Sarilumab for previously treated moderate or severe rheumatoid arthritis: A Single Technology Appraisal

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Iñigo Bermejo, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Shijie Ren, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Emma Simpson, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Mark Clowes, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK David L Scott, Department of Rheumatology, King's College Hospital NHS Foundation Trust, London, UK Adam Young, Department of Rheumatology, West Hertfordshire Hospital NHS Trust, Hertfordshire, UK Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK
Corresponding Author	Iñigo Bermejo, Research Associate, ScHARR, University of Sheffield, Sheffield, UK
Date completed	Date completed (19/07/2017)

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 19/04/17.

Declared competing interests of the authors

David Scott has no conflicts of interest relating to sarilumab. He has undertaken work for the following companies in rheumatology and related areas in the last 3 years:

1. Eli Lilly And Co. Autumn 2014: Advisory Board Baricitinib, summer 2015: Educational meeting on rheumatoid arthritis
2. Roche Products Ltd. Summer 2014: Advisory Board Biologics in Arthritis
3. Napp Pharmaceuticals. Summer 2014: Advisory Board Biosimilars in Arthritis
4. Baxalta. Autumn 2015: Advisory Board Biosimilars in Arthritis
5. Novartis. Spring 2016: Advisory Board Assessment of Multiple Sclerosis

He was paid between £1000 and £3300 for these various activities.

None of the other authors have any conflicts of interest to declare. None of the authors have any conflicts of interest to declare.

Acknowledgements

We would like to thank Paul Tappenden, ScHARR, for providing comments on the draft report and Andrea Shippam, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Bermejo I, Ren S, Simpson E, Clowes M, Scott DL, Young A, and Stevenson M. Sarilumab for previously treated moderate or severe rheumatoid arthritis: A Single Technology Appraisal: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2017.

Contributions of authors

Inigo Bermejo acted as the project lead and critiqued the health economic analysis submitted by the company. Emma Simpson summarised and critiqued the clinical effectiveness data reported within the company's submission. Shijie Ren critiqued the company's network meta-analysis. Ruth Wong critiqued the company's search strategy. David Scott and Adam Young provided clinical advice to the

team. Matt Stevenson critiqued and summarised the cost-effectiveness model. All authors were involved in drafting and commenting on the final report.

CONTENTS

List of Tables

List of Figures

ABBREVIATIONS

ABT	Abatacept
AC	Appraisal Committee
ACR	American College of Rheumatology
ACR20	20% improvement in the ACR score
ACR50	50% improvement in the ACR score
ACR70	70% improvement in the ACR score
ADA	Adalimumab
AE	Adverse event
AG	Assessment Group
AiC	Academic-in-confidence
AIC	Akaike Information Criterion
ALT	Alanine transaminase
AST	Aspartate transaminase
AZA	Azathioprine
AUC	Area under the curve
BAR	Baricitinib
BD	Twice per day
bDMARD	Biologic disease-modifying antirheumatic drug
BIC	Bayesian Information Criterion
BIW	Twice weekly
BNF	British National Formulary
BSC	Best supportive care
BSRBR	British Society for Rheumatology Biologics Register
cDMARD	Conventional disease-modifying antirheumatic drug
CDAI	Clinical Disease Activity Index
CG	Clinical Guideline
CI	Confidence interval
CIC	Commercial-in-confidence
CODA	Convergence diagnostic and output analysis
CRD	Centre for Reviews and Dissemination
CRP	C-reactive protein
CrI	Credible interval
CS	Company's submission
CSR	Clinical study report
CTZ	Certolizumab pegol

DAS28	Disease Activity Score 28
DAS-CRP	Disease Activity Score C-reactive protein
DES	Discrete event simulation
DSU	Decision Support Unit
DMARD	Disease-modifying antirheumatic drug
eGFR	Estimated glomerular filtration rate
EAIR	Exposure adjusted incidence rate
EMA	European Medicines Agency
EQ-5D	EuroQol 5 Dimensions
EQ-5D-5L	EuroQol 5 Dimensions 5 levels
ERAS	Early RA Study
ERG	Evidence Review Group
ESR	Erythrocyte Sedimentation Rate
ETN	Etanercept
ETN-b	Etanercept biosimilar
EULAR	European League Against Rheumatism
FAD	Final Appraisal Determination
GDG	Guideline Development Group
GLD	Gold injections
GOL	Golimumab
HAQ	Health Assessment Questionnaire
HAQ-DI	Health Assessment Questionnaire Disability Index
HCQ	Hydroxychloroquine
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IFX	Infliximab
IFX-b	Infliximab biosimilar
IL-6	Interleukin-6
INB	Incremental net benefit
IPS	Individual patient simulation
ITT	Intention-to-treat
IV	Intravenous infusion
JAK	Janus kinase
LDA	Low disease activity
LDL	Low-density lipoprotein

LOCF	Last observation carried forward
MACE	Major adverse cardiovascular event
MD-HAQ	Multidimensional HAQ
mITT	Modified intention-to-treat
MJS	Morning joint stiffness
MTA	Multiple Technology Appraisal
mTSS	Modified Total Sharp Score
MTX	Methotrexate
NBT	Non-biologic treatment
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NOAR	Norfolk Arthritis Register
NR	Not reported
NRI	Non-responder imputation
NSAIDs	Non-steroidal anti-inflammatory drugs
OLE	Open-label extension
ONS	Office for National Statistics
PAS	Patient Access Scheme
PASLU	Patient Access Schemes Liaison Unit
PBO	Placebo
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life year
QD	Once a day
QW	Once weekly
Q2W	Every two weeks
Q4W	Every four weeks
Q8W	Every eight weeks
RA	Rheumatoid arthritis
RCT	Randomised controlled trial
RF	Rheumatoid factor
RR	Rate ratio
RTX	Rituximab
SAE	Serious adverse event
SAR	Sarilumab

SC	Subcutaneous
SD	Standard deviation
SDAI	Simplified Disease Activity Index
SE	Standard error
SF-36	Short Form (36) Health Survey
SIR	Sirukumab
SMC	Scottish Medicines Consortium
SmPC	Summary of Product Characteristics
STA	Single Technology Appraisal
SSZ	Sulfasalazine
SW28	Swelling 28 joints
TA	Technology Appraisal
TCZ	Tocilizumab
TEN28	Tenderness 28 joints
TNF	Tumour necrosis factors
TNFi	Tumour necrosis factors inhibitor
TNFi-IR	TNFi inadequate response
TOF	Tofacitinib
TSD	Technical Support Document
VARA	Veteran’s Affairs Rheumatoid Arthritis Registry
WPAI-RA	Work Productivity and Activity Index-Rheumatoid Arthritis

1 SUMMARY

1.1 Critique of the decision problem in the company's submission

The Evidence Review Group (ERG) considers the company's description of the underlying health problem in the company's submission (CS) to be appropriate, mostly up-to-date and relevant to the decision problem set out in the final National Institute for Health and Care Excellence (NICE) scope. The decision problem assesses sarilumab (SAR) for previously treated moderate-to-severe active rheumatoid arthritis (RA). SAR is a new interleukin-6 (IL-6) pathway inhibitor,

1.2 Summary of clinical effectiveness evidence submitted by the company

The key clinical effectiveness evidence for SAR was based on five randomised controlled trials (RCTs). Additionally, one long-term extension study was included. There were three RCTs in methotrexate (MTX) intolerant or inadequate response (MTX-IR) patients with RA (MOBILITY-A, MOBILITY-B, MONARCH). Two RCTs (TARGET and ASCERTAIN) were in patients with RA who had had an inadequate response or were intolerant to biologic disease-modifying antirheumatic drug (bDMARD-IR). One RCT (ASCERTAIN) compared SAR with tocilizumab (TCZ), another study (MONARCH) compared it against adalimumab (ADA), and the remainder compared SAR against placebo (PBO).

Three RCTs had 20% improvement in the American College of Rheumatology (ACR) score (ACR20) as their primary endpoint (MOBILITY-A, MOBILITY-B, TARGET). In the MTX-IR population, the RCTs showed a significant advantage ($p \leq 0.05$) in ACR responses for licensed doses of SAR with concomitant MTX (SAR+MTX) over PBO + MTX (MOBILITY-A, MOBILITY-B), and a significant advantage ($p<0.01$) for SAR monotherapy over ADA monotherapy (MONARCH). In the bDMARD-IR population, TARGET reported a significant advantage for SAR with a concomitant conventional disease-modifying antirheumatic drug (cDMARD) over PBO+cDMARD on ACR20 ($p<0.0001$), ACR50 ($p \leq 0.005$) and ACR70 ($p \leq 0.005$).

[redacted] The most frequent serious adverse event (SAEs) were [redacted] The most common adverse events (AEs) [redacted]

Network meta-analyses (NMA) were performed to assess the relative efficacy and safety of SAR versus the relevant comparators in patients with moderate-to-severe RA who were inadequate responders to cDMARDs (cDMARD-IR) or to tumour necrosis factor inhibitors (TNFi-IR). The efficacy outcome

measures used in the NMA were ACR responses, the Health Assessment Questionnaire Disability Index (HAQ-DI), European League Against Rheumatism (EULAR) responses, Disease Activity Score 28 (DAS28) remission and modified Total Sharp Score (mTSS)). The safety outcome measures included in the NMA were serious infections (SI) and SAEs. In the cDMARD-IR population, separated networks were used for the combination therapies and monotherapies.

In the base case NMA for the cDMARD-IR population, SAR 200mg in combination with cDMARD demonstrated statistically superiority to cDMARD for all the efficacy outcome measures at 24 weeks. SAR 200mg combination therapy was comparable to other bDMARD combination therapies on ACR responses, DAS28 remission and HAQ-DI (subcutaneous [SC] TCZ combination therapy was not included in the HAQ-DI network). SAR 200mg combination therapy showed statistical superiority to ABT combination, infliximab (IFX) combination and intravenous (IV) TCZ 4mg/kg, rituximab (RTX) and SAR 150mg on good EULAR response at 24 weeks, and was comparable to golimumab (GOL) and TCZ IV 8mg/kg, all in combination with cDMARDs. SAR 200mg combination therapy was statistically inferior to certolizumab pegol (CTZ) combination on moderate to good EULAR response at 24 weeks, but comparable to GOL, IFX, TCZ IV 4mg/kg and 8mg/kg, RTX and SAR 150mg all in combination with cDMARDs. For mTSS at 24 weeks, SAR 200mg combination therapy was statistically superior to baricitinib (BAR) 2mg, tofacitinib (TOF) and CTZ all in combination with cDMARDs, and comparable to BAR 4mg, ADA, GOL, TCZ SC 162mg all in combination with cDMARD. For mTSS at 52 weeks, SAR 200mg combination therapy was comparable to ABT, ADA, CTZ and etanercept (ETN) all in combination with cDMARDs, and superior to SAR 150mg combination therapy.

In the NMA evaluating monotherapies in the cDMARD-IR population, all outcome measures were assessed at 24 weeks. SAR 200mg monotherapy showed statistically superiority to placebo and cDMARDs for all efficacy outcome measures, except that it was comparable to cDMARDs on HAQ-DI, and an analysis of DAS28 remission was not performed for placebo. SAR 200mg monotherapy was also statistically superior to ADA on all ACR responses, and sirukumab (SIR) 50mg on ACR20 and ACR50 responses. SAR 200mg was comparable to CTZ, ETN, SIR 100mg, TCZ IV 8mg/kg and TOF on all ACR responses. SAR 200mg was statistically superior to ADA and SIR 50mg on DAS28 remission, and comparable to SIR 100mg and TCZ IV 8mg/kg. SAR 200mg was statistically superior to ADA on HAQ-DI, and comparable to CTZ, ETN and TCZ IV 8mg/kg. SAR 200mg was statistically superior to ADA on EULAR responses, and comparable to TCZ 8mg/kg.

In the NMA for the TNFi-IR population, the outcome measures were all assessed at 24 weeks. SAR 200mg combination therapy showed statistically superiority to cDMARDs for all efficacy outcome measures. SAR 200mg combination was statistically superior to BAR 2mg combination, SIR 50mg combination on ACR50, and comparable to other bDMARD combination therapies on all ACR

responses. SAR 200mg combination therapy was statistically superior to ABT, BAR 2mg, GOL, SIR 50mg, TCZ IV 4mg/kg, and RTX combination therapies on DAS28 remission, and comparable to other bDMARD combination therapies. SAR 200mg was not statistically significantly different to other bDMARDs on changes in HAQ-DI. For good EULAR response, SAR 200mg combination therapy was statistically superior to RTX combination therapy, and comparable to ABT and SAR 150mg combination therapies. For at least a moderate EULAR response, SAR 200mg combination therapy was statistically inferior to TCZ 8mg/kg and RTX combination therapies, and comparable to ABT, GOL and SAR 150mg combination therapies.

Regarding safety, SAR 200mg combination therapy was associated with significantly higher odds of SAEs at 52 weeks when compared to cDMARDs in the cDMARD-IR population. All other outcomes were not statistically significant.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The ERG believes that all available RCTs informing on the clinical effectiveness of SAR were included in the CS. The eligibility criteria applied in the selection of evidence for the clinical effectiveness review were considered by the ERG to be reasonable and consistent with the decision problem outlined in the final NICE scope. The quality of the included RCTs was assessed using well-established and recognised criteria.

The ERG believes that the results presented in the NMAs of clinical effectiveness should be treated with caution, as the statistically significant results of SAR 200mg compared with other bDMARD treatments (both as combination therapy and monotherapy) may be a consequence of underestimating the uncertainty in treatment effects resulting from the use of a fixed effect model. The ordered categorical ACR response and EULAR response data were dichotomised in the NMA, which ignores the natural ordering and correlations between the categories within the outcome measure. When a risk difference model was used for binary data, the probability of response was not constrained to be below or equal to 1, potentially producing invalid probability values. Furthermore, the MOBILITY B and TARGET trial designs allowed patients who did not achieve a ≥20% improvement from baseline at two consecutive assessments in the swollen joint count or tender joint count to switch to open-label SAR 200mg at 16 and 12 weeks, respectively. Non-responder imputation was carried out for the control arm, assuming none of the non-responders in the cDMARD control group would become responders at 24 weeks, which may overestimate the relative treatment effect of SAR combination therapy versus cDMARDs.

1.4 Summary of cost effectiveness submitted evidence by the company

The company supplied a *de novo* individual patient-level Markov model constructed in Microsoft Excel®. The model, which has a cycle length of 6 months, simulates patients' disease progressions through the sequences of treatments being compared. For each treatment, patients may achieve good, moderate or no EULAR response: this is assessed at 6 months. The EULAR response rates for each treatment are based on the ACR response rates calculated using the company's NMA. Patients who achieve moderate or good EULAR response are assumed to have an improvement in Health Assessment Questionnaire (HAQ) score and remain on treatment until loss of efficacy (as assessed by a clinician), or until they experience an AE or death. Patients who fail to achieve a moderate or good EULAR response discontinue treatment at 6 months and initiate the next treatment in the sequence. HAQ progression whilst on treatment is assumed to be constant on bDMARDs and SAR; conversely, whilst on cDMARDs and best supportive care (BSC), HAQ progression is assumed to be linear. Time to treatment discontinuation for responders is dependent on the type of treatment (TNFi, IL-6, others) and is modelled using survival curves fitted to treatment discontinuation data from the Canadian observational database RHUMADATA. Upon treatment discontinuation, patients are assumed to experience a rebound in HAQ equal to that achieved on treatment initiation and then start on the next treatment in the sequence. The mortality rate is assumed to be affected by the HAQ score of a patient at treatment initiation. The model estimates the costs and quality-adjusted life years (QALYs) accrued over patients' remaining lifetimes. EuroQol 5 Dimensions (EQ-5D) values are estimated based on a mapping algorithm from HAQ scores and patient characteristics. Hospitalisation costs and resource use estimates were based on HAQ score bands as in previous NICE technology appraisals. Unit costs were taken from the British National Formulary and NHS Reference Costs. Serious infection were the only AE included in the analyses.

The company's analyses relate to seven different populations of rheumatoid arthritis patients: (1) cDMARD-IR patients with severe RA who can tolerate MTX; (2) cDMARD-IR patients with severe RA for whom MTX is contraindicated or not tolerated; (3) TNFi-IR patients with severe RA and who are rituximab (RTX) eligible; (4) TNFi-IR patients with severe RA for whom RTX is not an option; (5) TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated; (6) TNFi-IR patients with severe RA after treatment with RTX+MTX; and, (7) a subgroup of cDMARD-IR patients with moderate RA whose DAS28 scores are between 4.0 and 5.1. The definition of severe RA was a DAS28 score higher than 5.1, whilst moderate RA was defined as a DAS28 > 3.2 and ≤ 5.1. Baseline characteristics of patients are based on the relevant clinical SAR trials.

The company presented analyses in the CS and in the clarification response as per the ERG's request. The ERG believes that the analyses presented by the company in the clarification responses are closer to the company's intended base case than those reported in the CS. According to the company's revised

probabilistic analysis, in the cDMARD-IR population of patients with severe RA who could tolerate MTX, SAR with concomitant MTX (SAR+MTX) dominated both indications of TCZ with concomitant MTX and the incremental cost-effectiveness ratios (ICERs) for SAR+MTX compared with a weighted average of TNFi-s (TNFi bundle) and ABT (SC) + MTX were £69,884 and £117,482 per QALY gained respectively. In cDMARD-IR patients with severe RA who could not tolerate MTX, the deterministic ICER for SAR monotherapy compared with the TNFi bundle was estimated to be £17,123 per QALY gained, whilst the ICER for both TCZ indications compared with SAR was in excess of £1,000,000 per QALY gained. In TNFi-IR patients for whom RTX+MTX was an option, the ICER for SAR+MTX compared with RTX+MTX was estimated to be £130,691 per QALY gained. In patients for whom RTX is not an option, the ICER for the comparators versus SAR+MTX in TNFi-IR patients was greater than £60,000 per QALY. For TNFi-IR patients who cannot tolerate MTX, the ICER for SAR monotherapy compared with a TNFi bundle was estimated to be £17,794 per QALY gained. In patients who have received RTX+MTX, the ICER for both indications of TCZ compared with SAR+MTX were estimated to be greater than £130,000 per QALY gained. Finally, in cDMARD-IR patients with moderate RA and a DAS28 score higher than 4.0, the ICER for SAR+MTX was estimated to be £38,254 per QALY gained.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The company's model was based on the model developed by the Assessment Group (AG) in NICE Technology Appraisal 375 (TA375) but was an individual patient level Markov model rather than a discrete event simulation (DES). The ERG believes that the conceptual model was broadly appropriate.

After an initial evaluation of the company's analyses, the ERG requested that the company perform new analyses after addressing a number of issues. The company presented new analysis after addressing the following issues: (i) inadequate treatment sequences that did not reflect NICE recommendations; (ii) omission of the possibility of patients with moderate RA to progress to the severe state; (iii) use of Malottki *et al.* instead of Hernandez *et al.* for the mapping of HAQ scores to EQ-5D; (iv) limitations in the company's NMA explained in Section 1.3; (v) using percentages of improvement of HAQ instead of absolute mean changes; (vi) omission of rounding to the nearest valid HAQ score; (vii) use of an implausible extrapolation of time to treatment discontinuation; (viii) using independent samples for the probabilities of ACR responses in the PSA instead of correlated samples from the CODA of the NMA; (ix) assuming 9 free doses of CTZ instead of 10; and, (x) the inclusion of the speculative Patient Access Scheme (PAS) discount of 15% applied to TCZ and ABT.

The main issue remaining in the company's analyses after these amendments is the assumption that the HAQ score of patients on cDMARDs and BSC follow a linear trajectory. The ERG notes that there is extensive evidence that shows that the HAQ trajectory for these patients is not linear and that the appraisal committee for TA375 accepted the non-linear trajectory of HAQ scores using the latent class

approach used by the AG. The ERG notes that the company's assumption of linear HAQ increase is likely to lead to lower ICER estimations for SAR+MTX in the moderate RA population with a DAS28 score between 4.0 and 5.1 compared with a non-linear trajectory approach.

A further issue in the company's amended model is the inadequate implementation of the transition from moderate to severe RA. The ERG notes that patients should progress to the severe sequences at the point when their DAS28 score increases above 5.1, without waiting until they have reached the end of the moderate sequence.

The company did not present analyses comparing SAR to all other recommended bDMARDs independently; instead, the company created a blended comparator grouping all the TNFi-s together. The ERG believes that presenting analyses including the TNFi-s independently would have been more informative, given the differences in cost and efficacy of different TNFi-s and the fact that their market shares are currently changing.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

The ERG believes that all available RCTs of SAR were included in the CS. The five SAR RCTs included were considered to be of good methodological quality in terms of randomisation, blinding and performing intention-to-treat analyses.

The ERG notes that the model and analyses submitted after the clarification process appears to be conceptually appropriate with only two relevant limitations.

1.6.2 Weaknesses and areas of uncertainty

The ERG notes that the natural comparators for SAR are the other biologic agents specifically targeting the IL-6 pathway, e.g. TCZ, SIR, olokizumab, and clazakizumab. Of these IL-6 pathway inhibitors, only TCZ was approved by the European Medicines Agency at the time of writing, and prior to the start of the SAR trials. Therefore, TCZ should be the main comparator of SAR, but only one head-to-head comparison study (ASCERTAIN) was identified. The ERG notes that the primary endpoint of this study relates to safety, rather than efficacy outcomes. In addition, the ERG notes that some patients on the TCZ arm of ASCERTAIN received only half the recommended dose according to UK prescription guidance (4mg/kg instead of 8mg/kg). Only one head-to-head efficacy RCT was identified against another bDMARD, a monotherapy study against ADA. The ERG notes that in the ADACTA study, ADA monotherapy had previously been shown to be statistically significantly inferior to TCZ monotherapy in terms of DAS28 and ACR and EULAR responses.

The ERG believes that the DES paradigm is more appropriate to represent the disease than the individual patient Markov model approach used by the company.

In the company's amended model HAQ progression is still assumed to be linear for patients on cDMARDs and on BSC. This approach was used in previous appraisals but there has been since extensive evidence published against the appropriateness of this assumption. In line with this evidence, the AG in TA375 used a latent class approach of non-linear trajectories.

The company used a blended comparator grouping all the TNFi-s together, which may obscure the cost-effectiveness of SAR. This weakness has been alleviated by the company to a certain extent by including sensitivity analyses where TNFi-s have been considered separately in some of the populations in the CS but not in the clarification response.

In some populations the company used effectiveness estimates calculated from similar but different populations due to lack of available evidence: in TNFi-IR patients who could not tolerate MTX, the company used the effectiveness of therapies in combination with MTX; and in TNFi-IR patients who had received RTX + MTX, the effectiveness estimates calculated from the general TNFi-IR population were used.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG undertook exploratory analyses after implementing the latent class approach of non-linear HAQ trajectories used in TA375 and amending the transition of moderate RA patients to the severe state.

According to the ERG's exploratory analyses, in cDMARD-IR patients with severe RA who can tolerate MTX, SAR + MTX was estimated to dominate both indications of TCZ with concomitant MTX and the ICERs for TNFi bundle + MTX and ABT (SC) + MTX compared with SAR + MTX are estimated to be in excess of £150,000 per QALY gained. In cDMARD-IR patients with severe RA for whom MTX was contraindicated or not tolerated, the ICER for SAR monotherapy compared with TNFi bundle monotherapy was estimated to be £34,422 per QALY gained, whilst the ICERs for both indications of TCZ compared with SAR monotherapy where estimated to be in excess of £1,500,000 per QALY gained. In TNFi-IR patients with severe RA who can tolerate RTX and MTX the ICER for SAR+MTX compared with RTX+MTX was estimated to be £171,466 per QALY gained. In TNFi-IR patients with severe RA for whom RTX is not an option, SAR + MTX was estimated to result in an ICER of £34,979 per QALY gained compared with TNFi bundle whilst the ICER for both TCZ indications with concomitant MTX compared with SAR + MTX was estimated to be in excess of £195,000. In TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated, the ICER for SAR monotherapy compared with TNFi bundle was estimated to be £31,433 per QALY

gained. In TNFi-IR patients who have already received RTX+MTX, the ICERs for both indications of TCZ with concomitant MTX compared with SAR+MTX were estimated to be in excess of £200,000 per QALY gained. In cDMARD-IR patients moderate RA and a DAS28 higher than 4.0, a sequence starting with SAR+MTX compared with MTX was estimated to result in an ICER of £63,438 per QALY gained.

The ERG notes that the confidential PASs in place for ABT and TCZ were not included in these analyses. The ERG presents the analyses including the confidential PASs in a confidential appendix.

2 BACKGROUND

2.1 Critique of company's description of underlying health problem

The Evidence Review Group (ERG) considers the company's description of the underlying health problem in the company's submission (CS)[2] to be appropriate, mostly up-to-date and relevant to the decision problem set out in the final National Institute for Health and Care Excellence (NICE) scope.[3] The ERG provides a brief summary of the underlying health problem. Epidemiological numbers provided by the ERG may differ from those presented in the CS but do not affect the broad messages.

Clinical features of rheumatoid arthritis

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterised by: progressive, irreversible, joint damage; impaired joint function; pain and tenderness caused by swelling of the synovial lining of joints.[4] The condition is associated with increasing disability and reduced health-related quality of life.[4] The primary symptoms are: pain; morning stiffness; swelling; tenderness; loss of movement; redness of the peripheral joints; and fatigue.[5, 6] RA is associated with substantial direct costs (including drug acquisition and hospitalisation) and indirect costs (including reduced productivity).[7] RA has long been reported as being associated with increased mortality,[8, 9] particularly due to cardiovascular events.[10]

Epidemiology

NICE estimates that there are 400,000 people in the UK with RA,[11] based on a prevalence of 0.8% reported by Symmons *et al*.[12] The incidence of RA is greater in females (3.6 per 100,000 per year) than in males (1.5 per 100,000 per year).[13] For both genders, the peak age of incidence in the UK is in the eighth decade of life, but all ages can develop the disease.[13]

Aetiology

There is no identified specific cause for RA, but there seems to be a variety of contributing factors such as genetic and environmental influences. Genetic factors have a substantial contribution to RA. The heritability of RA is estimated to be between 53 and 65%[14] and family history of RA has a corresponding risk ratio of 1.6 compared with the general population.[15] Many genes associated with RA susceptibility are concerned with immune regulation. Infectious agents have been suspected but no consistent relationship with an infective agent has been proven. Similarly, sex hormones have been suspected due to the higher prevalence of RA in women and a tendency for the disease to improve during pregnancy. However, a precise relationship has not been identified. There is no proof of any causal link with lifestyle factors such as diet, smoking, or occupation.

Management of rheumatoid arthritis

Traditionally, patients have been treated with conventional disease-modifying anti-rheumatic drugs (cDMARDs) which include methotrexate (MTX), sulfasalazine (SSZ), hydroxychloroquine (HCQ), leflunomide (LEF), and gold injections as well as corticosteroids, analgesics and non-steroidal anti-inflammatory drugs (NSAIDs). However, more recently, a group of biologic immunosuppressant drugs have been developed that specifically modify the disease process by blocking key protein messenger molecules (such as cytokines) or cells (such as B-lymphocytes).[11] Such drugs have been labelled as biologic disease-modifying anti-rheumatic drugs (bDMARDs): certolizumab pegol (CTZ); adalimumab (ADA); etanercept (ETN); golimumab (GOL); and infliximab (IFX) are tumour necrosis factor (TNF) inhibitors (or antagonists) (TNFi). Of the remaining bDMARDs, tocilizumab (TCZ) is a cytokine interleukin-6 (IL-6) inhibitor; abatacept (ABT) is a selective modulator of the T lymphocyte activation pathway; and rituximab (RTX) is a monoclonal antibody against the CD20 protein. For patients who have exhausted all NICE recommended treatments, palliative care, also known as best supportive care (BSC), is the final treatment option.

Assessment of response to therapy

The initial response criteria for RA were produced in 1987 by the American College of Rheumatology[16] (ACR). NICE Clinical Guideline (CG) 79 provides a summary of the ACR criteria, namely that patients must have at least four of seven criteria: (i) morning stiffness lasting at least 1 hour; (ii) swelling in three or more joints; (iii) swelling in hand joints; (iv) symmetric joint swelling; (v) erosions or decalcification on X-ray of hand; (vi) rheumatoid nodules; (vii) and abnormal serum rheumatoid factor. For the first four criteria, these must have been present for a period of at least six weeks. However, in NICE CG 79, the Guideline Development Group (GDG) preferred a clinical diagnosis of RA rather than the ACR criteria referencing recommendations from the European League Against Rheumatism (EULAR)[17] stating that "*an early persistent synovitis where other pathologies have been ruled out needs to treated as if it is RA to try to prevent damage to joints. Identification of persistent synovitis and appropriate early management is more important than whether the disease satisfies classification criteria*".

In 2010, the ACR and EULAR jointly published RA classification criteria, which focussed on the features at earlier stages of disease which are associated with persistent and/or erosive disease, rather than defining the disease by its late stage features.[18] The classification criteria allocate scores to characteristics of joint involvement, serology, acute-phase reactants, and duration of symptoms, to produce a score between 0 and 10 inclusive. Those patients scoring 6 or greater and with obvious clinical synovitis being defined as having "definite RA" in the absence of an alternative diagnosis that better explains the synovitis.

Two classifications have dominated the measurement of improvement in RA symptoms: ACR responses[19] and EULAR responses.[20]

The initial ACR response 'ACR20', required: a 20% improvement in tender joint counts; a 20% improvement in swollen joint counts; and a 20% improvement in at least three of the following five 'core set items': physician global assessment; patient global assessment; patient pain; self-reported disability (using a validated instrument), and; erythrocyte sedimentation rate (ESR) / C-reactive protein (CRP).

ACR response has been widely adopted in randomised controlled trials (RCTs) although studies have shown that the value of the measure can vary between trials due to the timing of the response.[21] Since the inception of the ACR20, two further response criteria (ACR50 and ACR70) have become widely used. These are similar to ACR20 and differ only in the level of percentage improvements required to be classified as a responder. These are nested responses, thus patients who achieve ACR70 will also achieve ACR20 and ACR50.

In the UK, monitoring the progression of RA is often undertaken using the disease activity score of 28 joints (DAS28) in terms of swelling (SW28) and of tenderness to the touch (TEN28). The DAS28 score incorporates measures of the ESR and a subjective assessment on a scale of 0-100 made by the patient regarding disease activity in the previous week.

The equation for calculating DAS28 is as follows:[22]

$$DAS28 = 0.56* TEN28^{0.5} + 28* SW28^{0.5} + 0.70 * \ln(ESR) + 0.014 * \text{subjective assessment}$$

The DAS28 can be used to classify both the disease activity of the patient and the level of improvement estimated within the patient.

A second version of DAS28, using C-reactive protein (CRP) rather than ESR exists. However, as the majority of studies have used DAS28 ESR, this is the metric used by the company in assessing comparative effectiveness between interventions.

The EULAR response criteria use the individual change in DAS28 and the absolute DAS28 score to classify a EULAR response as: good; moderate; or none.[20] The EULAR response criteria and the ACR20 improvement criteria were found to have reasonable agreement in the same set of clinical trials, although van Gestel *et al.* state that the EULAR response criteria showed better construct and discriminant validity than ACR20.[23] EULAR response has been reported less frequently in RCTs than ACR responses,[24] although EULAR is much more closely aligned to the treatment continuation rules

stipulated by NICE for treatment in England. These rules require either a moderate or good EULAR response or a DAS28 improvement of more than 1.2 to continue treatment, with the latter criterion applying to RTX. The relationship between change in DAS28 and the absolute DAS28 score and EULAR response is shown in Table 1.

Table 1: Determining EULAR response based on DAS28[23]

	Improvement in DAS 28		
DAS28 at endpoint	>1.2	>0.6 and ≤1.2	≤0.6
≤ 3.2	Good	Moderate	None
>3.2 and ≤5.1	Moderate	Moderate	None
>5.1	Moderate	None	None

Patients with a DAS28 ≤3.2 are regarded as having low disease activity, those with a DAS28 > 3.2 and ≤ 5.1 are regarded as having moderate disease and >5.1 as having very active disease.[22] Within NICE Technology Appraisal (TA) 375, patients with a DAS28 > 3.2 and ≤ 5.1 were considered as having moderate to severe disease whilst those with a DAS28 > 5.1 were denoted as having severe disease.[25]

A widely used measure of patient disability is the Health Assessment Questionnaire (HAQ). The HAQ score is a patient completed disability assessment which has established reliability and validity.[26] HAQ scores range from 0 to 3, with higher scores indicating greater disability, and is a discrete scale with step values of 0.125, resulting in the HAQ scale containing 25 points. The HAQ has been used in many published RCTs in RA.[24]

2.2 Critique of company's overview of current service provision

The company's overview of current service but appropriate and relevant to the decision problem set out in the final NICE scope. The ERG provides a summary of current service provision below.

Clinical guidelines

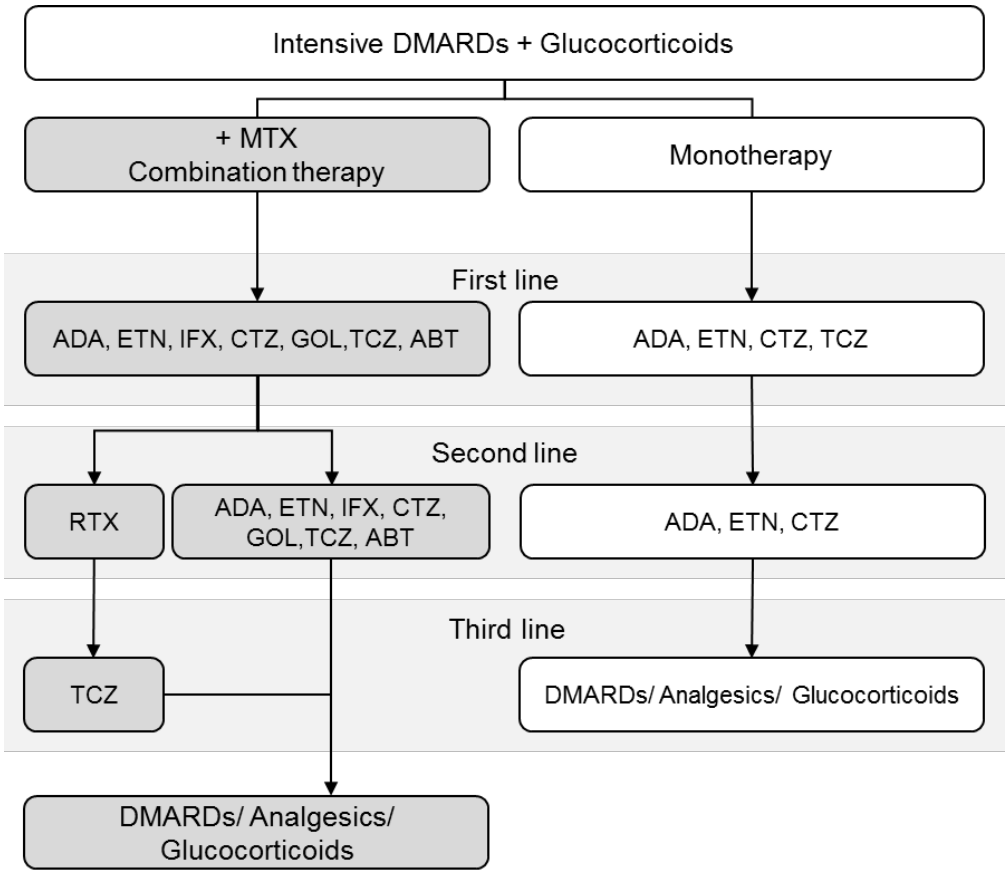
For people with newly diagnosed RA, NICE CG79[11] recommends a combination of cDMARDs (including MTX and at least one other cDMARD plus short-term glucocorticoids) as first-line treatment, ideally beginning within 3 months of the onset of persistent symptoms. Where combination therapies are not appropriate, for example, where there are comorbidities or pregnancy, cDMARD monotherapy is recommended. Where cDMARD monotherapy is used, efforts should be made to increase the dose quickly to obtain best disease control. For the purposes of this assessment, the term "intensive cDMARDs" has been used to denote that this involves treatment with multiple cDMARDs simultaneously.

NICE guidance (TA375)[25] recommends the use of ABT, ADA, CTZ, ETN, GOL, IFX, and TCZ in combination with MTX in people with RA after the failure to respond to intensive cDMARD treatment and who have severe active RA (defined as a DAS28 score > 5.1). For people who meet these criteria but cannot take MTX because it is contraindicated or because of intolerance, TA375[25] recommends the following bDMARDs as monotherapy options: ADA; CTZ; ETN; or TCZ.

After the failure of the first TNF-inhibitor, TA195[27] recommends RTX in combination with MTX for the treatment of severe active RA. If RTX is contraindicated or withdrawn because of an adverse event (AE), TA195 recommends ABT, ADA, ETN, or IFX in combination with MTX. If MTX is contraindicated, or withdrawn because of an AE, TA195 recommends ADA or ETN as monotherapy. TA247[28] recommends TCZ, and TA415[29] recommends CTZ as alternatives to TNF-inhibitors in the same circumstances as TA195, that is, after the failure of a TNF-inhibitor in patients with severe active RA, in combination with MTX when RTX is contraindicated or withdrawn and as monotherapy if MTX is contraindicated or withdrawn. In addition, TA247 recommends TCZ in combination with MTX in patients in whom TNF-inhibitors and RTX have not worked.

The summary of the NICE recommended treatment pathway for RA presented in the CS is reproduced in Figure 1.

Figure 1: Treatment pathway presented in the CS modified by the ERG



ABT=abatacept; ADA=adalimumab; CTZ=certolizumab pegol; DMARD=disease-modifying anti-rheumatic drug; ETN=etanercept; GOL=golimumab; IFX=infliximab; MTS=methotrexate; RTX=rituximab; TCZ=tocilizumab.

A key pathway for those who are RTX and MTX tolerant is a sequence of an initial bDMARD+MTX, followed by RTX+MTX and then TCZ+MTX, which typically uses three different classes of intervention. However, for those that cannot receive RTX+MTX, it is possible that a second TNFi would be used. In Figure 3.2 (p43) of the CS the company report evidence that the effectiveness of TNFi in terms of EULAR response diminishes as the number of prior TNFis used increases. Clinical advice provided to the ERG states that this result is not unexpected and that clinicians would try to avoid using a second TNFi where possible.

NICE criteria for continuing treatment

NICE TA375[25] states that for patients to continue treatment with their first bDMARD treatment they must maintain at least a moderate EULAR response. TA195,[27] which for all bDMARDs excluding RTX was updated in TA375[25], states that bDMARD treatment after the failure of a TNFi should be continued only if there is an adequate response (defined as an improvement in the DAS28 score of ≥ 1.2 points) at initiation of treatment and as long as this adequate response is maintained. If the criterion of having

at least a moderate EULAR response at six months has not been met, then treatment should be stopped and the next intervention in the sequence should be initiated.

3 CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM

3.1 Population

Sarilumab (Kevzara®) is licensed in the UK for the treatment of moderate to severe active RA in adult patients who have responded inadequately to, or who are intolerant to one or more DMARDs. Treatment can be provided with MTX or as a monotherapy if a patient is intolerant to MTX or where treatment with MTX is inappropriate. The target population in the company's decision problem aligns with the populations described in the final scope issued by NICE, although only a subgroup of patients with moderate RA have been evaluated. The company describe the patient populations analysed in the model in Table 5.3 of the CS; the company have assigned alpha-numeric codes for each population, although the ERG did not find these overly helpful and have renamed the populations. The populations are as follows:

- Patients with severe RA who have had an inadequate response to cDMARDs (cDMARD-IR) who can tolerate MTX
- Patients with severe RA who are cDMARD-IR who cannot tolerate MTX
- Patients with severe RA who have are TNFi-IR who can tolerate RTX
- Patients with severe RA who are TNFi-IR who cannot tolerate RTX
- Patients with severe RA who are TNFi-IR who cannot tolerate MTX
- Patients with severe RA who are RTX+MTX-IR
- Patients with moderate RA with DAS28> 4.0 who are cDMARD-IR who can tolerate MTX

Within the categorisation of patients, the company have assumed that severe disease is represented by a DAS28 >5.1. The ERG comments that this is in line with TA375.[25] The company have presented analyses for a subgroup of moderate patients, determined by having a DAS28 score of 4.0, considered by the company of being "*at risk of rapid progression*". The company argue that there is an unmet need in this population, where biologics are not recommended. The ERG notes that although the wording is similar, the definition of this subgroup is unrelated to the "fast progressors" discussed in work by the NICE Decision Support Unit (DSU).[30] The ERG also notes that moderate patients at risk of rapid progression are likely to progress to the severe state, where sequences of bDMARDs are recommended.

3.2 Intervention

Sarilumab (SAR) is a fully human immunoglobulin G1 monoclonal antibody that inhibits interleukin-6 (IL-6) mediating signalling. SAR is administered subcutaneously (SC) every other week (Q2W) and has two doses (200mg and 150mg). The contraindications to SAR listed in the draft Summary of Product Characteristics (SmPC) are hypersensitivity to the active substance or any of the excipients—histidine,

arginine, polysorbate 20 and sucrose. Reduction of dose from 200mg Q2W to 150mg Q2W is recommended for management of neutropenia, thrombocytopenia, and liver enzyme elevations. The efficacy and safety of SAR has not been studied in patients with hepatic impairment or in children aged up to 18 years of age, although paediatric studies are ongoing.

The list price for SAR is [redacted] per pen or syringe, the prices for both the 150mg and 200mg doses being the same. A Patient Access Scheme (PAS) in the form of a simple discount ([redacted]) has been agreed with the Department of Health which reduces the cost per pen or syringe to [redacted]. The cost-effectiveness results presented by the company are based on the PAS price.

3.3 Comparators

The comparators considered in the decision problem are shown in Table 2.

Table 2: Comparators to SAR considered in the CS (adapted from Table 5.5 of the CS)

Drug	Dose	Frequency
SAR SC[b]	200mg	Every other week
ABT IV[a]	500mg if <60 kg, 750mg if 60–100 kg, 1,000mg if > 100 kg	Week 0, 2, 4, then every 4 weeks
ABT SC[a]	125mg SC injections	Once per week
GOL SC[a]	50mg	Once per month
ETN SC[a]	25mg	Twice weekly
ETN biosimilar SC[a]	50mg	Every week
ADA SC[a]	40mg	Every other week
RTX IV[c]	2,000mg	Two 1,000mg IV infusions separated by 2 weeks (one course) every 9 months
CTZ SC[a]	400mg induction dose, 200mg maintenance dose	400mg dose at week 0, 2, and 4, followed by maintenance dose every other week
TCZ IV[a]	8mg/kg	Every 4 weeks
TCZ SC[a]	162mg SC	Every week
IFX IV[a]	3mg/kg	Week 0, 2 and 6, then every 8 weeks.
IFX biosimilar IV[a]		

a https://www.medicines.org.uk/emc/ b Draft SmPC c. TA375
ABT= abatacept; ADA= adalimumab; CTZ= certolizumab pegol; ETN: etanercept; GOL= golimumab; IFX= infliximab; IV = intravenous; RA = rheumatoid arthritis; Q2W, once every 2 weeks; RTX= rituximab; SAR= sarilumab; SC = subcutaneous; TCZ= tocilizumab

The comparators in the CS are largely in line with the final scope issued by NICE although biosimilars for ADA and RTX were not considered. MTX alone was not included as a comparator in the cDMARD-IR with severe RA population or the TNFi-IR population with severe RA, presumably because bDMARDs are recommended by NICE in this population. BSC was used as a comparator in the cDMARD-IR with moderate RA population and included at the end of every sequence in all populations but was omitted from Table 2.

3.4 Outcomes

The outcomes contained in the final scope issued were all addressed in the CS with the exception of extra-articular manifestations of disease where no data related to SAR were identified.

4 CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

This chapter presents a review of the clinical effectiveness evidence provided in the CS for SAR for treating RA. The clinical evidence provided in the CS comprised a systematic review of SAR and comparators for treating RA, used to provide effectiveness and safety data for SAR, and to populate a network meta-analysis (NMA).

4.1.1 Searches

The search strategies are reproduced in the Appendices (Section 5) of the CS. The company's searches were well-designed and are appropriately structured to include population, interventions of interest (SAR and comparator drugs) and study types (using a recognised RCT filter). An appropriate range of databases was searched (Medline; EMBASE and the Cochrane Library) in accordance with (NICE) guidelines. The ERG queried whether any additional steps had been taken to capture the latest evidence and the company confirmed that Medline In Process had been included in the search (see clarification response[31] – Literature searching, Q1). Initial searches covered the period from inception to March 2015 whilst update searches undertaken in December 2016 covered the period since March 2015. The ERG is broadly confident that all relevant published studies have been identified by the search process.

4.1.2 Inclusion criteria

The CS review (Table 3) was carried out in two stages, based on initial and update searches. The update differed from the initial search in being restricted to "investigational drugs to those likely to be relevant future comparators for" SAR (page 54 of the CS). In practice, these were baricitinib and sirukumab (CS Table 4.2). The update was also limited to studies of 12 weeks or more duration (see clarification response[31] – question A5). These restrictions were appropriate given the decision problem. Inclusion/exclusion criteria are shown in the CS as Table 4.1 for the initial search, and CS Table 4.2 for the update search. The intervention (technology of interest) was SAR monotherapy or in combination with cDMARDs. Other interventions / comparators of cDMARDs and bDMARDs were included to populate the NMA. Selection criteria were in accordance with the decision problem in the final NICE scope. Inclusion and exclusion criteria are shown in Table 3 of the ERG report.

Table 3: CS Review inclusion/exclusion criteria (Reproduced from CS Table 4.4)

Criteria		Inclusion	Exclusion
STUDY DESIGN	Abstract selection	RCTs above Phase I	• Case series/reports, letters to editor, commentary, editorials • Observational and registry studies • Non-English publications • Preclinical/Pharmacokinetic/ Pharmacogenomic studies • Animal or in vitro studies • Literature review/meta-analysis[a] • Phase I study • Prognostic study • Retrospective study • Open-label extension and extended access studies • *Post hoc* studies and pooled analyses[a] • Any other type of non-randomised study
	Full-text selection	RCT above Phase I	
POPULATION	Abstract and full-text selection	• Adult patients (≥18 years) with moderately to severely active RA who have had inadequate response to one or more cDMARDs • Adult patients (≥18 years) with moderately to severely active RA who have had inadequate response to one or more bDMARDS (TNFi or another MoA) • Adult patients (≥18 years) intolerant to MTX or for whom continued MTX is inappropriate	• Patients without RA • Patients with diseases other than RA • Patients with rheumatic diseases other than RA • Patients not being treated with an intervention of interest • Patients naïve for cDMARD
TREATMENT / INTERVENTION	Abstract and full-text selection	The following interventions are of interest at any dosage or administration type: • Sarilumab (REGN88, sarilumab153191) • Etanercept (Enbrel) • Tocilizumab (RoActemra/Actemra) • Adalimumab (Humira) • Abatacept (Orencia) • Infliximab (Remicade) • Rituximab	Other treatments

Criteria		Inclusion	Exclusion
		(MabThera/Rituxan) • Tofacitinib (Xeljanz) • Anakinra (Kineret) • Certolizumab (Cimzia) • Golimumab (Simponi) • Biosimilar DMARDs (CS Appendix 5.3) • Investigational drugs (CS Appendix 5.4)	
COMPARATOR	Abstract and full-text selection)	Placebo or any of the above listed treatments as monotherapy or in combination with a cDMARD(s) (i.e. MTX, leflunomide, hydroxychloroquine, minocycline, sulfasalazine, azathioprine, sodium aurothiomalate, and auranofin) or cDMARD as monotherapy or in combination with other cDMARD(s)	Other treatments not in the above listed treatments
OUTCOMES	Abstract and full-text selection	No selection was made on outcomes. *After the screening phase top-line data extraction was performed to detect which outcomes were selected for data extraction*	None[b]
Timepoint		No start limit – 31st March 2015[c]	
Language		English language	Non-English language

[a]Systematic literature reviews and meta-analyses (2010 – present) will be noted in a separate "study design" exclusion column; using this list of reviews, we will select the most recent and relevant systematic literature reviews/meta-analyses and check the reference lists of the reviews for relevant studies. For *post hoc* and pooled analyses, the reference list was also checked for relevant studies.

[b]Studies were not excluded based on the outcomes at the screening phase. Outcomes were selected during the top-line data extraction phase.

PICOS-T = population, intervention, comparison, outcomes, study, and time horizon.

bDMARD= biological disease-modifying anti-rheumatic drug; cDMARD=conventional disease-modifying anti-rheumatic drug; MoA=mode of action; MTX=methotrexate; RA=rheumatoid arthritis; RCT=randomised controlled trial

Note: These exclusion criteria, along with the PICOS-T criteria noted in Table 4.1 were applied during the abstract and full-text screening process to select appropriate studies.

[c]Update searches up to 6th December 2016

The population was adults with moderate to severe, active rheumatoid arthritis, whose disease has not responded adequately to, or who are intolerant of cDMARDs or bDMARDs. The intervention was SAR as monotherapy or in combination with cDMARDs.

Comparators included were: etanercept (Enbrel); tocilizumab (RoActemra/Actemra); adalimumab

(Humira); abatacept (Orencia); infliximab (Remicade); rituximab (MabThera/Rituxan); tofacitinib (Xeljanz); anakinra (Kineret); certolizumab (Cimzia); golimumab (Simponi); biosimilars. Investigational drugs were sought in the initial search, this was restricted to baricitinib and sirukumab in the update search. Investigational drugs were not mentioned in the final NICE scope. All comparators mentioned in the final NICE scope were included in the CS inclusion criteria.

Study designs for effectiveness data were restricted to RCTs and their long-term extension studies. This was appropriate given the availability of RCTs meeting the inclusion criteria.

The study selection process described in the CS (Section 4.1.3 of the CS) describes study selection by two reviewers, as is good practice in systematic reviews. A third and fourth reviewer were employed to resolve discrepancies.

4.1.3 Critique of data extraction

Data were extracted by two reviewers (CS Section 4.1.3) as is good practice. Data extracted for the SAR trials by the CS, and reported below, were checked by the ERG against published trial papers where available (MOBILITY-A Huizinga 2014,[32] MOBILITY-B Genovese 2015,[33] TARGET Fleischmann 2017,[34] MONARCH Burmester 2016[35]), and the clinical study report (CSR) for ASCERTAIN.[36]

4.1.4 Quality assessment

The CS Section 4.1.3 states that a quality assessment was performed using the methods recommended in the current NICE specification. Quality items assessed were taken from the Centre for Reviews and Dissemination (CRD) guidelines for undertaking reviews in health care.[37] These are standard and appropriate criteria for assessing the risk of bias in RCTs.

Table 4 of the ERG report shows quality assessment for the SAR trials (reproduced from Table 4.13 of the CS). The ERG checked the quality assessment against the publications of the trials where available, and the CSR for ASCERTAIN, and agreed with the assessment in the CS.

Table 4: Quality assessment of included SAR trials (Reproduced from CS Table 4.13)

Trial name	MOBILITY-A [32]	MOBILITY-B [33]	TARGET [34]	ASCERTAIN [36]	MONARCH [35]
Was randomisation carried out appropriately?	YES	YES	YES	YES	YES
Was the concealment of treatment allocation adequate?	YES	YES	YES	YES	YES
Were the groups similar at the outset of the study in terms of prognostic factors?	YES	YES	YES	YES	YES
Were the care providers, participants and outcome assessors blind to treatment allocation?	YES	YES	YES	YES	YES
Were there any unexpected imbalances in drop-outs between groups?	NO	NO	NO	NO	NO
Is there any evidence to suggest that the authors measured more outcomes than they reported?	NO**	NO**	NO**	Not yet published***	NO***
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	YES	YES	YES (mITT)	YES (mITT)	YES

Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination)

**from CS Appendix 6, some results not yet published, but in the process of being published

***from CS Appendix 6, some results not yet analysed

The five included SAR RCTs were generally at low risk of bias, in the view of both the company and the ERG. All five RCTs used central allocation generated by interactive voice response system (MOBILITY-A, MOBILITY-B, TARGET, MONARCH, ASCERTAIN).

In all five RCTs, randomisation was stratified by region, and there was also stratification by prior bDMARD use in MOBILITY-A and MOBILITY-B, number of prior bDMARDs in TARGET, and screening value of absolute neutrophil count in ASCERTAIN. All five RCTs were blinded.

Three of the RCTs planned an intent-to-treat (ITT) analysis with all randomised patients (MOBILITY-A, MOBILITY-B, MONARCH). The other two RCTs (TARGET and ASCERTAIN) planned a modified ITT, analysing all randomised patients who received at least one dose of study drug,[34 36] however in practice all randomised patients were treated and included in the analysis. All five included RCTs used a non-responder imputation for categorical data, in which patients who discontinued, received rescue therapy or otherwise had missing data were assumed to be failures.

Quality assessments of the trials in the NMA are presented in Appendix 8.7 of the CS. The same quality assessment items were used as above, as is appropriate for RCTs. There was some variation in quality, but the majority were blinded and reported ITT analyses.

4.2 Critique of trials of the technology of interest, their analysis and interpretation

SAR trials included in the CS

Five RCTs and one long-term extension study of SAR were included in the CS (Table 5). Three RCTs had populations of cDMARD experienced RA patients (MOBILITY-A, MOBILITY-B, MONARCH), and two RCTs had TNFi experienced RA patients (TARGET, ASCERTAIN).

Table 5: Included SAR trials (Reproduced from CS Table 4.3)

Study	Study population	Intervention	Comparator	Reference
MOBILITY A NCT01061736	MTX-IR	SAR + MTX	PBO + MTX	*Huizinga 2014*[32] *Sanofi Genzyme Data on File*[38]
MOBILITY B NCT01061736	MTX-IR	SAR + MTX	PBO + MTX	*Genovese 2015*[33] *Sanofi Genzyme Data on File*[39]
TARGET NCT01709578	TNFi-IR/ intolerant	SAR + cDMARD	PBO + cDMARD	*Fleischmann 2017*[34] *Sanofi Genzyme Data on File*[40]
MONARCH NCT02332590	MTX-IR/ intolerant	SAR	ADA	*Burmester 2016*[35] *Sanofi Genzyme Data on File*[41]
ASCERTAIN (safety study) NCT01768572	TNFi-IR/ intolerant	SAR + cDMARD	TCZ + cDMARD	*Sanofi Genzyme Data on File* [36]
EXTEND (Long-term extension safety study) NCT01146652	cDMARD/TNFi-IR/ intolerant	SAR + cDMARD, SAR monotherapy	NA, Extension study	*Sanofi Genzyme Data on File*[42]

ADA=adalimumab; cDMARD=conventional disease-modifying anti-rheumatic drugs; cDMARD = conventional disease-modifying anti-rheumatic drug; IR=inadequate response; MTX=methotrexate; NA=not applicable; PBO=placebo; SAR= sarilumab; TCZ= tocilizumab; TNFi =tumour necrosis factor inhibitor.

SAR trials excluded from the CS

Two trials of SAR were terminated early: NCT01764997 with comparators ADA and ETN, and NCT01217814 with comparator GOL. The company's clarification response[31] (question A2) states that these trials were terminated due to study delays. The company's clarification response to question A2 states that for NCT01217814 no effectiveness analyses were conducted, and that safety analyses were conducted but no conclusions were drawn due to the small sample size (16 patients randomised).

Two studies were excluded from the CS for being uncontrolled, ONE **(**NCT02121210) and EASY (NCT02057250). However, the company's clarification response[31] (question A3) states that safety data from these studies were included in the pooled safety analysis of SAR for the EMA license application.

Two studies were excluded from the CS for having exclusively Japanese populations. KAKEHASI (NCT02293902) was conducted in an MTX-IR population, and compared SAR+MTX with placebo + MTX. HARUKA (NCT02373202) compared SAR monotherapy with SAR +cDMARDs (non-MTX) in a population of MTX-IR, MTX intolerant, or non-MTX cDMARD experienced. The company's clarification response[31] states that this is due to the trials not being generalisable to the UK population, and suggests that

"[redacted]".

SAR trials included in the CS – trial characteristics

Trial characteristics for the studies included in the CS are shown in Table 6.

Table 6: Included SAR trials (Adapted from CS Tables 1.3, 4.3 and 4.7)

Study	Study population	Intervention	Comparator	Follow-up (weeks)	Primary endpoint	Used in NMA?
MOBILITY A NCT01061736	N=306 MTX-IR [24.5% prior bDMARDs]	SAR + MTX SAR doses: 100mg QW, 150mg QW, 100mg Q2W, 150mg Q2W, 200mg Q2W	PBO + MTX	12	ACR20 Week 12	No (12 week study)
MOBILITY B NCT01061736	N=1197 MTX-IR [27.9% prior bDMARDs]	SAR + MTX SAR doses: 150mg Q2W, 200mg Q2W	PBO + MTX	52	ACR20 at Week 24 Change in HAQ-DI from baseline to Week 16 Change in mTSS from baseline to Week 52	cDMARD-IR combination therapy
MONARCH NCT02332590	N=369 MTX-IR/ intolerant	SAR monotherapy SAR dose 200mg Q2W	ADA monotherapy ADA dose 40mg Q2W	24	Change in DAS28-ESR from baseline to Week 24	cDMARD-IR monotherapy
TARGET NCT01709578	N=546 TNFi-IR/ intolerant	SAR + cDMARD SAR doses: 150mg Q2W, 200mg Q2W	PBO + cDMARD	24	ACR20 response at Week 24 Change in HAQ-DI from baseline to Week 12	TNFi-IR
ASCERTAIN (safety study) NCT01768572	N=202 TNFi-IR/ intolerant	SAR + cDMARD SAR doses: 150mg Q2W, 200mg Q2W	TCZ + cDMARD TCZ dose 4-8mg/kg	24	Safety	TNFi-IR
EXTEND (Long-term extension safety study) NCT01146652	N=2023 cDMARD/TNFi-IR/ intolerant	SAR + cDMARD, SAR monotherapy	NA, Extension study	264 (at least)	Safety	No (extension study)

ADA: adalimumab; cDMARD: conventional disease-modifying antirheumatic drug; IR: inadequate response; MTX: methotrexate; PBO: placebo; SAR: sarilumab; TCZ: tocilizumab; TNFi: tumour necrosis factor inhibitor

All trials recruited adult populations. ASCERTAIN and EXTEND assessed safety, whereas the other included trials assessed effectiveness and safety. EXTEND was an open-label extension study, MOBILITY A was a Phase II RCT, and the other trials were Phase III RCTs. The RCTs were all international, multi-centre studies including centres in the US, South America and Europe. Two of the RCTs had centres in the UK (1 centre for Monarch and six centres for ASCERTAIN, see CS Section 4.5) however only ASCERTAIN recruited patients from the UK (n=14) (see clarification response[31] – question A4).

EXTEND was an open-label extension study of SAR (either monotherapy or in combination with cDMARDs). Patients in the EXTEND study were recruited from MOBILITY A and B ([redacted]), TARGET ([redacted]), ASCERTAIN ([redacted]), ONE ([redacted]) and [redacted] ([redacted]).

Of the five RCTs, three had a PBO comparator (MOBILITY-A, MOBILITY-B, TARGET). The monotherapy trial, MONARCH, compared SAR with ADA at its licensed UK dose. The ASCERTAIN trial compared SAR with TCZ 4-8mg/kg Q4W. The UK recommended dose of IV TCZ for adults is "8mg/kg every 4 weeks (max. per dose 800mg), for dose adjustments in patients with liver enzyme abnormalities, or low absolute neutrophil or platelet count, consult product literature."[43] The primary endpoint of ASCERTAIN was safety rather than effectiveness. The company's clarification response[31] (question A6) states that comparative effectiveness data were not provided as ASCERTAIN was not powered for effectiveness endpoints.

Baseline characteristics of patients included in the RCTs are shown in Table 7. Across trials, patients had a mean DAS of >5.1, had a mean age of 52.2, and were mostly female and Caucasian. There were no imbalances within trials between treatment groups at baseline.

Table 7: Baseline characteristics of included SAR trials (reproduced from CS Table 4.12)

	MOBILITY A N=306* [32;38]	**MOBILITY B N=398**[33;39]	**TARGET N=546**[34;40]	**ASCERTAIN** [redacted][36]	**MONARCH N=369**[35;41]
Age, mean (SD)	52.2 (12.5)	50.8 (11.7)	52.9 (12.4)	[redacted]	52.2 (12.3)
Males, %	20.6	18.3	18.1	[redacted]	16.8
Race, %					
Caucasian/ White	93.8	86.4	71.1	[redacted]	90.8
Black	2.6	2.4	3.7	[redacted]	1.1
Asian/ Oriental	2.0	8.0	0.9	[redacted]	3.0
Other	1.6	3.2	24.4	[redacted]	5.1

	MOBILITY A N=306* [32;38]	MOBILITY B N=398[33;39]	TARGET N=546[34;40]	ASCERTAIN [redacted][36]	MONARCH N=369[35;41]
Weight kg, mean (SD)	74.86 (15.27)	74.39 (18.52)	78.22 (21.52)	[redacted]	72.05 (17.15)
BMI kg/m^2, mean (SD)	28.28 (5.64)	28.26 (6.34)	29.53 (7.17)	[redacted]	27.18 (6.05)
Duration of RA in years, mean (SD)	7.81 (8.08)	9.03 (7.85)	12.09 (9.40)	[redacted]	7.33 (7.99)
RA functional class, %					
I	6.2	11.7	9.5	[redacted]	17.9
II	70.3	67.2	57.7	[redacted]	65.0
III	23.5	21.2	32.8	[redacted]	17.1
IV	0	0	0	[redacted]	0
RF +ve, %	79.7	84.9	75.5	[redacted]	65.8
Anti-CCP +ve, %	82.0	86.9	78.1	[redacted]	76.0
TJC (0–68), mean (SD)	27.39 (14.93)	26.85 (14.07)	28.88 (15.22)	[redacted]	27.32 (13.41)
SJC (0–66), mean (SD)	17.38 (9.73)	16.82 (9.49)	19.93 (11.49)	[redacted]	18.04 (10.50)
CRP inmg/L, mean (SD)	2.78 (2.96)	22.23 (23.69)	26.82 (25.89)	[redacted]	20.71 (26.78)
HAQ-DI (0–3), mean (SD)	1.59 (0.62)	1.64 (0.64)	1.78 (0.63)	[redacted]	1.64 (0.60)
DAS28-CRP, mean (SD)	6.11 (0.84)	5.96 (0.90)	6.20 (0.91)	[redacted]	6.01 (0.89)
Prior cDMARD use, %	100	100	100	[redacted]	100
Number of cDMARDs, %					
0	0	0	0	[redacted]	0
1	92.8	NR	53.5	[redacted]	46.3
2	4.9	NR	27.5	[redacted]	31.2
≥3	2.3	NR	19.0	[redacted]	22.5
Prior bDMARD use, %	24.5	27.9	100	[redacted]	0
Prior TNFi use, %	NR	NR	100%	[redacted]	0
Number of TNFi, %					

	MOBILITY A N=306* [32;38]	**MOBILITY B N=398**[33;39]	**TARGET N=546**[34;40]	**ASCERTAIN** [redacted][36]	**MONARCH N=369**[35;41]
1	NR	NR	76.8	[redacted]	0
≥1	NR	NR	23.2	[redacted]	0

*includes groups with unlicensed doses of SAR

RF rheumatoid factor, CCP anti-cyclic citrullinated peptide, CRP high-sensitivity C reactive protein

Discontinuation rates are shown in Table 8 and Table 9. At 24 weeks discontinuation rates for SAR ranged from 10.3% to [redacted]. In MONARCH, there was a discontinuation rate of 15.1% for the licensed dose of ADA, and 10.3% for SAR 200mg Q2W.

[redacted]

Table 8: Discontinuation during cDMARD-IR trials[38, 39, 41 32 33 35]

	MOBILITY-A 12weeks			MOBILITY-B 52 weeks			MONARCH 24 weeks	
	PBO + MTX (n=52)	SAR 150mg Q2W + MTX (n=51)	SAR 200mg Q2W + MTX (n=52)	Placebo + MTX (n=428)	SAR 150mg Q2W + MTX (n=430)	SAR 200mg Q2W + MTX (n=427)	ADA 40mg Q2W (n=185)	SAR 200mg Q2W (n=184)
Discontinuation during double blind period, n (%)	3 (5.8)	3 (5.9)	6 (11.5)	62 (14.5)	78 (18.1)	88 (20.6)	28 (15.1)	19 (10.3)
Any AE leading to treatment discontinuation, n (%)	1 (1.9)	2 (3.8)	4 (7.8)	20 (4.7)	54 (12.5)	59 (13.9)	15 (8.1)	11 (6.0)

AE: adverse events; PBO: placebo: MTX: methotrexate; SAR: salirumab; ADA: adalimumab; Q2W: every other week

Table 9: Discontinuation during TNFi-IR trials at week 24 [34 36 40]

	TARGET			ASCERTAIN		
	PBO + cDMARD (n=181)	SAR 150mg Q2W + cDMARD (n=181)	SAR 200mg Q2W + cDMARD (n=181)	TCZ IV 4–8mg/kg Q4W + cDMARD (n=102)	SAR 150mg Q2W + cDMARD (n=49)	SAR 200mg Q2W + cDMARD (n=51)
Discontinuation, n (%)	17 (9.4)	31 (17.1)	25 (13.6)	[redacted]	[redacted]	[redacted]
Any AE leading to treatment discontinuation, n (%)	8 (4.4)*	14 (7.7)	17 (9.2)	[redacted]	[redacted]	[redacted]

*additionally 1 PBO and 4 SAR 150mg, abnormal laboratory values at baseline [34]

AE: adverse events; PBO: placebo: SAR: salirumab; TCZ: tocilizumab; IV: intravenous; Q2W: every 2 weeks; Q4W: every 4 weeks; cDMARD: conventional disease-modifying antirheumatic drug

Effectiveness results from the included SAR RCTs

ACR response data

The five included SAR RCTs reported ACR data. ACR20 was the primary outcome for MOBILITY-A, MOBILITY-B and TARGET.

ACR response data from the cDMARD-IR trials

ACR response data for the cDMARD-IR RCTs are shown in Table 10 to Table 12. At 12 weeks, MOBILITY-A showed a statistically significant advantage for both licensed doses of SAR+MTX over PBO+MTX in ACR20 and ACR50 ($p<0.05$). A significantly higher proportion of patients in the SAR 200mg Q2W group, than the PBO group, achieved ACR70. At 24 weeks, MOBILITY-B showed a significant advantage for SAR 200mg Q2W+MTX and SAR 150mg Q2W+MTX over PBO for ACR20 (66.4%, 58.0% and 33.4%), respectively ($p<0.0001$). Both licensed doses also showed a significant advantage over PBO in AC50 and ACR70 at 24 weeks, and ACR20 at 52 weeks ($p<0.0001$). MONARCH reported a significantly ($p<0.01$) higher proportion of patients in the SAR 200mg Q2W monotherapy group, than the ADA 40mg Q2W monotherapy group, achieved ACR20 (71.7% versus 58.4%).

Table 10: ACR response rates in MOBILITY-A at 12 weeks (adapted from CS Section 4.7.1[32])

	PBO +MTX (n=52)	**SAR 150mg Q2W +MTX (n=51)**	***p*-value**	**SAR 200mg Q2W +MTX (n=52)**	***p*-value**
ACR20 response	46**%**	67**%**	0.0363	65**%**	0.0426
ACR50 response	15**%**	35**%**	0.0163	40**%**	0.0038
ACR70 response	2**%**	12**%**	0.0574	17**%**	0.0078

Table 11: ACR response rates in MOBILITY-B data at week 24 (adapted from Table 4.16 of the CS)

	PBO + MTX (N=398)	**SAR 150mg Q2W + MTX (N=400)**	***p*-value**	**SAR 200mg Q2W + MTX (N=399)**	***p*-value**
ACR20 response, n (%)	133 (33.4)	232 (58.0)	< 0.0001	265 (66.4)	< 0.0001
ACR50 response, n (%)	66 (16.6)	148 (37.0)	< 0.0001	182 (45.6)	<0.0001
ACR70 response, n (%)	29 (7.3)	79 (19.8)	< 0.0001	99 (24.8)	<0.0001
ACR20 response at Week 52, n (%)	126 (31.7)	214 (53.5)	<0.0001	234 (58.6)	<0.0001

Table 12: ACR response rates in MONARCH at Week 24 (adapted from CS Table 4.22)

	ADA 40mg Q2W (N=185)	SAR 200mg Q2W (N=184)	*p*-value
ACR20 response, n (%)	108 (58.4)	132 (71.7)	0.0074
ACR50 response, n (%)	55 (29.7)	84 (45.7)	0.0017
ACR70 response, n (%)	22 (11.9)	43 (23.4)	0.0036

ACR response data from the TNFi-IR trials

ACR response data from the TARGET and ASCERTAIN TNFi-IR trials are shown in Table 13 and Table 14, respectively. TARGET reported a significant ($p<0.0001$) advantage for SAR 200mg Q2W+ cDMARD and SAR 150mg Q2W+ cDMARD over PBO+ cDMARD for ACR20 (60.9%, 55.8% and 33.7% respectively) at 24 weeks. At 24 weeks, TARGET also reported a significant advantage for the SAR+ cDMARD doses over PBO+ cDMARD in ACR50 and ACR70 ($p\leq0.0002$). [redacted]

Table 13: ACR response rates in TARGET at week 24 (adapted from CS Table 4.19)

	PBO + cDMARD (N=181)	SAR 150mg Q2W + cDMARD (N=181)	*p*-value	SAR 200mg Q2W + cDMARD (N=184)	*p*-value
ACR20 response n (%)	61 (33.7)	101 (55.8)	<0.0001	112 (60.9)	<0.0001
ACR50 response n (%)	33 (18.2)	67 (37.0)	<0.0001	75 (40.8)	<0.0001
ACR70 response n (%)	13 (7.2)	36 (19.9)	0.0002	30 (16.3)	0.0056

Table 14: ACR response rates in ASCERTAIN at week 24 (adapted from CS Table 4.2)

	TCZ Q4W + cDMARD (N=102)	SAR 150mg Q2W + cDMARD (N=49)	SAR 200mg Q2W + cDMARD (N=51)
ACR20 response %	[redacted]	[redacted]	[redacted]
ACR50 response %	[redacted]	[redacted]	[redacted]

ACR70 response %	[redacted]	[redacted]	[redacted]

EULAR response data from the cDMARD-IR trials

[redacted]

Table 15: EULAR response rates in MOBILITY-A at 12 weeks[38]

	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

*[redacted]

Table 16: EULAR response rates in MOBILITY-B at week 24 (adapted from CS Appendix Tables 8.23 and 8.24)

	PBO + MTX (N=398)	**SAR 150mg Q2W + MTX (N=400)**	**SAR 200mg Q2W + MTX (N=399)**
EULAR good response (%)	[redacted]	[redacted]	[redacted]
EULAR moderate to good response (%)	[redacted]	[redacted]	[redacted]

Table 17: EULAR response rates in MONARCH at week 24 (adapted from CS Appendix Table 8.33)

	ADA 40mg Q2W (N=185)	**SAR 200mg Q2W (N=184)**
EULAR good (%)	[redacted]	[redacted]
EULAR moderate to good response (%)	[redacted]	[redacted]

EULAR response data from the from the TNFi-IR trials

[redacted]

Table 18: EULAR response rates in TARGET at week 24 (adapted from CS Table 8.43 and 8.44)

	Placebo + cDMARD (N=181)	SAR 150mg Q2W + cDMARD (N=181)	SAR 200mg Q2W + cDMARD (N=184)
EULAR good response (%)	[redacted]	[redacted]	[redacted]
EULAR moderate to good response (%)	[redacted]	[redacted]	[redacted]

HAQ-DI, DAS28 and mTSS effectiveness outcomes

HAQ-DI, DAS28 and mTSS outcomes are shown in Tables 19-21 for the cDMARD-IR trials, and in Table 22 and Table 23 for the TNFi-IR trials. The MOBILITY-A, MOBILITY-B, MONARCH and TARGET trials reported significantly favourable results for licensed doses of SAR over their comparators for improvement in HAQ-DI ($p \leq 0.0037$). SAR had a significant advantage over its comparator for DAS28-CRP in the MOBILITY-B and TARGET trials ($p < 0.0001$), and for DAS28-ESR in the MONARCH trial ($p < 0.0001$). MOBILITY-B measured radiographic progression by mTSS, and reported a significantly lower deterioration from baseline for SAR over comparator ($p \leq 0.01$). Comparative statistics were not available for ASCERTAIN.

Table 19: Efficacy results from MOBILITY-A (adapted from CS Table 4.14)

	PBO (n=52) LS Mean (SE)	SAR 150mg Q2W (n=51) LS Mean (SE)	SAR200mg Q2W (n=52) LS Mean (SE)
HAQ-DI	−0.26 (0.07)	−0.62 (0.07)	−0.57 (0.07)
***p*-value vs. placebo**		0.0003	0.0019
CRP (mg/L)	−3.1 (2.8)	−21.9 (2.8)	−21.9 (2.8)
***p*-value vs. placebo**		<0.0001	<0.0001

Table 20: Efficacy results from MOBILITY-B (adapted from CS Table 4.16)

	Placebo + MTX (N=398)	**SAR 150mg Q2W + MTX (N=400)**	***p*-value**	**SAR 200mg Q2W + MTX (N=399)**	***p*-value**
HAQ-DI	−0.33 ± 0.03	−0.53 ± 0.03	<0.0001	−0.55 ± 0.03	<0.0001
CRP,mg/dL	−0.0 ± 0.12	−1.3 ± 0.12	<0.0001	−1.7 ± 0.12	<0.0001
Major clinical response (ACR70 response maintained for ≥24 weeks), n (%)[a]	12 (3.0)	51 (12.8)	<0.0001	59 (14.8)	<0.0001
DAS28-CRP, LS mean change from baseline to Week 24 (SE)	−1.17(0.080)	−2.45(0.076)	< 0.0001	−2.82(0.075)	<0.0001
DAS28-CRP response at Week 24, n (%)					
Score <2.6[b]	40 (10.1)	111 (27.8)	<0.0001	136 (34.1)	<0.0001
Score ≤3.2	67(16.8)	159 (39.8)	<0.0001	196 (49.1)	<0.0001
Physical function (HAQ-DI)					
HAQ-DI, adjusted mean change from baseline at Week 16, using MMRM[a]	−0.29 ± 0.03	−0.53 ± 0.03	<0.0001	−0.55 ± 0.03	<0.0001
HAQ-DI response (MCID ≥0.3), n (%)					
At Week 16	169 (42.5)	215 (53.8)	<0.01	229 (57.4)	<0.0001
At Week 24	133 (33.4)	204 (51.0)	<0.0001	205 (51.4)	<0.0001
At Week 52	104 (26.1)	188 (47.0)	<0.0001	190 (47.6)	<0.0001
Radiographic progression (mTSS)					
Mean change from baseline in mTSS at week 52, using rank ANCOVA[b]	2.78 ± 7.73	0.90 ± 4.66	<0.0001	0.25 ± 4.61	<0.0001
No radiographic progression, n (%)					
At Week 24	158 (39.7)	185 (46.3)	<0.0001	226 (56.6)	<0.0001
At Week 52[a]	154 (38.7)	191 (47.8)	<0.01	222 (55.6)	<0.0001

Table 21: Efficacy results from MONARCH at week 24 (adapted from CS Table 4.22)

	ADA 40mg Q2W (N=185)	SAR 200mg Q2W (N=184)	*p*-value
Disease activity			
DAS28-ESR, mean (SD)	4.5 (1.4)	3.5 (1.4)	
DAS28-ESR, LSM change from baseline (SE)	−2.20 (0.106)	−3.28 (0.105)	<0.0001
DAS28-ESR <2.6 (remission), n (%)	13 (7.0)	49 (26.6)	<0.0001
Physical function and PROs			
HAQ-DI, mean (SD)	1.2 (0.7)	1.0 (0.7)	
HAQ-DI, LSM change from baseline (SE)	−0.43 (0.05)	−0.61 (0.05)	0.0037

ACR20/50/70=American College of Rheumatology 20%/50%/70% improvement; DAS28-ESR=28-joint disease activity score-erythrocyte sedimentation rate; EQ-5D= EuroQol five dimensions' questionnaire; FACIT=Functional Assessment of Chronic Illness Therapy; HAQ-DI=Health Assessment Questionnaire Disability Index; LSM=least square mean; Q2W=every 2 weeks; SF-36=Medical Outcomes Short Form 36 Health Survey.

Table 22: Efficacy results from TARGET (adapted from CS Table 4.19)

	Placebo + cDMARD (N=181)	SAR 150mg Q2W + cDMARD (N=181)	*p*-value	SAR 200mg Q2W + cDMARD (N=184)	*p*-value
Physical function at Week 12					
HAQ-DI, LSM change from baseline (SE)	−0.26 (0.04)	−0.46 (0.04)	<0.001	−0.47 (0.04)	<0.001
Physical function at Week 24					
HAQ-DI, LSM mean change from baseline (SE)	−0.3 (0.05)	−0.5 (0.05)	0.0078	−0.6 (0.05)	0.0004

HAQ-DI change from baseline >3.0, n (%)	57 (31.5)	78 (43.1)	<0.05	87 (47.3)	<0.01
DAS28-CRP, LS mean change from baseline (SE)	−1.38 (0.119)	−2.35 (0.111)	<0.0001	−2.82 (0.108)	<0.0001
Disease activity and remission at Week 24					
DAS28-CRP<2.6, n (%)	13 (7.2)	45 (24.9)	<0.0001	53 (28.8)	<0.0001

Table 23: Efficacy results from ASCERTAIN (adapted from CS Table 4.21)

	TCZ Q4W + cDMARD (N=102)	SAR 150mg Q2W + cDMARD (N=49)	SAR 200mg Q2W + cDMARD (N=51)
HAQ-DI, LSM change from baseline (SE)	[redacted]	[redacted]	[redacted]
CRP (mg/dL)	[redacted]	[redacted]	[redacted]
DAS28 remission <2.6, %	[redacted]	[redacted]	[redacted]
DAS28-CRP	[redacted]	[redacted]	[redacted]

HRQoL

Table 24 and Table 25 show the HRQoL outcomes from cDMARD-IR trials. [redacted]. MONARCH found no significant treatment effect for SF-36 MCS, EQ-5D single index utility or EQ-5D VAS, but reported a significantly (p=0.006) greater improvement in SF-36 PCS in the SAR 200mg Q2W monotherapy group than in the ADA 40mg Q2W monotherapy group.

TARGET reported a significant (p<0.0001) advantage for SAR 200mg Q2W+ cDMARD and SAR 150mg Q2W+ cDMARD over PBO+ cDMARD for SF36-PCS at 12 weeks and 24 weeks, and an advantage (p<0.05) in MCS at 12 weeks. There was no statistically significant treatment effect for SF-36 MCS at week 24[redacted].[40]

Table 24: HRQoL results from MOBILITY-B (adapted from CS Table 4.17)

	Placebo + MTX (N= 398)	**SAR 150mg Q2W + MTX (N=400)**	***p*-value**	**SAR 200mg Q2W + MTX (N=399)**	***p*-value**
Week 24					
SF-36 Physical	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SF-36 Mental	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Week 52					
SF-36 Physical	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SF-36 Mental	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

SF-36=Medical Outcomes Short Form 36 Health Survey

Table 25: HRQoL results from MONARCH (adapted from Table 4.22)

	ADA 40mg Q2W (N=185)	**SAR 200mg Q2W (N=184)**	***p*-value**
SF-36 PCS, LSM change from baseline (SE)	6.1 (0.6)	8.7 (0.6)	0.0006
SF-36 MCS, LSM change from baseline (SE)	6.8 (0.8)	7.9 (0.8)	0.3319
EQ-5D single index utility, LSM change from baseline (SE)	0.26 (0.35)	0.32 (0.35)	0.0382

EQ-5D VAS, LSM change from baseline (SE)	19.94 (1.720)	24.22 (1.686)	0.0699

SF-36=Medical Outcomes Short Form 36 Health Survey, EQ-5D= EuroQol five dimensions' questionnaire

Table 26 shows HRQoL from the TNFi-IR trial TARGET.

Table 26: HRQoL results from TARGET (adapted from CS Table 4.20)

LSM change from baseline (SE)	**Week 12**					**Week 24**				
	Placebo + cDMARDs (N =181)	**SAR 150mg Q2W + cDMARDs (N=181)**	***p*-value**	**SAR 200mg Q2W + cDMARDs (N=184)**	***p*-value**	**Placebo + cDMARDs (N =181)**	**SAR 150mg Q2W + cDMARDs (N=181)**	***p*-value**	**SAR 200mg Q2W + cDMARDs (N=184)**	***p*-value**
SF-36 PCS	3.7±0.6	6.9±0.6	<0.0001	6.8±0.6	<0.0001	4.4±0.7	7.7±0.7	<0.001	8.5±0.6	<0.0001
SF-36 MCS	3.5±0.7	5.1±0.8		6.5±0.7	<0.05	4.7±0.9	6.3±0.8		6.8±0.8	

SF-36=Medical Outcomes Short Form 36 Health Survey

Effectiveness data from the EXTEND study

At the time of writing, the EXTEND study was ongoing (see Table 27 and Table 28). The CS provided results of an interim analysis of EXTEND (see CS, Section 4.11).

[redacted]

Table 27: ACR response and DAS28-CRP remission rates from the interim analysis in EXTEND (reproduced from CS Table 4.38)

	ACR20 (%)	ACR50 (%)	ACR70 (%)	DAS28 remission (%)
SAR + cDMARD				
Week 0,	[redacted]	[redacted]	[redacted]	[redacted]
Week 24,	[redacted]	[redacted]	[redacted]	[redacted]
Week 48	[redacted]	[redacted]	[redacted]	[redacted]
Week 96	[redacted]	[redacted]	[redacted]	[redacted]
Week 144	[redacted]	[redacted]	[redacted]	[redacted]
Week 192	[redacted]	[redacted]	[redacted]	[redacted]
Week 216	[redacted]	[redacted]	[redacted]	[redacted]
Week 240	[redacted]	[redacted]	[redacted]	[redacted]
Week 264	[redacted]	[redacted]	[redacted]	[redacted]
SAR monotherapy				
Week 0	[redacted]	[redacted]	[redacted]	[redacted]
Week 24	[redacted]	[redacted]	[redacted]	[redacted]
Week 48	[redacted]	[redacted]	[redacted]	[redacted]

Table 28: Changes from baseline in mTSS from the interim analysis in EXTEND (reproduced from CS Table 4.39)

	2-year analysis SAR + DMARD (n=889)	3-year analysis SAR + DMARD (n=796)
CFB in mTSS, mean (SD)		
Week 0 (52 weeks from baseline)	[redacted]	[redacted]
Week 48 (100 weeks from baseline)	[redacted]	[redacted]

	2-year analysis SAR + DMARD (n=889)	3-year analysis SAR + DMARD (n=796)
Week 96 (148 weeks from baseline)	[redacted]	[redacted]

CFB=change from baseline, mTSS=modified Total Sharp Score

Adverse events

Adverse event rates from the included trials are shown in Table 29 and Table 30. AE rates were higher in SAR than PBO groups. For the cDMARD-IR trials, AE rates in the SAR groups ranged from 53.8% to 78.1%. In the MONARCH trial, ADA and SAR had similar AE rates (63.6% and 64.1% respectively).

[redacted]

Table 29: AEs in cDMARD-IR trials (adapted from CS Tables 4.41, 4.42 and 4.45)[32 33 35 38, 39, 41]

	MOBILITY-A 12weeks			MOBILITY-B 52 weeks			MONARCH 24 weeks	
	PBO + MTX (n=52)	SAR 150mg Q2W + MTX (n=51)	SAR 200mg Q2W + MTX (n=52)	PBO + MTX (n=427)	SAR 150mg Q2W + MTX (n=431)	SAR 200mg Q2W + MTX (n=424)	ADA 40mg Q2W (n=184)	SAR 200mg Q2W (n=184)
Any AE, n (%)	24 (47.1)	28 (53.8)	33 (64.7)	263 (61.6)	321 (74.5)	331 (78.1)	117 (63.6)	118 (64.1)
Any SAE, n (%)	2 (3.9)	0	0	23 (5.4)	38 (8.8)	48 (11.3)	12 (6.5)	9 (4.9)
Any AE leading to treatment discontinuation, n (%)	1 (1.9)	2 (3.8)	4 (7.8)	20 (4.7)	54 (12.5)	59 (13.9)	15 (8.1)	11 (6.0)
Deaths, n	0	0	0	2 (0.5)	2 (0.5)	1 (0.2)	0	1 (0.5)

AE: adverse events; SAE: serious AE; PBO: placebo: MTX: methotrexate; SAR: salirumab; ADA: adalimumab; Q2W: every other week

Table 30: AEs in TNFi-IR trials[34 40 36]

	TARGET 24 weeks			ASCERTAIN 24 weeks		
	PBO + cDMARD (n=181)	SAR 150mg Q2W + cDMARD (n=181)	SAR 200mg Q2W + cDMARD (n=184)	TCZ IV 4–8mg/kg Q4W + cDMARD (n=102)	SAR 150mg Q2W + cDMARD (n=49)	SAR 200mg Q2W + cDMARD (n=51)
Any AE, n (%)	90 (49.7)	119 (65.7)	120 (65.2)	[redacted]	[redacted]	[redacted]
Any SAE, n (%)	6 (3.3)	6 (3.3)	10 (5.4)	[redacted]	[redacted]	[redacted]
Any AE leading to treatment discontinuation, n (%)	8 (4.4)	14 (7.7)	17 (9.2)	[redacted]	[redacted]	[redacted]
Deaths, n (%)	1 (0.6)	0	0	[redacted]	[redacted]	[redacted]

AE: adverse events; SAE: serious AE; PBO: placebo: SAR: salirumab; TCZ: tocilizumab; IV: intravenous; Q2W: every 2 weeks; Q4W: every 4 weeks; cDMARD: conventional disease-modifying antirheumatic drug

The SmPC (published by the EMA after company submission, but draft provided by company in CS Appendix 1.2) provides the tabulated summary of AEs of SAR (Table 31). Data were provided for 2,887 patients receiving SAR in combination with cDMARDs, and 467 patients receiving SAR monotherapy.

AEs for SAR+cDMARDs are shown in Table 32 (from the EPAR) and for SAR monotherapy in Table 33 (from the EPAR).[1] The most frequent SAEs were infections and laboratory abnormalities (changes in absolute neutrophil count and alanine aminotransferase).[1]

As of the data extraction dates for the EPAR, there were 27 deaths in the SAR treated patients, the most common causes were cardiovascular, infections and malignancies.[1] The most common AEs were infections: the most common of these were nasopharyngitis, bronchitis, upper respiratory tract infections, and urinary tract infections. [1]

Table 31: Summary of AEs in controlled clinical studies (as published in SmPC)

System Organ Class	Frequency	Adverse Reaction
Infections and Infestations	Common	Upper respiratory tract infection
		Urinary tract infection
		Nasopharyngitis
		Oral herpes
Blood and Lymphatic System Disorders	Very common	Neutropenia
	Common	Thrombocytopenia
Metabolism and Nutrition Disorders	Common	Hypercholesterolemia
		Hypertriglyceridemia
Hepatobiliary Disorders	Common	Transaminases increased
General Disorders and Administration Site Conditions	Common	Injection site erythema
		Injection site pruritus

Very common: ≥ 1/10; Common:≥ 1/100 to < 1/10

Table 32: Percentages of patients with AEs on SAR+cDMARDs (≥2% in at least one treatment group) (adapted from the EPAR)[1]

Primary System Organ Class **Preferred Term**	**SAR+DMARD** **(N=2887)** **n (%)**
Any class	2418 (83.8%)
Infections and infestations	1428 (49.5%)
Upper respiratory tract infection	325 (11.3%)
Urinary tract infection	252 (8.7%)
Nasopharyngitis	237 (8.2%)
Bronchitis	196 (6.8%)
Sinusitis	110 (3.8%)
Influenza	107 (3.7%)
Pharyngitis	104 (3.6%)
Cellulitis	85 (2.9%)
Pneumonia	80 (2.8%)
Gastroenteritis	76 (2.6%)
Blood and lymphatic system disorders	670 (23.2%)
Neutropenia	507 (17.6%)
Leukopenia	111 (3.8%)
Thrombocytopenia	80 (2.8%)
Metabolism and nutrition disorders	338 (11.7%)
Hypertriglyceridaemia	97 (3.4%)
Hypercholesterolaemia	79 (2.7%)
Dyslipidaemia	65 (2.3%)
Nervous system disorders	311 (10.8%)
Headache	115 (4.0%)
Vascular disorders	279 (9.7%)
Hypertension	204 (7.1%)
Gastrointestinal disorders	553 (19.2%)
Diarrhoea	135 (4.7%)
Nausea	83 (2.9%)
Musculoskeletal and connective tissue disorders	599 (20.7%)
Rheumatoid arthritis	175 (6.1%)

Back pain	116 (4.0%)
Arthralgia	68 (2.4%)
Osteoarthritis	66 (2.3%)
General disorders and administration site conditions	474 (16.4%)
Injection site erythema	214 (7.4%)
Injection site pruritus	105 (3.6%)
Investigations	571 (19.8%)
Alanine aminotransferase increased	289 (10.0%)
Transaminases increased	75 (2.6%)
Aspartate aminotransferase increased	53 (1.8%)
Injury, poisoning and procedural complications	644 (22.3%)
Accidental overdose	316 (10.9%)
Fall	98 (3.4%)

Table 33: Percentages of patients with AEs on SAR monotherapy (≥2% in at least one treatment group) (adapted from EPAR)[1]

Primary System Organ Class **Preferred Term**	**SAR monotherapy** **(N=467)** **N (%)**
Any class	285 (61.0%)
Infections and infestations	135 (28.9%)
Nasopharyngitis	28 (6.0%)
Bronchitis	16 (3.4%)
Upper respiratory tract infection	16 (3.4%)
Urinary tract infection	15 (3.2%)
Blood and lymphatic system disorders	82 (17.6%)
Neutropenia	73 (15.6%)
Nervous system disorders	32 (6.9%)
Headache	15 (3.2%)
Vascular disorders	18 (3.9%)

Hypertension	11 (2.4%)
Musculoskeletal and connective tissue disorders	51 (10.9%)
Rheumatoid arthritis	11 (2.4%)
General disorders and administration site conditions	49 (10.5%)
Injection site erythema	29 (6.2%)
Investigations	36 (7.7%)
Alanine aminotransferase increased	15 (3.2%)
Injury, poisoning and procedural complications	45 (9.6%)
Accidental overdose	22 (4.7%)

4.3 Critique of trials identified and included in network meta-analysis

Trials included in the NMA are listed in Table 34 and Table 35. Trial characteristics of these studies are included in the CS Appendix 8.6 and were considered appropriate by the ERG to permit inclusion in the NMA, with the exceptions of the Fleischmann 2009,[44] Choy 2012[45] and Go-FURTHER,[46,] studies where unlicensed doses were used and Kay 2008[47] where there were no eligible data owing to treatment crossover. The quality of the included RCTs was assessed using well-established and recognised criteria and is reported in Appendix 8.7 of the CS.

In the cDMARD-IR NMAs, multiple trials (ASSET[48], Chen 2009[49], Lan 2004[50], Weinblatt 1999[51], Taylor 2004[52], Maini 1998[53], Tam 2012[54], Tanaka 2011[55], Smolen, 2014[56] (part A) and Smolen, 2014[56] (part B)) were excluded because these trials included fewer than 30 patients per arm. In the TNF-IR NMAs, two trials (Schiff 2014[57] and Genovese 2014[58]) were excluded because of small sample sizes. The ERG argues that all evidence is relevant unless there is a reason to assume that a study has questionable quality. The company justified the exclusion of RACAT[59] and Machado 2014[60] stating that it was unable to link them in the network. The ERG disagrees with the decision because both trials had ETN 50mg every week plus MTX, which can be linked to the network. Three studies assessing monotherapy versus combination therapy (SURPRISE,[61-63] ACT-RAY,[64] and JESMR[65, 66]) were excluded because the company stated that they '*were not part of either of the population network diagrams*'. The ERG notes that if all studies had been included within one network, there would have been no need to exclude these trials.

Table 34: Studies included in the NMA for the cDMARD-IR population: Updated review (reproduced from Table 4.27 of the CS)

Intervention arm(s)	Control arm	Duration of study (weeks)	Number of patients	References
Monotherapy studies vs. (placebo or cDMARD)				
TNF studies				
ADA SC 20mg QW ADA SC 20mg Q2W ADA SC 40mg QW ADA SC 40mg Q2W	Placebo	26	544	ADA efficacy and safety study (van de Putte 2004[67])
ADA SC 20mg Q2W ADA SC 40mg Q2W ADA SC 80mg Q2W	Placebo	24	352	CHANGE (Miyasaka 2008[68])
CTZ SC 400mg Q4W	Placebo	24	220	FAST4WARD (Fleischmann 2009[44])
ETN SC 25mg BIW ETN SC 25mg BIW + SSZ	SSZ	104	254	ETN study 309 (Combe 2006,[49] Combe 2009[69, 70]
ETN SC 10mg BIW ETN SC 25mg BIW	Placebo	26	234	ETN monotherapy study (Moreland 1999[71])
IL-6 studies				
TCZ SC 8mg/kg Q4W	MTX	24	125	SARTORI (Nishimoto 2009[72])
TCZ IV 8mg/kg Q4W	ADA SC 40mg Q2W	32	325	ADACTA (GABAY 2013[73])
SIR SC 50mg Q4W SIR SC 100mg Q2W	ADA SC 40mg Q2W	24	559	SIRROUND-H (Taylor 2016[74])
SAR SC 200mg Q2W	ADA SC 40mg Q2W	24	369	MONARCH (Burmester 2016[35])
JAK inhibitors studies				
TOF oral 1mg BID TOF oral 3mg BID TOF oral 5mg BID	Placebo	24	384	Efficacy and safety of TOF vs. ADA (Fleischmann 2012[75])

Intervention arm(s)	Control arm	Duration of study (weeks)	Number of patients	References
TOF oral 10mg BID TOF oral 15mg BID ADA SC 40mg QW for 12 weeks followed by oral TOF 5mg BID for 12 weeks				
Combination studies vs. (placebo or cDMARD)				
TNF studies				
ADA SC 20mg QW + MTX ADA SC 40mg Q2W + MTX	MTX	52 (plus 10 year OLE)	619	DE019 (Keystone 2013,[76] Keystone 2011,[77] Keystone 2004[76])
ADA SC 40mg Q2W + MTX	MTX	24	128	ADA efficacy and safety study (Kim 2007[78])
ADA SC 20mg Q2W + MTX ADA SC 40mg Q2W + MTX ADA SC 80mg Q2W + MTX	MTX	24	271	ARMADA (Weinblatt 2003[79])
ADA SC 40mg Q2W + standard treatment	Placebo + standard treatment	24	636	STAR (Furst 2003[80])
CTZ SC 200mg Q2W + MTX CTZ SC 400mg Q2W + MTX	MTX	52	982	RAPID (Keystone 2008,[81] Strand 2009[82])
CTZ SC 100mg Q2W + MTX CTZ SC 200mg Q2W + MTX CTZ SC 400mg Q2W + MTX	MTX	24	316	J-RAPID (Yamamoto 2014[83])
CTZ SC 200mg Q2W + MTX CTZ SC 400mg Q2W + MTX	MTX	24	619	RAPID-2 (Smolen 2009[84])
CTZ SC 400mg Q2W + MTX	MTX	24	247	CTZ efficacy and safety study (Choy 2012[45])
CTZ SC 400mg Q2W + cDMARD	cDMARD	24	194	CERTAIN (Smolen 2015[85])

Intervention arm(s)	Control arm	Duration of study (weeks)	Number of patients	References
ETN SC 25mg BIW ETN SC 25mg BIW + SSZ	SSZ	104	254	ETN 309 study (Combe 2006,[70] Combe 2009[69])
ETN SC 25mg BIW + MTX	MTX	104	222	ENCOURAGE (Yamanka 2016[86])
GOL SC 50mg Q4W + MTX GOL SC 100mg Q4W + MTX	MTX	24	269	GO-FORTH (Tanaka 2012[87])
GOL SC 50mg Q2W + MTX GOL SC 50mg Q4W + MTX GOL SC 100mg Q2W + MTX GOL SC 100mg Q4W + MTX	MTX	52	172	GOL efficacy and safety study (Kay 2008[47])
GOL SC 50mg Q4W + MTX	MTX	52	264	GOL efficacy and safety study (Li 2016[88])
GOL SC 2mg/kg Q8W+ MTX	MTX	112	592	GO-FURTHER (Weinblatt 2014,[89] Bingham 2014,[46] Weinblatt 2013[90])
GOL SC 100mg Q4W GOL SC 50mg Q4W + MTX GOL SC 100mg Q4W + MTX	MTX	312	444	GO-FORWARD (Keystone 2016,[91] Keystone 2013,[92] Genovese 2012,[93] Keystone 2010,[94] Keystone 2009[95])
IFX IV 3mg/kg Q8W + MTX IFX IV 3mg/kg Q4W + MTX IFX IV 10mg/kg Q8W + MTX IFX IV 10mg/kg Q4W + MTX	MTX	54 (plus 1 year OLE)	428	ATTRACT (Maini 1999,[96] Lipsky 2000,[97] Maini 2004[98])
IFX IV 3mg/kg Q8W + MTX IFX IV 10mg/kg Q8W + MTX	MTX	54	1084	START (Westerhovens 2006[99])
IFX IV 3mg/kg Q8W + MTX ABT IV 8–10mg/kg + MTX	MTX	52	431	ATTEST (Schiff 2008[100])

Intervention arm(s)	Control arm	Duration of study (weeks)	Number of patients	References
IFX IV 3mg/kg Q8W + MTX	SSZ 1000mg (oral) BID + HCQ 400mg (oral) BID + MTX	104	245	SWEFOT (Karlsson 2013,[101] Rezaei 2013,[102] van Vollenhoven 2012,[103] van Vollenhoven 2009,[104] Eriksson 2013[105])
Non-TNFi studies				
ABT IV 8–10mg/kg Q4W + MTX	MTX	52	652	AIM (Russell 2007,[106] Kremer 2006[107])
ABT IV 2mg/kg Q4W + MTX ABT IV 10mg/kg Q4W + MTX	MTX	52	339	ABT efficacy and safety study (Emery 2006,[108] Kremer 2005,[109] Kremer 2003[110])
ABT IV 2mg/kg Q4W + MTX ABT IV 10mg/kg Q4W + MTX	MTX	32	194	ABT efficacy and safety study (Takeuchi 2013[111])
IFX IV 3mg/kg Q8W + MTX ABT IV 8–10mg/kg + MTX	MTX	52	431	ATTEST (Schiff 2008[100])
ABT IV 8–10mg/kg Q4W + cDMARD	cDMARD	52	1456	ASSURE (Weinblatt 2006[112])
RTX IV 2 x 500mg at days 1 and 15 + MTX RTX IV 2 x 1,000mg at days 1 and 15 + MTX	MTX	48	511	SERENE (Emery 2010[113])
RTX IV 2 x 500mg at days 1 and 15 + MTX RTX IV 2 x 1000mg at days 1 and 15 + MTX	MTX	24	367	DANCER (Mease 2008[114])
RTX IV 1,000mg days 1 and 15 RTX IV 1,000mg days 1 and 15 + MTX RTX IV 1,000mg days 1 and 15 + CYC 750mg days 3 and 17	MTX	104	161	RTX efficacy and safety study (Strand 2006,[115] Edwards 2004[116])
RTX IV 500mg + MTX RTX IV 1,000mg + MTX	MTX	52	185	RA-SCORE (Peterfy 2016[117])

Intervention arm(s)	Control arm	Duration of study (weeks)	Number of patients	References
RTX IV 1,000mg + LEF	LEF	52	140	AMARA (Behrens 2016[118])
IL-6 studies				
SAR SC 150mg Q2W + MTX SAR 200mg Q2W + MTX	MTX	52	1,197	MOBILITY B (Genovese 2015[33])
TCZ IV 4mg/kg Q4W + MTX TCZ IV 8mg/kg Q4W + MTX	MTX	24	623	OPTION (Smolen 2008[119])
TCZ IV 8mg/kg Q4W + MTX	MTX	24	132	MEASURE (McInnes 2015,[120] Mirjafari 2013[121])
TCZ IV 4mg/kg Q4W + MTX TCZ IV 8mg/kg Q4W + MTX	MTX	104	1,196	LITHE (Fleischmann 2013,[122] Kremer 2011[123])
TCZ SC 162mg Q2W + cDMARD	cDMARD	24	656	BREVACTA (Kivitz 2014,[124] Kivitz 2013[125])
TCZ IV 8mg/kg Q4W + cDMARD	cDMARD	24	1,220	TOWARD (Genovese 2008[126])
TCZ IV 8mg/kg Q2W + cDMARD	cDMARD	24	619	ROSE (Yazici 2012[127])
JAK inhibitors studies				
TOF oral 1mg BID + MTX TOF oral 3mg BID + MTX TOF oral 5mg BID + MTX TOF oral 10mg BID + MTX TOF oral 15mg BID + MTX TOF oral 20mg BID + MTX	MTX	24	509	TOF efficacy and safety study (Kremer 2012[128])
TOF oral 5mg BID + MTX TOF oral 10mg BID + MTX	MTX	104	797	Oral Scan (van der Heijde 2013[129])
TOF oral 5mg BID + MTX TOF oral 10mg BID + MTX TOF oral 40mg BID + MTX	MTX	52	717	Oral Standard (Van Vollenhoven 2012[130])

Intervention arm(s)	Control arm	Duration of study (weeks)	Number of patients	References
ADA SC 40mg Q2W + MTX				
TOF oral 5mg BID + cDMARD TOF oral 10mg BID + cDMARD	cDMARD	53	636	TOF efficacy and safety study (Kremer 2013[131])
BAR oral 2mg OD + cDMARD BAR oral 10mg OD + cDMARD	cDMARD	24	684	RA-BUILD (Dougados 2017[132])
Biologic vs. same biologic				
Comparisons of different routes of administration				
TCZ SC 162mg QW+ cDMARDs	TCZ IV 162mg Q4W+ cDMARDs	104	1,262	SUMMACTA (Burmester 2014,[133, 134] Burmester 2013[135])
Head-to-head comparisons of bDMARDs				
TNFi vs. non-TNFi				
ADA SC 40mg Q2W + MTX	ABT SC 125mg QW + MTX	104	646	AMPLE (Schiff 2014,[136] Weinblatt 2013[137])
ADA SC 40mg Q2W + MTX	BAR oral 4mg OD + MTX	52	1307	RA-BEAM (Taylor 2017[138])
IL-6 vs. TNFi				
TCZ IV 8mg/kg Q4W	ADA SC 40mg Q2W	32	326	ADACTA (Gabay 2013[73])
SAR SC 200mg Q2W	ADA SC 40mg Q2W	24	396	MONARCH (Burmester 2016[35])

ABT=abatacept; ADA=adalimumab; BAR= baricitinib; BID=Twice a day; BIW=twice weekly; cDMARD= disease-modifying anti-rheumatic drugs; CTZ= certolizumab pegol; CYC= cyclophosphamide; ETN= etanercept; GOL= golimumab; HCQ= hydroxychloroquine; IFX=infliximab; IL-6=interleukin-6; IV=intravenous; MTX=methotrexate; OD=once daily; OLE=open labelled extension; QW=once a week; Q2W=every 2 weeks; Q4W=every 4 weeks; Q8W=every 8 weeks; RTX= rituximab; SAR= sarilumab; SC=subcutaneous; SIR= sirukumab; SSZ= sulfasalazine; TCZ= tocilizumab; TOF= tofacitinib

Table 35: Studies included in the NMA for the TNFi-IR population: Updated review (reproduced from Table 4.28 of the CS)

Intervention	Comparator(s)	Duration of study (weeks)	Number of patients	References
Monotherapy studies vs. placebo				
GOL SC 50mg Q4W +/- cDMARD	cDMARDs	24	461	GO-AFTER (Smolen 2009[139])

GOL SC 100mg Q4W +/- cDMARD					
SIR SC 500mg Q4W +/- cDMARD SIR SC 1000mg Q2W +/- cDMARD		cDMARD	NA	878	SIRROUND-T (Tanaka 2016[140])
Combination studies vs. cDMARD					
Non-TNFi studies					
ABT IV 10mg/kg Q4W + cDMARD		cDMARD	26	258	ATTAIN (Westhovens 2006,[141] Genovese 2005[142])
RTX IV 1,000mg at days 1 and 15 + MTX		MTX	104	520	REFLEX (Keystone 2009,[143] Keystone 2008,[144] Cohen, 2006[145])
TOF oral 5mg BID + MTX TOF oral 10mg BID + MTX		MTX	26	399	Oral Step (Strand 2015,[146] Burmester 2013[147])
IL-6 studies					
TCZ IV 4mg/kg Q4W + MTX TCZ IV 8mg/kg Q4W + MTX		MTX	24	489	RADIATE (Strand 2012,[148] Emery 2008[149])
SAR SC 150mg Q2W + cDMARD SAR SC 200mg Q2W + cDMARD		cDMARD	24	546	TARGET (Fleischmann 2017[34])
JAK inhibitors studies					
BAR oral 2mg OD + cDMARD BAR oral 4mg OD + cDMARD		cDMARD	24	527	RA-BEACON (Genovese 2016[150])
Head-to-head comparisons of bDMARDs					
SAR SC 150mg Q2W + cDMARD SAR SC 200mg Q2W + cDMARD		TCZ IV 4-8mg/kg Q4W + cDMARD	24	202	ASCERTAIN (Sanofi Genzyme[36])
ABT (dose/frequency not stated)	RTX (dose/frequency not stated)	TNFi (dose/frequency not stated)	52	143	Open-label study (Manders 2015[151])

ABT=abatacept; BAR= baricitinib; bDMARD= biological disease-modifying anti-rheumatic drugs; cDMARD=conventional disease-modifying anti-rheumatic drugs; IL-6=interleukin-6; GOL= golimumab; IR=irresponsive; IV=intravenous; MTX=methotrexate; QW=once a week; Q2W=every 2 weeks; Q4W=every 4 weeks; Q8W=every 8 weeks; RTX= rituximab; SAR= sarilumab; SC=subcutaneous; SIR= sirukumab; TNFi=tumour necrosis factor inhibitor; TCZ= tocilizumab; TOF= tofacitinib.

4.4 Critique of the network meta-analysis

NMAs were performed separately for the cDMARD-IR and TNFi-IR populations using a Bayesian approach for efficacy and safety outcome measures at either 24 weeks or 52 weeks (see Table 36). In the cDMARD-IR population, studies investigating combination therapies and monotherapies were separated into two different networks. The TNFi-IR network only had studies with combination therapy, as no studies were identified investigating bDMARDs as monotherapy in this population.

The ERG agrees with the decision to perform separate analyses for the two populations. In contrast, the ERG does not agree with the use of separate networks for combination therapy and monotherapy in the cDMARD-IR population as three studies had been excluded because of the use of two networks (see critique in Section 4.3).

For continuous outcomes, HAQ-DI and mTSS, a normal likelihood with identity link function model was used in the NMA. For ordered categorical outcomes, ACR and EULAR response, a binomial likelihood with either a logit link function in meta-regression on baseline risk or a risk difference model was used by dichotomising the data. For binary data, the efficacy outcome (DAS28 remission) used either meta-regression on the baseline risk model with a logit link function or a risk difference model; safety outcomes (serious infections and serious adverse events) used either a risk difference model or logit model.

Table 36: **Outcomes and models used in the NMA per population and time point for the combination therapy**

Outcome	cDMARD-IR (combination therapy)		cDMARD-IR (monotherapy)	TNFi-IR
	Model (24 weeks)	**Model (52 weeks)**	**Model (24 weeks)**	**Model (24 weeks)**
ACR20, 50 and 70	Random effects-baseline risk regression		Fixed effect-logit model	Fixed effect-risk difference
HAQ-DI CFB	Random effects-change from baseline		Fixed effects-change from baseline	Fixed effect-change from baseline
EULAR moderate-to-good, good	Fixed effect-risk difference		Fixed effect-risk difference	Fixed effect-risk difference
DAS28 remission	Random effects-baseline risk regression		Fixed effects-risk difference	Fixed effect-risk difference
mTSS CFB	Fixed effect-change from baseline	Fixed effect-change from baseline		
SIs		Random effects-risk difference	Fixed effect-risk difference	Fixed effect-logit model
SAE		Random effects-logit model	Fixed effect-logit model	Fixed effect-logit model

ACR20/50/70=American College of Rheumatology 20%/50%/70% improvement; CFB=change from baseline; DAS28=28-joint disease activity score; HAQ-DI=Health Assessment Questionnaire Disability Index; mTSS=modified Total Sharp Score; NMA=network meta-analysis; SAE=serious adverse event; SI=serious infections

The ERG disagrees with dichotomising ACR and EULAR response. The choice of the likelihood function/link function should be based on the data generating process. A multinomial likelihood with probit/logit link function is preferred to a binomial likelihood for the ordered categorical ACR or EULAR data, because it accounts for natural ordering and correlations between the categories within the outcome measure. This is important to the decision problem when these results are used to populate the economic model.

Meta-regression on the baseline risk is not very useful for decision-making as it does not explain the heterogeneity in terms of prognostic factors. When there were too few studies to perform a meaningful regression, a risk difference scale was used for all the efficacy outcomes rather than the most frequently applied odds ratio scale (a logit model). The company stated that this was because the observed treatment effect was statistically significantly correlated with the observed baseline risk when the effect was measured using an odds ratio scale, but was not statistically significantly correlated when the effect was measured using risk difference scale. The ERG disagrees with the model selection procedure.

Firstly, a *p*-value is not very useful as an estimate of the strength of an association, because it is influenced by the number of observations. Secondly, it is known that the sample estimate of the treatment effect has a negative association with the sample estimate of the baseline on the odds ratio scale, but the true magnitude of the association depends on the between-study variance in the true underlying risk, which is unknown. Finally, the treatment effect needs to be constrained when using a risk difference model in the NMA so that the probability of achieving an event is bounded between 0 and 1. In response to clarification question A11,[31] it is suggested that the constraint used was only to limit the probability not to be less than 0, but it still allowed it to exceed 1.

Random effects models were used to allow for heterogeneity when sufficient data were available, with fixed effect models used when data were sparse. The ERG disagrees with the rationale that too few studies would rule out a random effects analysis. If heterogeneity is expected, then a random effect model should be applied with careful consideration of the prior for the between-study variance. In response to clarification question A12,[31] the company stated that a less vague prior was used for the regression coefficient, the relative treatment effect *d* and the baseline effect *mu*. The ERG argues that the less vague prior should be applied on the between-study variance when data were sparse.

In response to clarification question A9,[31] the company stated that the baseline absolute effect was calculated by averaging all study effects with cDMARD/MTX in both cDMARD-IR and TNFi-IR networks. The code for generating the baseline effect was also provided. The ERG notes that averaging the effects for studies with cDMARD/MTX may not be appropriate, as this does not account for uncertainty in the baseline treatment effect of cDMARD/MTX properly.

In response to clarification question A15,[31] the company stated that "*the goodness-of-fit was estimated by calculating the mean residual deviance of the model (mean residual deviance close to 1 was considered to be a good model fit)*". It was also stated that mean total residual deviance compared to the number of fitted data points was considered in selecting the preferred model. However, no comments regarding the performance of each model fitting were provided by the company.

The I^2 statistic was used to assess the heterogeneity for the pairwise treatment comparisons. Heterogeneity was observed in the cDMARD-IR combination therapy NMAs, but not observed in the cDMARD-IR monotherapy NMAs and TNFi-IR NMAs. The ERG notes that both cDMARD-IR monotherapy and TNFi-IR NMAs had limited data and the I^2 statistic calculation may be biased due to too few studies. Inconsistency was checked using the Bucher method.[152] No inconsistency was found in most of the NMAs, except for ACR50 response (ADA combination, cDMARD and tofacitinib combination loop) in the cDMARD-IR population.

In the base case cDMARD-IR NMA, SAR 200mg combination therapy (in combination with cDMARD) demonstrated statistically superiority to cDMARD for all efficacy outcome measures (Table 4.30 in the CS). SAR 200mg combination therapy was comparable to other bDMARD combination therapies on ACR responses, DAS28 remission and HAQ-DI (TCZ SC combination therapy was not included in the HAQ-DI network) at 24 weeks (Table 4.30 in the CS). SAR 200mg combination therapy showed statistically superiority to ABT combination, IFX combination, TCZ 4mg IV, RTX and SAR 150mg on EULAR good response at 24 weeks, and was comparable with GOL and TCZ 8mg IV, each in combination with cDMARD. SAR 200mg combination therapy was statistically inferior to CTZ combination on EULAR at least moderate response at 24 weeks, but was comparable with GOL, IFX, TCZ 4mg IV and 8mg IV, RTX and SAR 150mg each in combination with cDMARD (Table 4.30 in the CS). For mTSS at 24 weeks, SAR 200mg combination therapy was statistically superior to baricitinib 2mg, tofacitinib and CTZ each in combination with cDMARD, and comparable with baricitinib 4mg, ADA, GOL, TCZ SC 162mg each in combination with cDMARD (Table 4.30 in the CS). For mTSS at 52 weeks, SAR 200mg combination therapy was comparable to ABT, ADA, CTZ and ETN each in combination with cDMARD, and superior to SAR 150mg combination therapy (Table 4.30 in the CS).

In the cDMARD-IR monotherapy NMA, the outcome measures were all assessed at 24 weeks and the results were provided in Table 4.31 of the CS. SAR 200mg monotherapy showed statistically superiority to placebo and cDMARD for all the efficacy outcome measures, except that it was comparable with cDMARD on HAQ-DI, and DAS28 remission was not analysed for placebo. SAR 200mg monotherapy was also statistically superior to ADA on all ACR responses, and sirukumab 50mg on ACR20 and ACR50 response. SAR 200mg was comparable with CTZ, ETN, sirukumab 100mg, TCZ 8mg and tofacitinib on all ACR responses. SAR 200mg was statically superior to ADA and sirukumab 50mg on DAS28 remission, and comparable with sirukumab 100mg and TCZ 8mg. SAR 200mg was statistically superior to ADA on HAQ-DI, and comparable with CTZ, ETN and TCZ 8mg. SAR 200mg was statistically superior to ADA on EULAR responses, and comparable with TCZ 8mg.

In the TNFi-IR population, the outcome measures were all assessed at 24 weeks and the results were provided in Table 4.32 of the CS. SAR 200mg combination therapy showed statistically superiority to cDMARD for all the efficacy outcome measures. SAR 200mg combination was statistically superior to baricitinib 2mg combination, sirukumab 50mg combination on ACR50, and comparable with other bDMARD combination therapies on all ACR responses. SAR 200mg combination was statistically superior to ABT combination, baricitinib 2mg combination, GOL combination, sirukumab 50mg combination, TCZ 4mg combination, and RTX combination on DAS28 remission, and comparable with other bDMARD combination therapies. None of the effect on HAQ-DI was statistically significant. For EULAR good response, SAR 200mg combination was statistically superior to rituximab combination,

and comparable to abatacept combination and SAR 150mg combination. For EULAR at least moderate response, SAR 200mg combination was statistically inferior to TCZ 8mg combination and RTX combination, and comparable with ABT, GOL and SAR 150mg all as combination therapies.

In relation to safety data, SAR 200mg combination therapy was associated with significantly higher odds of SAEs at 52 weeks when compared with cDMARD in the cDMARD-IR population. All other results were not statistically significant (Tables 4.30-4.32 in the CS).

Scenario analyses were conducted assuming TNFis had identical efficacy. These analyses showed that SAR 200mg combination therapy was statistically superior to cDMARD and comparable with all other combination therapies on ACR20 at 24 weeks. SAR 200mg monotherapy was statistically superior to placebo, cDMARD, sirukumab 50mg monotherapy, TNF monotherapy and tofacitinib monotherapy and comparable with sirukumab 100mg monotherapy and TCZ 8mg monotherapy on ACR 20 at 24 weeks (Tables 4.33-4.34 in the CS).

The ERG considers that the base case NMA results in the CS should be interpreted with caution. The statistically significant results of SAR 200mg compared with other bDMARD treatments (both as combination therapy and monotherapy) may be as a result of using a fixed effect model, which underestimates uncertainty in the treatment effects. The ordered categorical ACR response and EULAR response data were dichotomised in the NMA, which ignores the natural ordering and correlations between the categories within the outcome measure. When a risk difference model was used for binary data, the probability was not constrained to be below 1.0. Furthermore, the MOBILITY B and TARGET trial designs allowed patients who did not achieve a ≥20% improvement from baseline in the swollen joint count or tender joint count at two consecutive assessments to switch to open-label SAR 200mg at 16 and 12 weeks, respectively. Non-responder imputation was carried out for the control arm, assuming that none of the non-responders in the cDMARD control group would become responders at 24 weeks, which may overestimate the relative treatment effect of SAR combination therapy versus cDMARD.

The ERG requested that the company perform additional analysis for ACR and EULAR response in both populations (see clarification response[31] --question A7) with the following settings:

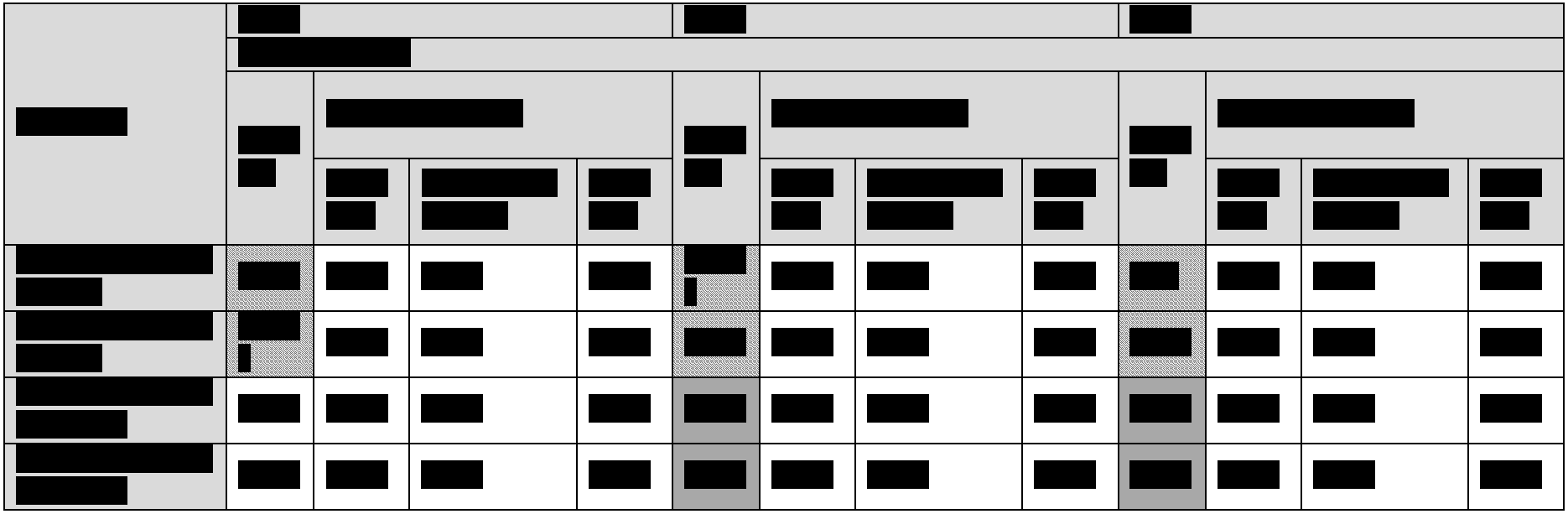
- Using a random effects probit model with an informative prior for the between-study variance (log normal with mean -2.56 and variance of 1.74*1.74, which is proposed by Turner *et al* 2012.[153] The log normal is truncated so that the odds ratio in one study would not be ≥50 times than in another, and re-scaled to match the probit scale).
- Keeping all treatments separate.

- Including combination therapy and monotherapy in a single network in order that trials comparing both regimens can provide evidence (including the studies in Appendix 8, Table 8.7 of the CS and HARUKA[154]).
- Including the studies which were excluded due to small sample size (CS Appendix 8 Table 8.1 and Table 8.9)
- Including the previously excluded studies that were included in TA375.
- Including the studies in Table 8.3 of the CS Appendix assuming that ETN 50mg once weekly was equivalent to ETN 25mg twice weekly.
- Incorporate the KAKEHASI[155] study for consistency with the main network, which includes studies in Asian patients.

The ERG also requested a sensitivity analysis for the requested NMA where TNFis were pooled into a 'TNFi-bundle'.

The company only provided the results for the cDMARD-IR population on ACR responses. The company justified not using a random effects probit model for the TNFi-IR population on ACR responses stating that "*this analysis produces results that are inconsistent with the observed head-to-head data from both the RADIATE and TARGET studies. The results from the random effects probit model both significantly under- and over-estimate relative treatment effect compared to trial data*" (clarification response[31] -- question A7). The comparison of ACR responder rate as observed data and the values estimated from the NMA using probit link at 24 weeks are reproduced in Table 37. The ERG disagrees with the approach that was taken by the company in determining that a random effects probit model was not a suitable model for ACR responses in the TNFi-IR population. This is because the absolute effect (the responder rate) was compared between the observed data and the values estimated from the NMA, and the estimated responder rate from NMA shown in Table 37 depends on how the baseline effect was estimated. The inconsistency observed in Table 37 was because the chosen baseline effect was different from the baseline effect in the RADIATE and TARGET studies. The relative effect should be used for comparison not the absolute effect as used by the company. In addition, the findings in the TNFi-IR population may not apply to the cDMARD-IR population.

Table 37: **Comparison ACR20/50/70 responder rate as observed (direct results) and estimated from NMA using probit link approach in a random effects model at 24 weeks in TNF-IR population**



White cells mean NMA predicts well, hatched cells that the NMA over predicts, grey cells that the NMA under predict

ACR: American College of Rheumatology; ACR20/50/70: 20%/50%/70% improvement in the ACR score; combi: combination therapy CrI: credible interval; SAR: sarilumab; TCZ: tocilizumab

In addition, the company did not conduct the requested analysis for EULAR responses. The company justified this omission stating that the data for EULAR outcomes were only available for two categories (EULAR no response and EULAR at least moderate response). The ERG notes that EULAR good response data are available from nine studies in the cDMARD-IR combination therapy network, four studies in the cDMARD-IR monotherapy network and three studies in the TNFi-IR network (Table 8.23, Table 8.32 and Table 8.43 in the CS Appendix). Although not all the studies have reported data in all three EULAR categories, the probit model is able to incorporate such data.

The company concluded that the results for ACR outcomes from the requested NMA were in line with the results in the original CS and the conclusion that SAR in combination with cDMARD showed comparable efficacy to other bDMARDs was unchanged. The ERG notes that the additional analyses performed by the company involved meta-regression on a baseline risk model with a probit link function, rather than the standard probit NMA model. All the results presented were the effects with covariate adjustment. To make the results more interpretable, the analyses used centred covariate values by subtracting the mean covariate value. However, the company did not report what this mean covariate value was. The estimates for the covariate coefficient suggested that there was not enough evidence for an interaction effect between the baseline risk and treatment effects on the probit scale. The ERG agrees with the company that the conclusion has not altered, but notes that the results from the requested NMA may be numerically different from the original NMA in the CS. The comparison of ACR responder rates between the two NMAs is presented in [redacted] to [redacted] for the combination therapies and [redacted] to [redacted] for the monotherapies.

[redacted]2[redacted]

[redacted]3[redacted]

[redacted]

[redacted]

[redacted] 6 [redacted]

[redacted] 7 [redacted]

4.5 Additional work on clinical effectiveness undertaken by the ERG

No additional work was undertaken by the ERG as the company performed the analyses requested by the ERG, albeit with some deviations.

4.6 Conclusions of the clinical effectiveness section

The key clinical effectiveness evidence for SAR was based on five RCTs. Additionally one long-term extension study was included. There were three RCTs in MTX-IR RA patients (MOBILITY-A, MOBILITY-B, MONARCH). Two RCTs were undertaken in a TNFi-IR RA population (TARGET and ASCERTAIN). One RCT had a comparator of TCZ (ASCERTAIN), one had a comparator of ADA (MONARCH); the other RCTs included a PBO comparator.

Three RCTs had ACR20 as their primary endpoint (MOBILITY-A, MOBILITY-B and TARGET). In the MTX-IR population, the RCTs showed a significant advantage in ACR responses for licensed doses of SAR+MTX over PBO+MTX ($p \leq 0.05$) (MOBILITY-A, MOBILITY-B), and a significant advantage for SAR monotherapy over ADA monotherapy ($p<0.01$) (MONARCH). In the TNFi-IR population, TARGET reported a significant advantage for SAR+cDMARD over PBO+cDMARD for ACR20 ($p<0.0001$), and ACR50 and ACR70 ($p \leq 0.005$).
[redacted]

[redacted]

[redacted]

[redacted]

The MOBILITY-A, MOBILITY-B, MONARCH and TARGET trials reported significantly favourable results for licensed doses of SAR over comparators for improvement in HAQ-DI. SAR had a significant advantage over comparator for DAS28-CRP in the MOBILITY-B and TARGET trials, and for DAS28-ESR in the MONARCH trial. MOBILITY-B measured radiographic progression by mTSS, and reported a significantly lower deterioration from baseline for SAR over comparator.

[redacted]

The ERG considers that the base case NMA results in the CS should be interpreted with caution. The statistically significant results of SAR 200mg compared with other bDMARD treatments (both as combination therapy and monotherapy) may be as a result of using a fixed effect model, which underestimates uncertainty in the treatment effects. The ordered categorical ACR response and EULAR response data were dichotomised in the NMA; this ignores the natural ordering and correlations between the categories within the outcome measure. When a risk difference model was used for binary data, the probability was not constrained to be below 1.0. Furthermore, the MOBILITY B and TARGET trial designs allowed patients who did not achieve a ≥20% improvement from baseline in the swollen joint count or tender joint count at two consecutive assessments to switch to open-label SAR 200mg at 16 and 12 weeks, respectively. Non-responder imputation was carried out for the control arm, assuming none of the non-responders in the cDMARD control group would become responders at 24 weeks; this may overestimate the relative treatment effect of SAR combination therapy versus cDMARD.

5 COST EFFECTIVENESS

This chapter presents a review of the cost-effectiveness evidence provided in the CS for SAR, with or without MTX, for treating moderate to severe, or severe RA. For brevity, the moderate to severe RA group is referred to as moderate RA The cost-effectiveness evidence comprised a systematic review of economic analyses that included SAR and the economic analysis based on the company's *de novo* model. Following the clarification round,[31] the company made a number of amendments to the model at the request of the ERG, which resulted in different ICERs to those presented in the CS; the broad conclusions remained unchanged for patients with severe RA, but are different for patients with moderate RA. The ERG report will discuss only the latest version of the model unless there is a clear reason to provide significant detail to the original version.

5.1 ERG's comment on company's review of cost-effectiveness evidence

5.1.1 Objective of cost effectiveness review

The company performed a literature search in order to identify cost-effectiveness evaluations of bDMARDs used to treat people with moderate or severe RA.

5.1.2 Inclusion and exclusion criteria used in the company's review

A full description of the company's search strategy is provided in Appendix 9 of the CS. The company's review was undertaken in two stages. An initial review was conducted in March 2014 searching MEDLINE, Embase, Health Economic Evaluation Database (HEED) and NHS Economic Evaluation Database (NHS EED). An update to this review was performed in December 2016, with the exceptions of HEED and NHS EED, whose coverage expired in December 2014 and March 2015 respectively. Conference proceedings were not included due to "*the limited reporting of methodologies in such publications.*"

The inclusion and exclusion criteria of the company's review are presented in Table 38.

Table 38: Inclusion and exclusion criteria of the company's review (reproduced from Table 5.1 of the CS)

Domain	Inclusion criteria	Exclusion criteria
• **Population**	• Adult patients with moderate-to-severe RA • Refractory to cDMARD or TNFi therapy • Or: • Intolerant to cDMARD or TNFi therapy	• Any patient population other than adult patients with moderate-to-severe RA • Studies that do not report separate results for moderate-to-severe RA patients
• **Intervention / comparators**	• bDMARDs	• Any treatment other than bDMARDs
• **Outcomes**	• Model characteristics • Costs/utilities/disutilities • LYs/QALYs • CERs/ICERs	• Epidemiologic outcomes • Clinical efficacy and safety outcomes • PROs • Other economic outcomes
• **Study designs**	• Economic evaluations: trial-based economic analyses and economic models • Cost-benefit analyses • Cost-effectiveness analyses • Cost-utility analyses	• The following study designs without an economic evaluation component • Cross-sectional studies • RCTs • Longitudinal observational studies • Economic evaluations: trial-based economic analyses and economic models • Cost-minimisation analyses • Cost-consequence analyses • Budget impact analyses
Geography	No limitation in regards to geography	
Time period	No date restrictions were applied	
Language	English language	Non-English language

bDMARD=biologic disease-modifying anti-rheumatic drug; cDMARD= disease-modifying anti-rheumatic drug; CER = cost-effectiveness ratio; DMARD = disease-modifying anti-rheumatic drugs; ICER = incremental cost-effectiveness ratio; LYs = life years; PRO=patient-reported outcomes; RA=rheumatoid arthritis; RCTs=randomised controlled trials; TNFi=tumour necrosis factor inhibitor; QALYs = quality-adjusted life years

The ERG has some concerns about what has been excluded in the company's review (every instance of the word "review" in searchable fields – this would include its use in a figurative sense as well as in reference to a type of evidence synthesis). Additionally, the ERG queries whether it was necessary to exclude all secondary evidence relating to the cost-effectiveness of SAR in this fashion; particularly

given the company's statement in their response to the clarification letter that "the references of any systematic literature review identified in the searches were reviewed for studies matching the inclusion criteria" (clarification response[31] -- Literature searching, Q2).[31]

In response to the ERG's query on the use of limits in its clarification letter,[31] the company justified its decision by citing several other NICE TAs in RA as evidence that it was unlikely that there were any more published cost-effectiveness studies the review had missed. In spite of concerns regarding the method of retrieval, the ERG considers it unlikely that any significant cost-effectiveness studies have been overlooked by this systematic literature review.

5.1.3 Findings of the cost-effectiveness review

A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram is presented by the company in Figure 5.1 of the CS. A total of 76 records were identified, of which 50 were economic evaluations and 26 were health technology assessment reports with economic models. A description of the identified studies are provided in Section 5.1.2 of the CS with further information provided in Appendix 10 and Appendix 11 of the CS. None of the identified studies considered SAR.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

Given that none of the identified records considered SAR, the company constructed a *de novo* model to address the cost-effectiveness of SAR, as monotherapy or in combination with MTX. The company state that the parameters in the *de novo* model "*were largely informed by previous models with special consideration to the independent assessment group model in TA375"*. This model has been published in a peer-reviewed journal.[24]

5.2.1 NICE Reference Case checklist

A summary of the key features of the company's *de novo* model relating to the NICE Reference Case[156] is provided in Table 39.

Table 39: Adherence of the company's economic analysis to the NICE Reference Care

Element	Reference case	Satisfactorily addressed within the CS?	ERG Comments
Defining the decision problem	The scope developed by NICE	Yes	-
Comparators	As listed in the scope developed by NICE	Mostly	Some comparators have been excluded from the decision problem including: biosimilars for ADA and RTX; and MTX alone where other bDMARDs have been recommended by NICE.
Perspective on costs	NHS and Personal Social Services (PSS)	Yes	-
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Yes	Health gains for patients are modelled in terms of QALYs gained.
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	Yes	-
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes	The time horizon of the analysis is 100 years, which is assumed to be representative of patients' remaining lifetimes.
Synthesis of evidence on health effects	Based on systematic review	Mostly	The ERG has concerns with the NMA (see Section 4.4).
Measure and valuation of health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	Yes	Health gains were valued in terms of QALYs. HAQ scores were mapped using three methods: (i) Malottki *et al.*[157] used in TA195;[27] Hernández-Alava *et al*[158] accepted by the Appraisal Committee in TA375;[25] and, (iii) Bansback *et al.*[159]
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes	Resource use estimates associated with HAQ categories were based on data from the Norfolk Arthritis Register (NOAR) database[160] and were inflated to 2016 values.
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	Yes	-
Equity considerations	An additional QALY has the same weight regardless of the other characteristics	Not applicable	No additional equity weighting is applied to the estimated QALY gains.

Element	Reference case	Satisfactorily addressed within the CS?	ERG Comments
	of the individuals receiving the health benefit		

5.2.2 Population

Patient-level data from three SAR trials (MOBILITY B,[33] TARGET[34] and MONARCH[35]) were used to populate the company's model. Table 5.3 of the CS uses alpha-numeric coding for each patient group. This convention was not intuitive for the ERG who have renamed the population groups for the purposes of this report. Further, the group labelled C4 was not clear, but the ERG has attempted to interpret this based on the sequences evaluated in this group.

The groups as renamed by the ERG are:

- cDMARD-IR patients with severe RA who can tolerate MTX (CS denoted A1);
- cDMARD-IR patients with severe RA who cannot tolerate MTX (CS denoted B);
- TNFi-IR patients with severe RA who can tolerate RTX and MTX (CS denoted C2);
- TNFi-IR patients with severe RA who cannot tolerate RTX (CS denoted C1);
- TNFi-IR patients with severe RA who cannot tolerate MTX (CS denoted C3);
- TNFi-IR patients who have received RTX and MTX (CS denoted C4); and
- cDMARD-IR patients with moderate RA and DAS28 between 4.0 and 5.1 who can tolerate MTX (CS denoted A2).

The data sources for the modelled population differ from the approach used in TA375[25] in which data from the British Society for Rheumatology Biologics Register (BSRBR) were used. The CS states that the baseline characteristics of the MOBILITY-B study (used for cDMARD-IR patients with severe RA who can tolerate MTX and for cDMARD-IR patients with moderate RA who can tolerate MTX), the MONARCH study (used for cDMARD-IR patients with severe RA who cannot tolerate MTX) and the TARGET study (used for all remaining populations) were found to be similar to data from the BSRBR. Data on the baseline patient characteristics are provided in Table 4.12 (p93-94) of the CS. Data used in the model are provided in Table 40.

Table 40: Population characteristics used in the model

	MOBILITY B	MONARCH	TARGET
Age (Years) (SD)	50.8 (12.5)	52.2 (12.3)	52.9 (12.3)
Proportion Male	18.3%	16.8%	18.1%
Weight (Kg) (SD)	74.39 (18.52)	72.05 (17.15)	78.8 (21.52)
HAQ score	1.64 (0.64)	1.64 (0.60)	1.78 (0.63)

5.2.3 Interventions and comparators

Descriptions of the intervention and the comparators are provided in Sections 3.2 and 3.3. The company did not include biosimilars for ADA or RTX. The ERG also notes that CTZ used as a monotherapy has not been included in the decision problem.

The model compares sequences of treatments that for simplicity include a 'TNFi bundle' in the base case. This TNFi bundle used the pooled efficacy of TNFis with the price weighted according to the estimated market share of each TNFi. The market share assumed by the company (Table 5.7, p211 of the CS) has been reproduced in Table 41, although the company have marked the data as commercial-in-confidence. These data were estimated from a freedom of information request to all UK hospital trusts asking for the number of RA patients treated with each named bDMARD between September and December 2016. The ERG comment that these data are likely to change as based on clinical advice provided to the ERG, clinicians are advised to start patients requiring bDMARDs on a biosimilar.

Table 41: Assumed market share of TNFis

TNFi	Market share
Etanercept (Enbrel®)	[redacted]
Etanercept biosimilar (Benepali ®)	[redacted]
Adalimumab (Humira®)	[redacted]
Infliximab (Remicade®)	[redacted]
Infliximab biosimilar (Remsima®/Inflectra®)	[redacted]
Golimumab (Simponi ®)	[redacted]
Certolizumab pegol (Cimzia®)	[redacted]
Total	**100%**

Different treatment sequences were evaluated for each of the populations. The sequences evaluated in the CS are reproduced in Appendix 1 of this report (Table 63 to Table 69). The ERG were concerned that these sequences were not consistent with those accepted in TA375 and requested that the company perform analyses using an alternative set of sequences. Following clarification,[31] the company evaluated the set of sequences requested by the ERG, which are provided in Table 42 to Table 48. The ERG notes that it erroneously included a second line of biologics in some sequences for the TNFi-IR RTX-ineligible population as indicated in Table 45. These sequences have been used in the company's analyses but have been amended in the ERG's exploratory analyses. A particularly significant change is for patients with moderate RA, where a strategy that incorporates patients becoming severe and then receiving bDMARDs has been added.

Table 42: Treatment sequences for a cDMARD-IR population with severe RA who can tolerate MTX

	SAR+MTX	**TCZ IV + MTX**	**TCZ SC + MTX**	**TNFi bundle + MTX**	**ABT SC + MTX**
1	SAR + MTX	TCZ IV + MTX	TCZ SC + MTX	TNFi bundle + MTX	ABT SC + MTX
2	RTX + MTX	RTX + MTX	RTX + MTX	RTX + MTX	RTX + MTX
3	MTX	MTX	MTX	TCZ IV + MTX	TCZ IV + MTX
4	BSC	BSC	BSC	MTX	MTX
5				BSC	BSC

ABT, abatacept; BSC, best supportive care; MTX, methotrexate; RTX, rituximab; SAR, sarilumab; TCZ, tocilizumab; IV, intravenous; SC, subcutaneous; TNFi, tumour necrosis factors inhibitor

Table 43: Treatment sequences for a cDMARD-IR population with severe RA who cannot tolerate MTX

	SAR	**TCZ IV**	**TCZ SC**	**TNFi bundle**
1	SAR	TCZ IV	TCZ SC	TNFi bundle
2	TNFi bundle	TNFi bundle	TNFi bundle	TNFi bundle
3	SSZ	SSZ	SSZ	SSZ
4	BSC	BSC	BSC	BSC

BSC, best supportive care; SAR, sarilumab; SSZ, sulfasalazine; TCZ, tocilizumab; IV, intravenous; SC, subcutaneous; TNFi, tumour necrosis factors inhibitor

Table 44: Treatment sequences for a TNFi-IR population with severe RA who can tolerate RTX and MTX

	SAR	**RTX**	**SAR,TCZ**	**RTX,TCZ**
1	SAR + MTX	RTX + MTX	SAR + MTX	RTX + MTX
2	MTX	MTX	TCZ IV + MTX	TCZ IV + MTX

3	BSC	BSC	MTX	MTX
4			BSC	BSC

BSC, best supportive care; MTX, methotrexate; RTX, rituximab; SAR, sarilumab; TCZ IV, intravenous tocilizumab; TNFi, Tumour necrosis factors inhibitor

Table 45: Treatment sequences for a TNFi-IR population with severe RA for whom RTX is not an option

	SAR + MTX	**TCZ IV + MTX**	**TCZ SC + MTX**	**TNFi bundle + MTX**	**ABT SC + MTX**
1	SAR + MTX	TCZ IV + MTX	TCZ SC + MTX	TNFi bundle + MTX	ABT SC + MTX
2	MTX	MTX	ABT SC + MTX*	TCZ IV + MTX*	TCZ IV + MTX*
3	BSC	BSC	MTX	MTX	MTX
4			BSC	BSC	BSC

ABT, abatacept ; BSC, best supportive care; MTX, methotrexate; SAR, sarilumab; TCZ, tocilizumab; IV, intravenous; SC, subcutaneous; TNFi, tumour necrosis factors inhibitor
*Erroneously included in the sequences requested by the ERG and in the analyses presented by the company in their clarification response but excluded from the ERG's exploratory analyses.

Table 46: Treatment sequences for a TNFi-IR population with severe RA who cannot tolerate MTX

	SAR	**TNFi bundle**
1	SAR	TNFi bundle
2	SSZ	SSZ
3	BSC	BSC

BSC, best supportive care; SAR, sarilumab; SSZ, sulfasalazine; TCZ, tocilizumab; TNFi, tumour necrosis factors inhibitor

Table 47: Treatment sequences for a TNFi-IR population with severe RA who have already received RTX + MTX

	SAR + MTX	**TCZ IV + MTX**	**TCZ SC + MTX**
1	SAR + MTX	TCZ IV + MTX	TCZ SC + MTX
2	MTX	MTX	MTX
3	BSC	BSC	BSC

BSC, best supportive care; MTX: methotrexate; SAR, sarilumab; TCZ, tocilizumab; IV, intravenous; SC, subcutaneous

Table 48: Treatment sequences for the cDMARD-IR population with moderate RA

	Moderate sequences	
	SAR + MTX	**MTX**
1	SAR + MTX	MTX
2	MTX	BSC
3	BSC	
	Severe sequences	
1	TNFi bundle + MTX	
2	RTX + MTX	
3	TCZ IV + MTX	
4	SSZ*	
5	BSC	

BSC, best supportive care; MTX, methotrexate; RTX, rituximab; SAR, sarilumab; SSZ, sulfasalazine; TCZ IV, intravenous tocilizumab; TNFi, tumour necrosis factors inhibitor

*The ERG notes that MTX could replace SSZ in this position

5.2.4 Perspective, time horizon and discounting

The model takes the perspective of the NHS and PSS. The time horizon is 100 years, which is assumed to be representative of a lifetime horizon. All costs and benefits were discounted at 3.5% per annum in line with the NICE Reference Case.[156]

5.2.5 Model structure

The company used a Markov model approach which differed from the discrete event simulation (DES) method used by the AG in TA375.[25] A Markov model requires the definition of time cycles and half-cycle correction. The company selected cycle lengths of six months to "*mirror the frequent of treatment decisions in the UK as per NICE guidance*". The ERG comments that a DES approach is more appropriate than a Markov approach, as fixed time cycles have limitations when costs are not apportioned equally through the cycle, for instance when 3 months' of intervention may be provided on day 1 of the cycle, and when patients discontinue treatment during a cycle which misaligns all subsequent six-month response periods.

The model structure presented by the company is reproduced in Figure 8 with a flow schematic shown in Figure 9.

Figure 8: Model structure presented by the company (reproduced from Figure 5.4 of the CS)

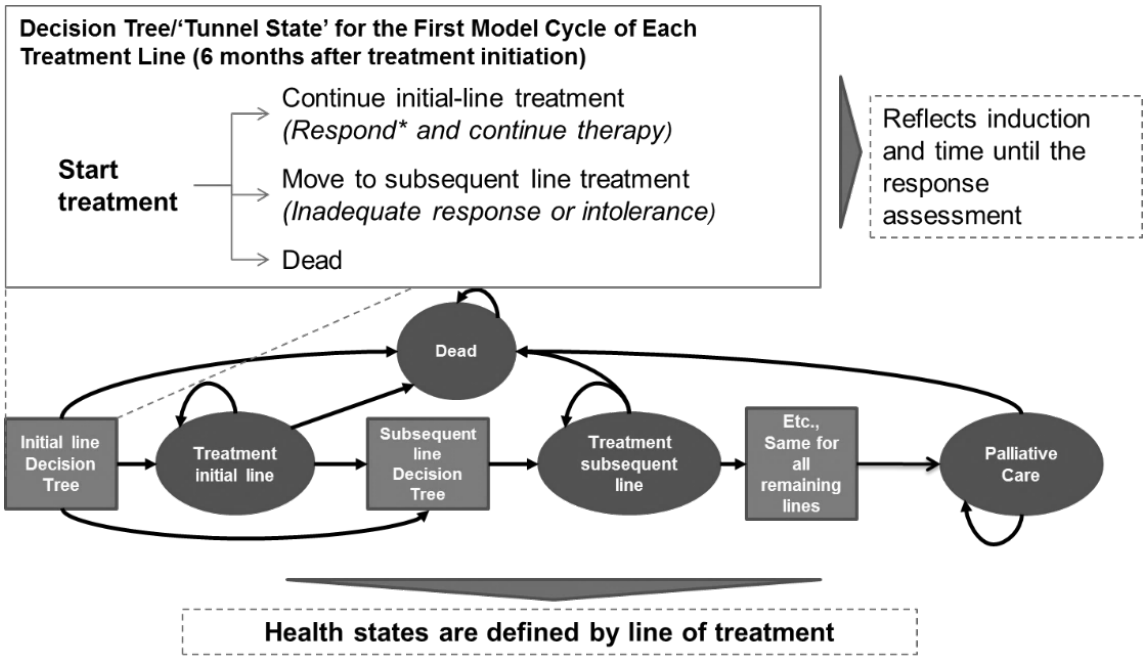
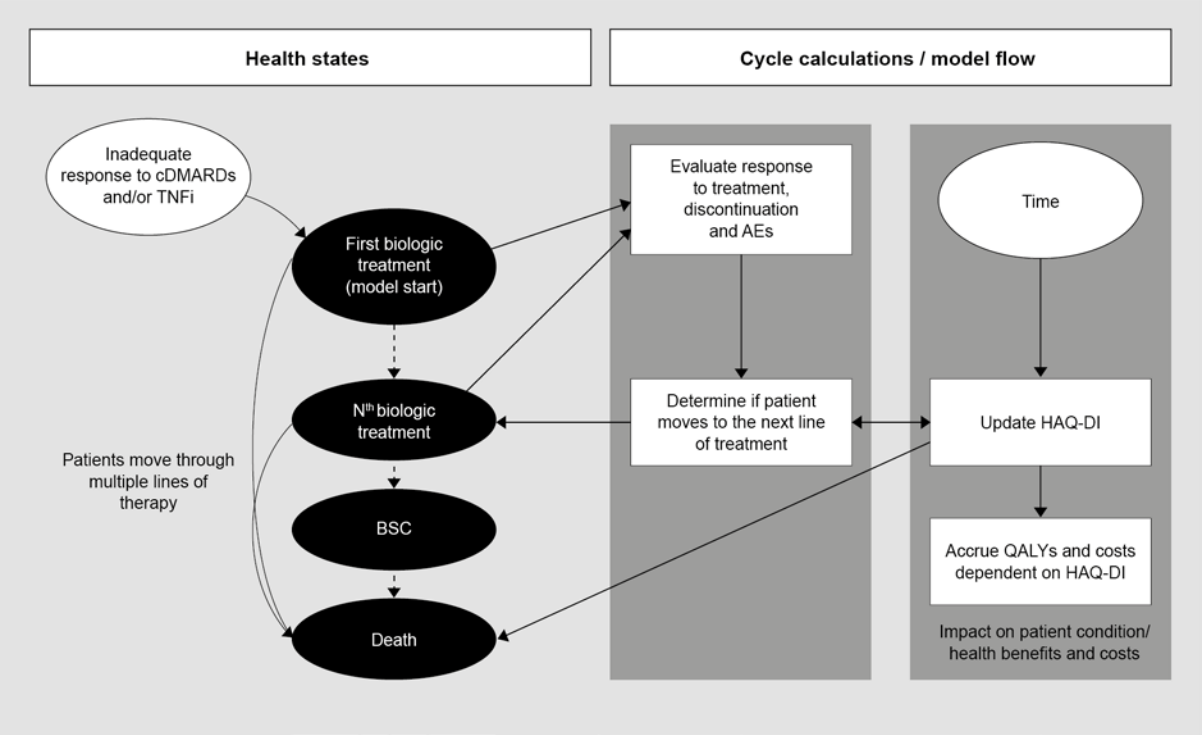


Figure 9: Model flow schematic presented by the company (reproduced from Figure 5.5 of the CS)



BSC=best supportive care; HAQ-DI=Health Assessment Questionnaire Disability Index; TNFi=tumour necrosis factor inhibitor

Within the model a clinical response in terms of EULAR (good, moderate or none) is estimated at six months. Patients who experience either a good or a moderate EULAR response remain on treatment; those who experience no response have their treatment withdrawn and move on to the next treatment in the sequence, unless the patient was already receiving BSC. Throughout the model, the costs incurred and the utility of the patient were assumed to be related to HAQ score.

5.2.6 Treatment effectiveness, extrapolation and discontinuation

The company estimated the probabilities of EULAR responses for SAR and competitors by initially undertaking an NMA of ACR responses, applying odds ratios for each intervention to predicted ACR responses on cDMARD to obtain estimated absolute ACR responses for each intervention. The ACR responses were then transformed to EULAR responses using a simple mapping that was used by the AG in TA375 based on data within the Veteran's Affairs Rheumatoid Arthritis (VARA) registry. The ERG comments that this mapping was not particularly robust due to small sample sizes (for example, only two patients had an ACR70 response in the VARA registry, one of whom had a moderate EULAR response and one who had a good EULAR response). The mapping was used in TA375 with the sole purpose of providing a secondary validation to the ICERs generated when using EULAR data directly. In TA375 the ICERs were fairly consistent regardless of whether the direct EULAR data, or ACR mapped to EULAR data were used.

The absolute EULAR responses estimated by the company are reproduced in Table 49 for cDMARD-IR patients, in Table 50 for cDMARD-IR patients who cannot receive MTX and in Table 51 for patients who are TNFi-IR. The company acknowledged the lack of evidence for TNFi-IR patients who cannot receive MTX and justified its assumption that the estimates for this population would be equal to those of combination therapies in TNFi-IR patients. The ERG notes that MTX alone, or alternative cDMARDs for those who cannot receive MTX have not been included in these efficacy tables, as cDMARDs not in combination with bDMARDs were excluded from the sequences evaluated. [redacted]

Table 49: Absolute EULAR responses estimated by the company in cDMARD-IR patients

Intervention	At least moderate EULAR response, % (95% CI)	Good EULAR response, % (95% CI)
SAR + MTX	[redacted]	[redacted]
TNFi bundle + MTX	[redacted]	[redacted]
TCZ (IV) + MTX	[redacted]	[redacted]
TCZ (SC) + MTX	[redacted]	[redacted]
ABT (SC) + MTX	[redacted]	[redacted]

ABT= abatacept; CI=confidence interval; EULAR= European League Against Rheumatism; IV=intravenous; MTX=methotrexate; TNFi=Tumour Necrosis Alpha inhibitor; SAR=sarilumab; SC=subcutaneous; TCZ=tocilizumab;

Table 50: Absolute EULAR responses estimated by the company in cDMARD-IR patients who cannot receive MTX

Intervention	At least moderate EULAR response, % (95% CI)	Good EULAR response, % (95% CI)
SAR	[redacted]	[redacted]
TNFi bundle	[redacted]	[redacted]
TCZ (IV)	[redacted]	[redacted]
TCZ (SC)	[redacted]	[redacted]

CI=confidence interval; EULAR= European League Against Rheumatism; TNFi=Tumour Necrosis Alpha inhibitor; IV=intravenous; SAR=sarilumab; SC=subcutaneous; TCZ=tocilizumab;

Table 51: Absolute EULAR responses estimated by the company in TNFi-IR patients

Intervention	At least moderate EULAR response, % (95% CI)	Good EULAR response, % (95% CI)
SAR	[redacted]	[redacted]
TNFi bundle + MTX	[redacted]	[redacted]
TCZ (IV) + MTX	[redacted]	[redacted]
TCZ (SC) + MTX	[redacted]	[redacted]
ABT (SC) + MTX	[redacted]	[redacted]
RTX (IV) + MTX	[redacted]	[redacted]

ABT= abatacept; CI=confidence interval; EULAR= European League Against Rheumatism; TNFi=Tumour Necrosis Alpha inhibitor; IV=intravenous; MTX=methotrexate; RTX= rituximab; SAR=sarilumab; SC=subcutaneous; TCZ=tocilizumab;

HAQ improvement upon treatment response

After six months, patients are assumed to be assessed for response. Patients who achieved a moderate or good EULAR response were assumed to have an associated reduction in HAQ score which is assumed independent of treatment. This value was taken from MOBILITY-B[33] and is reproduced in Table 52. The ERG notes that these values are percentage reductions, whereas fixed reductions in HAQ score conditional on EULAR response were used by the AG in TA375.[25] Following the clarification

round,[31] the company provided results using the values in TA375: these were reductions of 0.672 for patients who experienced a good EULAR response, and 0.317 for patients who experienced a moderate EULAR response.

Table 52: Changes in HAQ score conditional on EULAR response

	Original company submission	Post clarification
Treatment response	**% change in HAQ score (95% CI)**	**Change in HAQ score**
EULAR— No response	−7.17% (−15.98%, 1.63%)	0
EULAR—Moderate response	−22.63% (−28.27%, −16.99%)	-0.317
EULAR—Good response	−47.28% (−55.70%, −38.86%)	-0.672

CI=confidence interval; EULAR= European League Against Rheumatism

HAQ trajectory following initial response

In the base case, patients on bDMARD treatment are assumed to have zero HAQ progression in line with assumptions made in the AG model for TA375.[25] Supportive data were provided for the assumption for SAR using data from EXTEND,[42] an open-label study which recruited people from the MOBILITY B[33] and TARGET[34] RCTs. The company states that HAQ scores "*remained constant after the initial Week 24 improvement*" which the ERG acknowledges to appear to be correct. Further data from the BSRBR database and RHUMADATA, a large clinical database and registry in Canada, were presented to support the assumption of a constant HAQ score whilst on bDMARDs.

For patients on best supportive care, the company's base case assumes that HAQ scores progress at a rate of 0.06 per year; after clarification response,[31] the company assumed that the HAQ score of patients on cDMARDs would progress at a rate of 0.045 per year. The ERG believes that these analyses are inappropriate as HAQ progression has been proven to be non-linear[30] with the Appraisal Committee in TA375 in favour of a non-linear approach advocated by the AG.[25] This method used a modified version of the latent class approach of Norton *et al*., [161] which identifies four classes of HAQ trajectory: low, moderate, high and severe. Norton *et al*. report a regression model to calculate each patient's probability of belonging to each class based on the patient's baseline characteristics. The ERG comments that the linear method is not likely to significantly affect the conclusions in the comparison of SAR with bDMARDs, due to similar efficacy levels as shown in Section 4.4, but could have a significant effect, favourable to SAR, when the comparator is cDMARDs.

Adjustments to HAQ scores to consider initiation and discontinuation of treatments

In order to take into account the gradual improvement in HAQ score upon treatment initiation and the gradual deterioration in HAQ score prior to discontinuation the company adjust the HAQ score in the first and last cycle of a treatment. In both cases, the HAQ score in the cycle is calculated as the average

of two values: the HAQ score prior to treatment and the HAQ score following response for the initiation cycle; and the HAQ score following response and the HAQ score on treatment discontinuation. As the model assumes that HAQ score remains constant on bDMARDs, this means that the value is equal in both amended time cycles. The ERG believes that this adjustment is reasonable.

After applying changes to HAQ scores, the resulting values were rounded in the original CS to the nearest valid HAQ score (which is a multiple of 0.125). The ERG notes that this approach can lead to inaccurate results and contrasts with the approach used in TA375[25] where scores are rounded to either the higher or the lower valid HAQ score with a probability proportional to their distance to each (e.g. a value twice closer to the upper HAQ score would be twice as likely to be simulated as the upper score than simulated as the lower score). Following the clarification process,[31] the company provided results using the method employed in TA375 rather than the original method.

Treatment duration

Patients who fail to achieve moderate or good EULAR response at 6 months discontinue the current treatment and start the next treatment in the sequence. In contrast, patients who achieve moderate or good EULAR response stay on treatment until loss of efficacy. The company estimate time to treatment discontinuation from RHUMADATA and used the method employed by the AG in TA375 in a sensitivity analysis. The CS states that they have used RHUMADATA because "*it takes into account differences in retention among different classes of therapy*" and can therefore allow a time to discontinuation to be estimated for different types of bDMARDs (TNFi; IL-6; and other modes of action). These data (post amendments for typographical errors) are reproduced in Figure 10 along with the curve fits.

[redacted]

In the clarification process, the ERG asked the company to use a generalised gamma distribution. The company presented analyses in their clarification response[31] using the generalised gamma where the key conclusions did not change. However, the company's approach does not model discontinuation conditional on EULAR response, which is captured in the AG method used in TA375. The ERG is satisfied that the company have presented results also using the approach proposed by the AG in TA375 in their sensitivity analyses.

[redacted] **10**

5.2.7 Mortality

The company applied the mortality ratios per HAQ score at baseline used in TA375[25] to the life tables from the Office for National Statistics (ONS).[162] The company adopted the assumption that only baseline HAQ score, and not changes to the HAQ, affected mortality, as was the case in the AG's model in TA375.[25] This implies that the life expectancy of patients is independent of the treatment option. The CS states that this "*is considered to be a conservative approach because it does not acknowledge mortality benefits for improvements in disease severity*."

5.2.8 Health-related quality of life

The literature review detailed in Section 5.1 was used to inform health-related quality of life (HRQoL) values for patients with RA. The CS reports that the studies identified were contained in Appendix 16. The ERG, however, comments that key papers or reports appear not to have been identified, such as Hernandez *et al.*, which estimates EQ-5D based on patient characteristics (HAQ score, pain on a visual analogue scale, age and sex),[158] and Stevenson *et al*.,[24] which reviewed and critiqued the mapping of HAQ to utility undertaken in the companies' submissions for TA375. The CS comments that "*During early development of the model, the method used in TA375 was considered. However, in the Advisory Board, expert clinical opinion noted that it may double count the effects of pain since the HAQ-DI assessment already includes pain*" which resulted in the mapping of Malottki et al.[157] being used in the base case of the company submission. The ERG does not agree with the views of the company's expert clinical advisors, but notes that the company did use the method proposed by the AG in TA375, and accepted by the appraisal committee in sensitivity analyses, alongside a mapping reported by Bansback *et al*.[159] Following the clarification process, the company have used the mapping of Hernandez *et al.*[158]

In addition to HAQ-related utility, the company considered the impacts of serious infections on HRQoL. The rates of serious infections for SAR and BSC were taken from the pivotal studies: MOBILITY-B[33] for cDMARD-IR patients who could receive MTX (4.0% and 2.3% per cycle respectively); MONARCH[35] for cDMARD-IR patients who could not receive MTX (1.1% and 2.3% per cycle respectively); and TARGET[34] for the remaining patients (1.1% and 1.1% per cycle respectively). The company assumed that the rates for SAR were applicable to other bDMARDs. Within sensitivity analyses, the company employed the rates used in the AG model in TA375 (3.5% per cycle for bDMARDs and 2.6% for cDMARDs in cDMARD-IR patients). TA375 did not consider TNFi-IR patients and the company have assumed this population to have a rate of zero serious infections. The ERG believes that the use of rates equivalent to those for cDMARD-IR patients would have reflected a more reasonable assumption. Further the ERG notes that the method used by the AG in TA375 assumed only one serious infection per intervention and used the difference in incidences reported by Singh *et al.*[163] (35 per 1000 patient years on bDMARDs and 26 per 1000 patient years on cDMARDs) to calculate

the relative effect of bDMARDs. As the level of serious infections was shown not to be a key driver of the ICER in TA375, the impact related to the limitations of the company's approach will be minimal.

QALYs losses due to serious infections were stated to have been estimated based on the method used in the AG model for TA375 whereby serious infections were assumed to be of 28 days' duration and incur a disutility of 0.156, both taken from Oppong *et al*.[164] The company have translated this into a QALY loss of 0.024 per cycle.

5.2.9 Resources and costs

The company used the literature review previously described to identify economic evaluations with that deemed most appropriate selected. The company's model includes costs associated with drug acquisition, drug administration and monitoring, hospitalisation and serious infections. A detailed estimate of the price of each intervention is provided in Table 53.

There is a PAS for CTZ that provides the first 12 weeks of treatment free of charge; this was incorporated into the first year's acquisition costs. The PAS for GOL, where 100mg is provided at the same price of 50mg was also incorporated. The confidential PAS for ABT and TCZ were not excluded, as recommended by NICE, but were assumed to be equal to 15%. The ERG comments that this is not appropriate and such exploratory analyses should not be included in the base case. The ERG notes that biosimilars are available for both RTX and ADA and that these have not been included in the company's analyses.

Table 53: Drug acquisition costs

Drug	Package	Cost	Indicative annual cost (1st year)
SAR	200mg syringe x 1	[redacted]	[redacted]
ABT	125mg syringe x 4	£1,209.40[a]	£15,776
	250mg vial x 1	£302.40[a]	£11,834 (£12,741) ‡
GOL	50mg syringe x 1	£762.97[a]	£9,156*
	100mg syringe x 1	£1,525.94[a]*	
ETN	50mg syringe x 4	£715.00[a]	£9,327
	25mg syringe x 4	£357.50[a]	
ETNb	50mg syringe x 4	£656.00[a]	£8,557
ADA	40mg syringe x 2	£704.28[a]	£9,187
RTX	500mg vial x 1	£873.15[a]	£4,657
	100mg vial x 2	£349.25[a]	
CTZ	200mg syringe x 2	£715.00[a]**	£9,327 (£6,824)
TCZ IV	80mg vial x 1	£102.40[a]	£10,018 ‡

	200mg vial x 1	£256.00[a]	
	400mg vial x 1	£512.00 [a]	
TCZ SC	162mg x 4	£913.12[a]	£11,911
IFX	100mg vial x 1	£419.62[a]	£8,211 (£9,784) ‡
IFXb[165]	100mg vial x 1	£377.66[a]	£7,390 (£8,806) ‡
MTX[†]	2.5mg tablet x 28	£1.79 [a]	£42
	10mg tablet x 100	£37.89 [a]	

[¥]Sanofi Genzyme PASLU Application, [a]http://www.mims.co.uk/, , *PAS makes 100mg dose available at same price as 50mg dose applied in all analysis, **12 weeks free PAS applied in all analysis, [†]Adjuvant therapy added to all bDMARDs in combination analyses. ‡based on a weight of 74.3 Kg
ABT: abatacept; TCZ: tocilizumab; ADA: adalimumab; CTZ: certolizumab pegol; ETN: etanercept; ETNb: etanercept biosimilar; GOL: golimumab; IFX: infliximab; IFXb: infliximab biosimilar; RTX: rituximab; MTX: methotrexate;

The cost of the TNFi bundle was calculated using a weighted average of the individual agents informed by market share provided in Table 41. The retreatment interval for RTX was assumed to be 9 months and the cost of a RTX biosimilar has not been incorporated. BSC was costed at £360 per 6 months, and for the PSA the company put uncertainty on this value with an assumption that the standard error was 20% of the mean value. The company's base case assumes vial wastage with a sensitivity analysis exploring the impact of vial sharing.

Administration costs were based on TA375[166] and were inflated to 2015/16 prices using the hospital & community health services (HCHS) index in the PSSRU report.{Personal Social Services Research Unit, 2015 #31} This resulted in estimated costs of infusion of £170 and costs of a nurse visit (required by 10% of patients receiving a SC injection) of £77. The company states that this may be an overestimation "*since Sanofi Genzyme provides and funds a homecare service for sarilumab patients at no cost to the NHS. This is thought to be similar to comparator product manufacturers with SC bDMARDs therefore minimal impact is expected on the results.*" The ERG notes that in TA375, the time required by a district nurse was 30 minutes rather than the hour assumed by the company, although the ERG agrees that this limitation will have no impact on the conclusions.

Monitoring costs were also based on TA375[166] and included full blood count, erythrocyte sedimentation rate (ESR), biochemical profile, and chest x-ray prior to treatment with the addition of lipid profiles for TCZ and SAR. Full blood counts, biochemical profile and lipid profiles for TCZ and SAR were assumed to occur ten times in the first six months, and monthly thereafter. After the initial six months, monthly monitoring costs were assumed to be low: £7 for SAR and TCZ and £5 for all other bDMARDs although all interventions were associated with a monthly outpatient attendance assumed to cost £143 per visit, based on NHS Reference Costs.[167]

Hospitalisation costs were based on those within the AG's model in TA375,[166] inflated to 2015/2016 prices. In these estimates, hospitalisation costs were dependent on HAQ score band and were calculated

based on data from the NOAR database on inpatient days, joint replacements and NHS Reference Costs. The costs used in the model are provided in Table 54.

Table 54: Annual hospitalisation costs used in the company's model

HAQ-DI score	Annual costs
(0 - 0.5]	£180
(0.5 - 1.0]	£110
(1.0 - 1.5]	£391
(1.5 - 2.0]	£562
(2.0 - 2.5]	£1,338
(2.5 - 3.0]	£2,885

The cost per serious infection was assumed to be that used in the AG model for TA375[166] (£1479); this was uplifted to 2014/15 prices resulting in a cost of £1588 per episode.

5.2.10 Methods of the analysis

The company undertook analyses on the following groups:

- cDMARD-IR patients with severe RA who can tolerate MTX (CS denoted A1);
- cDMARD-IR patients with severe RA who cannot tolerate MTX (CS denoted B);
- TNFi-IR patients with severe RA who can tolerate RTX and MTX (CS denoted C2);
- TNFi-IR patients with severe RA who cannot tolerate RTX (CS denoted C1);
- TNFi-IR patients with severe RA who cannot tolerate MTX (CS denoted C3);
- TNFi-IR patients who have received RTX and MTX (CS denoted C4); and
- cDMARD-IR patients with moderate RA and DAS28 between 4.0 and 5.1 who can tolerate MTX (CS denoted A2).

The company used baseline characteristics of patients from the SAR trials for the patients simulated in the model. Instead of sampling with replacement from the patient pool, the model simulated each patient once in what the company called a replication. The deterministic results in the base case were produced by running enough replications to exceed 5,000 simulations. The company ran the model using a wide range of patient numbers and concluded that 5,000 patients provided the best trade-off between stability of the results and computation time. For sensitivity analyses, the number of replications for each patient was set so that the number of simulations was approximately 1,000. The company stated that this number of simulation provides a level of stability of results that ensures that the effect of changes in model parameters can be properly examined. Graphs and standard errors were presented to support the company's conclusion. The model generated a pool of random numbers that were used across sequences to alleviate differences stemming from Monte Carlo sampling error.

The company presented results of the probabilistic sensitivity analyses (PSA) in the CS for cDMARD-IR patients with severe RA who could tolerate MTX, for cDMARD-IR patients with severe RA for whom MTX was contraindicated or not tolerated, and for TNFi-IR RTX-ineligible patients. In the clarification response,[31] the company only presented results for the PSA for the cDMARD-IR patients with severe RA who could tolerate MTX. The company determined the number of PSA simulations required to obtain stable results analysing the convergence of mean incremental net benefit (INB) of SAR versus TCZ using the population from the TARGET trial. The company concluded that convergence occurs after approximately 200 runs and used 300 simulations in the PSA. Originally, independent draws from distributions were used for the probabilities of ACR response. However, at the ERG's request in the clarification letter,[31] draws from the joint posterior distribution (i.e. CODA) of the NMA were used instead.

5.2.11 Cost effectiveness results

The company presented results for their analyses in the CS, in which SAR+MTX was estimated to either dominate its comparators, result in ICERs lower than £20,000 per QALY gained or in cost savings per QALY lost higher than £60,000 in all populations except in cDMARD-IR patients with moderate RA and a DAS28 score higher than 4.0 and the TNFi-IR patients for whom RTX was an option. The ICER for SAR+MTX compared with BSC was estimated to be £22,275 per QALY gained in cDMARD-IR patients with moderate RA and a DAS28 score higher than 4.0. In TNFi-IR patients SAR+MTX compared with RTX+MTX results in an ICER of £104,012 per QALY gained.

However, the ERG identified as part of its initial assessment a series of issues in the company's analyses, described in Section 5.3, and asked the company in the clarification letter to provide analyses which addressed these problems. The ERG believes that these analyses better reflect the revised company's base case even if the company might disagree with some of the assumptions preferred by the ERG, such as the choice of the survival curve for time to treatment discontinuation. Therefore, the results presented below are the ones included in the company's clarification response.[31]

5.2.11.1 cDMARD-IR patients with severe RA who can tolerate MTX

Table 55 and Table 56 present the results for cDMARD-IR patients with severe RA who can tolerate MTX using the deterministic and probabilistic versions of the model respectively. SAR+MTX dominated both indications of TCZ both in the probabilistic and deterministic analyses. The ICERs of ABT (SC)+MTX and TNFi bundle + MTX compared with SAR+MTX were higher than £69,199 per QALY gained in both analyses.

Table 55: Results for cDMARD-IR patients with severe RA who can tolerate MTX (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)	ICER vs SAR (per QALY)
TCZ (SC) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
TCZ (IV) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
SAR + MTX	[redacted]	[redacted]	-	-	-	-
TNFi bundle + MTX	[redacted]	[redacted]	[redacted]	[redacted]	£79,199	£79,199
ABT (SC) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	£206,188	£126,110†

ABT: abatacept; MTX: methotrexate; TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 42
#Does not include confidential PAS
†Approximate ICER calculated by the ERG based on incrementals

The results of the PSA (see Table 56) were very similar to those of the deterministic analysis and the ranking of the treatments by effectiveness remained the same. Figure 1 of the company's clarification response[31] showed that the probability of SAR + MTX being cost-effective at willingness-to-pay (WTP) thresholds of £20,000 and £30,000 per QALY was close to 1.0.

Table 56: Results for cDMARD-IR patients with severe RA who can tolerate MTX (probabilistic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)	ICER vs SAR (per QALY)
TCZ (SC) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
TCZ (IV) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
SAR + MTX	[redacted]	[redacted]	-	-	-	-
TNFi bundle + MTX	[redacted]	[redacted]	[redacted]	[redacted]	£69,884	£69,884
ABT (SC) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	£203,809	£117,482†

ABT: abatacept; MTX: methotrexate; TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 42
#Does not include confidential PAS
†Approximate ICER calculated by the ERG based on incrementals

5.2.11.2 cDMARD-IR patients with severe RA for whom MTX is contraindicated or not tolerated

In cDMARD-IR patients with severe RA for whom MTX is contraindicated or not tolerated, SAR monotherapy resulted in an ICER of £17,123 per QALY gained compared with the TNFi bundle and

the ICERs for both indications of TCZ monotherapies compared with SAR monotherapy were higher than £1,000,000 per QALY gained (see Table 57).

Table 57: Results for cDMARD-IR patients with severe RA for whom MTX is contraindicated or not tolerated (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)	ICER vs SAR (per QALY)
TNFi bundle	[redacted]	[redacted]	[redacted]	[redacted]	-	£17,123‡
SAR	[redacted]	[redacted]	[redacted]	[redacted]	£17,123	-
TCZ (SC) [#]	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£2,596,000†
TCZ (IV) [#]	[redacted]	[redacted]	[redacted]	[redacted]	£1,578,976	£1,578,976

TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 43
[#]Does not include confidential PAS
†Approximate ICER calculated by the ERG based on incrementals
‡ICER in the south western quadrant representing cost savings per QALY lost

5.2.11.3 TNFi-IR patients with severe RA who can tolerate RTX and MTX

In TNFi-IR patients with severe RA who can tolerate RTX and MTX, the ICER for a sequence of SAR+MTX followed by TCZ+MTX (SAR,TCZ) compared with the currently recommended sequence (RTX,TCZ) was estimated to be £130,691 per QALY gained (see Table 58). A sequence including only SAR+MTX as biologic therapy was extendedly dominated by a sequence having only RTX + MTX and by the currently recommended sequence (RTX,TCZ).

Table 58: Results for TNFi-IR patients with severe RA who can tolerate RTX and MTX (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)
RTX	[redacted]	[redacted]	-	-	-
SAR	[redacted]	[redacted]	[redacted]	[redacted]	Extendedly dominated
RTX,TCZ‡[#]	[redacted]	[redacted]	[redacted]	[redacted]	£39,994
SAR,TCZ[#]	[redacted]	[redacted]	[redacted]	[redacted]	£130,691

RTX: rituximab; SAR: sarilumab;
*Sequences as defined in Table 44
[#]Does not include confidential PAS for TCZ
†Approximate ICER calculated by the ERG based on incrementals
‡Currently recommended sequence

5.2.11.4 TNFi-IR patients with severe RA for whom RTX is not an option

As shown in Table 59, in TNFi-IR patients with severe RA for whom RTX is not an option, the ICERs for all comparators versus SAR+MTX were higher than £60,000 per QALY gained.

Table 59: Results for TNFi-IR patients with severe RA for whom RTX is not an option (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)	ICER vs SAR (per QALY)
SAR + MTX	[redacted]	[redacted]	-	-	-	
TCZ (IV) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	Extendedly dominated	£141,995†
TNFi Bundle + MTX	[redacted]	[redacted]	[redacted]	[redacted]	£64,602	£64,602
ABT (SC) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£80,889†
TCZ (SC) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	£69,306	£69,306

ABT: abatacept; MTX: methotrexate; TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 45
#Does not include confidential PAS
†Approximate ICER calculated by the ERG based on incrementals

5.2.11.5 TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated

In TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated, SAR monotherapy compared with the TNFi bundle was estimated to result in an ICER of £17,794 per QALY gained (see Table 60).

Table 60: Results for TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)
TNFi Bundle	[redacted]	[redacted]	-	-	-
SAR	[redacted]	[redacted]	[redacted]	[redacted]	£17,794

TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab;
*Sequences as defined in Table 46
†Approximate ICER calculated by the ERG based on incrementals

5.2.11.6 TNFi-IR patients who have received RTX + MTX

As shown in Table 61, the ICER for TCZ (IV) + MTX and TCZ (SC) + MTX compared with SAR + MTX was estimated to be £141,995 and £133,548 respectively in TNFi-IR patients after receiving RTX + MTX.

Table 61: Results for TNFi-IR patients who have received RTX + MTX (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)	ICER vs SAR +MTX (per QALY)
SAR + MTX	[redacted]	[redacted]	[redacted]	[redacted]	-	
TCZ (IV) + MTX	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£141,995†
TCZ (SC) + MTX	[redacted]	[redacted]	[redacted]	[redacted]	£133,548	£133,548

ABT: abatacept; MTX: methotrexate; TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 47
#Does not include confidential PAS
†Approximate ICER calculated by the ERG based on incrementals

5.2.11.7 cDMARD-IR patients with moderate RA (DAS28 between 4.0 and 5.1) who can tolerate MTX

In cDMARD-IR patients with moderate RA (DAS28 between 4.0 and 5.1) who can tolerate MTX, the ICER for SAR+MTX compared with MTX was estimated to be £38,254 per QALY gained (see Table 62).

Table 62: Results for cDMARD-IR patients with moderate RA (DAS28 between 4.0 and 5.1) who can tolerate MTX (deterministic)

Sequences*#	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (£/QALY)
MTX	[redacted]	[redacted]	[redacted]	[redacted]	-
SAR + MTX	[redacted]	[redacted]	[redacted]	[redacted]	£38,254

ABT: abatacept; MTX: methotrexate; TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 48
#Does not include confidential PAS of TCZ

5.2.12 Model validation and face validity check

The ERG adopted a number of approaches to explore, interrogate and critically appraise the company's submitted economic evaluation and the underlying health economic model upon which this was based. These approaches included:

- Consideration of key items contained within published economic evaluation and health economic modelling checklists to critically appraise the company's model and analysis.[157, 168, 169]
- Scrutiny of the company's model by health economic modellers including:
    - White-box validation: checking of inputs, code and formulae
    - Black-box testing: changing inputs to check whether the output matches expectations
    - Face-validity testing: checking model results match expectations
    - Comparison of deterministic and probabilistic ICERs.
- Replication of the base case results, PSA and scenario analysis presented within the CS.[170]
- Where possible, checking parameter values used in the company's model against the original data sources.
- Examination of concordance between the description of the model reported within the CS[170] and the company's executable model.
- The use of expert clinical input to judge the clinical robustness of the company's economic evaluation and of the assumptions underpinning the model.

5.3 Summary of key limitations identified within the critical appraisal

The main limitations identified within the ERG's initial critical appraisal of the company's economic analysis that were corrected by the company are the following:

1.	Inadequate treatment sequences
2.	Omission of the possibility of patients with moderate RA to progress to the severe state
3.	Use of Malottki *et al.*[157] instead of Hernandez *et al.*[158] for the mapping of HAQ to EQ-5D
4.	Limitations in the company's NMA
5.	Using percentages of improvement of HAQ instead of absolute mean changes
6.	Omission of rounding to the nearest valid HAQ score
7.	Use of an inappropriate time to treatment discontinuation
8.	Using independent samples from beta distributions for the probabilities of ACR responses in the PSA instead of correlated samples from the CODA of the NMA
9.	Using 9 free doses of CTZ instead of 10.
10.	Inclusion of a speculative PAS of 15% applied to TCZ and ABT.

Based on the new analyses presented by the company following the clarification round, the key remaining limitations are as follows:

1.	Linear progression of HAQ score for patients on cDMARDs and BSC
2.	Incorrect implementation of transition from moderate to severe RA

3.	Assuming same efficacy for monotherapies as for combination therapies in TNFi-IR patients
4.	Assuming same efficacy for second and third lines of bDMARDs

The issues identified by the ERG and corrected by the company in the revised model presented along the clarification responses are further explained below:

1. Inadequate treatment sequences

The sequences used by the company suffer mainly from two issues:

- The omission of one cDMARD treatment (MTX or SSZ) after biologics and before BSC.
- The inclusion of ABT+MTX after RTX+MTX in cDMARD-IR patients or after SAR+MTX in TNFi-IR patients who are RTX-eligible.

After clarification, the company provided analyses using the sequences requested by the ERG, which addressed these limitations.

2. Omission of the possibility of patients with moderate RA to progress to the severe state

The company's model assumed that patients with moderate RA and a DAS28 score higher than 4.0 would never progress to severe RA. The ERG acknowledges that the independent analysis by the AG in TA375[24] also omitted this possibility. However, the ERG believes that including the possible transition of these patients to the severe RA state and subsequently to the recommended treatment sequences for severe patients provides a more accurate representation of clinical practice. In their clarification response,[31] the company presented analyses where patients with moderate RA could progress to the severe state.

3. Use of Malottki et al.[157]*. instead of Hernandez et al.*[158] *for the mapping of HAQ to EQ-5D*

For their base case analysis, the company adopted the approach taken by Malottki *et al.*[157] to mapping HAQ scores to EQ-5D and used the approach proposed by Hernandez *et al.*[158] in a scenario analysis. The company justified their choice referring to expert clinical opinion obtained during an advisory board, which noted that Hernandez *et al.*'s approach may double count the effects of pain since the HAQ-DI assessment already includes pain. The ERG disagrees with this view and notes that double counting is avoided by taking HAQ-DI and pain jointly into account. Most analyses previous to Hernandez *et al.*[158] had excluded pain. However, a substantially better estimate of EQ-5D is obtained by the inclusion of pain alongside HAQ than via HAQ alone, because HAQ and pain are not perfectly correlated.[158] It is therefore important to include pain as an explanatory variable in estimating EQ-5D.

4. Limitations with the company's NMA

The ERG identified a number of limitations in the NMA presented in the CS, such as the use of a fixed effect model, which have been described in Section 4.4. As requested by the ERG in its clarification

letter, the company undertook an NMA addressing some of these limitations and presented economic analyses using its results.

5. Using percentages of improvement of HAQ instead of absolute mean changes

In the CS, the company applied improvements in the HAQ score upon response in terms of percentage of improvement instead of applying a fixed reduction. The ERG notes that this approach differs from that accepted by the AC in TA375 and that percentage improvement is prone to vary depending on the patient mix. The company adopted absolute HAQ score improvements upon response in the analyses presented in the clarification response.[31]

6. Omission of rounding to the nearest valid HAQ score

HAQ scores range from 0 to 3, with higher scores indicating greater disability. HAQ scores lie on a discrete scale with step values of 0.125, resulting in 25 points. In the model, patients start with a baseline HAQ score and the HAQ progression of patients is modified reflecting treatment response, loss of treatment efficacy or disease progression over time. Changes applied to the HAQ score are usually estimates based on average changes observed in trials or registries and therefore are rarely exact multiples of 0.125. Thus, after applying such a change, the resulting HAQ score of a patient has to be assigned to a valid HAQ score. However, the company did not round the values to the nearest valid HAQ score. The ERG requested that the company implement a stochastic rounding of HAQ scores analogous to that used by the AG in TA375[25] i.e. rounding up values a probability inversely proportional to the distance of the value to the closest valid HAQ score, and rounding down otherwise. For example, a change of 0.4 would have a 0.8 probability of being rounded down to 0.375 and a probability of 0.2 of being rounded up to 0.5. The company correctly implemented this stochastic HAQ rounding after clarification.

7. Use of an inappropriate time to treatment discontinuation

Patients who achieve moderate or good EULAR response stay on treatment until loss of efficacy. The company estimated time to treatment discontinuation by fitting different survival curves to time to treatment discontinuation data from RHUMADATA. The company chose the Gompertz distribution for their base case because it provided a good statistical fit (both Akaike Information Criterion [AIC] and Bayesian Information Criterion [BIC]) and a good visual fit. The ERG notes that the generalised gamma resulted in better AIC and BIC scores than the Gompertz in most of the curves. Following a request for clarification from the ERG, the company justified their preference for the Gompertz curve stating that towards the tail of the Kaplan-Meier curve, the generalised gamma under-predicted the proportion of patients still on treatment for the IL-6 class and the class grouping other mechanisms of action whilst the Gompertz provided a good visual fit for all treatment classes (see clarification response [31] -- question B4). The ERG notes that the estimates at the tail of the Kaplan-Meier are most uncertain

[redacted]

Therefore, the ERG believes that the generalised gamma provides a more plausible extrapolation for time to treatment discontinuation. The company presented analyses using the generalised gamma curve for time to treatment discontinuation in their clarification responses.

8. Using independent samples from beta distributions for the probabilities of ACR responses in the PSA instead of correlated samples from the CODA of the NMA

Within the PSA, the company used independent samples from beta distributions to model the uncertainty around the ACR response rates. The ERG notes that this approach ignores the existing correlations. The ERG asked the company to provide analyses where samples from the CODA of the NMA were used instead, which the company did in their revised version of the model.

9. Using 9 free doses of CTZ instead of 10.

The company assumed that the PAS for CTZ comprised 9 free doses, instead of the 10 free doses established in the NICE guidance produced in TA375 and TA415. The company also varied the number of free doses for CTZ in the PSA, as if it were an uncertain value. The company adopted 10 free doses of CTZ and fixed the value in the PSA in the analyses presented in the clarification response.

10. Inclusion of a speculative PAS of 15% applied to TCZ and ABT.

The ERG notes that including a speculative PAS discount for TCZ and ABT is misleading and that NICE recommends the use of the list price in such cases. The ERG provides a confidential appendix which presents a set of analyses where the confidential PAS for TCZ and ABT have been included. The company removed the speculative PAS of 15% from the analyses presented in the clarification response.[31]

The company implemented these changes in their model and produced the results summarised in Section 5.2.11. The issues remaining in the revised version of the model and therefore in the analyses presented by the company in their clarification response[31] are further explained below:

1. Linear progression of HAQ score for patients on cDMARDs and BSC

The company applied a linear annual increase of 0.06 in HAQ score to BSC in the analyses presented in the CS. The ERG believes that these analyses are inappropriate as HAQ progression has been proven to be non-linear[30] with the Appraisal Committee in TA375 favouring the non-linear approach advocated by the AG.[25] The ERG comments that the linear method is not likely to significantly affect the conclusions in the comparison of SAR with bDMARDs, due to similar efficacy levels, but could have a significant favourable effect for SAR when the comparator is cDMARDs. The company

acknowledged this issue (see clarification response[31] -- question B2) but was unable to implement the latent class approach in the revised model due to time constraints, and instead used the linear HAQ increment also for patients on cDMARDs.

2. Incorrect implementation of transition from moderate to severe RA

The ERG requested from the company to implement the possibility for moderate patients to progress to the severe state and consequently to transition to the treatment sequences recommended for patients with severe RA. In their implementation, the company assumed patients would go through the moderate sequence and only once they would start on best supportive care they would transition to the sequences recommended for patients with severe RA, only if their HAQ was above a certain threshold that was calculated through a regression as being related to a DAS28 score of 5.1. The ERG believes there are two main issues with the company's implementation. First, the relationship between changes in HAQ and DAS28 scores should have been calculated instead of the absolute scores. The relationship between these scores is far from being linear and by applying it to the changes in these scores instead of the absolute values, the error in the extrapolation is minimised. The company acknowledged that their regression resulted in a DAS28 score of 5.1 being predictive of an implausibly low HAQ score of 0.375. Second, patients should progress to the severe sequences at the point when their DAS28 score increases above 5.1, without waiting until they have reached the ebd of the moderate sequence.

3. Assuming same efficacy for monotherapies as for combination therapies in TNFi-IR patients

In light of the absence of evidence of the efficacy of monotherapies in TNFi-IR patients, the company assumed that the effectiveness of SAR monotherapy and its comparators would be the same as for the respective combination therapies. The company did not identify any RCTs that reported the efficacy of bDMARDs in this population and assumed that the effectiveness of monotherapies would be equal to that of combination therapies in TNFi-IR patients. The ERG notes that even though such an assumption is reasonable in light of the lack of evidence, the true effectiveness of bDMARD monotherapies is in TNFi-IR patients is still uncertain.

4. Assuming same efficacy for different lines of bDMARDs

The company assumed that the effectiveness of interventions in terms of ACR response rates and HAQ score improvements upon response would be the same, whether it was first-line or in subsequent therapy lines. In practice, as can be seen by comparing ACR rates in cDMARD-IR and TNFi-IR patients, the efficacy of treatment is reduced for subsequent treatment lines. This issue is mostly cancelled out when comparing sequences of equal length but might produce inaccurate results when comparing sequences of different lengths.

5.4 Additional exploratory analyses undertaken by the ERG

The ERG undertook exploratory analyses based on the company's revised model after applying the following two changes:

- The implementation of the non-linear HAQ progression based on the latent classes' approach described by Norton *et al.*[161] as implemented in the model developed by the AG in TA375.
- The amendment of the mechanism by which patients with moderate RA transition to the severe state and consequently to the treatment sequence recommended for patients with severe RA:
    - Calculating the DAS28 score of the patient at each cycle based on their DAS28 score at baseline, the change in HAQ score from baseline and the coefficient for HAQ score calculated by the company in their regression and used in their amended model.
    - Assuming that patients would transition to the sequence recommended for patients the moment their estimated DAS28 score is above 5.1.

Due to the time constraints and the running times of the company's model, the ERG only presents results of deterministic analyses. The ERG believes the probabilistic results would be very similar to the deterministic ones based on the similarity of the deterministic and probabilistic results presented for the cDMARD-IR population with severe RA presented in Table 55 and Table 56 respectively.

5.4.1 cDMARD-IR patients with severe RA who can tolerate MTX

In cDMARD-IR patients with severe RA who can tolerate MTX, SAR + MTX was estimated to dominate both indications of TCZ with concomitant MTX and the ICERs for TNFi bundle + MTX and ABT (SC) + MTX compared with SAR + MTX were estimated to be in excess of £150,000 per QALY gained – see Table 63.

Table 63: Results for cDMARD-IR patients with severe RA who can tolerate MTX (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)	ICER vs SAR (per QALY)
TCZ (SC) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
TCZ (IV) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
SAR + MTX	[redacted]	[redacted]	[redacted]	[redacted]	-	
TNFi bundle + MTX	[redacted]	[redacted]	[redacted]	[redacted]	£151,563	£151,563
ABT (SC) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	£311,453	£214,071

ABT: abatacept; MTX: methotrexate; TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 42
#Does not include confidential PAS

5.4.2 cDMARD-IR patients with severe RA for whom MTX is contraindicated or not tolerated

As shown in Table 64, the ICER for SAR monotherapy compared with TNFi bundle monotherapy in cDMARD-IR patients with severe RA for whom MTX was contraindicated or not tolerated was estimated to be £34,422 per QALY gained, whilst the ICERs for both indications of TCZ compared with SAR monotherapy were estimated to be in excess of £1,500,000 per QALY gained. The ERG notes that the effectiveness of TCZ SC and TCZ IV is assumed to be the same and therefore the differences in the estimated total QALYs are the result of Monte Carlo sampling error.

Table 64: Results for cDMARD-IR patients with severe RA for whom MTX is contraindicated or not tolerated (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)	ICER vs SAR (per QALY)
TNFi bundle	[redacted]	[redacted]	[redacted]	[redacted]	-	£34,422‡
SAR	[redacted]	[redacted]	[redacted]	[redacted]	£34,422	-
TCZ (SC) #	[redacted]	[redacted]	[redacted]	[redacted]	Extendedly dominated	£2,541,618
TCZ (IV) #	[redacted]	[redacted]	[redacted]	[redacted]	£1,676,280	£1,676,280

TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 43
#Does not include confidential PAS
‡ICER in the south western quadrant representing cost savings per QALY lost

5.4.3 TNFi-IR patients with severe RA who can tolerate RTX and MTX

Table 65 shows the results of the analysis in patients with severe RA who can tolerate RTX and MTX: a sequence where SAR+MTX replaced RTX+MTX was estimated to result in an ICER of £171,466 per QALY gained.

Table 65: Results for TNFi-IR patients with severe RA who can tolerate RTX and MTX (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)
RTX	[redacted]	[redacted]	[redacted]	[redacted]	
SAR	[redacted]	[redacted]	[redacted]	[redacted]	Extendedly dominated
RTX,TCZ‡#	[redacted]	[redacted]	[redacted]	[redacted]	£69,947
SAR,TCZ#	[redacted]	[redacted]	[redacted]	[redacted]	£171,466

RTX: rituximab; SAR: sarilumab;
*Sequences as defined in Table 44
#Does not include confidential PAS for TCZ
‡Currently recommended sequence

5.4.4 TNFi-IR patients with severe RA for whom RTX is not an option

In TNFi-IR patients with severe RA for whom RTX is not an option, SAR + MTX was estimated to result in an ICER of £34,979 per QALY gained compared with TNFi bundle whilst the ICER for both TCZ indications with concomitant MTX compared with SAR + MTX was estimated to be in excess of £195,000 – see Table 66. The ERG notes that the sequences evaluated in the company's analyses and the ERG's exploratory analyses differ as explained in Table 45.

Table 66: Results for TNFi-IR patients with severe RA for whom RTX is not an option (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)	ICER vs SAR (per QALY)
TNFi bundle+ MTX	[redacted]	[redacted]	[redacted]	[redacted]	-	£34,979
ABT (SC) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	Dominated
SAR + MTX	[redacted]	[redacted]	[redacted]	[redacted]	£34,979	-
TCZ (IV) + MTX#	[redacted]	[redacted]	[redacted]	[redacted]	£198,863	£198,863
TCZ (SC)+MTX#	[redacted]	[redacted]	[redacted]	[redacted]	£777,770	£205,638

ABT: abatacept; MTX: methotrexate; TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 45
#Does not include confidential PAS

5.4.5 TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated

In TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated, the ICER for SAR monotherapy compared with TNFi bundle was estimated to be £31,433 per QALY gained – see Table 67. The ERG notes that this analysis is subject to considerable uncertainty given that the effectiveness estimates for the monotherapies were assumed to be equal to those in combination with MTX due to lack of evidence. The ERG also notes that TCZ, the only IL-6 recommended by NICE for severe RA, is not recommended in this population.

Table 67: Results for TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)
TNFi bundle	[redacted]	[redacted]			
SAR	[redacted]	[redacted]	[redacted]	[redacted]	£31,433

TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab;
*Sequences as defined in Table 46

5.4.6 TNFi-IR patients who have received RTX + MTX

As shown in Table 68, in TNFi-IR patients who have already received RTX+MTX, the ICERs for both indications of TCZ with concomitant MTX compared with SAR+MTX were estimated to be in excess of £200,000 per QALY gained.

Table 68: Results for TNFi-IR patients who have received RTX + MTX (deterministic)

Sequences*	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (per QALY)	ICER vs SAR +MTX (per QALY)
SAR + MTX	[redacted]	[redacted]	[redacted]	[redacted]	-	-
TCZ (IV) + MTX	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£245,465
TCZ (SC) + MTX	[redacted]	[redacted]	[redacted]	[redacted]	£219,153	£219,153

ABT: abatacept; MTX: methotrexate; TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 47
#Does not include confidential PAS

5.4.7 cDMARD-IR patients with moderate RA (DAS28 between 4.0 and 5.1) who can tolerate MTX

In cDMARD-IR patients moderate RA and a DAS28 higher than 4.0, a sequence starting with SAR+MTX compared with MTX was estimated to result in an ICER of £63,438 per QALY gained – see Table 69.

Table 69: Results for cDMARD-IR patients with moderate RA (DAS28 between 4.0 and 5.1) who can tolerate MTX (deterministic)

Sequences*#	Total QALYs	Total costs	Incr. QALYs	Incr. costs	ICER (£/QALY)
MTX	[redacted]	[redacted]			
SAR + MTX	[redacted]	[redacted]	[redacted]	[redacted]	£63,438

ABT: abatacept; MTX: methotrexate; TCZ: tocilizumab; TNFi: tumour alpha necrosis inhibitor; SAR: sarilumab; IV: intravenous; SC: subcutaneous
*Sequences as defined in Table 48
#Does not include confidential PAS of TCZ

5.5 Discussion

The CS includes a systematic review of economic evaluations of treatments for moderate and severe RA together with a *de novo* model-based economic evaluation of SAR + MTX versus currently recommended treatments in adult moderate and severe RA, cDMARD-IR and TNFi-IR patients. The company's systematic review of existing economic evaluations did not identify any studies that estimated the cost effectiveness of SAR + MTX.

The company's *de novo* economic model was largely based on the model developed by the AG in TA375.[25] Costs and health outcomes for SAR + MTX and its comparators were estimated from the

perspective of the NHS and PSS over a lifetime horizon. The analyses presented in the CS relate to seven different populations of RA patients: (1) cDMARD-IR patients with severe RA who can tolerate MTX; (2) cDMARD-IR patients with severe RA for whom MTX is contraindicated or not tolerated; (3) TNFi-IR patients with severe RA for whom RTX is an option; (4) TNFi-IR patients with severe RA for whom RTX is not an option; (5) TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated; (6) TNFi-IR patients with severe RA after treatment with RTX+MTX; and, (7) a subgroup of cDMARD-IR patients with moderate RA whose DAS28 scores are between 4.0 and 5.1. The definition of severe RA was a DAS28 score higher than 5.1, whilst moderate RA was defined as a DAS28 > 3.2 and ≤ 5.1. Baseline characteristics of patients are based on the relevant clinical SAR trials.

The company presented analyses in the CS and in the clarification response as per the ERG's request. The ERG believes that the analyses presented by the company in the clarification responses are closer to the company's intended base case than those in the CS. According to the company's revised probabilistic analysis, in cDMARD-IR patients with severe RA who could tolerate MTX, SAR+MTX dominated both indications of TCZ with concomitant MTX and the ICERs for SAR+MTX versus the TNFi bundle and ABT(SC) + MTX were £69,884 and £117,482 per QALY gained respectively. In the cDMARD-IR population with severe RA who could not tolerate MTX, the estimated ICER for SAR monotherapy versus the TNFi bundle was £17,123 per QALY gained, whilst the ICER for both TCZ indications compared with SAR was higher than £1,000,000 per QALY gained. In TNFi-IR patients for whom RTX+MTX was an option, the ICER for SAR+MTX compared with RTX+MTX was estimated to be £130,691 per QALY gained. If RTX was not an option, the ICER for the considered comparators versus SAR+MTX in TNFi-IR patients was higher than £60,000 per QALY gained. For TNFi-IR patients who could not tolerate MTX, the ICER for SAR monotherapy compared with a TNFi bundle was estimated to be £17,794 per QALY gained. In patients who have received RTX+MTX, the ICER for both indications of TCZ compared with SAR+MTX were estimated to be greater than £130,000 per QALY gained. Finally, in cDMARD-IR patients with moderate RA and a DAS28 score higher than 4.0, the ICER for SAR+MTX was estimated to be £38,254 per QALY gained.

The ERG undertook exploratory analyses after amending the transition from moderate to severe RA and implementing the non-linear HAQ trajectory based on the latent class approach for patients on cDMARDs and BSC. According to the ERG's exploratory analyses, in cDMARD-IR patients with severe RA who can tolerate MTX, SAR + MTX was estimated to dominate both indications of TCZ with concomitant MTX and the ICERs for TNFi bundle + MTX and ABT (SC) + MTX compared with SAR + MTX are estimated to be in excess of £150,000 per QALY gained. In cDMARD-IR patients with severe RA for whom MTX was contraindicated or not tolerated, the ICER for SAR monotherapy compared with TNFi bundle monotherapy was estimated to be £34,422 per QALY gained, whilst the

ICERs for both indications of TCZ compared with SAR monotherapy where estimated to be in excess of £1,500,000 per QALY gained. In TNFi-IR patients with severe RA who can tolerate RTX and MTX the ICER for SAR+MTX compared with RTX+MTX was estimated to be £171,466 per QALY gained. In TNFi-IR patients with severe RA for whom RTX is not an option, SAR + MTX was estimated to result in an ICER of £34,979 per QALY gained compared with TNFi bundle whilst the ICER for both TCZ indications with concomitant MTX compared with SAR + MTX was estimated to be in excess of £195,000. In TNFi-IR patients with severe RA for whom MTX is contraindicated or not tolerated, the ICER for SAR monotherapy compared with TNFi bundle was estimated to be £31,433 per QALY gained. In TNFi-IR patients who have already received RTX+MTX, the ICERs for both indications of TCZ with concomitant MTX compared with SAR+MTX were estimated to be in excess of £200,000 per QALY gained. In cDMARD-IR patients moderate RA and a DAS28 higher than 4.0, a sequence starting with SAR+MTX compared with MTX was estimated to result in an ICER of £63,438 per QALY gained.

The ERG presents the results of the analyses using the company's model whilst incorporating the confidential PAS currently in place for TCZ and ABT in a confidential appendix.

There remain two potentially important areas of uncertainty:

- *The effectiveness of SAR monotherapy and its comparators in TNFi-IR patients.* The company did not identify any RCTs that reported the efficacy of bDMARDs in this population and assumed that the efficacy of monotherapies would be equal to that of combination therapies in TNFi-IR patients. The ERG notes that even though such an assumption is reasonable in light of the lack of evidence, the true effectiveness of bDMARD monotherapies is in TNFi-IR patients is still uncertain.
- *The effectiveness of SAR + MTX as third line biologic*. The company assumed that the same efficacy estimate for SAR + MTX in TNFi-IR patients would apply before and after RTX + MTX. However, only 23.2% of patients enrolled in the TARGET trial, from which the effectiveness of SAR + MTX was estimated, had more than one previous TNFi and the company did not provide a subgroup analysis of the efficacy of SAR + MTX in these patients. The ERG notes that considering that the efficacy of SAR + MTX is reduced in TNFi-IR patients compared with cDMARD-IR patients, it is reasonable to believe that its efficacy will be further reduced in subsequent treatment lines. However, the ERG acknowledges that this issue is unlikely to have an important impact in the cost-effectiveness of SAR + MTX because it also applies to the comparators and any potential reduction in the efficacy of SAR + MTX would probably be cancelled out by similar reductions in its comparators.

6 END OF LIFE

NICE end of life supplementary advice should be applied in the following circumstances and when both the criteria referred to below are satisfied:

- The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;
- There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment, and;

The company did not include any claim or justification in the CS for SAR to be considered as an end of life treatment. The ERG believe that neither criterion would be met as patients receiving treatment would be expected to have a life expectancy considerably longer than 24 months and there is little robust evidence to suggest that SAR would provide an additional 3 months of life compared with its comparators.

7 OVERALL CONCLUSIONS

The key clinical effectiveness evidence for SAR was based on five RCTs and one long-term extension study. Three RCTs had ACR20 as their primary endpoint (MOBILITY-A, MOBILITY-B and TARGET). In the MTX-IR population, the RCTs showed a significant advantage in ACR responses for licensed doses of SAR+MTX over PBO+MTX, and a significant advantage for SAR monotherapy over ADA monotherapy. In the TNFi-IR population, there was a significant advantage for SAR+cDMARD over PBO+cDMARD. The MOBILITY-A, MOBILITY-B, MONARCH and TARGET trials reported significantly favourable results for licensed doses of SAR over comparators for improvement in HAQ-DI. SAR had a significant advantage over comparator for DAS28-CRP in the MOBILITY-B and TARGET trials, and for DAS28-ESR in the MONARCH trial. MOBILITY-B measured radiographic progression by mTSS, and reported a significantly lower deterioration from baseline for SAR over PBO+MTX.

SAEs were [redacted]. The most common AEs [redacted] The ASCERTAIN trial reported a similar safety profile for SAR to that of TCZ.

The company presented results of analyses based on a *de novo* economic model that was largely based on the model developed by the AG in TA375. In their clarification response the company presented a new set of analyses after addressing a multitude of issues identified by the ERG. The ERG considers these to be closer to the company's intended base case than those presented in the CS. The ERG undertook exploratory analyses after addressing two remaining issues: the HAQ trajectories of patients on cDMARDs and BSC; and the timing of the transition of patients with moderate RA to severe RA.

In cDMARD-IR patients with severe RA who could tolerate MTX, according to the analyses presented by the company in their clarification response and the ERG's exploratory analyses, SAR + MTX dominates both indications of TCZ in combination with MTX and the estimated ICER of the other comparators versus SAR +MTX is in excess of £75,000 per QALY gained.

In cDMARD-IR patients with severe RA for whom MTX was contraindicated or not tolerated, the ICER of SAR monotherapy compared with the TNFi bundle monotherapy is estimated to be £17,123 and £34,422 per QALY gained based on the company's analyses and the ERG's analyses respectively. The difference between the ICERs can be explained by the comparatively lower long-term HAQ progression whilst on cDMARDs based on the non-linear HAQ progression. In both analyses, the ICER of TCZ monotherapies (SC and IV) is estimated to be in excess of £1,500,000 per QALY gained.

In TNFi-IR patients who can tolerate MTX and for whom RTX is an option, a sequence of SAR + MTX followed by TCZ + MTX results in ICERs ranging from £130,691 to £171,466 per QALY gained compared with a sequence of RTX + MTX followed by TCZ + MTX.

In TNFi-IR patients who can tolerate MTX but for whom RTX is not an option, according to the company's analyses the ICERs of all comparators versus SAR + MTX are in excess of £64,602 per QALY gained. Contrastingly, according to the ERG's analyses the ICER of SAR+MTX compared with TNFi bundle + MTX is £34,979 but SAR + MTX dominates ABT + MTX and the ICERs of both indications of TCZ in combination with MTX versus SAR + MTX are in excess of £195,000. The difference in the results is partly explained by differences in the sequences used by the ERG and the company: the company's include bDMARDs in the second line of the sequence that are not recommended by NICE.

In TNFi-IR patients for whom MTX is contraindicated or not tolerated, the ICER for SAR monotherapy compared with TNFi bundle monotherapy is estimated to be £17,794 per QALY gained according to the company's analyses and £31,433 according to the ERG's analyses. The ERG notes that the difference between the ICERs is likely to be due to the lower benefit estimated from bDMARDs compared with cDMARDs/BSC when assuming a non-linear HAQ progression.

In TNFi-IR patients who have had RTX + MTX, the ICERs of both indications of TCZ are estimated to be in excess of £130,000 per QALY gained according to both the company's and the ERG's analyses.

In cDMARD-IR patients with moderate RA, a sequence starting with SAR + MTX compared with the currently recommended sequence starting with MTX is estimated to result in an ICER of £38,254 per QALY gained according to the company's analyses and £63,438 per QALY gained according to the ERG's analyses.

The ERG notes that the confidential PASs in place for ABT and TCZ were not included in these analyses. The ERG presents the analyses including the confidential PASs in a confidential appendix.

8 REFERENCES

1. Agency EM. Summary of product characteristics - Xeljanz (tofacitinib); 2017.
2. Pfizer I. Rheumatoid arthritis - tofacitinib citrate – Company Evidence Submission. London: NICE; 2017.
3. National Institute for Health and Care Excellence. Tofacitinib for treating moderate to severe active rheumatoid arthritis after the failure of disease-modifying anti-rheumatic drugs - Final scope. 2017.
4. Scott DL, Steer S. The course of established rheumatoid arthritis. *Best Practice & Research Clinical Rheumatology* 2007;21:943-67.
5. Pincus T, Fuchs HA, Callahan LF, Nance EP, Kaye JJ. Early radiographic joint space narrowing and erosion and later malalignment in rheumatoid arthritis: a longitudinal analysis. *The Journal of Rheumatology* 1998;25:636-40.

6. Drossaers-Bakker KW, Kroon HM, Zwinderman AH, Breedveld FC, Hazes JMW. Radiographic damage of large joints in long-term rheumatoid arthritis and its relation to function. *Rheumatology* 2000;39:998-1003.
7. Allaire S, Wolfe F, Niu J, LaValley MP, Zhang B, Reisine S. Current risk factors for work disability associated with rheumatoid arthritis: Recent data from a US national cohort. *Arthritis Care & Research* 2009;61:321-8.
8. Naz SM, Symmons DPM. Mortality in established rheumatoid arthritis. *Best Practice & Research Clinical Rheumatology* 2007;21:871-83.
9. Dadoun S, Zeboulon-Ktorza N, Combescure C, Elhai M, Rozenberg S, Gossec L, *et al.* Mortality in rheumatoid arthritis over the last fifty years: Systematic review and meta-analysis. *Joint Bone Spine* 2013;80:29-33.
10. Meune C, Touzé E, Trinquart L, Allanore Y. Trends in cardiovascular mortality in patients with rheumatoid arthritis over 50 years: a systematic review and meta-analysis of cohort studies. *Rheumatology* 2009;48:1309-13.
11. National Collaborating Centre for Chronic Conditions. Rheumatoid arthritis: national clinical guideline for management and treatment in adults: Royal College of Physicians (UK); 2009.
12. Symmons D, Turner G, Webb R, Asten P, Barrett E, Lunt M, *et al.* The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology* 2002;41:793-800.
13. Symmons DPM, Barrett EM, Bankhead CR, Scott DGL, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *Rheumatology* 1994;33:735-9.
14. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, *et al.* Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis & Rheumatism* 2000;43:30-7.
15. Jones MA, Silman AJ, Whiting S, Barrett EM, Symmons DP. Occurrence of rheumatoid arthritis is not increased in the first degree relatives of a population based inception cohort of inflammatory polyarthritis. *Ann Rheum Dis* 1996;55:89-93.
16. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, *et al.* The american rheumatism association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis & Rheumatism* 1988;31:315-24.
17. Combe B, Landewe R, Lukas C, Bolosiu HD, Breedveld F, Dougados M, *et al.* EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2007;66:34-45.
18. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, *et al.* 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69:1580-8.
19. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, *et al.* American college of rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis & Rheumatism* 1995;38:727-35.
20. van Gestel AM, Prevoo MLL, van't Hof MA, van Rijswijk MH, van de Putte LBA, van Riel PL. Development and validation of the european league against rheumatism response criteria for rheumatoid arthritis: Comparison with the preliminary american college of rheumatology and the world health organization/international league against rheumatism criteria. *Arthritis & Rheumatism* 1996;39:34-40.
21. Felson DT. Assessing the efficacy and safety of rheumatic disease treatments: Obstacles and proposed solutions. *Arthritis & Rheumatism* 2003;48:1781-7.
22. Prevoo MLL, Van'T Hof MA, Kuper HH, Van Leeuwen MA, Van De Putte LBA, Van Riel PLCM. Modified disease activity scores that include twenty-eight-joint counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis & Rheumatism* 1995;38:44-8.
23. van Gestel AM, Anderson JJ, van Riel PL, Boers M, Haagsma CJ, Rich B., *et al.* ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. American

College of Rheumatology European League of Associations for Rheumatology. *Journal of Rheumatology* 1999;26:705-11.
24. Stevenson M, Archer R, Tosh J, Simpson E, Everson-Hock E, Stevens 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* 2016;20.
25. National Institute for Health and Clinical Excellence (NICE). Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with DMARDs or after conventional DMARDs only have failed; 2016.
26. Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23:S14-S8.
27. National Institute for Health and Clinical Excellence (NICE). Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor. London: NICE; 2010.
28. National Institute for Health and Clinical Excellence (NICE). Tocilizumab for the treatment of rheumatoid arthritis.; 2012.
29. National Institute for Health and Care Excellence. Certolizumab pegol for treating rheumatoid arthritis after inadequate response to a TNF inhibitor – Company Submission. London: NICE; 2016.
30. Gibson L, Hernandez Alava M, Wailoo A. Progression of Disease in People with Rheumatoid Arthritis Treated with Non-Biologic Therapies. Sheffield; 2015.
31. Sanofi Genzyme. Sarilumab for treating moderate-to-severe rheumatoid arthritis [ID994]. Clarification response; 2017.
32. Huizinga TW, Fleischmann RM, Jasson M, Radin AR, van Adelsberg J, Fiore S, *et al.* 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. *Ann Rheum Dis* 2014;73:1626-34.
33. Genovese MC, Fleischmann R, Kivitz AJ, Rell-Bakalarska M, Martincova R, Fiore S, *et al.* Sarilumab Plus Methotrexate in Patients With Active Rheumatoid Arthritis and Inadequate Response to Methotrexate: Results of a Phase III Study. *Arthritis rheumatol* 2015;67:1424-37.
34. Fleischmann R, van Adelsberg J, Lin Y, da Rocha Castelar-Pinheiro G, Brzezicki J, Hrycaj P, *et al.* Sarilumab and Nonbiologic Disease-Modifying Antirheumatic Drugs in Patients With Active Rheumatoid Arthritis and Inadequate Response or Intolerance to Tumor Necrosis Factor Inhibitors. *Arthritis rheumatol* 2017;69:277-90.
35. Burmester GR, Lin Y, Patel R, van Adelsberg J, Mangan EK, Graham NM, *et al.* 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. *Ann Rheum Dis* 2016;76:840-7.
36. Sanofi Genzyme. EFC11072. MOBILITY Part B. A randomized, double-blind, placebo-controlled, multicenter, two-part, dose ranging and confirmatory study with an operationally seamless design, evaluating efficacy and safety of SAR153191 on top of methotrexate (MTX) in patients with active rheumatoid arthritis who are inadequate responders to MTX therapy. Data on file.
37. Centre for R, Dissemination. Systematic review: CRD's guidance for undertaking reviews in health care. CRD, University of York; 2009. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]; 2011.
38. Sanofi Genzyme. EFC14092. MONARCH. A randomized, double-blind, parallel-group study assessing the efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy in patients with rheumatoid arthritis. Data on file.
39. Sanofi Genzyme. Sarilumab. Summary of Product Characteristics. 2017.
40. Sanofi Genzyme. EFC13752. ONE. An Open-label, Randomized, Parallel Group Study Assessing the Immunogenicity and Safety of Sarilumab Administered as Monotherapy in Patients with Active Rheumatoid Arthritis. Data on file.

41. Sanofi Genzyme. EFC14059. KAKEHASI. A randomized, double-blind, multicenter study with a placebo-controlled period assessing the efficacy and safety of Sarilumab added to methotrexate (MTX) in Japanese patients with moderately to severely active rheumatoid arthritis who are inadequate responders to MTX therapy Data on file.
42. Sanofi Genzyme. EFC10832. TARGET. A randomized, double-blind, parallel, placebo-controlled study assessing the efficacy and safety of sarilumab added to non-biologic DMARD therapy in patients with rheumatoid arthritis who are inadequate responders to or intolerant of TNF-α antagonists. Data on file.
43. British National Formulary. British National Formulary (BNF) 71. 2016.
44. Fleischmann R, Vencovsky J, van Vollenhoven RF, Borenstein D, Box J, Coteur G, *et al.* 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* 2009;68:805-11.
45. Choy E, McKenna F, Vencovsky J, Valente R, Goel N, Vanlunen B, *et al.* Certolizumab pegol plus MTX administered every 4 weeks is effective in patients with RA who are partial responders to MTX. *Rheumatology (Oxford, England)* 2012;51:1226-34.
46. Bingham CO, Weinblatt M, Han C, Gathany TA, Kim L, Lo KH, *et al.* The Effect of Intravenous Golimumab on Health-related Quality of Life in Rheumatoid Arthritis: 24-week Results of the Phase III GO-FURTHER Trial. *The Journal of Rheumatology* 2014;41:1067-76.
47. Kay J, Matteson EL, Dasgupta B, Nash P, Durez P, Hall S, *et al.* Golimumab in patients with active rheumatoid arthritis despite treatment with methotrexate: a randomized, double-blind, placebo-controlled, dose-ranging study. *Arthritis and Rheumatism* 2008;58:964-75.
48. Conaghan PG, Durez P, Alten RE, Burmester GR, Tak PP, Klareskog L, *et al.* 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. *Ann Rheum Dis* 2013;72:1287-94.
49. Chen DY, Chou SJ, Hsieh TY, Chen YH, Chen HH, Hsieh CW, *et al.* 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 = Taiwan yi zhi* 2009;108:310-9.
50. Lan JL, Chou SJ, Chen DY, Chen YH, Hsieh TY, Young M, Jr. 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 = Taiwan yi zhi* 2004;103:618-23.
51. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI, *et al.* A Trial of Etanercept, a Recombinant Tumor Necrosis Factor Receptor:Fc Fusion Protein, in Patients with Rheumatoid Arthritis Receiving Methotrexate. *N Engl J Med* 1999;340:253-9.
52. Taylor PC, Steuer A, Gruber J, Cosgrove DO, Blomley MJ, Marsters PA, *et al.* 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* 2004;50:1107-16.
53. Maini RN, Breedveld FC, Kalden JR, Smolen JS, Davis D, Macfarlane JD, *et al.* 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* 1998;41:1552-63.
54. Tam LS, Shang Q, Li EK, Wang S, Li RJ, Lee KL, *et al.* Infliximab is associated with improvement in arterial stiffness in patients with early rheumatoid arthritis -- a randomized trial. *J Rheumatol* 2012;39:2267-75.
55. Tanaka Y, Suzuki M, Nakamura H, Toyoizumi S, Zwillich SH. Phase II study of tofacitinib (CP-690,550) combined with methotrexate in patients with rheumatoid arthritis and an inadequate response to methotrexate. *Arthritis Care Res (Hoboken)* 2011;63:1150-8.
56. Smolen JS, Weinblatt ME, Sheng S, Zhuang Y, Hsu B. Sirukumab, a human anti-interleukin-6 monoclonal antibody: a randomised, 2-part (proof-of-concept and dose-finding), phase II study

in patients with active rheumatoid arthritis despite methotrexate therapy. *Ann Rheum Dis* 2014;73:1616-25.
57. Schiff MH, von Kempis J, Goldblum R, Tesser JR, Mueller RB. 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. *Ann Rheum Dis* 2014;73:2174-7.
58. Genovese MC, Fleischmann R, Furst D, Janssen N, Carter J, Dasgupta B, *et al.* Efficacy and safety of olokizumab in patients with rheumatoid arthritis with an inadequate response to TNF inhibitor therapy: outcomes of a randomised Phase IIb study. *Ann Rheum Dis* 2014;73:1607-15.
59. O'Dell JR, Mikuls TR, Taylor TH, Ahluwalia V, Brophy M, Warren SR, *et al.* Therapies for active rheumatoid arthritis after methotrexate failure. *N Engl J Med* 2013;369:307-18.
60. Machado DA, Guzman RM, Xavier RM, Simon JA, Mele L, Pedersen R, *et al.* 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. *J Clin Rheumatol* 2014;20:25-33.
61. Kaneko Y, Atsumi T, Tanaka Y, Inoo M, Kobayashi-Haraoka H, Amano K, *et al.* 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). *Ann Rheum Dis* 2016;75:1917-23.
62. Takeuchi T, Kaneko Y, Atsumi T, Tanaka Y, Inoh M, Kobayashi H, *et al.* OP0040 Adding Tocilizumab or Switching to Tocilizumab Monotherapy in RA Patients with Inadequate Response to Methotrexate: 24-Week Results from a Randomized Controlled Study (Surprise Study). *Ann Rheum Dis* 2013;72:A62-A3.
63. Takeuchi T, Kaneko Y, Atsumi T, Tanaka Y, Inoh M, Kobayashi H, *et al.* SAT0257 Clinical and Radiographic Effects after 52-Week of Adding Tocilizumab or Switching to Tocilizumab in RA Patients with Inadequate Response to Methotrexate: Results from A Prospective Randomized Controlled Study (Surprise Study). *Ann Rheum Dis* 2014;73:686-.
64. Huizinga TW, Conaghan PG, Martin-Mola E, Schett G, Amital H, Xavier RM, *et al.* Clinical and radiographic outcomes at 2 years and the effect of tocilizumab discontinuation following sustained remission in the second and third year of the ACT-RAY study. *Ann Rheum Dis* 2015;74:35-43.
65. Kameda H, Kanbe K, Sato E, Ueki Y, Saito K, Nagaoka S, *et al.* Continuation of methotrexate resulted in better clinical and radiographic outcomes than discontinuation upon starting etanercept in patients with rheumatoid arthritis: 52-week results from the JESMR study. *J Rheumatol* 2011;38:1585-92.
66. Kameda H, Ueki Y, Saito K, Nagaoka S, Hidaka T, Atsumi T, *et al.* Etanercept (ETN) with methotrexate (MTX) is better than ETN monotherapy in patients with active rheumatoid arthritis despite MTX therapy: a randomized trial. *Modern rheumatology / the Japan Rheumatism Association* 2010;20:531-8.
67. van de Putte LB, Atkins C, Malaise M, Sany J, Russell AS, van Riel PL, *et al.* 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* 2004;63:508-16.
68. Miyasaka N. Clinical investigation in highly disease-affected rheumatoid arthritis patients in Japan with adalimumab applying standard and general evaluation: the CHANGE study. *Modern rheumatology* 2008;18:252-62.
69. Combe B, Codreanu C, Fiocco U, Gaubitz M, Geusens PP, Kvien TK, *et al.* 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. *Ann Rheum Dis* 2009;68:1146-52.
70. Combe B, Codreanu C, Fiocco U, Gaubitz M, Geusens PP, Kvien TK, *et al.* Etanercept and sulfasalazine, alone and combined, in patients with active rheumatoid arthritis despite receiving sulfasalazine: a double-blind comparison. *Ann Rheum Dis* 2006;65:1357-62.

71. Moreland LW, Schiff MH, Baumgartner SW, Tindall EA, Fleischmann RM, Bulpitt KJ, *et al.* Etanercept therapy in rheumatoid arthritis. A randomized, controlled trial. *Annals of internal medicine* 1999;130:478-86.
72. Nishimoto N, Miyasaka N, Yamamoto K, Kawai S, Takeuchi T, Azuma J, *et al.* 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 / the Japan Rheumatism Association* 2009;19:12-9.
73. Gabay C, Emery P, van Vollenhoven R, Dikranian A, Alten R, Pavelka K, *et al.* Tocilizumab monotherapy versus adalimumab monotherapy for treatment of rheumatoid arthritis (ADACTA): a randomised, double-blind, controlled phase 4 trial. *Lancet (London, England)* 2013;381:1541-50.
74. Taylor P, Schiff M, Wang Q, Jiang Y, Kurrasch R, Daga S, *et al.* Efficacy and Safety of Monotherapy with Sirukumab, an Anti–IL-6 Cytokine Monoclonal Antibody, Compared with Adalimumab Monotherapy in Biologic-Naive Patients with Active Rheumatoid Arthritis: Results of a Global, Randomized, Double-Blind, Parallel-Group, Phase 3 Study [abstract]. *Arthritis rheumatol* 2016;68 (suppl 10).
75. Fleischmann R, Cutolo M, Genovese MC, Lee EB, Kanik KS, Sadis S, *et al.* Phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) or adalimumab monotherapy versus placebo in patients with active rheumatoid arthritis with an inadequate response to disease-modifying antirheumatic drugs. *Arthritis and Rheumatism* 2012;64:617-29.
76. Keystone EC, van der Heijde D, Kavanaugh A, Kupper H, Liu S, Guerette B, *et al.* Clinical, functional, and radiographic benefits of longterm adalimumab plus methotrexate: final 10-year data in longstanding rheumatoid arthritis. *J Rheumatol* 2013;40:1487-97.
77. Keystone EC, Kavanaugh A, Weinblatt ME, Patra K, Pangan AL. Clinical consequences of delayed addition of adalimumab to methotrexate therapy over 5 years in patients with rheumatoid arthritis. *J Rheumatol* 2011;38:855-62.
78. Kim H-Y, Lee S-K, Song YW, Yoo D-H, Koh E-M, Yoo B, *et al.* 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* 2007;10:9-16.
79. Weinblatt ME, Keystone EC, Furst DE, Moreland LW, Weisman MH, Birbara CA, *et al.* 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* 2003;48:35-45.
80. Furst DE, Schiff MH, Fleischmann RM, Strand V, Birbara CA, Compagnone D, *et al.* 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). *J Rheumatol* 2003;30:2563-71.
81. Keystone E, Heijde D, Mason D, Jr., Landewe R, Vollenhoven RV, Combe B, *et al.* 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* 2008;58:3319-29.
82. Strand V, Mease P, Burmester GR, Nikai E, Coteur G, van Vollenhoven R, *et al.* 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 Res Ther* 2009;11:R170.
83. Yamamoto K, Takeuchi T, Yamanaka H, Ishiguro N, Tanaka Y, Eguchi K, *et al.* 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 / the Japan Rheumatism Association* 2014;24:715-24.

84. Smolen J, Landewé RB, Mease P, Brzezicki J, Mason D, Luijtens K, *et al.* Efficacy and safety of certolizumab pegol plus methotrexate in active rheumatoid arthritis: the RAPID 2 study. A randomised controlled trial. *Ann Rheum Dis* 2009;68:797-804.
85. Smolen JS, Emery P, Ferraccioli GF, Samborski W, Berenbaum F, Davies OR, *et al.* Certolizumab pegol in rheumatoid arthritis patients with low to moderate activity: the CERTAIN double-blind, randomised, placebo-controlled trial. *Ann Rheum Dis* 2015;74:843-50.
86. Yamanaka H, Nagaoka S, Lee SK, Bae SC, Kasama T, Kobayashi H, *et al.* Discontinuation of etanercept after achievement of sustained remission in patients with rheumatoid arthritis who initially had moderate disease activity-results from the ENCOURAGE study, a prospective, international, multicenter randomized study. *Modern rheumatology / the Japan Rheumatism Association* 2016;26:651-61.
87. Tanaka Y, Harigai M, Takeuchi T, Yamanaka H, Ishiguro N, Yamamoto K, *et al.* Golimumab in combination with methotrexate in Japanese patients with active rheumatoid arthritis: results of the GO-FORTH study. *Ann Rheum Dis* 2012;71:817-24.
88. Li Z, Zhang F, Kay J, Fei K, Han C, Zhuang Y, *et al.* 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* 2016;19:1143-56.
89. Weinblatt ME, Westhovens R, Mendelsohn AM, Kim L, Lo KH, Sheng S, *et al.* 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. *Ann Rheum Dis* 2014;73:2152-9.
90. Weinblatt ME, Bingham CO, Mendelsohn AM, Kim L, Mack M, Lu J, *et al.* 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. *Ann Rheum Dis* 2013;72:381-9.
91. Keystone EC, Genovese MC, Hall S, Bae SC, Han C, Gathany TA, *et al.* Safety and Efficacy of Subcutaneous Golimumab in Patients with Active Rheumatoid Arthritis despite Methotrexate Therapy: Final 5-year Results of the GO-FORWARD Trial. *J Rheumatol* 2016;43:298-306.
92. Keystone EC, Genovese MC, Hall S, Miranda PC, Bae SC, Palmer W, *et al.* Golimumab in patients with active rheumatoid arthritis despite methotrexate therapy: results through 2 years of the GO-FORWARD study extension. *J Rheumatol* 2013;40:1097-103.
93. Genovese MC, Han C, Keystone EC, Hsia EC, Buchanan J, Gathany T, *et al.* Effect of golimumab on patient-reported outcomes in rheumatoid arthritis: results from the GO-FORWARD study. *J Rheumatol* 2012;39:1185-91.
94. Keystone E, Genovese MC, Klareskog L, Hsia EC, Hall S, Miranda PC, *et al.* Golimumab in patients with active rheumatoid arthritis despite methotrexate therapy: 52-week results of the GO-FORWARD study. *Ann Rheum Dis* 2010;69:1129-35.
95. Keystone EC, Genovese MC, Klareskog L, Hsia EC, Hall ST, Miranda PC, *et al.* 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. *Ann Rheum Dis* 2009;68:789-96.
96. Maini R, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M, *et al.* 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 (London, England)* 1999;354:1932-9.
97. Lipsky PE, van der Heijde DM, St Clair EW, Furst DE, Breedveld FC, Kalden JR, *et al.* Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. *N Engl J Med* 2000;343:1594-602.
98. Maini RN, Breedveld FC, Kalden JR, Smolen JS, Furst D, Weisman MH, *et al.* Sustained improvement over two years in physical function, structural damage, and signs and symptoms

among patients with rheumatoid arthritis treated with infliximab and methotrexate. *Arthritis and Rheumatism* 2004;50:1051-65.

99. Westhovens R, Yocum D, Han J, Berman A, Strusberg I, Geusens P, *et al.* The safety of infliximab, combined with background treatments, among patients with rheumatoid arthritis and various comorbidities: A large, randomized, placebo-controlled trial. *Arthritis & Rheumatism* 2006;54:1075-86.
100. Schiff M, Keiserman M, Codding C, Songcharoen S, Berman A, Nayiager S, *et al.* 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. *Ann Rheum Dis* 2008;67:1096-103.
101. Karlsson JA, Neovius M, Nilsson JA, Petersson IF, Bratt J, van Vollenhoven RF, *et al.* Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in early rheumatoid arthritis: 2-year quality-of-life results of the randomised, controlled, SWEFOT trial. *Ann Rheum Dis* 2013;72:1927-33.
102. Rezaei H, Saevarsdottir S, Geborek P, Petersson IF, van Vollenhoven RF, Forslind K. 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 Musculoskelet Disord* 2013;14:79.
103. van Vollenhoven RF, Geborek P, Forslind K, Albertsson K, Ernestam S, Petersson IF, *et al.* Conventional combination treatment versus biological treatment in methotrexate-refractory early rheumatoid arthritis: 2 year follow-up of the randomised, non-blinded, parallel-group Swefot trial. *Lancet (London, England)* 2012;379:1712-20.
104. van Vollenhoven RF, Ernestam S, Geborek P, Petersson IF, Cöster L, Waltbrand E, *et al.* 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. *The Lancet* 2009;374:459-66.
105. Eriksson JK, Neovius M, Bratt J, Petersson IF, van Vollenhoven RF, Geborek P, *et al.* Biological vs. conventional combination treatment and work loss in early rheumatoid arthritis: a randomized trial. *JAMA internal medicine* 2013;173:1407-14.
106. Russell AS, Wallenstein GV, Li T, Martin MC, Maclean R, Blaisdell B, *et al.* Abatacept improves both the physical and mental health of patients with rheumatoid arthritis who have inadequate response to methotrexate treatment. *Ann Rheum Dis* 2007;66:189-94.
107. Kremer JM, Genant HK, Moreland LW, et al. Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: A randomized trial. *Annals of Internal Medicine* 2006;144:865-76.
108. Emery P, Kosinski M, Li T, Martin M, Williams GR, Becker JC, *et al.* Treatment of rheumatoid arthritis patients with abatacept and methotrexate significantly improved health-related quality of life. *J Rheumatol* 2006;33:681-9.
109. Kremer JM, Dougados M, Emery P, Durez P, Sibilia J, Shergy W, *et al.* 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* 2005;52:2263-71.
110. Kremer JM, Westhovens R, Leon M, Di Giorgio E, Alten R, Steinfeld S, *et al.* Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. *N Engl J Med* 2003;349:1907-15.
111. Takeuchi T, Matsubara T, Nitobe T, Suematsu E, Ohta S, Honjo S, *et al.* Phase II dose-response study of abatacept in Japanese patients with active rheumatoid arthritis with an inadequate response to methotrexate. *Modern rheumatology / the Japan Rheumatism Association* 2013;23:226-35.
112. Weinblatt M, Combe B, Covucci A, Aranda R, Becker JC, Keystone E. 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* 2006;54:2807-16.
113. Emery. 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]). *Ann Rheum Dis* 2010;69:1629-35.

114. Mease PJ, Revicki DA, Szechinski J, Greenwald M, Kivitz A, Barile-Fabris L, *et al.* 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. *J Rheumatol* 2008;35:20-30.
115. Strand V, Balbir-Gurman A, Pavelka K, Emery P, Li N, Yin M, *et al.* Sustained benefit in rheumatoid arthritis following one course of rituximab: improvements in physical function over 2 years. *Rheumatology (Oxford, England)* 2006;45:1505-13.
116. Edwards JC, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR, *et al.* Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. *N Engl J Med* 2004;350:2572-81.
117. Peterfy C, Emery P, Tak PP, Ostergaard M, DiCarlo J, Otsa K, *et al.* 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. *Ann Rheum Dis* 2016;75:170-7.
118. Behrens F, Rossmanith T, Köhm M, Alten R, Aringer M, Backhaus M, *et al.* FRI0199 Rituximab in Combination with Leflunomide: Results from A Multicenter Randomized Placebo Controlled Investigator Initiated Clinical Trial in Active Rheumatoid Arthritis (Amara-Study). *Ann Rheum Dis* 2016;75:502-.
119. Smolen JS, Beaulieu A, Rubbert-Roth A, Ramos-Remus C, Rovensky J, Alecock E, *et al.* Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial. *The Lancet* 2008;371:987-97.
120. McInnes IB, Thompson L, Giles JT, Bathon JM, Salmon JE, Beaulieu AD, *et al.* Effect of interleukin-6 receptor blockade on surrogates of vascular risk in rheumatoid arthritis: MEASURE, a randomised, placebo-controlled study. *Ann Rheum Dis* 2015;74:694-702.
121. Mirjafari H, Klearman M, Wang J, Sattar N, Giles J. Adiponectin levels are increased after tocilizumab therapy among adults with rheumatoid arthritis: Results from the measure study. *Arthritis and Rheumatism* 2013;54.
122. Fleischmann RM, Halland A-M, Brzosko M, Burgos-Vargas R, Mela C, Vernon E, *et al.* Tocilizumab Inhibits Structural Joint Damage and Improves Physical Function in Patients with Rheumatoid Arthritis and Inadequate Responses to Methotrexate: LITHE Study 2-year Results. *The Journal of Rheumatology* 2013;40:113-26.
123. Kremer JM, Blanco R, Brzosko M, Burgos-Vargas R, Halland A-M, Vernon E, *et al.* 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 & Rheumatism* 2011;63:609-21.
124. Kivitz A, Olech E, Borofsky M, Zazueta BM, Navarro-Sarabia F, Radominski SC, *et al.* Subcutaneous tocilizumab versus placebo in combination with disease-modifying antirheumatic drugs in patients with rheumatoid arthritis. *Arthritis Care & Research* 2014;66:1653-61.
125. Kivitz A, Olech E, Borofsky M, et al. The safety and efficacy of tocilizumab subcutaneous in combination with traditional DMARDS in patients with moderate to severe rheumatoid arthritis up to 48 weeks (BREVACTA). *Arthritis and Rheumatism* 2013;65.
126. Genovese MC, McKay JD, Nasonov EL, Mysler EF, da Silva NA, Alecock E, *et al.* 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* 2008;58:2968-80.
127. Yazici Y, Curtis JR, Ince A, Baraf H, Malamet RL, Teng LL, *et al.* 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. *Ann Rheum Dis* 2012;71:198-205.

128. Kremer JM, Cohen S, Wilkinson BE, Connell CA, French JL, Gomez-Reino J, *et al.* A phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) versus placebo in combination with background methotrexate in patients with active rheumatoid arthritis and an inadequate response to methotrexate alone. *Arthritis and Rheumatism* 2012;64:970-81.
129. van der Heijde D, Tanaka Y, Fleischmann R, Keystone E, Kremer J, Zerbini C, *et al.* Tofacitinib (CP-690,550) in patients with rheumatoid arthritis receiving methotrexate: twelve-month data from a twenty-four-month phase III randomized radiographic study. *Arthritis and Rheumatism* 2013;65:559-70.
130. van Vollenhoven RF, Fleischmann R, Cohen S, Lee EB, Garcia Meijide JA, Wagner S, *et al.* Tofacitinib or adalimumab versus placebo in rheumatoid arthritis. *The New England journal of medicine* 2012;367:508-19.
131. Kremer J, Li Z-G, Hall S, Fleischmann R, Genovese M, Martin-Mola E, *et al.* Tofacitinib in combination with nonbiologic disease-modifying antirheumatic drugs in patients with active rheumatoid arthritis: a randomized trial. *Annals of internal medicine* 2013;159:253-61.
132. Dougados M, van der Heijde D, Chen YC, Greenwald M, Drescher E, Liu J, *et al.* Baricitinib in patients with inadequate response or intolerance to conventional synthetic DMARDs: results from the RA-BUILD study. *Ann Rheum Dis* 2017;76:88-95.
133. Burmester GR, Rubbert-Roth A, Cantagrel A, Hall S, Leszczynski P, Feldman D, *et al.* Efficacy and safety of subcutaneous tocilizumab versus intravenous tocilizumab in combination with traditional DMARDs in patients with RA at week 97 (SUMMACTA). *Ann Rheum Dis* 2014;73.
134. Burmester GR, Rubbert-Roth A, Cantagrel A, Hall S, Leszczynski P, Feldman D, *et al.* 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). *Ann Rheum Dis* 2014;73:69-74.
135. Burmester GR, Rubbert-Roth A, Cantagrel A, Hall S, Leszczynski P, Feldman D, *et al.* Summacta: A randomized, double-blind, parallel group study of the safety and efficacy of tocilizumab SC versus tocilizumab IV, in combination with traditional dmards in patients with moderate to severe RA. *Annals of the Rheumatic Diseases Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR* 2013;72.
136. Schiff M, Weinblatt ME, Valente R, van der Heijde D, Citera G, Elegbe A, *et al.* Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: two-year efficacy and safety findings from AMPLE trial. *Ann Rheum Dis* 2014;73:86-94.
137. Weinblatt ME, Schiff M, Valente R, van der Heijde D, Citera G, Zhao C, *et al.* Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: findings of a phase IIIb, multinational, prospective, randomized study. *Arthritis and Rheumatism* 2013;65:28-38.
138. Taylor PC, Keystone EC, van der Heijde D, Weinblatt ME, Del Carmen Morales L, Reyes Gonzaga J, *et al.* Baricitinib versus Placebo or Adalimumab in Rheumatoid Arthritis. *N Engl J Med* 2017;376:652-62.
139. Smolen JS, Kay J, Doyle MK, Landewe R, Matteson EL, Wollenhaupt J, *et al.* 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 (London, England)* 2009;374:210-21.
140. Tanaka Y, Bingham CI, Aletaha D, Agarwal P, Popik S, Kurrasch R, *et al.* Sirukumab, an Anti–IL-6 Cytokine Monoclonal Antibody, Significantly Improves Physical Function and Reduces Morning Stiffness in Patients with Active Rheumatoid Arthritis Despite Anti-TNF Therapy: Results from a Global, Randomized, Placebo-Controlled, Phase 3 Trial [abstract]. *Arthritis Rheumatol* 2016;68.
141. Westhovens R, Cole JC, Li T, Martin M, Maclean R, Lin P, *et al.* 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 (Oxford, England)* 2006;45:1238-46.

142. Genovese MC, Becker JC, Schiff M, Luggen M, Sherrer Y, Kremer J, *et al.* Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. *N Engl J Med* 2005;353:1114-23.
143. Keystone E, Emery P, Peterfy CG, Tak PP, Cohen S, Genovese MC, *et al.* Rituximab inhibits structural joint damage in patients with rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitor therapies. *Ann Rheum Dis* 2009;68:216-21.
144. Keystone E, Burmester GR, Furie R, Loveless JE, Emery P, Kremer J, *et al.* 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* 2008;59:785-93.
145. Cohen SB, Emery P, Greenwald MW, Dougados M, Furie RA, Genovese MC, *et al.* 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* 2006;54:2793-806.
146. Strand V, Burmester GR, Zerbini CA, Mebus CA, Zwillich SH, Gruben D, *et al.* Tofacitinib with methotrexate in third-line treatment of patients with active rheumatoid arthritis: patient-reported outcomes from a phase III trial. *Arthritis Care Res (Hoboken)* 2015;67:475-83.
147. Burmester GR, Blanco R, Charles-Schoeman C, Wollenhaupt J, Zerbini C, Benda B, *et al.* Tofacitinib (CP-690,550) in combination with methotrexate in patients with active rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitors: a randomised phase 3 trial. *The Lancet*;381:451-60.
148. Strand V, Burmester GR, Ogale S, Devenport J, John A, Emery P. 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* 2012;51:1860-9.
149. Emery P, Keystone E, Tony HP, Cantagrel A, van Vollenhoven R, Sanchez A, *et al.* 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. *Ann Rheum Dis* 2008;67:1516-23.
150. Genovese M, Kremer J, Zamani O, et al. Baricitinib an oral janus kinase (JAK)1/JAK2 inhibitor, in patients with active rheumatoid arthritis (RA) an an inadequate response to TNF inhibitors: Results of the Phase 3 RA-BEACON study. *The Journal of Rheumatology* 2016;43:1188.
151. Manders SH, Kievit W, Adang E, Brus HL, Moens HJ, Hartkamp A, *et al.* Cost-effectiveness of abatacept, rituximab, and TNFi treatment after previous failure with TNFi treatment in rheumatoid arthritis: a pragmatic multi-centre randomised trial. *Arthritis Res Ther* 2015;17:134.
152. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology* 1997;50:683-91.
153. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JPT. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *International Journal of Epidemiology* 2012;41:818-27.
154. ClinicalTrials.gov. A Study Assessing the Safety and Efficacy of Sarilumab Added to Non-MTX DMARDs or as Monotherapy in Japanese Patients With Active Rheumatoid Arthritis (SARIL-RA-HARUKA). NCT02373202.; 2016. https://clinicaltrials.gov/ct2/show/NCT02373202 (accessed July 2017). (Accessed
155. ClinicalTrials.gov. A Study Assessing the Efficacy and Safety of Sarilumab Added to MTX in Japanese Patients With Moderately to Severely Active Rheumatoid Arthritis (SARIL-RA-KAKEHASI). NCT02293902. 2016, https://clinicaltrials.gov/ct2/show/NCT02293902?term=kakehasi&rank=1 (accessed April 2017).
156. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal; 2013.
157. Malottki K, Barton P, Tsourapas A, Uthman A, Liu Z. Adalimumab, etanercept, infliximab, 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* 2011;15:278.
158. Hernández Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. *Rheumatology* 2013;52:944-50.
159. Bansback N, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. *Ann Rheum Dis* 2005;64:995-1002.
160. Wiles NJ, Symmons D. Resource Use Within the Norfolk Arthritis Register (NOAR) Cohort During the First Five Years of Disease. Report for Roche. London: NICE; 2005.
161. Norton S, Sacker A, Dixey J, Done J, Williams P, Young A. Trajectories of functional limitation in early rheumatoid arthritis and their association with mortality. *Rheumatology* 2013;52:2016-24.
162. Office for National Statistics. National Life Tables, United Kingdom: 2011-2013. 2014. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2014-09-25 (Accessed 20 June 2017).
163. Singh JA, Wells GA, Christensen R, Tanjong Ghogomu E, Maxwell L, Macdonald JK, *et al.* Adverse effects of biologics: a network meta-analysis and Cochrane overview. *Cochrane Database Syst Rev* 2011; 10.1002/14651858.CD008794.pub2:Cd008794.
164. Oppong R, Kaambwa B, Nuttall J, Hood K, Smith RD, Coast J. 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. *Eur J Health Econ* 2013;14:197-209.
165. MIMS. Infliximab (Remsima). 2016. http://www.mims.co.uk/drugs/musculoskeletal-disorders/rheumatoid-arthritis-other-autoimmune-disorders/remsima (Accessed December).
166. National Institute for Health and Care Excellence. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with DMARDs or after conventional DMARDs only have failed. 2016
167. Department of Health. National Schedule of Reference Costs 2015-16 for NHS Trusts. Total average cost of outpatient attendance for Service Code 410 (Rheumatology). Sheet: Total Outpatient Attendances.; 2016.
168. Drummond M, Sculpher M, Torrance G, O'Brien B, G S. Methods for the economic evaluation of health care programmes. 3rd edn. Oxford: Oxford University Press; 2005.
169. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, *et al.* Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices — Modeling Studies. *Value in Health* 2014;6:9-17.
170. Eli Lilly. Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis [ID979]; 2017.

9 APPENDICES

Appendix 1: The sequences evaluated in the original company submission.

Table 70: Treatment Sequences compared for cDMARD-IR patients with severe RA who can receive MTX

Line 1	Line 2	Line 3	Line 4
Sarilumab + MTX	> Rituximab + MTX	> Abatacept IV + MTX	> BSC
Tocilizumab IV + MTX	> Rituximab + MTX	> Abatacept IV + MTX	> BSC
Tocilizumab SC + MTX	> Rituximab + MTX	> Abatacept IV + MTX	> BSC
TNFi Bundle + MTX	> Rituximab + MTX	> Abatacept IV + MTX	> BSC
Abatacept SC + MTX	> Rituximab + MTX	> Tocilizumab IV + MTX	> BSC

BSC=best supportive care; IV=intravenous; MTX=methotrexate; SC=subcutaneous; TNFi=tumour necrosis factor inhibitor

Table 71: Treatment Sequences compared for cDMARD-IR patients with severe RA who cannot receive MTX

Line 1	Line 2	Line 3
Sarilumab	> TNFi Bundle	> BSC
Tocilizumab IV	> TNFi Bundle	> BSC
Tocilizumab SC	> TNFi Bundle	> BSC
TNFi Bundle	> TNFi Bundle	> BSC

BSC=best supportive care; IV=intravenous; SC=subcutaneous; TNFi=tumour necrosis factor inhibitor

Table 72: Treatment Sequences compared for TNFi-IR patients with severe RA who can receive RTX and MTX

Line 1	Line 2	Line 3
Sarilumab + MTX	> Abatacept IV + MTX	> BSC
Rituximab + MTX	> Abatacept IV + MTX	> BSC

BSC=best supportive care; MTX=methotrexate; TNFi=tumour necrosis factor inhibitor

Table 73: Treatment Sequences compared for TNFi-IR patients with severe RA who cannot receive but can receive MTX

Line 1	Line 2	Line 3
Sarilumab + MTX	> Abatacept IV + MTX	> BSC
Tocilizumab IV + MTX	> Abatacept IV + MTX	> BSC
Tocilizumab SC + MTX	> Abatacept IV + MTX	> BSC
TNFi Bundle + MTX	> Abatacept IV + MTX	> BSC
Abatacept SC + MTX	> Tocilizumab IV + MTX	> BSC

BSC=best supportive care; IV=intravenous; SC=subcutaneous; TNFi=tumour necrosis factor inhibitor

Table 74: Treatment Sequences compared for TNFi-IR patients with severe RA who cannot receive MTX

Line 1	Line 2
Sarilumab	> BSC
TNFi Bundle	> BSC

BSC=best supportive care; MTX=methotrexate; TNFi=tumour necrosis factor inhibitor

Table 75: Treatment Sequences compared for TNFi-IR patients with severe RA who have received RTX + MTX

Line 1	Line 2
Sarilumab + MTX	> BSC
Tocilizumab IV + MTX	> BSC
Tocilizumab SC + MTX	> BSC
BSC	-

BSC=best supportive care; IV=intravenous; SC=subcutaneous

Table 76: Treatment Sequences compared for cDMARD-IR patients with moderate RA (DAS28 > 4.0) who can receive MTX

Line 1	Line 2
Sarilumab + MTX	> BSC
BSC	

BSC=best supportive care