intervention analyzed.

2.7 Model Analysis

Lifetime cost-effectiveness will be estimated using the incremental cost-effectiveness ratios, with incremental analyses comparing the JAK inhibitors to each other and to comparator treatments, using a health care sector perspective in the base-case analyses. These will include cost per LY gained, cost per QALY, and cost per evLYG outcomes. Additionally, we will present results from a cost per consequence analysis, such as incremental cost per remission.

Sensitivity Analyses

We will conduct one-way sensitivity analyses to identify the impact of parameter uncertainty and key drivers of model outcomes. Probabilistic sensitivity analyses will also be performed by jointly varying all model parameters over 5,000 simulations, then calculating 95% credible range estimates for each model outcome based on the results. We will also perform threshold price analyses across a range of incremental cost-effectiveness ratios (from $50,000 to $150,000 per QALY).

Scenario Analyses

In addition to the base-case analysis, we plan to conduct the following scenario analyses, pending data availability.

1) Modified/restricted societal perspective that includes components such as productivity loss.
2) Including a population that is intolerant to cDMARDs, thus comparing monotherapy with JAK inhibitors to monotherapy with adalimumab.
3) Including a population that has previously failed a TIM therapy.
4) Shorter time horizon, such as one and five years.

Model Validation

We will use several approaches to validate the model. First, we will share preliminary methods with manufacturers, as well as share a write-up of our methods with clinicians and health economics with relevant expertise in the field of RA. Based on feedback from these groups, we will refine data inputs used in the model as needed. Second, we will vary model input parameters to evaluate face validity of changes in results. We will perform model verification for model calculations using internal reviewers. As part of ICER's efforts toward modeling transparency, we will also share the model with the manufacturers for external review around the time of publishing the draft report for this review. Finally, we will compare results to other cost-effectiveness models in this therapy area. The outputs from the model will be validated against the trial/study data of the interventions and also any relevant observational datasets.

References

1. Canhão H, Rodrigues AM, Gregório MJ, et al. Common Evaluations of Disease Activity in Rheumatoid Arthritis Reach Discordant Classifications across Different Populations. *Front Med (Lausanne).* 2018;5:40-40.
2. Incerti D, Jansen J. *A Description of the IVI-RA Model.* Innovation and Value Initiative;2017.
3. Carlson JJ, Ogale S, Dejonckheere F, Sullivan SD. Economic evaluation of tocilizumab monotherapy compared to adalimumab monotherapy in the treatment of severe active rheumatoid arthritis. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2015;18(2):173-179.
4. Stevenson M, Archer R, Tosh J, et al. 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 (Winchester, England).* 2016;20(35):1-610.
5. Stevenson MD, Wailoo AJ, Tosh JC, et al. 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.* 2017;44(7):973-980.
6. Smolen JS, Breedveld FC, Burmester GR, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Annals of the rheumatic diseases.* 2016;75(1):3-15.
7. Gibson L, Alava MH, Wailoo A. *Progression of Disease in People with Rheumatoid Arthritis Treated with Non-Biologic Therapies: Report by the Decision Support Unit.* School of Health and Related Research, University of Sheffield;2015.
8. Michaud K, Wallenstein G, Wolfe F. Treatment and nontreatment predictors of health assessment questionnaire disability progression in rheumatoid arthritis: a longitudinal study of 18,485 patients. *Arthritis Care Res (Hoboken).* 2011;63(3):366-372.
9. Wolfe F, Michaud K. The loss of health status in rheumatoid arthritis and the effect of biologic therapy: a longitudinal observational study. *Arthritis research & therapy.* 2010;12(2):R35-R35.
10. Curtis JR, Jain A, Askling J, et al. A comparison of patient characteristics and outcomes in selected European and U.S. rheumatoid arthritis registries. *Semin Arthritis Rheum.* 2010;40(1):2-14.e11.
11. Strand V, Miller P, Williams SA, Saunders K, Grant S, Kremer J. Discontinuation of Biologic Therapy in Rheumatoid Arthritis: Analysis from the Corrona RA Registry. *Rheumatology and therapy.* 2017;4(2):489-502.
12. Zhang J, Shan Y, Reed G, et al. Thresholds in disease activity for switching biologics in rheumatoid arthritis patients: experience from a large U.S. cohort. *Arthritis Care Res (Hoboken).* 2011;63(12):1672-1679.
13. Human Mortality Database. 2016. https://usa.mortality.org/. Accessed 07/15/2018.
14. Wolfe F, Michaud K, Gefeller O, Choi HK. Predicting mortality in patients with rheumatoid arthritis. *Arthritis and rheumatism.* 2003;48(6):1530-1542.
15. Singh JA, Wells GA, Christensen R, et al. Adverse effects of biologics: a network meta-analysis and Cochrane overview. *The Cochrane database of systematic reviews.* 2011(2):Cd008794.
16. Wailoo AJ, Bansback N, Brennan A, Michaud K, Nixon RM, Wolfe F. Biologic drugs for rheumatoid arthritis in the Medicare program: a cost-effectiveness analysis. *Arthritis and rheumatism.* 2008;58(4):939-946.

17. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Medical care.* 2005;43(3):203-220.
18. Hernandez Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. *Rheumatology (Oxford, England).* 2013;52(5):944-950.
19. SSR Health L. Data on File. In:2019.
20. Redbook. 2019. Accessed July 28, 2019.
21. Physician Fee Schedule Search. 2019. https://www.cms.gov/apps/physician-fee-schedule/search/search-criteria.aspx. Accessed August 14, 2018.
22. Medicare National HCPCS Aggregate Summary Table CY2016. 2016. https://data.cms.gov/Medicare-Physician-Supplier/Medicare-National-HCPCS-Aggregate-Summary-Table-CY/jtra-d83c/data.
23. Medicare Provider Utilization and Payment Data: Physician and Other Supplier. 2016. https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/medicare-provider-charge-data/physician-and-other-supplier.html.