NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Baricitinib for treating moderate to severe rheumatoid arthritis [ID979]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Baricitinib for treating moderate to severe rheumatoid arthritis [ID979]

Contents:

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

Pre-meeting briefing Baricitinib for treating moderate to severe rheumatoid arthritis

Contains AIC and CIC

This slide set is the pre-meeting briefing for this appraisal. It has been prepared by the technical team with input from the committee lead team and the committee chair. It is sent to the appraisal committee before the committee meeting as part of the committee papers. It summarises:

- the key evidence and views submitted by the company, the consultees and their nominated clinical experts and patient experts and
- the Evidence Review Group (ERG) report.

It highlights key issues for discussion at the first appraisal committee meeting and should be read with the full supporting documents for this appraisal.

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies.

The lead team may use, or amend, some of these slides for their presentation at the Committee meeting.

Abbreviations (shaded rows contain comparator technologies)

ABA	Abatacept
ACPA	Anti-citrullinated protein antibody
ACR20/50/70	20%/50%/70% improvement in American College of Rheumatology Criteria
ADA	Adalimumab
BARI	Baricitinib
bDMARD	Biological DMARD
BSRBR	British Society for Rheumatology Biologics Register
CDAI	Clinical Disease Activity Index
cDMARD	Conventional DMARD
CTZ	Certolizumab pegol
DAS28/44	Disease activity score in 28/44 Joints
DMARD	Disease-modifying anti-rheumatic drug
ESR	Erythrocyte sedimentation rate
ETN	Etanercept
EULAR	European League against Rheumatism
FACIT-F	Functional Assessment of Chronic Illness Therapy-Fatigue
GOL	Golimumab
HAQ-DI	Health Assessment Questionnaire–Disability Index
HCQ	Hydroxychloroqine
HsCRP	High sensitivity C-reactive protein
IFX	Infliximab
IR	Insufficient response
JAK	Janus kinase
LDA	Low disease activity
MJS	Morning joint stiffness
mTSS	Modified Total Sharp Score
MTX	Methotrexate
Q2W	Every 2 weeks
NRI	Non-responder imputation
QD	Once daily
QW	Weekly
RA	Rheumatoid arthritis
RF	Rheumatoid factor
RTX	Rituximab
SDAI	Simplified Disease Activity Index
SSZ	Sulfasalazine
TCZ	Tocilizumab
TNF	Tumour necrosis factor
TNFi	Tumour necrosis factor inhibitor
TOFA	Tofacitinib

2

Note: Throughout the presentation the term '+ cDMARDs' includes '+ MTX'

Key issues: Clinical effectiveness

- Innovation, including that BARI is oral rather than subcutaneous or i.v. administration
- Is BARI comparable to the bDMARDs in clinical effectiveness in moderate and severe RA?
- Is BARI effective as a monotherapy?
- The ERG considered that the company's NMA results should be treated with caution. The ERG carried out an updated NMA
  - Are the Committee comfortable that the conclusions of the company NMA and the ERG NMA are broadly similar?

3

Key issues: Cost effectiveness

- Is BARI comparable to the bDMARDs in both clinical effectiveness and cost effectiveness?
- Has the case for BARI monotherapy been proven?

4

See section 5.5 of ERG report for a detailed critique of cost-effectiveness issues

Background to rheumatoid arthritis

- Autoimmune disease characterised by chronic systemic inflammation and progressive inflammation of the synovial joints resulting in pain and stiffness. Can lead to joint damage, deformities and loss of function
  - Hands, wrists and feet most commonly affected
- Initial symptoms are reversible but joint damage is not
- Disease severity measured using the composite disease activity score (DAS28), includes assessment of 28 joints for swelling/tenderness, the patient's assessment of health and erythrocyte sedimentation rate or C-reactive protein
  - DAS28 <3.2 indicates low disease activity, DAS28 ≥3.2 and ≤5.1 indicates moderate activity, and DAS28 >5.1 indicates high activity
- For the majority of patients, disease remains mild with occasional flare-ups of higher disease activity. However, for some patients the disease may be active and progressive, significantly compromising quality of life
- Management of RA aims to suppress disease activity and induce remission, prevent the development of irreversible joint damage and, in more severe disease, maintain quality of life and address comorbidities associated with the condition
- Affects ~450,000 people in the UK, with a prevalence of 0.86% and incidence of 0.47 per 1,000 person-years. Around 12,000 new cases are diagnosed each year. The National Rheumatoid Arthritis Society estimates up to 690,000 people are living with RA in the UK
  - In the UK, approximately 15% (~60,000) of RA patients have severe disease
  - More prevalent in women than men, with 2–3 times as many cases in women
  - Can develop at any age but the typical age of onset in the UK is approximately 40–70 years, with most diagnoses made when people are in their 70s

5

Relevant NICE technology appraisals		
TA	**Treatment**	**Population**
415 2016	CTZ + MTX	Adults whose disease has responded inadequately to, or who cannot tolerate, other DMARDs including at least 1 TNF inhibitor, only if: • disease activity is severe and RTX is contraindicated or not tolerated
	CTZ monotherapy	As above but only if: • RTX therapy cannot be given because MTX is contraindicated or not tolerated
375 2016	ADA, ETN, IFX, CTZ, GOL, TCZ, ABA (all + MTX)	Disease is severe (disease activity score [DAS28] >5.1) and has not responded to intensive therapy with a combination of cDMARDs
	ADA, ETN, CTZ, TCZ monotherapy	As above but for people who cannot have MTX because of contraindications or intolerance
247 2012	TCZ + MTX	Disease has responded inadequately to DMARDs and a TNF inhibitor and the person cannot have RTX because it is contraindicated or not tolerated, and TCZ is used as described for TNF inhibitor treatments in TA195, specifically the recommendations on disease activity or the disease has responded inadequately to 1 or more TNF inhibitor treatments and to RTX
225 2011	GOL + MTX	Adults whose RA has responded inadequately to other DMARDs, including a TNF inhibitor, if it is used as described for other TNF inhibitor treatments in TA195
195 2010	RTX + MTX	Adults with severe active RA with an inadequate response to, or are intolerant of, other DMARDs, including at least 1 TNF inhibitor.
	ADA, ETN, IFX, ABA (all + MTX)	As for RTX + MTX but for people who cannot have RTX because of contraindications or intolerance
	ADA, ETN monotherapy	As for RTX + MTX but for people who cannot have RTX because they have a contraindication to, or intolerance of MTX

6

Stopping rules

TA415 - Continue treatment only if there is at least a moderate response measured using EULAR criteria at 6 months. After an initial response within 6 months, withdraw treatment if at least a moderate EULAR response is not maintained.

TA375 - Continue treatment only if there is a moderate response measured using EULAR criteria at 6 months after starting therapy.

TA247 - As described for TA195

TA225 - GOL + MTX – As described for TA195

TA195 – RTX + MTX - Treatment with rituximab in combination with methotrexate should be continued only if there is an adequate response following initiation of therapy and if an adequate response is maintained following retreatment with a dosing interval of at least 6 months. An adequate response is defined as an improvement in DAS8 of 1.2 points or more.

TA195 - ADA, ETN, IFX, ABA (all + MTX) - Treatment with adalimumab, etanercept, infliximab and abatacept should be continued only if there is an adequate response (an improvement in DAS28 of 1.2 points or more) 6 months after initiation of therapy. Treatment should be monitored, with assessment of DAS28, at least every 6 months and continued only if an adequate response is maintained.

Details of the technology		
Technology	Baricitinib (Olumiant; Lilly)	
Marketing authorisation	Treatment of moderate to severe active RA in adult patients • who have responded inadequately to, or • who are intolerant to one or more DMARDs – used as monotherapy or in combination with MTX	
Mechanism of action	Reversible janus kinase (JAK) inhibitor; selective for JAK1 and JAK2. Baricitinib disrupts cytokine signalling, reducing inflammation, cellular activation and proliferation of key immune cells involved in RA	
Administration	Oral, 4 mg once daily. 2 mg once daily for people aged ≥ 75 years and may be appropriate for people with a history of chronic or recurrent infections. Treatment is continuous (no stopping rule), but dose reduction to 2 mg once daily may be considered for people with sustained control of disease activity	
Acquisition cost	List price per pack: 2 or 4 mg x 28 tab: £805.56 2 or 4 mg x 84 tab: £2,416.68 Annual per patient: £10,501	PAS price per pack: 2 or 4 mg x 28 tab: £[redacted] 2 or 4 mg x 84 tab: £[redacted] Annual per patient: £[redacted]

Confidential

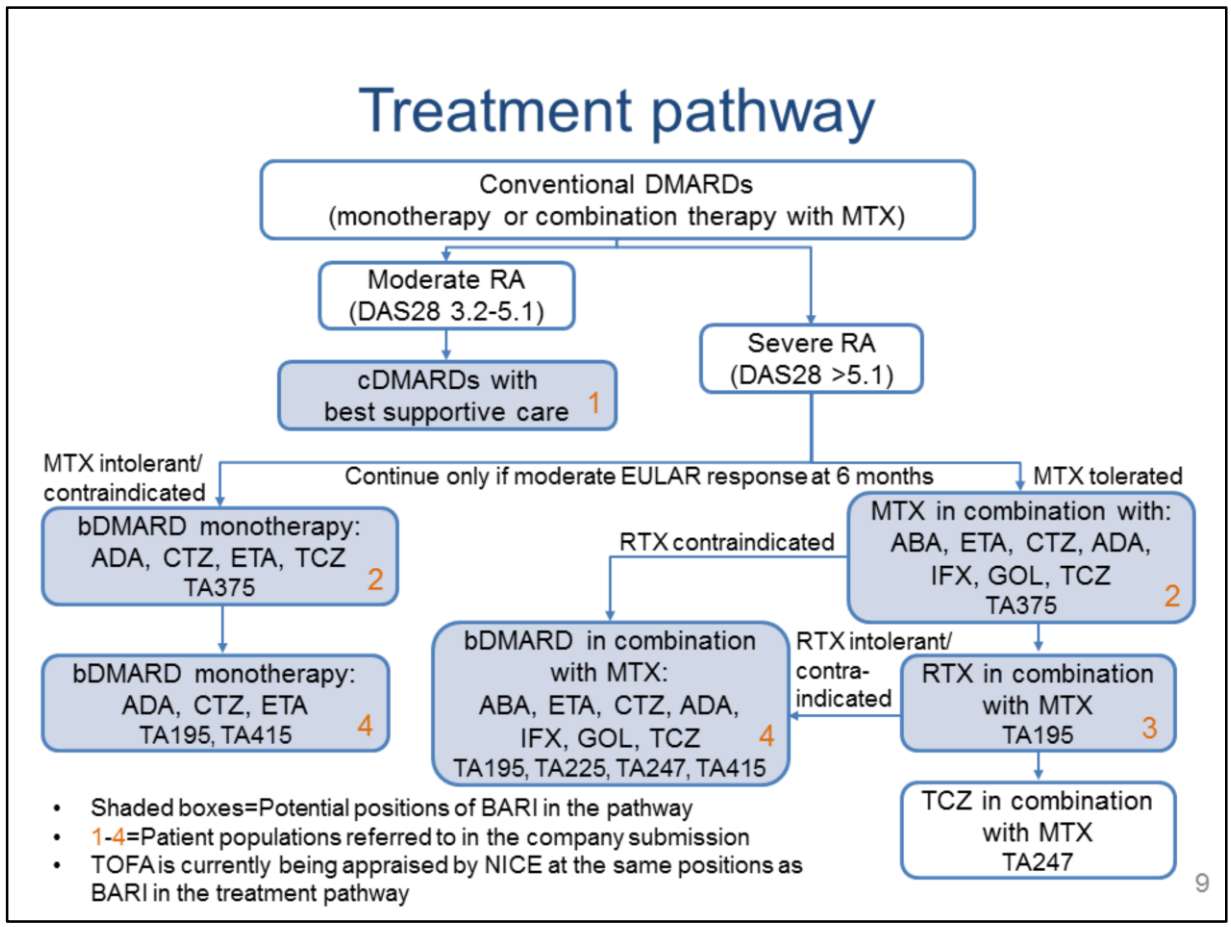
7

Source: Information extracted from table 7 of the company submission

Innovation

- First JAK1/2 inhibitor licenced in Europe
- Oral rather than subcutaneous or i.v. administration
  - Eliminates injection site reactions which can result in discontinuation of bDMARDs
  - Offers treatment for people who don't like needles
- Small molecule rather than an biologic
  - Does not induce the production of anti-drug antibodies seen with TNFi biologics, which cause efficacy to decline over time
- Selective for JAK1/2 with low affinity for JAK3
  - Off-target effects limited

8



Source: Adapted from Figure 3 (page 49) of the company submission

EULAR response criteria

- Good EULAR response=change of >1.2 in DAS28 from baseline AND a DAS28 of ≤3.2 at endpoint
- Moderate EULAR response=change >0.6 and ≤1.2 in DAS28 from baseline AND DAS28 >3.2 and ≤5.1 or DAS2 ≤3.2 at endpoint OR if change of >1.2 in DAS28 from baseline AND DAS28 >3.2 at baseline
- No EULAR response=change ≤0.6 in DAS28 from baseline OR if change of >0.6 and ≤1.2 in DAS28 from baseline AND DAS28 >5.1 at baseline

Potential positions for baricitinib in the treatment pathway

Population	Comparators
1. Moderately active RA that has not responded adequately to therapy with cDMARDs (Moderate RA cDMARD-IR)	• Combination therapy with cDMARDs (including MTX and at least 1 other DMARD, such as sulfasalazine and leflunomide) • cDMARD monotherapy with dose escalation • Best supportive care (only where cDMARDs are not appropriate due to intolerance)
2. Severely active RA that has not responded adequately to therapy with cDMARDs (Severe RA cDMARD-IR)	• ADA, ETN, CTZ or TCZ only (each as monotherapy) • Biologic DMARDs in combination with MTX (ADA, ETN, IFX, CTZ, GOL, TCZ, ABA)
3. Severely active RA that has not responded adequately to therapy with bDMARDs, including at least 1 TNFi (Severe RA bDMARD-IR RTX-eligible)	• RTX in combination with MTX
4. As in 3, when RTX is contraindicated or withdrawn due to adverse events (Severe RA bDMARD-IR RTX-ineligible)	• ADA, ETN and CTZ (each as monotherapy) • ADA, ETN, IFX, ABA, TCZ or CTZ each in combination with MTX

10

Source: Adapted from table 8 of the company submission

Decision problem: Population & intervention

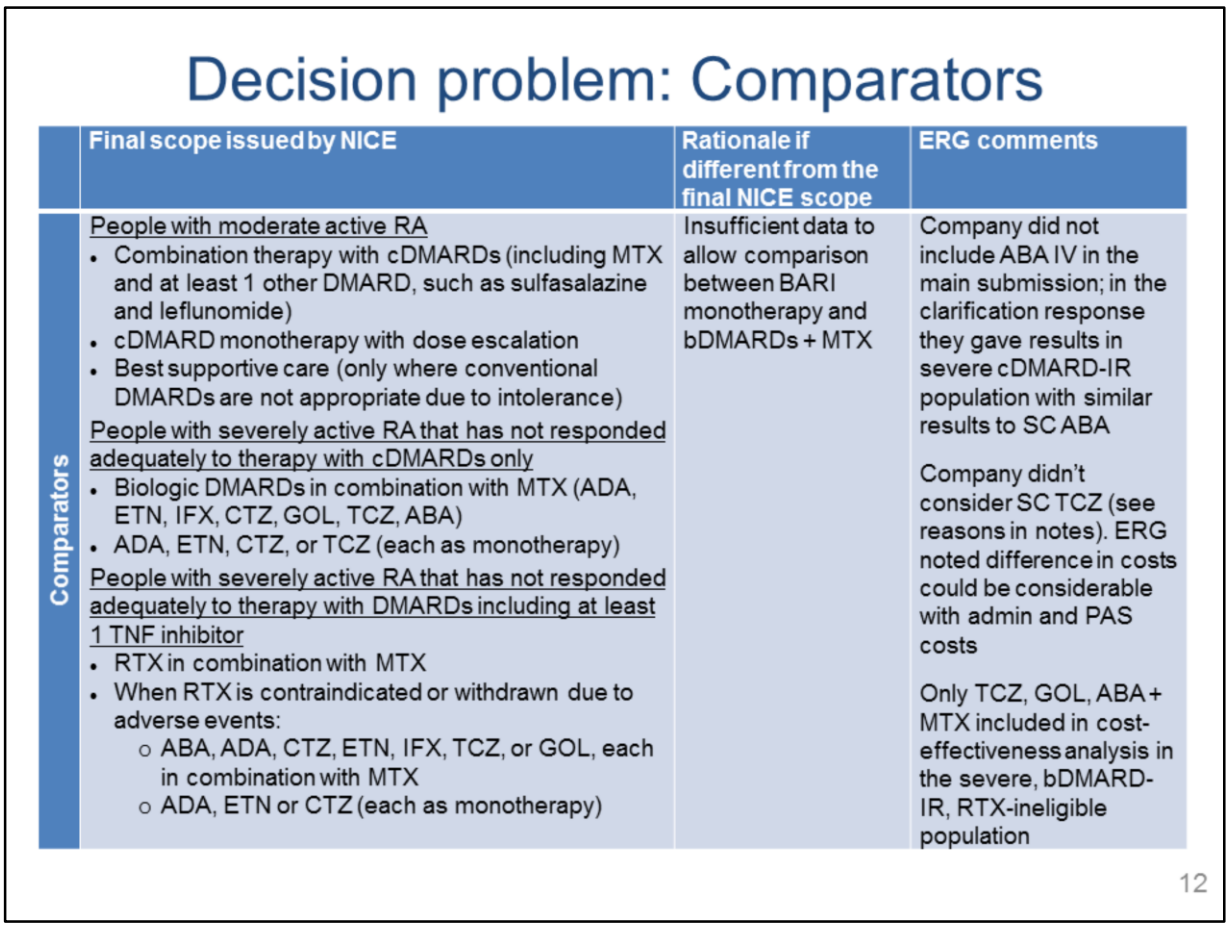
	Final scope issued by NICE	Rationale if different from the final NICE scope	ERG comments
Population	Adults with moderate to severe, active RA whose disease has responded inadequately to, or who are intolerant of 1 or more DMARDs, including conventional or biologic DMARDs	NA	Division of the population is appropriate. However, no analyses are presented for patients who cannot take MTX and for whom BARI would be used as monotherapy
Intervention	Baricitinib monotherapy or in combination with MTX	Clinical data is also provided for BARI 2 mg and a scenario analysis in which the dose is tapered from 4 mg to 2 mg	-

11

Source: Adapted from table 1 of the company submission and section 3 of the ERG report

The ERG considered the company's description of the underlying health problem to be appropriate, mostly up-to-date and relevant to the decision problem set out in the NICE scope.

The company submission generally adhered to the scope. Exceptions related to the exclusion of the SC formulation of tocilizumab and the IV formulation of abatacept as comparators.



Decision problem: Comparators

	Final scope issued by NICE	Rationale if different from the final NICE scope	ERG comments
Comparators	People with moderate active RA • Combination therapy with cDMARDs (including MTX and at least 1 other DMARD, such as sulfasalazine and leflunomide) • cDMARD monotherapy with dose escalation • Best supportive care (only where conventional DMARDs are not appropriate due to intolerance) People with severely active RA that has not responded adequately to therapy with cDMARDs only • Biologic DMARDs in combination with MTX (ADA, ETN, IFX, CTZ, GOL, TCZ, ABA) • ADA, ETN, CTZ, or TCZ (each as monotherapy) People with severely active RA that has not responded adequately to therapy with DMARDs including at least 1 TNF inhibitor • RTX in combination with MTX • When RTX is contraindicated or withdrawn due to adverse events: o ABA, ADA, CTZ, ETN, IFX, TCZ, or GOL, each in combination with MTX o ADA, ETN or CTZ (each as monotherapy)	Insufficient data to allow comparison between BARI monotherapy and bDMARDs + MTX	Company did not include ABA IV in the main submission; in the clarification response they gave results in severe cDMARD-IR population with similar results to SC ABA Company didn't consider SC TCZ (see reasons in notes). ERG noted difference in costs could be considerable with admin and PAS costs Only TCZ, GOL, ABA + MTX included in cost-effectiveness analysis in the severe, bDMARD-IR, RTX-ineligible population

12

Source: Adapted from table 1 of the company submission and section 3 of the ERG report

Company justification/ERG comments

- Insufficient data: In TA375, the committee agreed that the minority of (cDMARD-IR) patients with severe active rheumatoid arthritis who could not tolerate MTX should not be treated differently from other people with severe disease. The committee concluded that biologic DMARDs should be recommended as a cost-effective use of NHS resources when used as monotherapy for severe active disease previously treated with DMARDs, where the marketing authorisation of the bDMARD allows for this recommendation to be made.

ABA IV was included in the NMA, but was excluded from the analyses. The company stated it was 'a pragmatic decision … to limit the number of sequences … where … different administration routes was unlikely to be informative.' In response to a clarification request by the ERG, the company presented the results of ABA IV only for the cDMARD-IR population with severe RA, which led to similar results compared with ABA SC (£[redacted] versus £[redacted] and [redacted] versus [redacted] QALYs respectively).

- The company stated that it excluded TCZ SC because: 1. the available evidence for TCZ SC was limited; 2. it provided a lower efficacy estimate than for TCZ IV; and 3. the cost difference between the 2 formulations was relatively small. The company stated 'including 'IV tocilizumab only was felt to be a reasonable choice, with it likely to be representative of the costs and outcomes associated with the SC version.' The ERG noted that the difference in costs might be considerable taking into account the administration costs and the confidential

PAS. ABA IV was included in the NMA, but was excluded from the analyses. In response to a clarification request by the ERG, the company presented the results of ABA IV only for the cDMARD-IR population with severe RA, which led to similar results compared with ABA SC (£[redacted] versus £[redacted] and [redacted] versus [redacted] QALYs respectively).

- The ERG broadly believes the company has evaluated the correct comparators but they make 2 comments:

1. The company have not explicitly modelled BARI used as a monotherapy. The rationale stated by the company for this is 'the paucity of efficacy data in the baricitinib clinical trial programme for patients receiving baricitinib monotherapy, which would be insufficient to form a reliable estimate of efficacy in the modelled populations for baricitinib monotherapy. The ERG note in TA375, the committee agreed that the minority of (cDMARD-IR) patients with severely active rheumatoid arthritis who could not tolerate methotrexate should not be treated differently from other people with severe disease, as far as possible. The Committee concluded that biologic DMARDs should be recommended as a cost-effective use of NHS resources when used as monotherapy for severely active disease previously treated with DMARDs, where the marketing authorisation of the bDMARD allows for this recommendation to be made. The ERG consider that a similar rationale will be applied to baricitinib monotherapy. The lack of data for BARI when used as a monotherapy will increase the uncertainty in its ICER when compared with interventions with a larger evidence base. Clinical advice to the ERG suggests that there is no clearly defined relationship between the efficacy of a bDMARD in combination with MTX and in the bDMARD used as monotherapy. However, data from RA-BEGIN showed that the addition of MTX to BARI 4 mg did not produce a marked improvement over BARI monotherapy in a MTX-naïve population. This provides supportive evidence regarding the efficacy of BARI monotherapy.

2. In all comparisons, the biosimilar prices for IFX and ETN have been used rather than the prices of the original compounds. ABA (both IV and SC), and TCZ (both IV and SC) are subject to commercial-in-confidence (CIC) patient access schemes (PAS). Given this, the company has solely used list prices for these drugs, with the ERG incorporating the discounts for these interventions in a confidential appendix.

Decision problem: Outcomes and economic analyses

	Final scope issued by NICE		Rationale if different from the final NICE scope	ERG comment
Outcomes	The outcome measures to be considered include:		NA	-
	• disease activity • physical function • joint damage • pain • mortality	• fatigue • radiological progression • extra-articular manifestations • adverse effects of treatment • health-related quality of life		
Economic analysis	• Cost-effectiveness should be expressed in terms of incremental cost per QALY • Time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared • Costs will be considered from an NHS and Personal Social Services perspective • Patient access schemes for the intervention or comparator technologies will be taken into account • Availability and cost of biosimilar products should be taken into account		Costs were considered from an NHS perspective only, consistent with the AG's model in TA375	-

13

Source: Adapted from table 1 of the company submission and section 3 of the ERG report

Commentary from Prof. Peter Taylor (clinical expert):

The categorical ACR20 response metric (the primary endpoint) is widely used in RA clinical trials and will be familiar to UK rheumatologists, but it is not used in routine clinical practice. Clinical assessment at around 3 months after initiating a targeted therapy (1st assessment point in the trials) reflects clinical practice. Assessment of response in a routine setting is usually based on a composite score of disease activity of which one of the variants of DAS28 (based on either ESR or CRP as an acute phase response measure, and with or without inclusion of the patient global health assessment visual analogue scale). In all 3 pivotal phase III trials, DAS28CRP change at week 12 was included as a secondary endpoint, confirming clinical efficacy with a metric familiar to rheumatologists in routine clinical practice.

Other outcome measures important to physicians include those indicative of long term inhibition of structural damage to joints and preservation of function. RA-BEAM (MTX-IR, bDMARD-naïve) and RA-BUILD (cDMARD-IR, bDMARD-naïve RA) also demonstrated significant inhibition of structural damage to joints at 6 months. This is important information for rheumatologists but formal assessment of radiographic structural damage inhibition is not routinely measured at 6 months of treatment intervention.

Prof. Taylor also the noted considerable body of data that demonstrates statistically significant and clinically meaningful improvements in pain and the length and severity of early morning joint

stiffness, patient reported outcomes that are important to people with RA.

Decision problem: Subgroups and special considerations

	Final scope issued by NICE	Rationale if different from the final NICE scope	ERG comment
Subgroups to be considered	• If the evidence allows: people with moderate disease activity (DAS28 between 3.2 and 5.1) and severely active disease (DAS28 greater than 5.1).	The primary endpoint (ACR20 response at Week 12) of each of the 3 trials (RA-BEAM, RA-BUILD and RA BEACON) are presented for : • Moderate disease activity • Severe disease activity In the cDMARD-IR population, the economic analysis presents results separately for moderate patients and severe patients. In the bDMARD-refractory population, the economic analysis presents results for severe patients.	-
Special consid.	NA	NA There are no equality issues arising in relation to this technology.	-

14

Source: Adapted from table 1 of the company submission and section 3 of the ERG report

For RA-BEACON, results are also presented for:

- <3 previous bDMARDs used and ≥3 previous bDMARDs used

Submissions from patient/carer organisation (National Rheumatoid Arthritis Society) and patient expert

- Being diagnosed RA can be extremely distressing as it is life-changing, affects physical and emotional wellbeing, and impacts the whole family
- People can be diagnosed at any age over 16 years so it can have a major impact on life plans and expectations, dreams and aspirations affecting:
  - Personal confidence and future relationships in younger people
  - Working life and job security, and caring for young children
  - Retirement plans
- People simply want their life back. They want a reduction in pain and fatigue, to prevent permanent disability, and to maintain independence and ability to work
- The side effects of some drugs can be debilitating. Even with all the new treatments available, the heterogeneity of RA means that there remains unmet need
- Baricitinib offers a new class of therapy adding to the therapeutic options, especially for people with refractory RA who have been treated with all the available biologics
- Because it is a tablet, it eliminates the need for self-injection or hospital visits for infusions

15

Submissions from clinical experts

- Treatment aimed at controlling inflammatory disease activity as early and optimally as possible. cDMARDs are insufficient for a significant proportion of people. Many patients don't respond adequately to their first biologic; there are few tools available to predict response, or to help decide which biologic is to be used first
- Baricitinib is novel; there are no other JAK inhibitors available in the UK. The EULAR 2016 update recommends that JAK inhibitors are considered as an alternative to bDMARDs in poor prognosis patients after failure of cDMARDs. Suitable for people who don't like needles
- Trials show baricitinib (with and without MTX) is effective after DMARD failure and after anti-TNF failure, is superior to MTX, and gives similar results to currently available biologics. Fast clinical response. Use as a monotherapy is an advantage as many people don't tolerate MTX, which leads to poor adherence. In a direct head to head RCT baricitinib is superior to ADA in some (but not all) endpoints. There is no evidence to suggest greater or less effects in patient subgroups (e.g. by disease severity or antibody status). There are no new safety signals, and the overall benefit:risk profile is favourable and broadly comparable to that seen with bDMARDs
- The people in the trials broadly reflect those in the UK and are comparable to those in trials for other NICE approved RA treatments. Trial outcomes are appropriate and include measures relevant to routine clinical practice
- Patients will receive baricitinib under supervision of consultant rheumatologist, but it is expected to carry a lower administrative burden compared with subcut or i.v. therapies

16

No extra tests, staff education/training or equipment of facilities would be required

Clinical effectiveness systematic review and network meta-analysis

- Company systematic review identified 4 RCTs and 1 long-term safety and tolerability study
  - RA–BEAM RCT: MTX-treated, bDMARD-naïve *vs* placebo *vs* ADA
  - RA-BUILD RCT: cDMARD-IR, bDMARD-naïve *vs* placebo
  - RA-BEACON RCT: bDMARD-IR *vs* placebo
  - RA-BEGIN RCT: DMARD-naïve (unlicensed) *vs* MTX
  - RA-BEYOND long-term study: Included patients from RA-BEAM, RA-BUILD, RA-BEACON, a phase III study of BARI in MTX-naïve patients (RA-BEGIN; JADZ)) and a phase II study of BARI (JADA)
- Network meta-analysis assessed the relative efficacy of BARI in the cDMARD-IR and bDMARD-IR populations

17

Effectiveness results from the RA-BEYOND long term safety and tolerability study are shown in Table 20 of the ERG report (reproduced from Table 58 of the company submission).

The company notes that:

The cDMARD-IR NMA was performed from a global perspective, and, as a result, RTX, which is not normally considered for the treatment of the cDMARD-IR population in the UK, was included in the analysis. However, the RTX trials included in the network only provide information on RTX and MTX and therefore only increase the amount of evidence available in the network for the estimation of treatment effect for the RTX and MTX nodes. Therefore, the inclusion of the RTX studies is not expected to impact the validity of the treatment effect estimates for baricitinib versus its relevant comparators.

Treatments included in the network meta-analysis

cDMARD-IR population	bDMARD-IR population
• BARI (2 mg or 4 mg QD) + cDMARD • PBO • cDMARD • SSZ • SSZ + HCQ + cDMARD • MTX • ADA, ETN, RTX or TCZ monotherapy • ABA, ADA, CTZ, ETN, GOL, IFX, RTX, TCZ or TOFA + cDMARD • ETN + SSZ	• BARI (2 mg or 4 mg QD) + cDMARD • cDMARD • ABA , GOL or TCZ + cDMARD • RTX or TCZ + MTX

18

Source: Adapted from table 68 of the company submission

The company notes that:

The cDMARD-IR NMA was performed from a global perspective, and, as a result, RTX, which is not normally considered for the treatment of the cDMARD-IR population in the UK, was included in the analysis. However, the RTX trials included in the network only provide information on RTX and MTX and therefore only increase the amount of evidence available in the network for the estimation of treatment effect for the RTX and MTX nodes. Therefore, the inclusion of the RTX studies is not expected to impact the validity of the treatment effect estimates for baricitinib versus its relevant comparators.

Study characteristics

Trial name	Population and number enrolled	Intervention	Comparators	Primary outcome
RA-BEAM (JADV)	MTX-IR, bDMARD-naïve adult patients with moderate to severe RA 1307 randomised (1305 at least 1 dose, included in mITT); [redacted]	• BARI 4 mg, oral, QD (with background MTX)	• ADA 40 mg, SC injection, Q2W (with background MTX) • Placebo (with background MTX)	% of patients achieving an ACR20 response at week 12
RA-BUILD (JADX)	cDMARD-IR, bDMARD-naïve adult patients with moderate to severe active RA 684 randomised; [redacted]	• BARI 2 mg, oral, QD • Baricitinib 4 mg, oral, QD Patients on ≥1 cDMARDs (with or without MTX) continued to take background therapy during study	• Placebo Patients on ≥1 cDMARDs (with or without MTX) continued to take background therapy during study	% of patients achieving an ACR20 response at week 12
RA-BEACON (JADW)	bDMARD-IR adult patients with moderate to severe active RA 527 randomised; [redacted]	• BARI 2 mg, oral, QD (with background cDMARDs) • BARI 4 mg, oral, QD (with background cDMARDs)	• Placebo (with background cDMARDs)	% of patients achieving an ACR20 response at week 12
RA-BEGIN (JADZ)	DMARD-naïve adult patients with moderate to severe RA (outside MA) 588 randomised; 15 UK patients	• BARI 4 mg, oral, QD BARI 4 mg, oral, QD (with MTX)	• MTX oral, QW	% of patients achieving an ACR20 response at week 24
RA-BEYOND	Patients with moderate to severe RA who completed Phase 2b study JADA or Phase 3 studies (RA-BEAM, BUILD, BEACON, BEGIN) [redacted]	• BARI 2 mg, oral, QD • BARI 4 mg, oral, QD	N/A	Long-term safety and tolerability

- Baseline patient characteristics within trials were balanced across trial arms

 19

See section 4.2 (page 34) of the ERG report for more details

Summary results: ACR20 and EULAR at 12 weeks

MTX-IR, bDMARD-naïve, moderate to severe RA (RA-BEAM)

	PBO (n=488)	BARI 4 mg + cDMARD (n=487)	ADA (n=330)
ACR20 (%)	40	70***+	61***
ACR20 odds ratio (95% CI)	-	BARI *vs* PBO 3.6 (2.7 to 4.7) p=0.001	BARI *vs* ADA 1.5 (1.1 to 2.0) p=0.014
EULAR (good + moderate) response rate (%)	[redacted]	[redacted]	[redacted]
EULAR (good) response rate (%)	[redacted]	[redacted]	[redacted]
EULAR good and moderate response Odds ratio (95% CI)	-	BARI *vs* PBO [redacted]	BARI *vs* ADA [redacted]
EULAR good response Odds ratio (95% CI)	-	[redacted]	[redacted]

***p≤0.001 versus placebo, and +p≤0.05, ++p≤0.01 versus adalimumab using logistic regression, without control for multiple comparisons

 20

Source: Adapted from table 8 and table 11 of the ERG report

- BARI 4 mg also showed a significant improvement in:
  - ACR50 and ACR70 at 12 weeks compared with PBO and ADA
  - ACD20, ACR50 and ACR70 at 24 weeks compared with PBO
  - ACR20 and ACR70 at 24 weeks compared with ADA
  - ACR20 and ACR50 at 52 weeks compared with ADA
  - EULAR good and moderate response rate and good response rate at 24 weeks compared to PBO
  - EULAR good and moderate response rate at 24 weeks compared to ADA
  - EULAR good and moderate response rate and good response rate at 52 weeks compared to ADA
- BARI 4 mg also showed a significant (and where tested, long term) improvement over PBO in:
  - DAS28-hsCRP, HAQ-DI, mTSS (24 and 52 week follow-up only), and SDAI, CDAI, FACIT-F and MJS parameters
- BARI 4 mg had similar outcomes to ADA for DAS28-hsCRP, mTSS, and SDAI, CDAI, FACIT-F and MJS parameters although BARI 4 mg significantly improved SDAI low disease activity (≤11.0) response rate and FACIT-F change from baseline least squares mean at 52 weeks

compared with ADA

Analysis of ACR20 according to baseline DAS28-hsCRP (i.e. in the moderate and severe subgroups) showed no significant interaction with treatment group. ACR50 and EULAR were not assessed in this analysis.

Summary results: ACR20 and EULAR at 12 weeks

cDMARD-IR, bDMARD-naïve, moderate to severe RA (RA-BUILD)

	PBO (n=228)	BARI 2 mg + cDMARD (n=229)	BARI 4 mg + cDMARD (n=227)
ACR20 (%)	39.5	65.9***	61.7***
ACR20 odds ratio (95% CI)	-	BARI 2 mg *vs* PBO 3.0 (2.0 to 4.4) p=0.001	BARI 4 mg *vs* PBO 2.5 (1.7 to 3.7) p=0.001
EULAR (good + moderate) response rate (%)	53.5	79.0***	79.3***
EULAR (good) response rate (%)	15.4	34.1***	38.3***
EULAR good and moderate response Odds ratio (95% CI)	-	BARI 2 mg *vs* PBO 3.3 (2.2 to 5.0) p=0.001	BARI 4 mg *vs* PBO 3.5 (2.3 to 5.4) p=0.001
EULAR good response Odds ratio (95% CI)	-	2.9 (1.8 to 4.6) p=0.001	3.6 (2.3 to 5.7) p=0.001

***p≤0.001 versus placebo

21

Source: Adapted from table 9 and table 12 of the ERG report

- BARI 2 mg and 4 mg also showed a significant improvement in:
  - ACR50 and ACR70 at 12 weeks compared with PBO
  - ACD20, ACR50 and ACR70 at 24 weeks compared with PBO
  - EULAR good and moderate response rate and good response rate at 24 weeks compared to PBO
- BARI 2 mg and BARI 4 mg also showed a significant (and where tested, long term) improvement over PBO in:
  - DAS28-hsCRP, HAQ-DI, mTSS (24 week follow-up only), and SDAI, CDAI, FACIT-F and MJS parameters

Analysis of ACR20 according to baseline DAS28-hsCRP (i.e. in the moderate and severe subgroups) showed no significant interaction with treatment group. ACR50 and EULAR were not assessed in this analysis.

Summary results: ACR20 and EULAR at 12 weeks

bDMARD-IR moderate to severe RA (RA-BEACON)

	PBO (n=176)	BARI 2 mg + cDMARD (n=174)	BARI 4 mg + cDMARD (n=177)
ACR20 (%)	27.3	48.9***	55.4***
ACR20 odds ratio (95% CI)	-	BARI 2 mg *vs* PBO 2.7 (1.7 to 4.2) p=0.001	BARI 4 mg *vs* PBO 3.4 (2.2 to 5.4) p=0.001
EULAR (good + moderate) response rate (%)	42.6	66.1***	72.3***
EULAR (good) response rate (%)	8.5	24.1***	29.9***
EULAR good and moderate response Odds ratio (95% CI)	-	BARI 2 mg *vs* PBO 2.7 (1.8 to 4.2) p=0.001	BARI 4 mg *vs* PBO 3.6 (2.3 to 5.7) p=0.001
EULAR good response Odds ratio (95% CI)	-	3.6 (1.9, 6.8) p=0.001	4.8 (2.6, 9.0) p=0.001
***p≤0.001 versus placebo			

22

Source: Adapted from table 10 and table 13 of the ERG report

- BARI 2 mg and 4 mg also showed a significant improvement in:
  - ACR50 and ACR70 at 12 weeks compared with PBO
  - ACD20, ACR50 and ACR70 at 24 weeks compared with PBO
  - EULAR good and moderate response rate and good response rate at 24 weeks compared to PBO
- BARI 2 mg and 4 mg also showed a significant (and where tested, long term) improvement over PBO in:
  - DAS28-hsCRP, HAQ-DI, FACIT-F and MJS parameters, and some SDAI and CDAI parameters

The company did not carry out subgroup analysis on data from RA-BEAM because of the small number of responders in the PBO arm for patients with moderate disease activity at baseline.

Clinical effectiveness results: EULAR RA-BEAM, RA-BUILD, RA-BEACON

- Baricitinib 4 mg + cDMARD
  - 12 weeks follow-up: BARI + cDMARD superior to PBO (p≤0.001) in all 3 trials
    - RA-BEAM [redacted]% vs [redacted]%
    - RA-BUILD 79.0% vs 53.5%
    - RA-BEACON 66.1% vs 42.6%
  - 12 weeks follow-up: BARI + cDMARD superior to ADA for good and moderate response (p=0.002), and good response (p=0.010)
  - 24 weeks follow-up: BARI + cDMARD superior to PBO for the bDMARD-naïve studies (p<0.001; RA-BEAM and RA-BUILD) and the bDMARD-experienced population (p<0.05; RA-BEACON)
- Baricitinib 2 mg + cDMARD
  - 12 weeks follow up: BARI + cDMARD superior to PBO (p<0.001) for both the bDMARD-naïve (RA-BUILD) and bDMARD-experienced populations (RA-BEACON)
  - 24 weeks follow up, BARI + cDMARD superior to PBO in the bDMARD-naïve (p<0.001, RA-BUILD) and bDMARD-experienced (p<0.05; RA-BEACON) populations

23

Includes EULAR good or moderate response

Clinical effectiveness results: Other efficacy outcomes

- Baricitinib 4 mg + cDMARD
  - 12 weeks follow-up: BARI + cDMARD superior to PBO across the 3 RCTs for DAS28-CRP, HAQ-DI, SDAI low disease activity (LDA), CDAI LDA and MJS duration.
  - RA-BEAM planned a statistical comparison of BARI 4 mg + cDMARD and ADA + cDMARD at week 12 for DAS28-CRP, and this was found to significantly favour BARI 4 mg + cDMARD ($p \leq 0.01$)
  - 24 weeks follow-up: BARI + cDMARD superior to PBO on several measures (all 3 RCTs)
- Baricitinib 2 mg + cDMARD
  - BARI + cDMARD superior to PBO at 12 weeks and 24 weeks on several measures (RA-BUILD and RA-BEACON)

24

Clinical effectiveness results: Health-related quality of life

- BARI 4 mg + cDMARD superior to PBO for EQ-5D-5L ($p \leq 0.001$) at 12 weeks and 24 weeks follow-up in all 3 RCTs
- BARI 2 mg + cDMARD superior to PBO ($p \leq 0.01$) at 12 and 24 weeks follow-up (RA-BUILD and RA-BEACON)

25

See tables 17-19 of the ERG report

Summary of adverse events from weeks 0 to 24 (RA-BEAM, -BUILD, -BEACON)

Trial	RA-BEAM			RA-BUILD			RA-BEACON		
Treatment (n)	PBO (n=488)	BARI 4 mg QD + cDMARD (n=487)	ADA + cDMARD (n=330)	PBO (n=228)	BARI 2 mg QD + cDMARD (n=229)	BARI 4 mg QD + cDMARD (n=227)	PBO (n=176)	BARI 2 mg QD + cDMARD (n=174)	BARI 4 mg QD + cDMARD (n=177)
Treatment exposure, patient-years (total/group)	197.7	215.0	141.9	89.8	97.7	96.4	65.8	69.9	73.3
Overall AE, n (%) [EAIR]	[redacted]	[redacted]	[redacted]	161 (71)	154 (67)	162 (71)	112 (64)	123 (71)	137 (77)
Serious AE, n (%) [EAIR]	[redacted]	[redacted]	[redacted]	11 (5)	6 (3)	12 (5)	13 (7)	7 (4)	18 (10)
Withdrawal because of AE, n (%) [EAIR]	[redacted]	[redacted]	[redacted]	10 (4)	10 (4)	12 (5)	7 (4)	7 (4)	11 (6)
Temporary interruption due to AE, n [EAIR]	[redacted]	[redacted]	[redacted]	NR	NR	NR	NR	NR	NR
Death, n [EAIR]	[redacted]	[redacted]	[redacted]	2	0	0	0	0	1
Infection, n (%)	134 (27)	176 (36)	110 (33)	79 (35)	70 (31)	96 (42)	55 (31)	76 (44)	70 (40)
Serious infection, n (%)	7 (1)	5 (1)	2 (<1)	4 (2)	2 (<1)	4 (2)	5 (3)	4 (2)	6 (3)
Cancer, n (%)	3 (<1)	2 (<1)	0	0	0	1 (<1)	0	0	2 (1)
MACE	0	1 (<1)	0	2 (<1)	0	0	0	0	2 (1)

Confidential

26

Source: Table 24 of the ERG report, adapted from table 95 of the company submission and Taylor *et al*. 2017, Dougados *et al*. 2016 and Genovese *et al*. 2016.

- During preparation

  - The lead team were made aware that the FDA delayed approval for baricitinib citing that there was a need for more information to be collected on safety and the most appropriate dose

- Baricitinib is approved by EMA

• BARI was well-tolerated. A small proportion of patients discontinued from the baricitinib studies because of adverse events

• The most commonly reported adverse drug reactions in ≥2% of patients treated with BARI monotherapy or in combination with cDMARDs included increased LDL cholesterol, upper respiratory tract infections and nausea. However, the majority of all ADRs were mild to moderate in severity.

• The proportion of patients with serious adverse events (SAEs) (including serious infections) was similar across treatment groups in the phase III studies and integrated placebo-controlled analysis sets, except for RA-BEAM, where a higher proportion of SAEs were reported with placebo and BARI versus ADA.

• BARI was associated with a higher incidence of SAEs compared with ADA up to 52 weeks in

RA-BEAM, but the AE profiles were similar across clinically significant categories of risk including major adverse cardiovascular events (MACE), malignancies, hypercholesterolemia, serious infections and herpes zoster.

• Despite a higher risk of cardiovascular disease, infection, and malignancy in the RA population, treatment with BARI did not result in increased risk of malignancy, serious or opportunistic infections, or MACE.

• Non-serious herpes simplex and herpes zoster infections were more frequent in patients treated with BARI than placebo, but were not significantly higher than those seen with MTX or ADA. The majority of herpes zoster cases were mild to moderate in severity and complicated cases were uncommon.

• Increases in LDL cholesterol with BARI were concomitant with increases in HDL-C so the mean HDL/LDL ratio was unchanged. In addition, there was a significant decrease in the amount of small and very small LDL particles in RA-BEAM, which are considered the most atherogenic. Few MACE events were seen with BARI and there was no relationship between MACE and increased LDL.

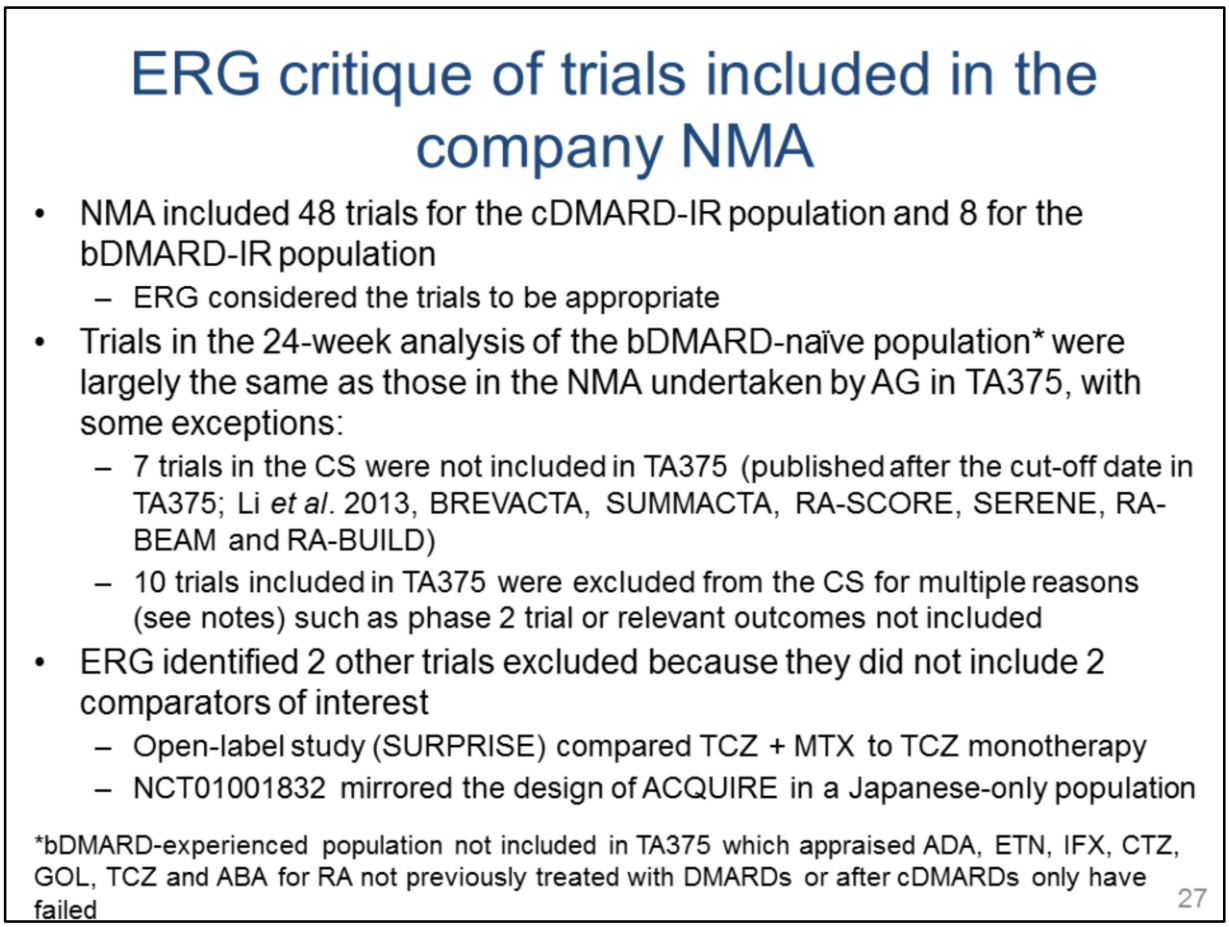
• Treatment with BARI resulted in changes in haematology and clinical chemistry markers. Some of the changes were greater magnitude than seen with the active comparators. The company asserts that these are likely to be related to the pharmacology of JAK inhibition (such as increases in lipids [including total cholesterol, low-density lipoprotein cholesterol and high-density lipoprotein cholesterol] and creatine phosphokinase).

In RA-BEAM, in 52 weeks of the RCT there were five deaths: 1 PBO, 1 PBO switched to BARI, 2 BARI, and 2 ADA. From week 0 to week 52, SAEs were experienced by 8% of BARI-treated, and 4% of ADA-treated patients.

The most common adverse events listed in the SmPC for baricitinib are increased LDL cholesterol (33.6%), upper respiratory tract infections (14.7%) and nausea (2.8%)

Herpes simplex; herpes zoster; acne; increased creatine phosphokinase; increased triglycerides; increased liver function tests (AST and ALT); neutropenia and thrombocytosis are also considered to be adverse events by the EMA

Clinical advice given to the ERG suggests ensuring arrangements are in place to identify safety signals



Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

Ten trials included in TA375 but excluded from the company submission:

- AUGUSTII, IIBCREATE, NCT00254293, and Kremer 2012 were excluded from the company submission as they were Phase II trials
- ACQUIRE. Excluded because the "study compared SC vs IV abatacept. The search strategy specified that studies were to include two different comparators of interest to be included" (see clarification question A4). This appears to be inconsistent with the inclusion of SUMMACTA, which compares IV and SC TCZ.
- ATTRACT. The company excluded this trial as it only provided data relating to ACR20. These data can be used within the NMA and should not be discarded.
- CERTAIN. Within the clarification response process (clarification question A4), the company stated that this trial was excluded as it included patients with low to moderate disease activity. The ERG considered baseline DAS28 in the treatment arms of 4.47 and 4.53 to be moderate, to severe, disease activity.
- SAMURAI. The company stated that 12 or 24-week data were not identified. However, data at 24 weeks from Nishimoto *et al*. 2007 were used in TA375.
- Swefot. The company stated that this trial focussed on patients with "early rheumatoid arthritis (less than a year since diagnosis) and was therefore excluded. Additionally, the infliximab arm, allowed an increase in dose frequency (to every 6 weeks) or a switch to ETN and it does not appear that reported results take this into account" (clarification question A4).

- TACIT – The company excluded this study as the intervention arm combined bDMARDs (Appendix 4 company submission).

cDMARD-IR median EULAR response at 24 weeks follow-up: Company NMA results

- The median EULAR moderate + good response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted], respectively. The median EULAR good response rates were [redacted] and [redacted], respectively.

Confidential 28

Source: Adapted from table 77 of the company submission

bDMARD-IR median EULAR response at 24 weeks follow-up: Company NMA results

- The median EULAR moderate or good response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted], respectively. The median EULAR good response rates were [redacted] and [redacted], respectively.

Confidential 29

Source: Adapted from table 80 of the company submission

ERG critique of the indirect comparison and/or multiple treatment comparison

- NMAs were performed separately for the cDMARD-IR and bDMARD-IR populations
- The ERG identified several issues with approaches taken by the company including:
  - The conversion of ACR data to EULAR data before synthesis
  - The use of simultaneous models for baseline and treatment effects
  - Not assessing goodness-of fit
  - Using a random effects model for the cDMARD-IR population and fixed effects model for the bDMARD-IR population
- The ERG noted that inappropriate pooling of the control arms means that all results should be treated with caution

30

See section 4.4 of the ERG report for further details

Additional work carried out by the ERG

- ACR and EULAR outcomes at week 24 re-analysed for both the cDMARD-IR and bDMARD-IR populations. All cDMARDs were assumed to have equivalent efficacy and were grouped together
- EULAR data from van de Putte *et al*. (2004) amended so the moderate EULAR responders did not include good EULAR responders
- The ERG's ACR NMA used the same studies included in the company submission. The ERG's EULAR NMA only included studies that reported EULAR outcomes
- The ERG used the same model for the relative treatment as the NICE DSU TSD which did not assume a random effects model for the baseline for each study
  - The baseline and relative treatment effect models were run separately to ensure information in the baseline model did not propagate to the relative treatment effect model
- Random effects model used for both ACR and EULAR outcomes in both the cDMARD-IR and bDMARD-IR populations

31

For EULAR outcomes in the bDMARD-IR population, and ACR outcome in the cDMARD-IR and bDMARD-IR population, the ERG computed the number of responses in each category using the data provided in percentages reported in the company submission (appendix 14) and in response to clarification request (question A6).

cDMARD-IR median EULAR response at 24 weeks follow-up: ERG NMA results

- BARI 4 mg + cDMARD associated with statistically significant beneficial treatment effects relative to PBO and cDMARD. No statistically significant differences were found versus any other comparator, with the exception of TCZ + cDMARD, which was associated with statistically beneficial treatment effects relative to BARI 4 mg + cDMARD

Confidential 32

Source: Figure 3 of the ERG report

ERG NMA result for ACR: BARI 4 mg + cDMARD associated with statistically significant beneficial treatment effects relative to PBO, cDMARD and ADA monotherapy. No statistically significant differences were found versus any other comparator, with the exception of CTZ + cDMARD, which was associated with a statistically significant beneficial treatment effect relative to BARI 4 mg + cDMARD

bDMARD-IR median EULAR response at 24 weeks follow-up: ERG NMA results

- BARI 4 mg + cDMARD associated with statistically significant beneficial treatment effects relative to cDMARD. No statistically significant differences were found versus RTX 1000 mg + cDMARDs with the effect favouring RTX 1000 mg + cDMARDs, which was the only other comparator in the network

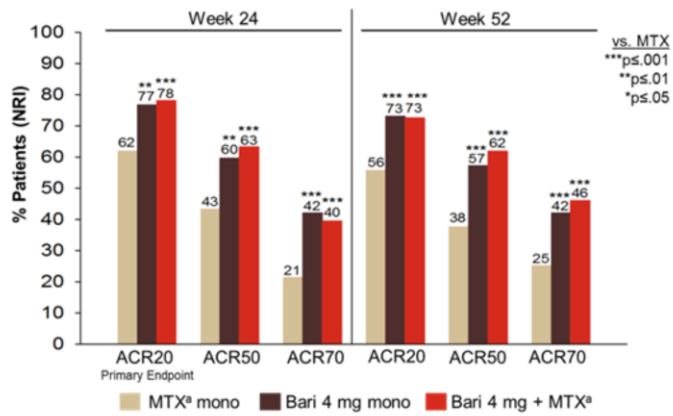
 33

Source: Figure 4 of the ERG report erratum

ERG NMA result for ACR: BARI 4 mg + cDMARD was associated with a statistically significant beneficial treatment effect relative to cDMARD. No statistically significant differences were found versus any other comparator.

BARI monotherapy

- The company did not make a case for BARI monotherapy
  - The ERG point out that data from RA-BEGIN showed that the addition of MTX to BARI 4 mg did not produce a marked improvement over BARI monotherapy



34

Source: Figure 4 of the company submission appendices

Why BARI mon was not included: The only available evidence of BARI monotherapy is in MTX-naïve patients and therefore considerable uncertainty exists about the efficacy of BARI monotherapy in cDMARD-IR and bDMARD-IR patients. See Figure 2 in ERG report.

See section 5.3 of the ERG report for further details

Key issues: Clinical effectiveness

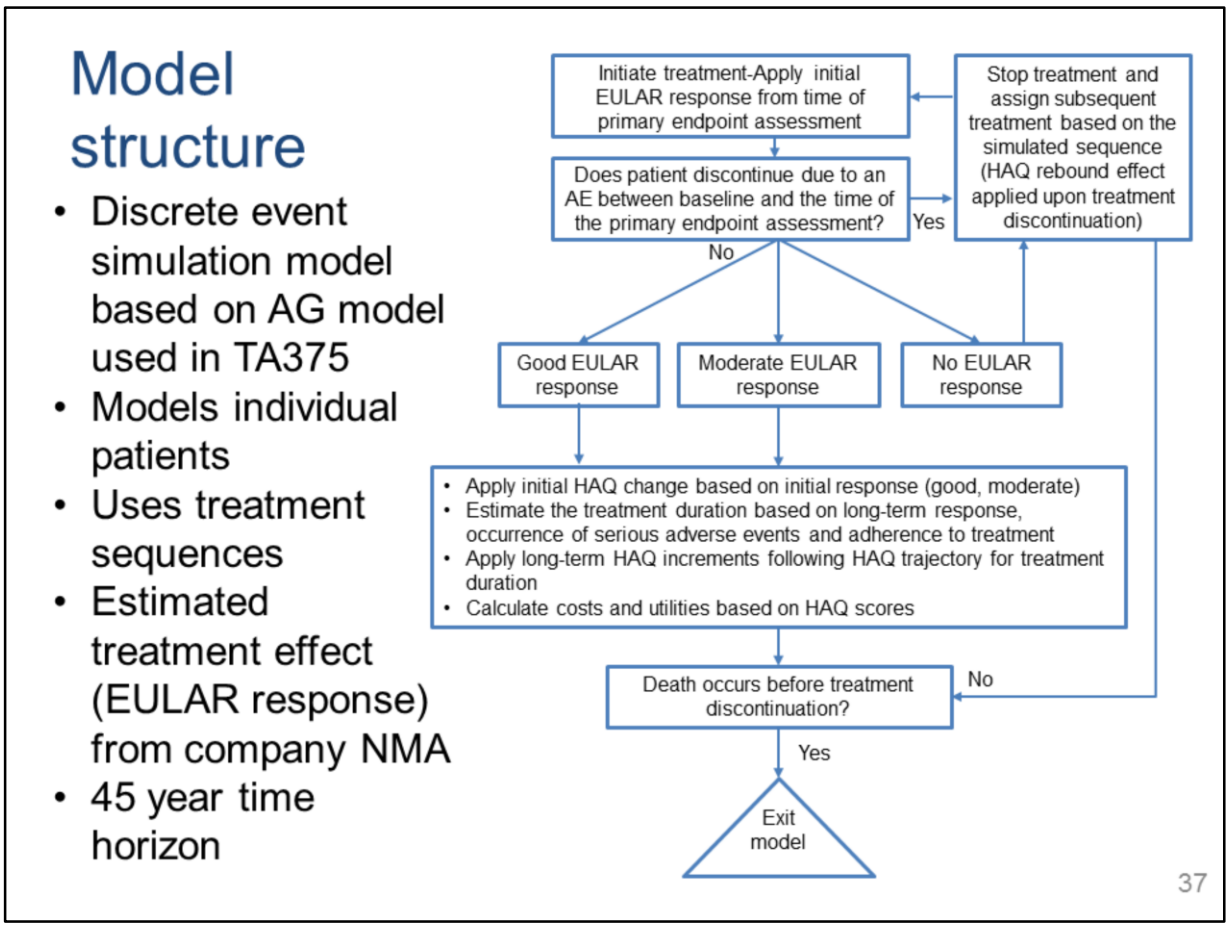
- Innovation, including that BARI is oral rather than subcutaneous or i.v. administration
- Is BARI comparable to the bDMARDs in clinical effectiveness in moderate and severe RA?
- Is BARI effective as a monotherapy?
- The ERG considered that the company's NMA results should be treated with caution. The ERG carried out an updated NMA
  - Are the Committee comfortable that the conclusions of the company NMA and the ERG NMA are broadly similar?

35

Cost effectiveness studies

- The company identified 9 UK cost-effectiveness studies
  - 8 models used in NICE technology appraisals
  - 1 independent published review
- The company did not identify any models that included BARI
  - Therefore the company developed a *de novo* health economic model to assess the cost effectiveness of BARI
- The company based the model on the AG model used in TA375

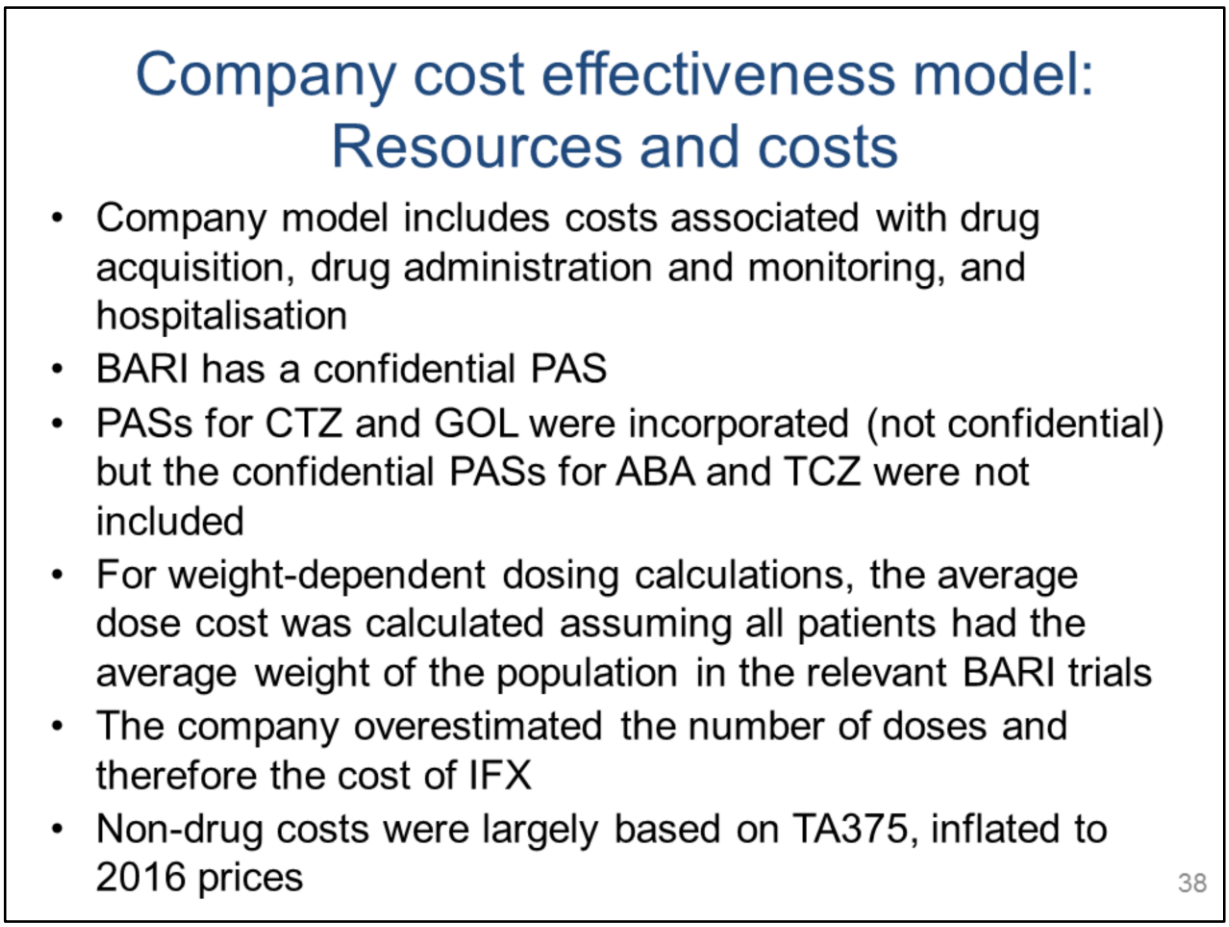
36



Source: Figure adapted from figure 47 of the company submission

Model inputs and assumptions are described in detail in section 5.6 of the company submission.

Although baricitinib is not a biologic treatment and there is no continuation rule in the SmPC for baricitinib, given that baricitinib would be used as an alternative to biologic therapies in the treatment pathway, the company anticipate that NICE would apply the same continuation rule for baricitinib as for comparator biologic therapies. Therefore, the base case cost-effectiveness analysis models the assessment of EULAR response at 24 weeks, with patients who exhibit no response being modelled to withdraw from baricitinib therapy and move on to the next treatment in the sequence, which is likely a reasonable assumption of how baricitinib may be used in clinical practice- i.e. if adequate patient response is not achieved, then therapy would be discontinued and an alternative treatment initiated. The time point of assessment is explored in a scenario analysis that considers assessment of EULAR response at 12 weeks using the 12 week NMA scenario analysis.



For weight-dependent dosing calculations, the average dose cost was calculated assuming all patients had the average weight of the population in the relevant BARI trials. The ERG notes this approach is not appropriate given that the relationship between weight and dosing cost is not linear (the average cost of a dose is not necessarily equal to the cost of the patient with the average weight, due to drug wastage and differences in cost per mg of some drugs). The ERG notes that the company should have calculated the average cost of a dose using the distribution of the weight of the modelled patient population instead of using the average weight.

For more details on the costs, see section 5.2.9 of the ERG report

Company cost effectiveness model: Utilities

- EQ-5D-5L questionnaire used to collect HRQOL data in all 3 RCTs
  - Baseline at week 1
  - Every 4 weeks from week 4 onwards
    - To week 52 for RA-BEAM
    - To week 24 for RA-BUILD and RA-BEACON
- Patient-level EQ-5D-5L responses converted to utility index-based HAQ scores using the UK-specific scoring algorithm as reported in Hernández Alava *et al*. (2012)
  - Approach not in line with TA375 which used the four-class mixture model by Hernández Alava *et al*. (2013)
  - The ERG does not consider that this changes the overall conclusions

39

Company base case: Treatment sequences for moderate RA, cDMARD-IR

Sequence	First-line treatment	Second-line treatment	Third-line treatment	Fourth-line treatment/ Rescue	Rescue
1	BARI (4 mg or 2 mg QD)	Combination of cDMARDs	MTX	PalI	NA
2	Combination of cDMARDs	MTX	PalI	NA	NA

40

Source: Table 103 of the company submission

The baricitinib economic model follows treatment sequences that reflect NICE guidance on the RA treatment pathway and TA375

The company submission notes that patients with moderate disease could progress to severe disease so the comparison sequence is potentially artificial but predicated by current NICE guidance restricting treatment beyond cDMARDs to severe patients (i.e. a DAS28 score >5.1). Therefore this sequence looks to assess cost-effectiveness of baricitinib in the moderate population assuming that patients do not become eligible for biologic treatment over time.

The company submission notes that the principle taken in determining the sequences was that they were consistent between interventions in order to avoid spurious cost-effectiveness estimates driven by having different treatments in the sequence. Due to lack of data, there is no later-line adjustment of efficacy (i.e. the NMA estimates are propagated through the model regardless of the position of the treatment in the sequence).

Company base case: Treatment sequences for severe RA, cDMARD-IR

Sequence	First-line treatment	Second-line treatment	Third-line treatment	Fourth-line treatment/ Rescue	Rescue
1	BARI (4 mg or 2 mg QD)+MTX	RTX+MTX	TCZ+MTX	MTX	Pall
2	bDMARDs (excluding TCZ)+MTX	RTX+MTX	TCZ+MTX	MTX	Pall
3	TCZ+MTX	RTX+MTX	ADA+MTX	MTX	Pall

41

Source: Table 104 of the company submission

For the cDMARD-IR severe population, the company has assumed that RTX would be the second treatment in the sequence for patients treated with BARI, in line with the use of RTX as the second-line treatment following comparator biologics. The principle taken in determining the sequences was that they were consistent between interventions in order to avoid spurious cost-effectiveness estimates driven by having different treatments in the sequence. Due to lack of data, there is no later-line adjustment of efficacy (i.e. the NMA estimates are propagated through the model regardless of the position of the treatment in the sequence).

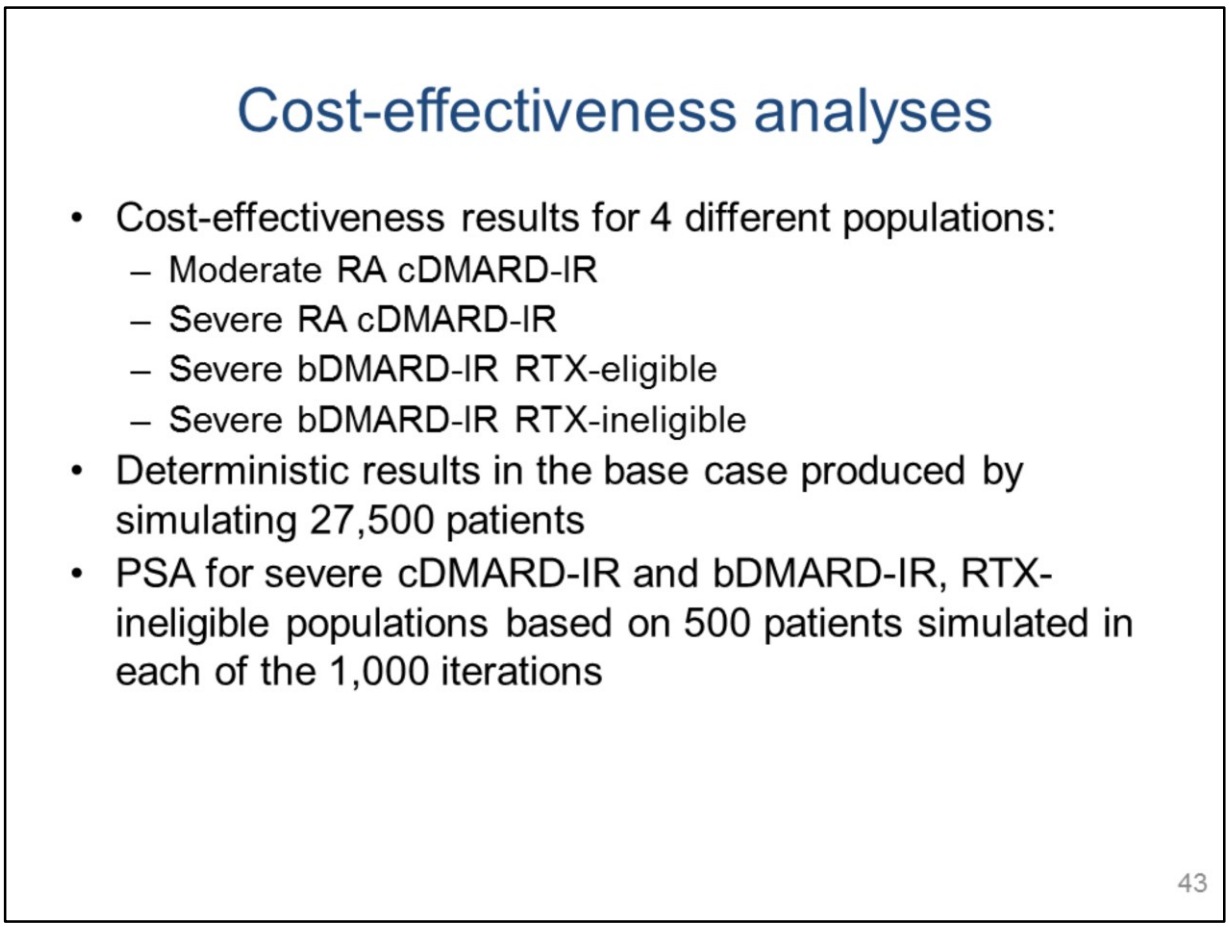
Company base case: Treatment sequences for severe RA, bDMARD-IR

Sequence	First-line treatment	Second-line treatment	Third-line treatment	Fourth-line treatment/ Rescue	Rescue
Rituximab-eligible patients					
1	BARI (4 mg or 2 mg QD)+MTX	TCZ+MTX	MTX	PalI	NA
2	RTX+MTX	TCZ+MTX	MTX	PalI	NA
Rituximab-ineligible patients					
1	BARI (4 mg or 2 mg QD)+MTX	TCZ+MTX	MTX	PalI	NA
2	bDMARDs	TCZ+MTX	MTX	PalI	NA
3	TCZ+MTX	ADA+MTX	MTX	PalI	NA

42

Source: Table 105 of the company submission

For the cDMARD-IR severe population, it has been assumed that rituximab would be the second treatment in the sequence for patients treated with baricitinib, in line with the use of rituximab as the second-line treatment following comparator biologics. The principle taken in determining the sequences was that they were consistent between interventions in order to avoid spurious cost-effectiveness estimates driven by having different treatments in the sequence. Due to lack of data, there is no later-line adjustment of efficacy (i.e. the NMA estimates are propagated through the model regardless of the position of the treatment in the sequence).



Why BARI mon was not included: The only available evidence of BARI monotherapy is in MTX-naïve patients and therefore considerable uncertainty exists about the efficacy of BARI monotherapy in cDMARD-IR and TNFi-IR patients. See Figure 2 in ERG report.

The ERG identified several limitations with the company's economic analysis, including:

1. Limitations with the company's NMA
2. Face validity and reproducibility of scenario analyses
3. Limitations of the PSA
4. Using the efficacy of treatments in cDMARD-IR population for all bDMARDs in the sequence
5. Rounding to nearest HAQ score
6. Incorrect implementation on the HAQ trajectory classes
7. HAQ improvement for responders assumed immediate
8. Averaging HAQ across large time periods
9. Exclusion of ABA IV and TCZ SC from the list of comparators
10. Using an older mapping from HAQ score to EQ-5D than the AG

11. Assuming BARI would be inserted before intensive cDMARDs for patients with moderate RA

12. Different life years gained across sequences

13. Lack of consideration of the distribution of weight for interventions where the dosage is weight based

14. Dosage of IFX

See section 5.3 of the ERG report for further details

Moderate RA cDMARD-IR

Company base-case cost effectiveness results: deterministic

Interventions	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)
Intensive cDMARDs→MTX →PALL	[redacted]	16.04	[redacted]	[redacted]	[redacted]	-
BARI+MTX→ intensive cDMARDs →MTX→PALL	[redacted]	16.03	[redacted]	[redacted]	[redacted]	37,420

- Providing BARI + MTX before cDMARDs results in [redacted] additional QALYs gained at an additional cost of £[redacted] resulting in an ICER of £37,420 per QALY gained compared with current practice.

Confidential

44

Source: Table 33 from the ERG report

The company did not provide probabilistic results

Severe RA cDMARD-IR

Company base-case cost effectiveness results: deterministic

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	Inc. analysis
IFX-b+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
ABA SC+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
GOL+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
ADA+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
ETN-b+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
TCZ IV+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
BARI+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Baseline
CTZ+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	£18,400

*All treatments followed by sequence RTX+MTX→TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by RTX+MTX→ADA+MTX→MTX→ PALL
Note: Does not include the confidential PASs for ABA and TCZ
Dominated: Treatment is less effective and more costly than an alternative

Confidential

45

Source: Table 34 of the ERG report

For patients with severe RA who are cDMARD-IR, BARI + MTX dominated all recommended comparators except for CTZ + MTX. CTZ + MTX was estimated to produce [redacted] additional QALYs more compared with BARI + MTX at an additional cost of £[redacted], resulting in an ICER of £18,400 per QALY gained. However, the ERG notes that the confidential PASs in place for ABA SC and TCZ IV were not included in the company's analysis.

Severe RA cDMARD-IR

Company base-case cost effectiveness results: probabilistic

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	Inc. analysis
IFX-b+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
ABA SC+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	Dominated
ADA+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
GOL+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	Dominated
ETN-b+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	Dominated
TCZ IV+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	Dominated
BARI+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	Baseline
CTZ+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	£18,414

*All treatments followed by sequence RTX+MTX→TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by RTX+MTX→ADA+MTX→MTX→ PALL
Note: Does not include the confidential PASs for ABA and TCZ

Confidential

46

Source: Table 35 of the ERG report

PSA results very similar to deterministic results

BARI + cDMARDs dominates all its comparators except for CTZ + cDMARDs; the ICER for CTZ + cDMARDs compared with BARI + cDMARDs is estimated to be £18,414 per QALY gained.

Severe RA bDMARD-IR RTX-eligible

Company base-case cost effectiveness results: deterministic

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)
BARI + MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated
RTX +MTX	[redacted]	13.51	[redacted]	[redacted]	[redacted]	-

*All treatments followed by sequence TCZ IV+MTX→MTX→PALL
Note: Confidential PAS for TCZ IV not included

Confidential

47

Source: Table 36 of the ERG report

BARI + MTX dominated by RTX + MTX

The company did not provide probabilistic results

Severe RA bDMARD-IR RTX-ineligible

Company base-case cost effectiveness results: deterministic

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)	ICER vs BARI + MTX (£/QALY)
GOL+MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	Dominated
BARI + MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Baseline	
ABA SC+MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	484,782
IFX-b+MTX†	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	34,942†
TCZ IV+MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	36,757
ADA+MTX†	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	27,008†
ETN-b+MTX†	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Extendedly dominated	19,874†
CTZ+MTX†	[redacted]	13.49	[redacted]	[redacted]	[redacted]	16,201	16,201†

*All treatments followed by sequence TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by ADA+MTX→MTX→ PALL

Note: Does not include the confidential PASs for ABA and TCZ

†Efficacy estimates assumed to be equal to those for the severe cDMARD-IR population

Extendedly dominated: The intervention has an ICER greater than an ICER of a more effective intervention

Confidential

8

Source: Table 37 of the ERG report

BARI + MTX was the least expensive and the second least effective intervention compared with the comparators. In the full incremental analysis, BARI + MTX was the baseline, as it dominated the only cheaper option (GOL + MTX). All other options were dominated or extendedly dominated except for CTZ + MTX, which was estimated to produce [redacted] additional QALYs compared with BARI + MTX at an additional cost of £[redacted], resulting in an ICER of £16,201 per QALY gained. The ICERs of ETN-b + MTX compared with BARI + MTX and ADA+MTX compared with BARI + MTX are also below £30,000 per QALY gained.

The company didn't identify any evidence on the effectiveness of ADA, CTZ, ETN and IFX in combination with MTX in severe bDMARD-IR patients. In the absence of data, the company used the same efficacy estimates for these treatments in severe cDMARD-IR patients instead. The ERG note that the EULAR responses are lower in the severe bDMARD-IR population compared with the severe cDMARD-IR for all treatments. Therefore, the efficacy of ADA, CTZ, ETN and IFX in combination with MTX in severe bDMARD-IR patients is likely to be overestimated in the company's base case analysis.

Severe RA bDMARD-IR RTX-ineligible

Company base-case cost effectiveness results: probabilistic

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)	ICER *vs* BARI + MTX (£/QALY) ‡
GOL+MTX	[redacted]	13.53	[redacted]	[redacted]	[redacted]	Dominated	20,824§¶
TCZ IV+MTX	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	19,962#§¶
ADA+MTX†	[redacted]	13.53	[redacted]	[redacted]	[redacted]	Dominated	19,947†§¶
ETN-b+MTX†	[redacted]	13.53	[redacted]	[redacted]	[redacted]	Baseline	19,457†§¶
IFX-b+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Extendedly dominated	5,367†¶
BARI + MTX	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Extendedly dominated	
ABA SC+MTX	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	442,044
CTZ+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	18,738	17,149†

*All treatments followed by sequence TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by ADA+MTX→MTX→ PALL
Note: Does not include the confidential PASs for ABA and TCZ; †Efficacy estimates assumed to be equal to those for the severe cDMARD-IR population;
‡ Approximate ICERs calculated by the ERG based on total costs and QALYs reported by the company; § These interventions are less effective than BARI + MTX and therefore the ICERs represent savings per QALY lost; ¶ Results affected by a programming error in the PSA

Confidential 9

Source: Table 38 of the ERG report

Some differences between the deterministic and the probabilistic results. These are mostly caused by an error in the company's model that affects the sequences starting with TCZ IV + MTX, ADA + MTX, ETN-b + MTX, GOL + MTX and IFX-b + MTX.

Cost effectiveness summary: Company estimates

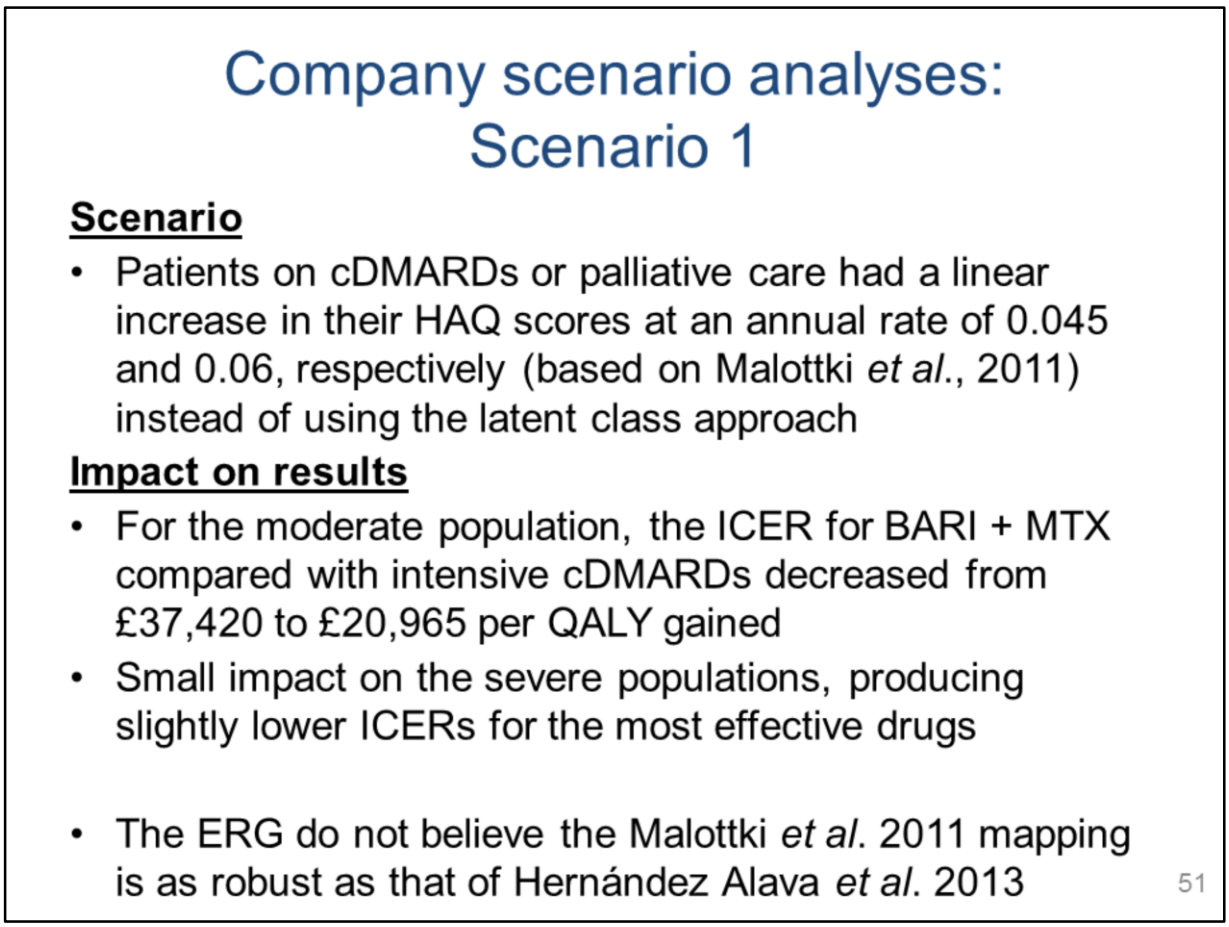
Population	Summary (ICERs)
1. Moderate RA cDMARD-IR	• BARI + MTX *vs* intensive cDMARDs = £37,420
2. Severe RA cDMARD-IR	• BARI + MTX dominated all comparators • **Except** BARI + MTX *vs* CTZ + MTZ = £18,400
3. Severe RA bDMARD-IR RTX-eligible	• BARI + MTX dominated by RTX + MTX
4. Severe RA bDMARD-IR RTX-ineligible	• BARI + MTX less effective and less expensive than all comparators • **Except** BARI + MTX dominated GOL + MTX

- The confidential PASs for ABA and TCZ were not included in these analyses.

50

For the severe RA bDMARD-IP RTX-ineligible population:

- BARI + MTX vs ETN-b + MTX & CTZ + MTX & ADA = ICER lower than £30,000
- BARI + MTX vs TCZ + MTX & ABA SC + MTX = ICER higher than £30,000



Company didn't carry out one-way sensitivity analyse due to the computational burden but presented some scenario analyses

The ERG noted a number of errors in the scenario analyses

1. The ERG do not believe the Malottki *et al*. 2011 mapping is as robust as that of Hernández Alava *et al*. 2013
2. In response to clarification, the company stated that due to a limitation in the model, this scenario analysis can only be run when patients on cDMARDs or palliative care are also assumed to suffer a linear HAQ increase.

Company scenario analyses: Other scenarios

- HAQ score for BARI + MTX deteriorates (increases) at half of rate assumed for cDMARDs
- HAQ score improvements for BARI calculated from trial data rather than BSRBR database
- Different time to treatment discontinuation for patients on BARI
- Alternative methods used to map HAQ scores to the EQ-5D
- Serious adverse events accounted for
- Tapering BARI from 4 mg QD to 2 mg QD
- Head-to-head comparison between BARI + MTX and ADA + MTX

- The ERG believes these scenarios are unlikely to change the conclusions of the cost-effectiveness analyses

52

For further details of the scenarios and their impact see section 5.2.11 of the ERG report

ERG exploratory analyses

- ERG undertook few exploratory analyses. The ERG believes [redacted]
- ERG identified 2 programming errors that affected the company's PSA results and re-ran the PSA
  - 1. Error resulted in patients on GOL + MTX, ETN + MTX, ADA + MTX and IFX + MTX never achieving a good or moderate EULAR response
    - Also affects the sequence starting with TCZ + MTX, given that ADA + MTX is included in the sequence
  - 2. Error in the calculations of the CODA samples for moderate response probability for BARI + MTX in the severe cDMARD-IR population

53

Severe RA cDMARD-IR population

ERG exploratory analyses: Error affected the PSA

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	Inc. analysis
IFX-b+MTX	[redacted]	14.72	[redacted]	[redacted]	[redacted]	Dominated
ABA SC+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
ADA+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
GOL+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
TCZ IV+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
ETN-b+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
BARI+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Baseline
CTZ+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	£18,135

*All treatments followed by sequence RTX+MTX→TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by RTX+MTX→ADA+MTX→MTX→PALL. Confidential PAS for TCZ IV not included.
Note: Does not include the confidential PASs for ABA and TCZ

- Minimal impact on the results

Confidential 54

Source: Table 41 of the ERG report

This is the ERG corrected version of the company bases case analysis shown in table 35 of the ERG report (slide 46 in this presentation)

Severe RA bDMARD-IR RTX-ineligible population
ERG exploratory analyses: Error affected the PSA

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)	ICER *vs* BARI+MTX (£/QALY)
GOL+MTX	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Baseline	18,805 §
BARI + MTX	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Extendedly dominated	
ABA SC+MTX#	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	454,225#
TCZ IV+MTX#	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	37,063#
ADA+MTX; †	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	21,494†
ETN-b+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	£15,527	10,197†
IFX-b+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	35,045†
CTZ+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	£20,170	16,962†

*All treatments followed by sequence TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by ADA+MTX→MTX→PALL. Confidential PAS for TCZ IV not included.
Note: Does not include the confidential PASs for ABA and TCZ
†Efficacy estimates assumed to be equal to those for the severe cDMARD-IR population
§ GOL+MTX is less effective than BARI+MTX; the ICER represents savings per QALY lost compared with BARI+MTX

- Important impact in the sequences effected
  - Markedly higher costs and QALYs gained for TCZ+MTX, ETN-b+MTX, IFX-b+MTX, GOL+MTX and ADA+MTX

Confidential

55

Source: Table 42 of the ERG report

This is the ERG corrected version of the company bases case analysis shown in table 38 of the ERG report (slide 49 in this presentation)

Cost effectiveness summary: ERG estimates

- In the moderate RA population, the AG in TA375 estimated that the median ICER of bDMARDs compared with cDMARDs was in the region of £50,000 per QALY gained. [redacted]
- In the severe cDMARD-IR population who can tolerate MTX, [redacted]
- In severe cDMARD-IR patients and in bDMARD-IR patients for whom RTX is contraindicated or not tolerated, [redacted]
- The ERG believe that the results will also apply to BARI monotherapy

Confidential

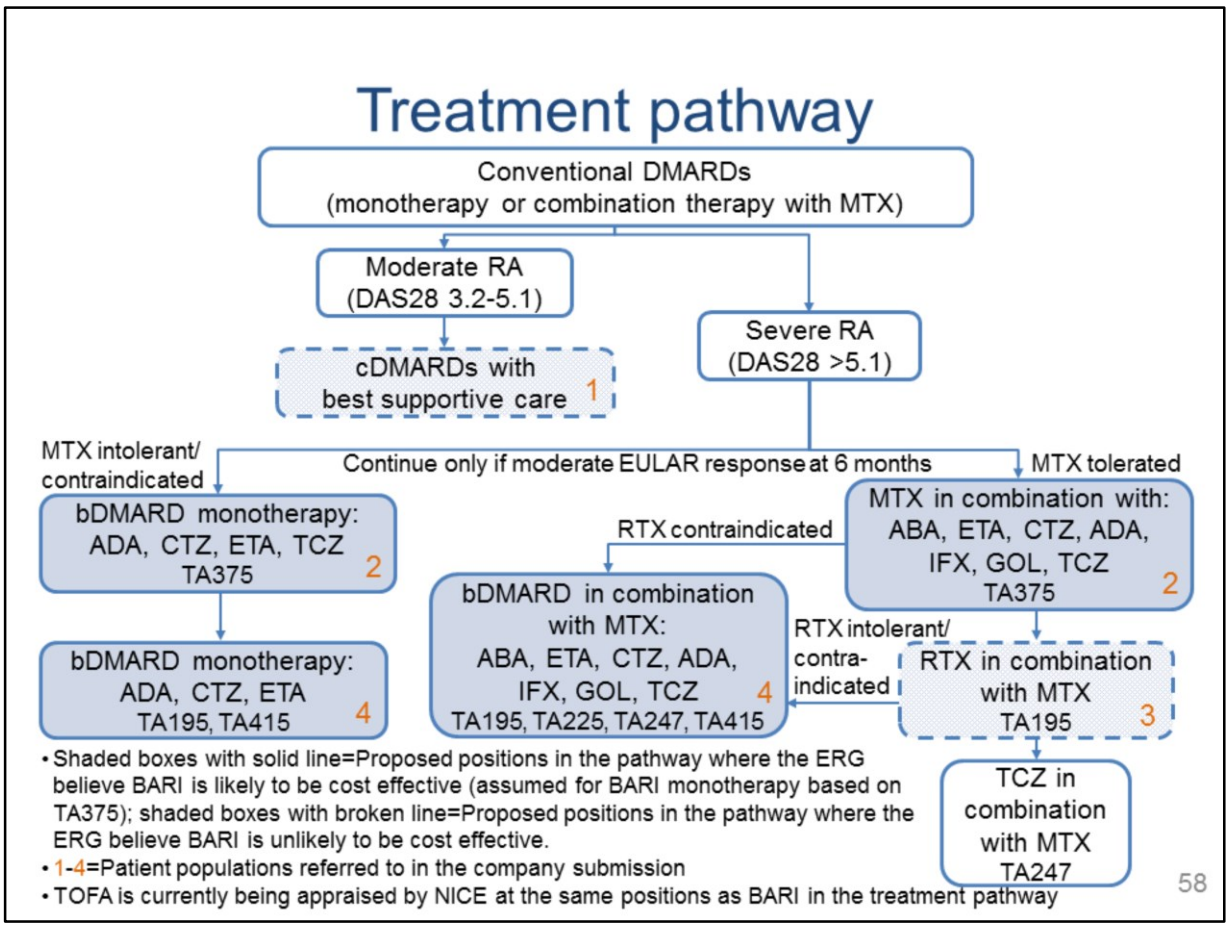
56

Key issues: Cost effectiveness

- Is BARI comparable to the bDMARDs in both clinical effectiveness and cost effectiveness?
- Has the case for BARI monotherapy been proven?

57

See section 5.5 of ERG report for a detailed critique of cost-effectiveness issues



Adapted from Figure 3 (page 49) of the company submission

Impact of the confidential PASs

- CTZ and GOL have non-confidential PASs
  - Incorporated into the previous analyses
- ABA and TCZ have confidential simple discount PASs
- ERG re-ran the analyses for all 3 severe populations: cDMARD-IR, bDMARD-IR RTX-eligible, and bDMARD-IR RTX-ineligible
  - Neither ABA nor TCZ are included as comparators in the severe, bDMARD-IR, RTX-eligible population
  - TCZ is included in the sequence of both the cDMARD-IR and bDMARD-IR RTX ineligible populations
- The analysis for the moderate population was not re-run, as it was not affected by the PAS of ABA or TCZ
- All other sequences are affected by the TCZ PAS, as it is included either as first or last bDMARD in every sequence
- All the analyses were run using the original submitted model

59

Authors

- **Helen Powell**
  Technical Lead
- **Alex Filby**
  Technical Adviser
- With input from the Lead Team (Andrew Renehan, Stephen O'Brien and David Chandler)

60

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Baricitinib for treating moderate to severe rheumatoid arthritis

Final scope

Remit/appraisal objective

To appraise the clinical and cost effectiveness of baricitinib within its marketing authorisation for treating moderate to severe active rheumatoid arthritis.

Background

Rheumatoid arthritis is an inflammatory autoimmune disease that typically affects the synovial tissue of the small joints of the hands and feet but can affect any synovial joint, causing swelling, stiffness, pain and progressive joint destruction. It is a systemic disease and can affect the whole body, including the lungs, heart and eyes. Rheumatoid arthritis is usually a chronic relapsing condition which has a pattern of flare-ups followed by periods of lower disease activity; however, for some people, the disease is constantly progressive. Rheumatoid arthritis is associated with increased mortality and increasing disability, which has a severe impact on quality of life. Severity of disease can be classified into 3 categories, based on the disease activity score (DAS28) scoring system. A DAS28 greater than 5.1 indicates high disease activity or severe disease, between 3.2 and 5.1 indicates moderate disease activity, and less than 3.2 indicates low disease activity.

The prevalence of rheumatoid arthritis in the UK is estimated to be 0.44% in males and 1.16% in females[1]; which is approximately 441,000 people in England (119,000 males and 322,000 females)[1,2]. There are approximately 17,500 people diagnosed with rheumatoid arthritis every year in England[3]. It can develop at any age, but the peak age of onset in the UK is about 45–75 years[1].

There is no cure for rheumatoid arthritis and treatment aims to improve quality of life and to prevent or reduce joint damage. The main aim of management in early disease is to suppress disease activity and induce disease remission, prevent loss of function, control joint damage, maintain pain control and enhance self-management. For people with newly diagnosed rheumatoid arthritis, NICE clinical guideline 79 ('Rheumatoid arthritis: management') recommends a combination of conventional disease modifying anti-rheumatic drugs (DMARDs; including methotrexate, leflunomide and sulfasalazine) 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) DMARD monotherapy is recommended. Where the disease has not responded to intensive

combination therapy with conventional DMARDs, NICE Technology appraisal guidance 375 recommends biological DMARDs (adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept) in combination with methotrexate for severe rheumatoid arthritis only. For those people with severe rheumatoid arthritis who cannot take methotrexate because it is contraindicated or because of intolerance, the guidance recommends that adalimumab, etanercept, certolizumab pegol or tocilizumab monotherapy can be used.

Where the disease has not responded adequately or in the case of intolerance to other DMARDs, including at least one TNF inhibitor (a subgroup of biological DMARDs), rituximab in combination with methotrexate is recommended for severe active disease only (NICE Technology appraisal guidance 195). Where rituximab is contraindicated or withdrawn because of an adverse event, adalimumab, etanercept, infliximab, abatacept, golimumab, tocilizumab and certolizumab pegol each in combination with methotrexate are recommended as options (NICE Technology appraisal guidance 195, 225, 247 and 415). Where rituximab therapy cannot be given because methotrexate is contraindicated or has been withdrawn due to an adverse event, adalimumab, etanercept and certolizumab pegol, each as a monotherapy, can be used (NICE Technology appraisal guidance 195 and 415).

The technology

Baricitinib (Olumiant, Eli Lilly and Company) is a human tyrosine kinase protein that inhibits Janus kinase 1 and 2 and thereby disrupts mediating signalling pathways involved in inflammatory diseases. It is administered orally.

Baricitinib does not currently have a marketing authorisation in the UK for rheumatoid arthritis. It has been studied in clinical trials as monotherapy or in combination with methotrexate. It has been compared with methotrexate in adults with moderate to severe rheumatoid arthritis who have not received treatment with conventional DMARDs or methotrexate. It has also been compared with placebo in adults with moderate to severe active rheumatoid arthritis, who have been treated with, and whose disease did not respond adequately to, methotrexate. It has also been studied in a clinical trial in combination with methotrexate, compared with adalimumab in combination with methotrexate in adults with moderate to severe active rheumatoid arthritis who have already had methotrexate. It has also been compared with placebo in adults whose disease did not respond adequately to a TNF inhibitor.

Intervention	Baricitinib monotherapy or in combination with methotrexate

Population	Adults with moderate to severe, active rheumatoid arthritis whose disease has responded inadequately to, or who are intolerant of one or more disease modifying anti-rheumatic drugs (DMARDs), including conventional or biologic DMARDs.
Comparators	People with moderate active rheumatoid arthritis: • Combination therapy with conventional DMARDs (including methotrexate and at least one other DMARD, such as sulfasalazine and leflunomide) • Conventional DMARD monotherapy with dose escalation • Best supportive care (only where conventional DMARDs are not appropriate due to intolerance) People with severe active rheumatoid arthritis that has not responded adequately to therapy with conventional DMARDs only: • Biological DMARDs in combination with methotrexate (adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab, abatacept) • Adalimumab, etanercept, certolizumab pegol, or tocilizumab (each as monotherapy) People with severe active rheumatoid arthritis that has not responded adequately to therapy with DMARDs including at least one TNF inhibitor: • Rituximab in combination with methotrexate • When rituximab is contraindicated or withdrawn due to adverse events: - Abatacept, adalimumab, certolizumab pegol, etanercept, infliximab, tocilizumab, or golimumab, each in combination with methotrexate - Adalimumab, etanercept or certolizumab pegol (each as monotherapy)

Outcomes	The outcome measures to be considered include: • disease activity • physical function • joint damage • pain • mortality • fatigue • radiological progression • extra-articular manifestations of the disease • adverse effects of treatment • health-related quality of life
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies will be taken into account. The availability and cost of biosimilar products of etanercept, infliximab and rituximab should be taken into account.
Other considerations	If the evidence allows the following subgroups will be considered. These include people with moderate disease activity (DAS28 between 3.2 and 5.1) and severe active disease (DAS28 greater than 5.1). Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.
Related NICE recommendations	**Related Technology Appraisals:** Certolizumab pegol for treating rheumatoid arthritis after

and NICE Pathways	inadequate response to a TNF-alpha inhibitor (2016) NICE Technology Appraisal TA415. Review date: October 2019. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with DMARDs or after conventional DMARDs only have failed (review of TA guidance 130, 186, 224, 234 and part review of TA guidance 225 and 247) (2016) NICE Technology Appraisal TA375. Review date: January 2019. Tocilizumab for the treatment of rheumatoid arthritis (rapid review of technology appraisal guidance 198) (2012) NICE technology appraisal TA247. Guidance on static list. Golimumab for the treatment of rheumatoid arthritis after the failure of previous disease modifying anti-rheumatic drugs (2011) NICE technology appraisal TA225. Guidance on static list. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor (2010) NICE technology appraisal TA195. Guidance on static list. **Appraisals in development (including suspended appraisals)** Rheumatoid arthritis (after the failure of conventional DMARDs) -rituximab. Technology Appraisal [ID333]. Status March 2011: Suspended – manufacturer is no longer seeking a licence for this indication. Rheumatoid arthritis - tofacitinib.Technology Appraisal [ID526]. Publication date: January 2018. **Related Guidelines:** Rheumatoid arthritis in adults: management (2009) NICE guideline CG79. Review date: August 2018. **Related Quality Standards:** Rheumatoid arthritis in over 16s (2013) Quality Standard QS33. **Related NICE Pathways:** Rheumatoid arthritis (2015) NICE Pathway.
Related National Policy	NHS England (2016) Manual for prescribed specialised services 2016/17. 5. Adult highly specialist rheumatology services.

	NHS England & BMJ Group (2012) Shared Decision Making Sheets: Rheumatoid Arthritis. NHS England (2013) A13. Specialised Rheumatology. National programmes of care and clinical reference groups. Department of Health (2016) NHS Outcomes Framework 2016 to 2017. Domains 1, 2, 3, 4 and 5.

References

1. Arthritis Research UK (2015) How common is rheumatoid arthritis? Accessed March 2016.

2. Office for National Statistics (2015) Population Estimates by Age and Sex. Accessed September 2015.

3. National Institute for Health and Care Excellence (2013) Rheumatoid arthritis: commissioning guide. Accessed November 2016.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Baricitinib for treating moderate to severe rheumatoid arthritis [ID979]

Matrix of consultees and commentators

Consultees	Commentators (no right to submit or appeal)
Company • Eli Lilly (baricitinib) Patient/carer groups • Action on Pain • Arthritis Action • Arthritis & Musculoskeletal Alliance • Arthritis Care • BackCare • Disability Rights UK • Leonard Cheshire Disability • Muslim Council of Britain • National Rheumatoid Arthritis Society • Pain Concern • Pain Relief Foundation • Pain UK • South Asian Health Foundation • Specialised Healthcare Alliance Professional groups • British Geriatrics Society • British Institute of Musculoskeletal Medicine • British Orthopaedic Association • British Pain Society • British Society for Rheumatology • British Society of Rehabilitation Medicine • Physiotherapy Pain Association • Primary Care Rheumatology Society • Royal College of General Practitioners • Royal College of Nursing • Royal College of Pathologists • Royal College of Physicians • Royal Pharmaceutical Society • Royal Society of Medicine • UK Clinical Pharmacy Association	General • Allied Health Professionals Federation • Board of Community Health Councils in Wales • British National Formulary • Care Quality Commission • Department of Health, Social Services and Public Safety for Northern Ireland • Healthcare Improvement Scotland • Medicines and Healthcare products Regulatory Agency • National Association of Primary Care • National Pharmacy Association • NHS Alliance • NHS Commercial Medicines Unit • NHS Confederation • Scottish Medicines Consortium Comparator companies • AbbVie (adalimumab) • Accord Healthcare (methotrexate) • Biogen Idec (etanercept biosimilar, infliximab biosimilar) • Bristol-Myers Squibb (abatacept) • Concordia International Rx (methotrexate) • Hameln Pharmaceuticals (methotrexate) • Hospira UK (methotrexate, infliximab biosimilar) • Medac (methotrexate, leflunomide) • Merck Sharp and Dohme (infliximab, golimumab) • Napp Pharmaceuticals (infliximab biosimilar) • Orion Pharma (methotrexate) • Pfizer (methotrexate, sulfasalazine, etanercept)

Consultees	**Commentators (no right to submit or appeal)**
Others • Department of Health • NHS Canterbury and Coastal CCG • NHS England • NHS High Weald Lewes Haven CCG • Welsh Government	• Roche (tocilizumab, rituximab) • Rosemont pharmaceuticals (methotrexate, sulfasalazine) • Sandoz (leflunomide, methotrexate) • Sanofi (leflunomide) • UCB Pharma (certolizumab pegol) • Zentiva (leflunomide) Relevant research groups • Arthritis Research UK • Chronic Pain Policy Coalition • Cochrane Musculoskeletal Group • MRC Clinical Trials Unit • National Institute for Health Research Associated Public Health Groups • Public Health England • Public Health Wales

NICE is committed to promoting equality, eliminating unlawful discrimination and fostering good relations between people who share a protected characteristic and those who do not. Please let us know if we have missed any important organisations from the lists in the matrix, and which organisations we should include that have a particular focus on relevant equality issues.

PTO FOR DEFINITIONS OF CONSULTEES AND COMMENTATORS

Definitions:

Consultees

Organisations that accept an invitation to participate in the appraisal; the company that markets the technology; national professional organisations; national patient organisations; the Department of Health and the Welsh Government and relevant NHS organisations in England.

The company that markets the technology is invited to make an evidence submission, respond to consultations, nominate clinical specialists and has the right to appeal against the Final Appraisal Determination (FAD).

All non-company consultees are invited to submit a statement[1], respond to consultations, nominate clinical specialists or patient experts and have the right to appeal against the Final Appraisal Determination (FAD).

Commentators

Organisations that engage in the appraisal process but that are not asked to prepare an evidence submission or statement, are able to respond to consultations and they receive the FAD for information only, without right of appeal. These organisations are: companies that market comparator technologies;
Healthcare Improvement Scotland; other related research groups where appropriate (for example, the Medical Research Council [MRC], National Cancer Research Institute); other groups (for example, the NHS Confederation, NHS Alliance and NHS Commercial Medicines Unit, and the British National Formulary.

All non-company commentators are invited to nominate clinical specialists or patient experts.

[1]Non-company consultees are invited to submit statements relevant to the group they are representing.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis [ID979]

Eli Lilly and Company Limited

File name	Version	Contains confidential information	Date
Baricitinib_rheumatoid arthritis_ID979_main submission_ACIC removed_21 April 2017	V1.0	No	21 April 2017

Please note all academic and commercial in confidence data has been redacted

Contents

List of tables

List of figures

Table of abbreviations

Acronym	Definition
ABA	Abatacept
ABTS	Abatacept subcutaneous
ABTSMTX	Abatacept subcutaneous + methotrexate
ACPA	Anti-citrullinated protein antibodies
ACR	American College of Rheumatology
ACR20	20% improvement in American College of Rheumatology Criteria
ACR50	50% improvement in American College of Rheumatology Criteria
ACR70	70% improvement in American College of Rheumatology Criteria
ADA	Adalimumab
ADAMTX	Adalimumab + methotrexate
ADCC	Antibody-dependent cell cytotoxicity
AE	Adverse event
AIC	Academic in confidence
ALC	Absolute lymphocyte count
ALT	Alanine transaminase
ANC	Absolute neutrophil count
ANCOVA	Analysis of covariance
AST	Aspartate transaminase
ATP	Adenosine triphosphate
AUC	Area under the curve
AWMSG	All Wales Medicines Strategy Group
AZA	Azathioprine
BAR4MTX	Baricitinib 4 mg + methotrexate
BARI	Baricitinib
BCP	Biochemical profile
BHPR	British Health Professionals in Rheumatology
BIC	Bayesian information criteria
bDMARD	Biologic disease-modifying antirheumatic drug
BID	Twice daily
BIM	Budget impact model
BIW	Twice weekly
BMI	Body mass index
BNF	British National Formulary
BRAM	Birmingham Rheumatoid Arthritis Model
BSC	Best supportive care

Acronym	Definition
BSR	British Society for Rheumatology
BSRBR	British Society for Rheumatology Biologics Register
CADTH	Canadian Agency for Drugs and Technologies in Health
CCS	Corticosteroids
CCP	Cyclic citrullinated peptide
CDAI	Clinical Disease Activity Index
CDC	Complement-dependent cytotoxicity
cDMARD	Conventional disease-modifying antirheumatic drug
CEA	Cost-effectiveness analysis
CEE	Central eastern Europe
CEM	Cost-effectiveness model
CFB	Change from baseline
CHMP	Committee for Medicinal products for Human Use
CI	Confidence interval
CIC	Commercial in confidence
COMORA	COMOrbidities in Rheumatoid Arthritis study
COPD	Chronic obstructive pulmonary disease
COX	Cyclooxygenase
CPK	Creatine phosphokinase
CRD	Centre for Reviews and Dissemination
CRP	C-reactive protein
CRU	Cost and resource use
CSR	Clinical study report
CTZ	Certolizumab pegol
CTZMTX	Certolizumab pegol + methotrexate
CXR	Chest X-ray
DALY	Disability-adjusted life year
DAS28	Disease Activity Score in 28 Joints
DAS44	Disease Activity Score in 44 Joints
DAS28-hsCRP	Disease Activity Score in 28 Joints high-sensitivity C-reactive protein
DAS28-ESR	Disease Activity Score in 28 Joints erythrocyte sedimentation rate
DES	Discrete event simulation
DCHS	Different health care settings
DIC	Deviance information criterion
DMARD	Disease-modifying antirheumatic
DMC	Data monitoring committee

Acronym	Definition
DNA	Deoxyribonucleic acid
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
DXR	Digital x-ray radiogrammetry
EAIR	Exposure-adjusted incidence rate
ECG	Electrocardiogram
EMA	European Medicines Agency
EPAR	European public assessment report
ERAN	Early Rheumatoid Arthritis Network
ERAS	Early Rheumatoid Arthritis Study
ERG	Evidence Review Group
ESPOIR	Etude et Suivi des Polyarthrites Indifferenciees Recentes
ESR	Erythrocyte sedimentation rate
ETN	Etanercept
ETNMTX	Etanercept + methotrexate
ETN-bMTX	Etanercept biosimilar + methotrexate
EULAR	European League Against Rheumatism
EQ-5D-5L	European Quality of Life-5 Dimensions-5 levels
FACIT-F	Functional Assessment of Chronic Illness Therapy-Fatigue
FBC	Full blood count
FSH	follicle-stimulating hormone
GBP	British Pound Sterling
GOL	Golimumab
GOLMTX	Golimumab + methotrexate
GSTT	Guys and St Thomas' NHS Foundation Trust
HAQ-DI	Health Assessment Questionnaire-Disability Index
HBV	Hepatitis B Virus
HCHS	Hospital and community health services
HCQ	Hydroxychloroquine
HCV	Hepatitis C Virus
HDA	High disease activity
HDL	High-density lipoprotein
HES	Hospital episode statistic
HIV	Human immunodeficiency virus
HRQOL	Health-related quality of life
HSUV	Health state utility values

Acronym	Definition
HTA	Health Technology Assessment
HUI	Health utility index
IBS	Irritable bowel syndrome
ICER	Incremental cost-effectiveness ratio
ICF	Informed consent form
IFX	Infliximab
IFXMTX	Infliximab + methotrexate
IFX-bMTX	Infliximab biosimilar + methotrexate
IQR	Interquartile range
IR	Insufficient response
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITT	Intent-to-treat
IVRS	Interactive Voice Response System
JAK	Janus kinase
IV	Intravenous
JAMA	Journal of the American Medical Association
JSN	Joint space narrowing
LDA	Low disease activity
LDAS	Low disease activity state
LDL	Low-density lipoprotein
LEF	Leflunomide
LOCF	Last observation carried forward
LS	Least squares
LSM	Least squares mean
LTE	Long-term extension
LYG	Life years gained
MACE	Major adverse cardiovascular event
MAE	Mean absolute error
MCID	Minimum clinically important difference
MDRD	Modification of Diet in Renal Disease
MHAQ	Multi-Dimensional Health Assessment Questionnaire
MJS	Morning joint stiffness
MRI	Magnetic resonance imaging
MTA	Multiple technology appraisal
MTC	Mixed-treatment analysis
mTSS	Modified Total Sharp Score

Acronym	Definition
MTX	Methotrexate
NCPE	National Centre for Pharmacoeconomics
NDB	National Data Bank
NHS	National Health Service
NICE	National Institute of Health and Care Excellence
NJR	National Joint Registry
NMA	Network meta-analysis
NMSC	Non-melanoma skin cancer
NNQ	Number needed per QALY gained
NOAR	Norfolk Arthritis Register
NRAS	National Rheumatoid Arthritis Society
NRI	Non-responder imputation
NRS	Numeric rating scale
NSAID	Non-steroidal anti-inflammatory
ONS	Office for National Statistics
OWSA	One-way sensitivity analysis
PAS	Patient Access Scheme
PASLU	Patient Access Scheme Liasion Unit
PASS	Patient acceptable symptom state
Pall	Palliative care
PBAC	Pharmaceutical Benefits Advisory Committee
PBO	Placebo
PFS	Pre-filled syringe
PGA	Physician's Global Assessment of Disease Activity
PICO	Population, Intervention, Comparator, Outcomes
PML	progressive multifocal leukoencephalopathy
PPD	Positive purified protein derivative
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PRO	Patient reported outcome
PSA	Probabilistic sensitivity analysis
PSL	Prednisolone
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PYE	Patient-years of exposure
QALY	Quality-adjusted life year
QIDS-SR$_{16}$	Quick Inventory of Depressive Symptomatology Self-Rated-16

Acronym	Definition
QD	Once a day
QOL	Quality of life
QOW	Every other week
RAF	Rheumatoid arthritis and falls
RA	Rheumatoid arthritis
RAPID	Routine Assessment of Patient Index Data
RCT	Randomised controlled trial
RMSE	Root mean square error
RTX	Rituximab
RTXMTX	Rituximab + methotrexate
RF	Rheumatoid factor
SAA	Serum amyloid A
SAP	Statistical analysis plan
SC	Subcutaneous
SD	Standard deviation
SDAI	Simplified Disease Activity Index
SE	Standard error
SJC	Swollen joint count
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SmPC	Summary of Product Characteristics
SSZ	Sulfasalazine
STA	Single Technology Appraisal
STAT	Signal transducers and activators of transcription
SUBCUT	Subcutaneous
SUPPL	Supplement
TCZ	Tocilizumab
TCZMTX	Tocilizumab + methotrexate
TEAE	Treatment-emergent adverse event
THR	Total hip replacement
TJC	Tender joint count
TJR	Total joint replacement
TKR	Total knee replacement
TNF	Tumour necrosis factor
TNFi	Tumour necrosis factor inhibitor
TSH	thyroid-stimulating hormone

Acronym	Definition
TTO	Time trade-off
ULN	Upper limit of normal
URTI	Upper respiratory tract infection
VARA	Veterans Affairs Rheumatoid Arthritis
VAS	Visual analogue scale
VAT	Value added tax
VBA	Visual basic analogue
VRS	Visual rating scale
WHO	World Health Organisation
WPAI-RA	Work Productivity and Activity Impairment -Rheumatoid Arthritis
WTP	Willingness to pay
WPL	Work productivity loss
WPS-RA	Work Productivity Survey-Rheumatoid Arthritis

1 Executive summary

- Baricitinib is anticipated to be the first oral Janus-kinase 1/2 (JAK1/2) inhibitor licensed in the European Union for the treatment of moderate to severe rheumatoid arthritis (RA), a chronic autoimmune disease characterised by progressive pain and stiffness of the joints.[1,2]
- Baricitinib is indicated for the treatment of moderate to severe active rheumatoid arthritis in adult patients who have responded inadequately to, or who are intolerant to one or more disease-modifying antirheumatic drugs. Baricitinib may be used as monotherapy or in combination with methotrexate
- Baricitinib 4 mg once daily (QD) (either as monotherapy or in combination with methotrexate) is considered for the full licensed population of adult patients with moderate to severe RA for whom methotrexate, conventional disease-modifying antirheumatic drugs (cDMARDs) or biologic DMARDs (bDMARDs) have been inadequately effective or not tolerated.[3]
- More than half of moderate to severe RA patients still do not achieve a satisfactory response following bDMARD treatment.[4] In patients treated with tumour necrosis factor inhibitor (TNFi) biologics in particular, efficacy may diminish over time due to the production of anti-drug antibodies to the injected therapeutic protein antibody.[5]
- Furthermore, all bDMARDs must be administered via subcutaneous injection or intravenous infusion, which can be associated with painful injection site reactions.[6,7]
- As a small molecule, baricitinib does not induce the production of anti-drug antibodies, and as an oral therapy baricitinib could have a substantial impact on patients who may experience painful side effects with, and potentially discontinue currently available bDMARDs.
- Baricitinib is the first JAK inhibitor to receive CHMP positive opinion in Europe. As such, baricitinib represents a therapy with a novel mechanism of action that may provide a valuable extension to the armamentarium available to clinicians and may fulfil a particular unmet need in patients who have not responded to the currently available treatment options.

Summary of Clinical Effectiveness

- Baricitinib was studied in a robust, comprehensive clinical trial programme reflective of the RA treatment pathway in the UK, including comparator therapies commonly used in UK clinical practice. Compared with active control or placebo, oral baricitinib 4 mg QD was associated with significant improvements in signs and symptoms, physical function, and patient reported outcomes in a broad patient population across the RA treatment spectrum, including:[8-10]
    - MTX-inadequate response (-IR) patients in RA-BEAM;
    - cDMARD-IR patients in RA-BUILD;
    - Anti-TNF-IR patients in RA-BEACON.
- Two additional studies of relevance are RA-BEYOND (long-term extension study) and RA-BEGIN (DMARD-naïve population- unlicensed indication, methodology and results presented in Appendix 1)
- The primary endpoint of a 20% improvement in American College of Rheumatology Criteria (ACR20) at Week 12 or 24 was met in all studies. In addition, baricitinib 4 mg (QD) showed

statistically significant improvements to all comparators for 50/70% improvement in American College of Rheumatology Criteria (ACR20/50/70) response rates at the primary timepoints.[9,11,12]

- In studies measuring radiographic outcomes (RA-BEAM and RA-BUILD), progression of joint damage, as measured by modified Total Sharp Score (mTSS) scores, was significantly reduced in patients treated with baricitinib 4 mg plus methotrexate compared to placebo and similar compared to adalimumab plus methotrexate.[9,11]
- Baricitinib in combination with methotrexate (MTX) demonstrated consistent improvements in European League Against Rheumatism (EULAR) response across trials.[9,11,12]
- Patient reported symptoms were statistically significantly improved versus placebo across all three trials, as determined by Health Assessment Questionnaire-Disability Index (HAQ-DI), duration/severity of morning joint stiffness and Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F).[9,11,12]
- In addition, the results obtained in the HAQ-DI minimum clinically important difference (MCID) and FACIT-F MCID response rates demonstrate the statistical superiority of baricitinib compared to placebo.[9,12]

Summary of Safety

- Treatment with baricitinib was well-tolerated and a small proportion of patients discontinued from the baricitinib studies because of adverse events (AEs).[2,13]
- The most commonly reported adverse drug reactions (ADRs) in ≥2% of patients treated with baricitinib monotherapy or in combination with cDMARDs included increased low-density lipoprotein (LDL) cholesterol, upper respiratory tract infections (URTI) and nausea. However, the majority of all ADRs were mild to moderate in severity. [2,13]
- The proportion of patients with SAEs (including serious infections) was similar across treatment groups in the phase III studies and integrated placebo-controlled analysis sets, except for RA-BEAM, where a higher proportion of SAEs were reported with placebo and baricitinib versus adalimumab. [2,13]
- Although baricitinib was associated with a higher incidence of serious adverse events (SAEs) compared with adalimumab through 52 weeks in RA-BEAM, their AE profiles were similar across clinically significant categories of risk including major adverse cardiovascular events (MACE), malignancies, hypercholesterolemia, serious infections and herpes zoster [2,13]
- Despite a higher risk of cardiovascular disease, infection, and malignancy in the RA population, treatment with baricitinib did not result in increased risk of malignancy, serious or opportunistic infections, or MACE.[2,13]
- Non-serious herpes simplex and herpes zoster infections were more frequent in patients treated with baricitinib than placebo, yet rates were not significantly higher than those seen with MTX or adalimumab.[2,13]
- The majority of herpes zoster cases were mild to moderate in severity and complicated cases were uncommon.[2,13]
- Increases in LDL cholesterol were one of the most commonly reported ADRs, yet increases in high-density lipoprotein (HDL) cholesterol were also seen with baricitinib so that the mean HDL/LDL ratio was unchanged. Furthermore, there was a significant decrease in the amount

of small and very small LDL particles in RA-BEAM, which are considered the most atherogenic. Few MACE were also observed in the baricitinib clinical programme and no relationship was seen between MACE and increased LDL.[2,13]

- Treatment with baricitinib also resulted in changes to haematology and clinical chemistry analytes. These included mean changes of greater magnitude for some analytes than seen with the active comparators, which are, therefore, likely to be related to the pharmacology of JAK inhibition (such as increases in lipids [including total cholesterol, low-density lipoprotein cholesterol and high-density lipoprotein cholesterol] and creatine phosphokinase).[2,13]

Summary of Network Meta-Analysis

- Network meta-analyses (NMA) were performed in order to assess the relative efficacy of baricitinib compared with the relevant comparators in cDMARD- (including MTX)-IR or anti-TNF-IR patients with moderate to severe RA.
- In the cDMARD-IR population base case analysis at week 24, baricitinib 4 mg QD was found to be associated with statistically significantly higher odds of ACR50 response compared to cDMARD, adalimumab 40 mg, placebo, etanercept and sulfasalazine. No statistically significant differences were found versus any other comparator for the ACR50 outcome, with the exception of certolizumab pegol, in which odds of ACR50 response was found to be significantly in favour of the comparator. Although statistically significant differences were not found versus most biologic comparators, in the majority of cases the point estimate of relative treatment effect was favourable to baricitinib 4 mg.
- In the anti-TNF-IR population base case analysis at week 24, baricitinib (4 mg QD) demonstrated significantly higher ACR50 response rates than the cDMARD comparator. No statistically significant differences were seen versus biologic comparators, with the exception of the comparison of baricitinib 4 mg and barictinib 2 mg to tocilizumab 8 mg, and the comparison of baricitinib 2 mg to rituximab 1000 mg, in which statistically significant treatment effects in favour of the comparator were observed. Versus the other comparators, point estimates in some cases favoured barictinib 4 mg and in other cases favoured the comparator treatment.

Summary of Cost-effectiveness

- A discrete event simulation (DES) economic model, similar to the cost-effectiveness model built by the Assessment Group in the recent NICE Multiple Technology Appraisal in RA (TA375), was developed to evaluate the cost-effectivess of baricitinib in the moderate and severe cDMARD-IR populations, and the rituximab-eligible and -ineligible severe TNFi-IR populations.
- The results of the base case analysis demonstrate that baricitinib represents a cost-effective treatment option in the severe, cDMARD-IR population and the severe, rituximab-ineligible population compared to comparators currently used in UK clinical practice for these patient populations.
- The budget impact of using baricitinib with the PAS resulted in a substantial estimated cost saving in Years 1–5 in both the severe, cDMARD-IR population and moderate to severe, anti-TNF-IR populations. In the former population, cost savings ranged from £[redacted] in Year 1 to £[redacted] in Year 5, and in the latter population £[redacted] in Year 1 to £[redacted] in Year 5.

Conclusions

- Baricitinib, a once daily, oral DMARD, is the first molecule in RA that has demonstrated superiority over a leading biologic, adalimumab, in a head-to-head comparison in patients with an inadequate response to MTX on background MTX.
- In all phase III studies, the primary endpoint of ACR20 was met (at both Week 12 and Week 24), as well as most secondary endpoints. Compared to placebo and active comparators, baricitinib 4 mg QD demonstrated rapid and durable improvements for relevant domains of efficacy across the RA treatment continuum and different patient populations. This leads to greater value when compared to branded biologics and cDMARDs, which supports its use in patients who can no longer be sufficiently controlled after initial cDMARDs.
- In the NMA, baricitinib 4 mg QD was found to have comparable efficacy to the majority of bDMARD comparators in both the cDMARD-IR and anti-TNF-IR populations.
- Economic analyses found baricitinib to be a cost-effective treatment option versus treatments currently used in UK clinical practice in the severe, cDMARD-IR population and rituximab-ineligible, severe, TNFi-IR populations.
- The results of the budget impact analysis demonstrate that the introduction of baricitinib may generate significant cost savings to the NHS, presenting further evidence of the value of baricitinib in addition to the evidence of clinical and cost-effectiveness.

1.1 *Statement of decision problem*

This submission addresses the clinical efficacy and safety, the comparative effectiveness and cost-effectiveness of baricitinib 4 mg QD, as monotherapy or in combination with methotrexate, in adult patients with moderate to severe RA for whom methotrexate, cDMARDs or bDMARDs have been inadequately effective or not tolerated. The decision problem addressed is consistent with the final NICE scope for this appraisal, as outlined in Table 1.

Table 1. The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with moderate to severe, active rheumatoid arthritis whose disease has responded inadequately to, or who are intolerant of one or more disease-modifying antirheumatic drugs (DMARDs), including conventional or biologic DMARDs.	This submission considers the use of baricitinib in four patient populations: 1. Patients with severely active RA who have been previously treated with and failed on cDMARDs (cDMARD-IR) including methotrexate; 2. Patients with severely active RA who have been previously treated with and failed on TNFis (TNFi-IR), who are ineligible for treatment with rituximab; 3. Patients with severely active RA who have been previously treated with and failed on TNFis (TNFi-IR) and who are eligible for treatment with rituximab; 4. Patients with moderately active RA who have been previously treated with and failed on cDMARDs (cDMARD-IR).	NA
Intervention	Baricitinib monotherapy or in combination with methotrexate	Baricitinib 4 mg QD, as monotherapy or in combination with methotrexate	Note that clinical data is also provided for baricitinib 2 mg and for the economic evaluation a scenario analysis is provided in which the dose is tapered from 4 mg to 2 mg

Comparator (s)	People with moderate active rheumatoid arthritis: • Combination therapy with conventional DMARDs (including methotrexate and at least one other DMARD, such as sulfasalazine and leflunomide) • Conventional DMARD monotherapy with dose escalation • Best supportive care (only where conventional DMARDs are not appropriate due to intolerance) People with severely active rheumatoid arthritis that has not responded adequately to therapy with conventional DMARDs only: • Biologic DMARDs in combination with methotrexate (adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab, abatacept) • Adalimumab, etanercept, certolizumab pegol, or tocilizumab (each as monotherapy) People with severely active	Moderate active RA: • Best supportive care (with continued cDMARDs) Severely active RA that has not responded adequately to therapy with cDMARDs only: • Adalimumab, etanercept biosimilar, infliximab biosimilar, abatacept, golimumab, tocilizumab, certolizumab pegol (all in combination with MTX) Severely active rheumatoid arthritis that has not responded adequately to therapy with DMARDs including at least one TNF inhibitor: • Rituximab-eligible patients: rituximab • Rituximab-ineligible patients: tocilizumab, abatacept, certolizumab pegol, golimumab, etanercept biosimilar, infliximab biosimilar and adalimumab.	Baricitinib monotherapy is not compared to bDMARDs in combination with MTX as insufficient baricitinib monotherapy data were available from the baricitinib clinical trial programme. It may be noted that in the recent MTA regarding the use of biologics in DMARD-naïve and cDMARD-IR patients (TA375), the Committee agreed that the minority of (cDMARD-IR) patients with severe active rheumatoid arthritis who could not tolerate methotrexate should not be treated differently from other people with severe disease, as far as possible. The Committee concluded that biologic DMARDs should be recommended as a cost-effective use of NHS resources when used as monotherapy for severe active disease previously treated with DMARDs, where the marketing authorisation of the bDMARD allows for this recommendation to be made.

	rheumatoid arthritis that has not responded adequately to therapy with DMARDs including at least one TNF inhibitor: • Rituximab in combination with methotrexate • When rituximab is contraindicated or withdrawn due to adverse events: ○ Abatacept, adalimumab, certolizumab pegol, etanercept, infliximab, tocilizumab, or golimumab, each in combination with methotrexate ○ Adalimumab, etanercept or certolizumab pegol (each as monotherapy)		
Outcomes	The outcome measures to be considered include: • disease activity • physical function • joint damage • pain • mortality • fatigue • radiological progression • extra-articular manifestations of the disease	• Disease activity (ACR20; ACR50; ACR70; EULAR Response; DAS28-hsCRP; DAS28-ESR; SDAI; CDAI) • Physical function (MJS, HAQ-DI) • Joint damage (mTSS) • Pain (captured as part of the ACR core set) • RA-related mortality • Fatigue (FACIT–F; Worst Tiredness Score) • Radiological progression	NA

	• adverse effects of treatment • health-related quality of life	(mTSS) • Extra-articular manifestations of the disease (captured under safety reporting) • Adverse effects of treatment • Health-related quality of life (EQ-5D-5L; SF-36v2) • WPAI-RA	
Economic analysis	• The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. • The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. • Costs will be considered from an NHS and Personal Social Services perspective. • The availability of any patient access schemes for the intervention or comparator technologies will be	• Cost-effectiveness analysis results presented as incremental cost-effectiveness ratios (ICERs) in terms of cost per QALY • Lifetime time horizon: a lifetime time horizon is consistent with the AG model in TA375.[14] As a progressive life-long condition with no cure, and taking into account the mean age of patients entering the model (52.11), a 45 year time horizon is appropriate to capture the lifetime of patients • Costs were considered from an NHS perspective • Patient access schemes (PAS) for baricitinib were taken into account. The PAS for certolizumab pegol was also taken into account as details of this PAS are publically available. • Biosimilar versions of etanercept and infliximab were considered in the analysis	• Costs were considered from an NHS perspective only, consistent with the Assessment Group's (AG) model in TA375[14]

	taken into account. • The availability and cost of biosimilar products of etanercept and infliximab should be taken into account.		
Subgroups to be considered	• If the evidence allows the following subgroups will be considered. These include people with moderate disease activity (DAS28 between 3.2 and 5.1) and severely active disease (DAS28 greater than 5.1).	Results for the primary endpoint of each of the three trials (ACR20 response at Week 12) are presented for the following subgroups: RA-BEAM (Section 4.8.2): • Moderate disease activity (DAS-hsCRP ≤5.1) and severe disease activity (DAS28-hsCRP >5.1) RA-BUILD (Section 4.8.1): • Moderate disease activity (DAS-hsCRP ≤5.1) and severe disease activity (DAS28-hsCRP >5.1) RA-BEACON (Section 4.8.3): • Moderate disease activity (DAS-hsCRP ≤5.1) and severe disease activity (DAS28-hsCRP >5.1) • <3 previous bDMARDs used and ≥3 previous bDMARDs used In the cDMARD-IR population, the economic analysis presents results separately for moderate patients and severe patients. In the bDMARD-refractory population, the economic analysis presents results for severe patients.	NA

Special considerations including issues related to equity or equality	NA	There are no equality issues arising in relation to this technology.	NA

Abbreviations: NICE = National Institute for Health and Care Excellence, DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, bDMARD = biologic DMARD, OD = once daily, TNF = tumour necrosis factor inhibitor, ACR = American College of Rheumatology, ACR20/50/70 = 20/50/70% improvement in ACR criteria, EULAR = European League Against Rheumatism, EULAR = EULAR response index, DAS28 = Disease Activity Score modified to include the 28 diarthrodial joint count, hsCRP = high-sensitivity C-reactive protein, ESR = erythrocyte sedimentation rate, SDAI = Simplified Disease Activity Index, CDAI = Clinical Disease Activity Index, MJS = morning joint stiffness, WJP = worst joint pain, FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, mTSS = modified Total Sharp Score, WPAI-RA = Work Productivity and Activity Index-Rheumatoid Arthritis, EQ-5D-5L = EuroQoL 5 dimensions–5 levels, HAQ-DI = Health Assessment Questionnaire-Disability Index, SF-36v2 = Medical Outcomes Study 36-Item Short Form Health Survey Version 2 Acute, NHS = National Health Service, QALY = quality-adjusted life year.

1.2 *Description of the technology being appraised*

A description of the technology being appraised (baricitinib [Olumiant®]) is provided in Table 2 below.

Table 2. Summary of baricitinib

UK approved name and brand name	Baricitinib (Olumiant®)
Marketing authorisation/CE mark status	Baricitinib is anticipated to receive marketing authorisation with the European Medicines Agency (EMA) and should therefore be licensed for marketing in the European Union.
Indications and any restriction(s) as described in the summary of product characteristics	Olumiant is indicated for the treatment of moderately to severely active RA in adult patients who have responded inadequately to, or who are intolerant to one or more DMARDs. Olumiant may be used as monotherapy or in combination with methotrexate.
Method of administration and dosage	The recommended dose of Olumiant is 4 mg QD. A dose of 2 mg QD is appropriate for patients such as those aged ≥ 75 years and may be appropriate for patients with a history of chronic or recurrent infections. The recommended dose is 2 mg QD in patients with creatinine clearance between 30 and 60 mL/min. Olumiant is not recommended for use in patients with creatinine clearance < 30 mL/min. A dose of 2 mg QD may also be considered for patients who have achieved sustained control of disease activity with 4 mg QD and are eligible for dose tapering. Olumiant is to be taken once daily with or without food and may be taken at any time of the day.

Abbreviations: EMA = European Medicines Agency, QD = once daily.
Source: Olumiant® Summary of Product Characteristics.[2]

1.3 *Summary of the clinical effectiveness analysis*

The clinical efficacy of baricitinib in patients with moderately to severely active RA with an inadequate response to, or who are intolerant of DMARDs, was assessed in four pivotal phase III studies in different RA patient populations (three of which are presented in this submission). All three phase III pivotal studies were randomised, double-blind, placebo- or active-controlled trials:

RA-BEAM in MTX-refractory patients (Section 4.7.1)[9]

RA-BUILD in cDMARD-refractory patients (Section 4.7.2)[11]

RA-BEACON in TNFi-refractory patients (Section 4.7.3)[12]

RA-BEGIN in cDMARD-naïve patients (*unlicensed indication, not presented in this submission*)

RA-BEYOND is a 4-year extension study that allows patients completing treatment in either RA-BEAM, RA-BUILD or RA-BEACON to continue treatment. RA-BEYOND is designed to evaluate the long-term safety and efficacy of baricitinib (Section 4.7.4).[15]

Together, RA-BEAM, RA-BUILD and RA-BEACON included 2,516 patients with moderately to severely active RA in a variety of patient populations across the RA treatment spectrum, including cDMARD-IR (including MTX) and bDMARD-IR patients. In all three trials, there were prespecified subgroup efficacy analyses for patients with either moderate disease activity (DAS-hsCRP ≤5.1) or severe disease activity (DAS28-hsCRP >5.1), whilst RA-BEACON also included a prespecified subgroup analysis for patients with either <3 previous bDMARDs used or ≥3 previous bDMARDs used.

Across all trials, a large number of relevant outcomes were investigated, spanning disease activity, physical function, joint damage, pain, radiological progression and health-related quality of life (HRQOL). The primary endpoint was the ACR20 criteria at Week 12, a multidimensional assessment of disease activity, pain and physical function, which is used in almost all published studies assessing the efficacy of RA interventions.[16] RA-BEAM randomised patients to baricitinib 4 mg QD, adalimumab or placebo, whilst RA-BUILD and RA-BEACON randomised patients to baricitinib 2 mg QD, baricitinib 4 mg QD or placebo.[9,11,12]

Compared to placebo and each active comparator, baricitinib 4 mg demonstrated rapid and durable improvements for relevant domains of efficacy across the RA treatment continuum and different patient populations. Furthermore, RA-BEAM was the first head-to head study where an RA therapy achieved statistical superiority over adalimumab in combination with MTX. Importantly, adalimumab was used at its approved dose on background MTX, which is the setting in which it is most effective, thereby reflecting the optimal use of adalimumab.[9]

- RA-BEAM met its primary objective (ACR20 response rate), with a response rate for baricitinib 4 mg QD of 69.6%, which was significantly higher than the placebo arm (40.2%, $p \leq 0.05$). Baricitinib was also statistically superior to adalimumab + MTX 61.2%, $p \leq 0.05$) and met all secondary objectives.[9]
- Similarly, RA-BUILD met its primary objective (ACR20 response rate), with a response rate for baricitinib 4 mg QD of 61.7%, which was statistically significantly higher than the placebo arm (39.5%, $p \leq 0.001$). Additionally, baricitinib 2 mg QD also had a statistically significantly higher response rate of 65.9% compared to placebo ($p \leq 0.001$). RA-BUILD also met all secondary objectives.[11]

- RA-BEACON met its primary objective (ACR20 response rate), with a response rate for baricitinib 4 mg QD of 55.4%, which was statistically significantly higher than the placebo arm (27.3%, $p \leq 0.001$). Additionally, baricitinib 2 mg QD also had a statistically significantly higher response rate of 48.9% compared to placebo ($p \leq 0.001$). RA-BEACON also met the first two secondary objectives (change in baseline of HAQ-DI at Week 12 and DAS28-hsCRP response rate at Week 12).[12]

In studies measuring radiographic outcomes (RA-BEAM and RA-BUILD), progression of joint damage, as measured by mTSS, was significantly reduced in patients treated with baricitinib compared to placebo, and with a similar profile to adalimumab.

Patient reported symptoms were statistically significantly improved versus placebo across all three trials, as determined by HAQ-DI, duration/severity of morning joint stiffness (MJS) and FACIT-F. Compared to adalimumab + MTX, baricitinib + MTX was statistically significantly superior for severity of MJS (as early as Week 4 and through Week 12), worst tiredness (as early as Week 8 and through Week 12) and fatigue (FACIT-F scores, as early as Week 8 and at Week 52).[9,11]

In addition to the results presented above for the whole trial populations, in all three trials subgroups of patients with either moderate of severe disease, and in RA-BEACON, patients who had received either ≥3 or <3 previous bDMARDs, were evaluated. The subgroup analyses demonstrate the efficacy of baricitinib in all patient subgroups, regardless of disease activity at baseline or number of previous bDMARDs used in the bDMARD-IR population.[9,11,12]

Comparative Effectiveness: Network Meta-analysis

Network meta-analyses (NMA) were performed in order to assess the relative efficacy of baricitinib compared with the relevant comparators in cDMARD- (including MTX)-IR or anti-TNF-IR patients with moderate to severe RA. Both ACR response and EULAR response endpoints were selected for inclusion in the NMA; both were analysed as ordinal outcomes, using probit models. Fixed-effects (FE) and random-effects (RE) simultaneous Bayesian models were fitted for both the cDMARD- and anti-TNF-IR populations, however, the RE models for the anti-TNF-IR population were unstable and did not converge, hence results from the fixed effects model are presented for this population.

In the cDMARD-IR population base case analysis at week 24, baricitinib 4 mg was found to be associated with statistically significantly higher odds of ACR50 response compared to cDMARD, adalimumab 40 mg, placebo, etanercept and sulfasalazine. No statistically significant differences were found versus any other comparator for the ACR50 outcome, with the exception of the comparison of baricitinib 4 mg to certolizumab pegol, in which odds of ACR50 response was found to be significantly in favour of the comparator. This pattern of results was also observed for baricitinib 2 mg. For baricitinib 4 mg, although statistically significant differences were not found versus most biologic comparators, in the majority of cases the point estimate of relative treatment effect was favourable to baricitinib 4 mg.

In the anti-TNF-IR population base case analysis at week 24, baricitinib demonstrated significantly higher ACR50 response rates than the cDMARD comparator. No statistically significant differences were seen versus biologic comparators, with the exception of the comparison of baricitinib 4 mg and barictinib 2 mg to tocilizumab 8 mg, and the comparison of baricitinib 2 mg to rituximab 1000 mg, in which statistically significant treatment effects in favour of the comparator were observed. Versus the other comparators, point estimates in some cases favoured barictinib 4 mg and in other cases favoured the comparator treatment.

Safety

In addition to the four Phase III trials described above, the overall safety and tolerability of baricitinib was evaluated in a number of integrated safety analyses comprising data from baricitinib Phase I–III RA trials, as well as trials of baricitinib in the indications of psoriasis and diabetic nephropathy. The analysis set used in this submission to provide an overview of the safety and tolerability of baricitinib in RA was the baricitinib 4 mg RA PC analysis set, which evaluated the safety of baricitinib 4 mg QD compared to placebo, and comprised data from three Phase II studies (JADA, JADC and JADN) and the three Phase III studies presented in this submission.[13]

Treatment with baricitinib was found to be generally well-tolerated, with similar incidences of the following outcomes between baricitinib 4 mg QD and placebo over 24 weeks: overall treatment-emergent adverse events (69.7% versus 61.6%), severe treatment-emergent adverse events (5.3% versus 4.0%), serious adverse events (5.3% versus 4.7%), temporary interruptions due to adverse events (10.9% versus 8.3%), permanent discontinuation (4.7% versus 3.3%), death (0.1% versus 0.1%).[2,13]

Rates of MACE, infection, malignancies and gastrointestinal perforation observed with baricitinib did not appear to exceed background rates in the RA population. The incidence of herpes zoster of baricitinib was increased compared to adalimumab in the trial, and historical data of TNF-inhibitors and tocilizumab.[2,13]

The most commonly reported adverse reactions in ≥2% of patients treated with baricitinib monotherapy or in combination with cDMARDs included increased LDL cholesterol, upper respiratory tract infections and nausea. However, the majority of all ADRs were mild to moderate in severity.[13]

Treatment with baricitinib also resulted in increases in lipids [including total cholesterol, low-density lipoprotein cholesterol and high-density lipoprotein cholesterol] and creatine phosphokinase). Overall, the LDL/HDL ratio remained unchanged after baricitinib treatment. Significant increases in the number of large LDL-C particles (considered less atherogenic than small particles) and statistically significant decreases in the number of small, medium-small and very small LDL-C particles were observed in baricitinib 4 mg (QD) and adalimumab as compared to placebo.[2]

As described in the SmPC,[2] baricitinib is associated with an increased rate of infections such as URTIs compared to placebo. In treatment naïve patients, combination with methotrexate resulted in increased frequency of infections compared to baricitinib monotherapy.

Although baricitinib was associated with a higher incidence of SAEs compared to adalimumab through 52 weeks in RA-BEAM, their AE profiles were similar across clinically significant categories of risk including MACE, malignancies, hypercholesterolemia, serious infections and herpes zoster.[13]

1.4 *Summary of the cost-effectiveness analysis*

The results of the base case analysis demonstrate that baricitinib represents a cost-effective treatment option in the severe, cDMARD-IR population and the severe, rituximab-ineligible population compared to comparators currently used in UK clinical practice for these patient populations.

The summary results of the base case analysis for each of the patient populations of interest can be found in the following tables:

Table 3: Severe, cDMARD-IR population

Table 4: Rituximab-ineligible, severe, anti-TNF-IR population

Table 5: Rituximab-ineligible, severe, anti-TNF-IR population

Table 6: Moderate, cDMARD-IR population

Table 3. Base case cost-effectiveness results for the severe, cDMARD-IR population

Technology sequence	Total costs (£)	Total QALYs	Total LYG	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (per QALY)	ICER (£) incremental (per QALY)
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Referent	Referent
ETN-bMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	£18,400	£18,400
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
IFX-bMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN-b = etanercept biosimilar, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX-b = infliximab biosimilar, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

Table 4. Base-case cost-effectiveness results for the severe, anti-TNF-IR (rituximab-ineligible) population

Technology sequence	Total costs (£)	Total QALYs	Total LYG	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (QALYs)	ICER (£) incremental (QALYs)
BAR4MTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	Referent	Referent
GOLMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	Dominated	Dominated
ETN-bMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£19,874	Ext Dominated
CTZMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£16,201	£16,201
ADAMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£27,008	Dominated
IFX-bMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£34,942	Dominated
TCZMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£36,757	Dominated
ABTSMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£484,782	Dominated

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.

Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN-b etanercept biosimilar, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX-b = infliximab biosimilar, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

Table 5. Base-case cost-effectiveness results for the severe, anti-TNF-IR (rituximab-eligible) population

Technologies	Total costs (£)	Total QALYs	Total LYG	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (per QALY)
RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	5.35	[redacted]	[redacted]	-
BAR4MTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	5.25	[redacted]	[redacted]	Dominated

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: BAR4 = baricitinib (4 mg QD), ICER = incremental cost-effectiveness analysis, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

Table 6. Base-case cost-effectiveness results for the moderate, cDMARD-IR population

Technology sequence	Total costs (£)	Total QALYs	Total LYG	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (QALYs)
COMB+MTX+Pall	[redacted]	[redacted]	16.04	[redacted]	[redacted]	-
BAR4MTX+COMB+MTX+Pall	[redacted]	[redacted]	16.03	[redacted]	[redacted]	£37,420

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: BAR4 = baricitinib 4 mg (QD), COMB = combination cDMARDs, ICER = incremental cost-effectiveness ratio, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, LYG = life years gained

Budget Impact

The budget impact of using baricitinib with the patient access scheme (PAS) resulted in a substantial estimated cost saving in Years 1–5 in both the severe, cDMARD-IR population and moderate to severe, TNFi-IR populations. In the former population, cost savings ranged from £[redacted] in Year 1 to £[redacted] in Year 5, and in the latter population £[redacted] in Year 1 to £[redacted] in Year 5. These results demonstrate the significant cost savings that could be achieved by the NHS through the introduction of baricitinib, presenting further evidence of the value of baricitinib in addition to the evidence of clinical and cost-effectiveness.

2 The technology

2.1 *Description of the technology*

Brand name: Olumiant®

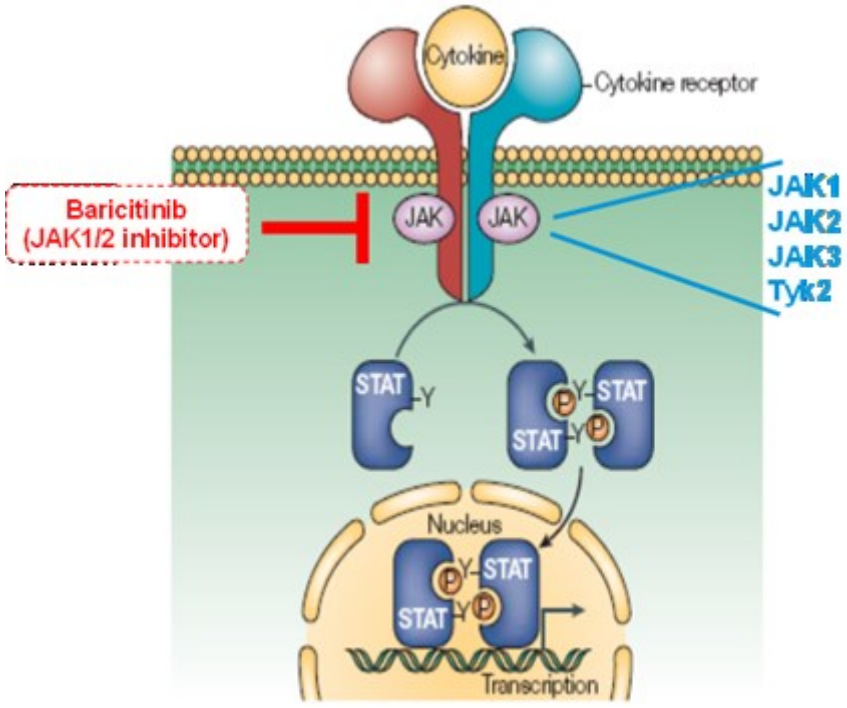
UK approved name: Baricitinib

Therapeutic and pharmacological class: Oral Janus kinase (JAK) inhibitor

Mechanism of action:

Janus kinases (JAKs) are a family of tyrosine kinases with four members: JAK1, JAK2, JAK3 and TYK2.[17] JAKs, and their associated signal transducers and activators of transcription (STATs), are the major intracellular pathway for the control of signalling via Type I and Type II receptor-binding cytokines (JAK-STAT pathway).[18] Transmembrane cytokine receptors lack intrinsic enzymatic activity and instead, upon binding of a cytokine to the receptor, STATs are activated by the phosphorylation of a single tyrosine residue by receptor-associated JAKs.[19] Activated STAT dimers translocate to the nucleus where they induce the expression of multiple genes important for immune cell activation, localisation, survival and proliferation (Figure 1).[18]

Figure 1. The JAK-STAT pathway



Abbreviations: JAK = Janus kinase, STAT = Signal Transducer and Activator of Transcription, P = phosphate group.

Baricitinib is an oral, reversible JAK inhibitor with selectivity for JAK1 and JAK2.[1] Baricitinib transiently occupies the ATP-binding pocket of JAK1 and JAK2, preventing the phosphorylation of STATs and thus disrupting cytokine signalling.[1] Different cytokines utilise varying combinations of JAKs for signal transduction, and many of the pro-inflammatory cytokines implicated in the pathogenesis of RA signal via the JAK/STAT pathway, including interleukin-6 (IL-6) (JAK1/JAK2), granulocyte macrophage colony-stimulating factor (GM-CSF) (JAK2/JAK2), and interferons (JAK1/JAK2, JAK1/TYK2). JAK3, on the other hand, is primarily involved in T cell and natural killer (NK) cell activation, maturation and immune function, whilst a specific role for TYK2 has not been definitively established. As such, inhibition of JAK1 and JAK2 signalling can thereby reduce inflammation, cellular activation and proliferation of key immune cells in patients with RA.[20-22] Baricitinib is the first JAK1/2 inhibitor licensed for the treatment of moderate to severe RA in the European Union (EU).

2.2 *Marketing authorisation/CE marking and health technology assessment*

Baricitinib is indicated for the treatment of moderate to severe active rheumatoid arthritis in adult patients who have responded inadequately to, or who are intolerant to one or more disease-modifying antirheumatic drugs. Baricitinib may be used as monotherapy or in combination with methotrexate

The recommended dose of baricitinib is 4 mg once daily. A dose of 2 mg once daily is appropriate for patients such as those aged ≥75 years and may be appropriate for patients with a history of chronic or recurrent infections. A dose of 2 mg once daily may also be considered for patients who have achieved sustained control of disease activity with 4 mg once daily and are eligible for dose tapering. For patients with renal impairment, as determined by an estimated glomerular filtration rate (eGFR) between 30 and 60 ml/min/1.73 m^2, the dose is 2 mg QD. For patients taking Organic Anion Transporter 3 (OAT3) inhibitors with strong inhibition potential (such as probenecid), the recommended dose is 2 mg QD.

An application for a marketing authorisation in this indication in Europe was submitted to the European Medicines Agency (EMA) in January 2016 and a positive opinion from the Committee for Medicinal Products for Human Use (CHMP) was adopted on 15th December 2016. The Summary of Product Characteristics (SmPC) for baricitinib, which details the anticipated licensed indication for baricitinib, is provided in the reference pack accompanying this submission.[2] The European Public Assessment Report (EPAR) will also be provided if available at the time of submission.

Baricitinib is not currently licensed for any other indications.

Baricitinib is anticipated to be launched in the UK on 4th April 2017. It is anticipated that Eli Lilly will prepare submissions to the Scottish Medicines Consortium (SMC; expected [redacted]) and National Centre for Pharmacoeconomics (NCPE; expected [redacted]).

2.3 *Administration and costs of the technology*

Details of the treatment regimen, including the method of administration, healthcare resource use and costs associated with the technology are provided in Table 7. A submission has been made to PASLU for a Simplified discount patient access scheme.

Table 7. Costs of the technology being appraised

Pharmaceutical formulation	Olumiant® 4 mg film-coated tablets Olumiant® 2 mg film-coated tablets
Acquisition cost (excluding VAT)*	List price per pack: 2 mg x 28 pack: £805.56 4 mg x 28 pack: £805.56 2 mg x 84 pack: £2,416.68 4 mg x 84 pack: £2,416.68 PAS price per pack: [redacted]
Method of administration	Olumiant® is to be taken orally with or without food and may be taken at any time of the day
Doses	The recommended dose of Olumiant is 4 mg once daily. A dose of 2 mg once daily is appropriate for patients such as those aged ≥ 75 years and may be appropriate for patients with a history of chronic or recurrent infections
Dosing frequency	Olumiant® is to be taken once daily (QD).
Average length of a course of treatment	RA is a chronic condition and therefore treatment is continuous; no stopping rule is specified in the SmPC
Average cost of a course of treatment	Annual cost per patient (to nearest whole Great British Pound Sterling): List price - £10,501 PAS price - [redacted]
Anticipated average interval between courses of treatments	N/A – continuous treatment
Anticipated number of repeat courses of treatments	N/A – continuous treatment
Dose adjustments	A dose of 2 mg once daily may be considered for patients who have achieved sustained control of disease activity with 4 mg once daily and are eligible for dose tapering
Anticipated care setting	Olumiant® treatment should be initiated and supervised by an experienced physician/rheumatologist. It is anticipated that Olumiant® maintenance treatment would be self-administered at home by patients

2.4 *Changes in service provision and management*

Changes to service provision and management are not expected.

The SmPC for baricitinib states that before initiating treatment with baricitinib patients should be screened for tuberculosis (TB) and that screening for viral hepatitis should be performed in accordance with clinical guidelines. The SmPC also provides monitoring guidance for laboratory measures. Absolute neutrophil count (ANC), absolute lymphocyte count (ALC), haemoglobin levels and hepatic transaminases should be measured prior to treatment initiation and thereafter according to routine patient management. Finally, lipid parameters should be measured 12 weeks after initiation of treatment and thereafter according to international clinical guidelines for hyperlipidaemia. Testing for tuberculosis and hepatitis in advance of therapy initiation is in line with testing requirements for bDMARDs licensed for the treatment of moderate to severe RA. The other tests and monitoring requirements would be expected to form a part of routine clinical management and would therefore also not be associated with requirements for new service provision over and above the current standard of care, as described in Section 5.5.

No additional infrastructure in the NHS is required over and above the current standard of care for baricitinib.

In line with several other bDMARDs licensed for the treatment of moderate to RA, baricitinib is indicated for the treatment of moderately to severely active rheumatoid arthritis in adult patients who have responded inadequately to, or who are intolerant to one or more DMARDs. Baricitinib may be used as monotherapy or in combination with methotrexate.[2]

2.5 *Innovation*

Despite the approvals of a number of novel biologic DMARDs (bDMARDs) in recent years, more than half of moderate to severe RA patients still do not achieve a satisfactory response following treatment.[4] In patients treated with tumour necrosis factor inhibitor (TNFi) biologics in particular, efficacy may diminish over time due to the production of anti-drug antibodies to the injected therapeutic protein antibody,[5] and a systematic literature review of TNFis found that dose escalation is frequently required to regulate disease levels.[23] Biologic DMARDs are also associated with an increased risk of opportunistic infection (specifically anti-TNF biologics[24]) and the reactivation of latent tuberculosis;[25] in many cases leading to treatment discontinuation.[26, 27] Furthermore, due to their molecular composition, all bDMARDs must be administered via the parenteral route (subcutaneous injection or intravenous infusion)—methods of administration that can be associated with painful side injection site reactions .[6,7] Furthermore, in a 2008 survey of RA patient treatment preferences, 15% of patients were found to be needle phobic.[28] In a survey of 380 RA patients in the U.S., route of administration was found to be an important factor in patient preference for RA treatments, with a majority of patients preferring the oral route of administration over parenteral routes.[29] Whilst the impact of oral therapy on adherence compared to parenteral administration is captured in the economic model presented in Section 5.2, and is therefore reflected in the QALY calculations, the QALY decrement associated with the side effects of injection such as burning or stinging is not. A recent study reported that 21–35% of patients treated with bDMARDs discontinued parenteral RA therapy within a year,[30] and an observational study of RA patients found that dropouts who discontinued use of intensive conventional DMARD (cDMARD) therapy were significantly more disabled, and more likely to develop associated conditions, than those who adhered to cDMARD therapy.[31]

Combined with the severe impact of RA on patient health-related quality of life (HRQOL) and economic burden on society (see Section 3.2), there is a clear unmet need for new therapies offering robust efficacy and safety profiles with more convenient routes of administration. As an orally-administered small molecule, baricitinib does not induce the production of anti-drug antibodies, which cause efficacy to decline over time, as supported by sustained ACR categorical outcomes in Section 4.7. Baricitinib has selectivity for JAK1 and JAK2, which are associated with many of the pro-inflammatory cytokines implicated in the pathogenesis of RA, including IL-6, GM-CSF and interferons, whilst having a low affinity for JAK3.[1] As a once-daily oral therapy, baricitinib could have a substantial impact on the large number of patients who may experience painful side effects with, and potentially discontinue currently available bDMARD therapies.

Four phase III trials have demonstrated that baricitinib, compared with standards of care, is associated with significant improvements in the signs and symptoms of RA, physical function, patient reported outcomes and HRQOL outcomes. These benefits have been demonstrated across four different RA patient populations: DMARD-naïve (*unlicensed and hence not considered in this submission*),[10] and inadequate responders to: methotrexate,[9] cDMARDs[11] and anti-TNF bDMARDs.[12] Baricitinib is the first JAK inhibitor to be licensed in Europe. As such, baricitinib represents a therapy with a novel mechanism of action that may provide a valuable extension to the armamentarium available to clinicians compared to currently available treatment options and may fulfil a particular unmet need in patients who have not responded to the currently available treatment options, in particular as an oral option for escalation therapy in cDMARD-IR patients. Therefore, baricitinib has the potential to be considered an innovative step-change in the RA treatment paradigm that will fulfil many of the unmet needs of patients with RA.

3 Health condition and position of the technology in the treatment pathway

- RA is a chronic autoimmune disease characterised by progressive pain and stiffness of the joints.[1,2]
- RA places a significant burden on patient HRQOL, with severe impacts on employment and self-esteem, whilst also contributing to a range of comorbidities such as cardiovascular and gastrointestinal diseases.[32-34]
- Additionally, all therapies for moderate-to-severe RA are administered via subcutaneous injection or IV infusion, which may be unpleasant for patients due to common adverse reactions, such as burning or stinging.[6,7]
- RA also decreases patient life expectancy by 3–5 years in those on bDMARDs,[35] and up to 10–15 years in those with the most severe forms of the disease.[32]
- Relevant NICE guidance and pathways include CG79 (Rheumatoid arthritis in adults: management) and multiple technology appraisal TA375 (Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with DMARDs or after conventional DMARDs only have failed).[14,36]
- The established NICE treatment pathway for RA begins with cDMARD monotherapy or combination therapy with MTX, followed by bDMARD monotherapy or combination therapy with MTX in those patients who do not respond adequately to treatment.[37]
- Additionally, established international guidelines from EULAR and ACR describe recommended treatment practices, and are broadly in line with the NICE treatment pathway.[38,39]
- In accordance with the NICE scope, potential positions in the RA treatment pathway for baricitinib include cDMARD-IR patients with moderate RA, cDMARD-IR patients with severe RA, bDMARD-IR (including at least one TNFi) and in patients for whom rituximab is contraindicated or withdrawn due to adverse events.[40]
- Issues with current clinical practice include patients with moderate disease activity who do not respond adequately to cDMARDs, for whom under current NICE guidelines receive no further treatment beyond best supportive care, unless their disease becomes severe.
- There are no equality considerations arising in relation to baricitinib.

3.1 *Overview of Rheumatoid Arthritis*

Rheumatoid arthritis (RA) is an autoimmune disease characterised by chronic systemic inflammation. It primarily affects the synovial joints, such as the hands, wrists and feet, resulting in pain and stiffness. Progressive joint damage can lead to deformities and loss of function.[41] In some cases, RA can contribute to an increased risk of cardiovascular disease, infections, malignancies and mental health disorders.[32]

RA is variable in its severity, ranging from mild, limited disease to severely active disease. Mild RA typically results in minimal joint destruction, whilst moderate-to-severe RA causes systemic inflammation and associated fatigue, pain and joint stiffness.[42] Disease severity is measured using the disease activity score (DAS28), a composite measure that includes an assessment of 28 joints for swelling/tenderness, the patient's assessment of health and erythrocyte sedimentation rate/C-reactive protein (DAS28(ESR) and DAS28(hsCRP), respectively). A DAS28 of less than 3.2 indicates low disease activity, between 3.2 and 5.1 indicates moderate activity, and a score greater than 5.1 indicates high activity.[43] For the majority of patients, their disease remains mild with occasional flare-ups of higher disease activity. However, for some patients the disease may be active and progressive.[32] As such, HRQOL may be significantly compromised by RA.

Early signs and symptoms of RA include warm, swollen joints and stiffness, particularly in the morning. Left untreated, synovial membranes become inflamed and cause damage to the surrounding cartilage, ligaments, tendons and bone—ultimately resulting in joint weakness and deformity.[44] The initial signs and symptoms of RA are reversible, but joint damage and the associated disability are not. RA typically progresses through four stages: early, moderate, severe and end stage/terminal. Patients with moderate RA typically present some bone density loss and experience painful joint stiffness, particularly in the morning, as a result of the systemic inflammation. They may also present with fatigue and have evidence of anaemia. For patients with severe RA the systemic inflammation is widespread and destructive: bone density loss is significant, as is joint destruction, and they typically experience significant morning joint stiffness, anaemia and joint swelling and pain.[45] Management of RA aims to suppress disease activity and induce remission, prevent the development of irreversible joint damage and, in more severe disease, maintain quality of life and address comorbidities associated with the condition.[38]

RA affects approximately 450,000 people in the UK as a whole, with a prevalence of 0.86% and incidence of 0.47 per 1,000 person-years.[46,47] An alternative estimate from the National Rheumatoid Arthritis Society (NRAS) suggests that there may be as many as 690,000 people living with RA in the UK.[48] Global estimates from cohort studies suggest that 47–53% of patients with early RA develop moderate to severe disease over a period of 5 years.[49] In the UK, approximately 15% (~60,000) of RA patients have severe disease.[14,47] Around 12,000 new cases are diagnosed each year in the UK.[36] The disease is more prevalent in women than men, with 2–3 times as many cases in women.[50] Whilst the disease can develop at any age, the typical age of onset in the UK is approximately 40–70 years, peaking in the 70s.[36] Therefore, moderate to severe RA, with its impacts on fatigue and joint stiffness, has a particularly disruptive effect on the working lives of patients, with many seeking early retirement due to the condition.[51]

3.2 *Impact of Rheumatoid Arthritis on Patients, Carers and Society*

RA has a highly detrimental impact on patient HRQOL. The unpredictability of symptoms and the daily presence of pain and fatigue can result in disruption to the ability to perform normal activities and attend work.[52] The occurrence of a flare-up in disease activity can result in difficulty maintaining normal working hours and physically demanding work, where there is repetitive stress on joints, can be challenging to maintain for RA patients.[53] A 2007 survey of 782 RA patients in the UK by the National Rheumatoid Arthritis Society (NRAS) found that approximately one-third of people stop working within two years of disease onset, with this figure rising to almost two-thirds within six years. At least 10% of RA patients become severely disabled even with full progression through the recommended RA treatment pathway, with severe impacts on their personal and working lives.[52]

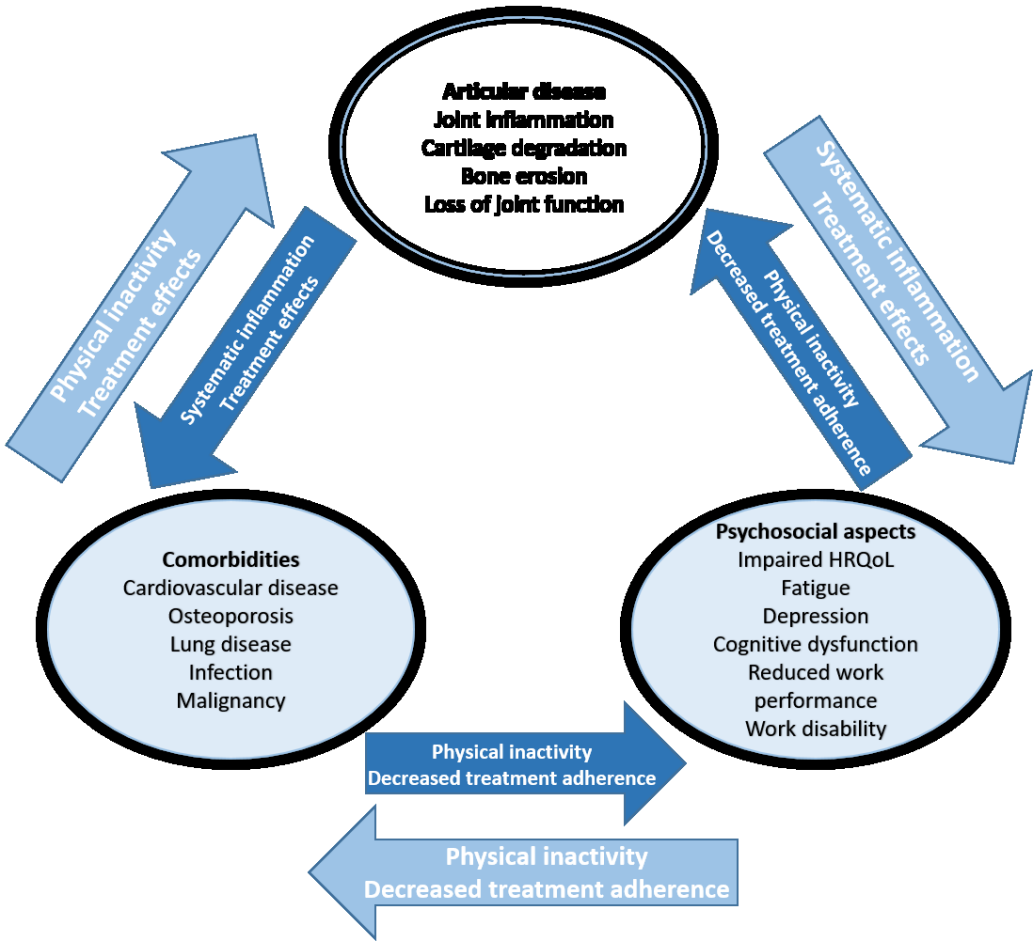
As a lifelong disease, RA also imposes a significant psychological burden on patients who know they face the prospect of having to deal with the condition in the long-term. Knowledge that effective treatment options that can be taken in a sustainable manner are available is therefore valuable to patients. The pain and fatigue of RA can frequently result in anxiety and depression, whilst negative body image and low self-esteem–due to swelling and deformity–is common and associated with low HRQOL.[33,34] A 2013 NRAS survey concerning emotions, relationships and sexuality in 1,343 RA patients found that RA inflicts a pervasive and persistent impact on many aspects of emotional life, with 92% of responders describing frustration about the activities they cannot do because of RA. Furthermore, the vast majority of responders said that RA negatively affected their confidence, mood, self-esteem and anxiety.[54]

RA-associated morbidities impose an additional burden on the patient. Cardiovascular, respiratory and gastrological disorders, as well as infections and cancers, are observed more frequently in RA patients.[32] RA patients have an approximately two-fold higher risk of cardiovascular disease than the general population,[55] whilst respiratory conditions account for approximately 10–20% of deaths in these patients. Frequent use of non-steroidal anti-inflammatory drugs (NSAIDs) may also increase the incidence of gastrointestinal problems,[56] and commonly used treatments for moderate-to-severe RA may contribute to HRQOL reductions (e.g. issues with the parenteral administration route).

Currently, all therapies recommended by NICE for the treatment of severe RA are administered via the parenteral route (subcutaneous injection or IV infusion). This may be inconvenient or unpleasant for patients, as these treatment options are commonly associated with injection or infusion site reactions, such as burning or stinging.[6,7] It is reported that 21–35% of patients discontinue injectable RA therapy within a year.[30] A majority of RA patients prefer oral therapies over injectables,[29] a preference which is currently unmet in severe RA therapies.[30]

A 2014 study by Cutolo et al[57] suggests that the burden of RA extends beyond the joints to include other tissues and organs, involving multiple comorbidities, and psychosocial manifestations that impact patient quality of life, as shown in Figure 2. The interdependency of these three aspects of RA is an increasing focus of RA research, and the impact of RA treatments on HRQOL is an important factor in disease management.

Figure 2. Schematic of RA disease burden



Abbreviation: RA = rheumatoid arthritis.
Source: Adapted from Cutolo et al 2014.[57]

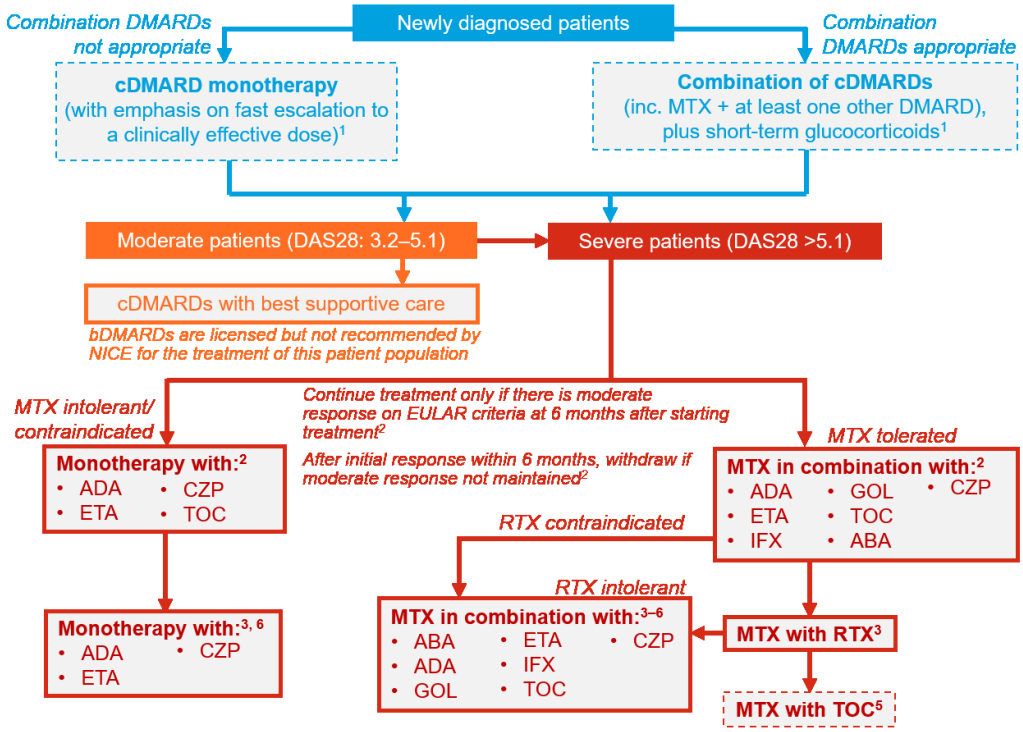
RA imposes direct costs on the NHS of approximately £560 million annually.[47] However, the majority of the costs of RA to society are indirect; disability associated with RA results in substantial productivity losses, both paid and unpaid.[58] Patients report that their earning potential is reduced by having to reduce their working hours, hampered career progression and needing to change occupations to accommodate their RA.[51] In addition, nearly two-thirds of RA patients report early retirement due to their condition.[51] Reduced productivity whilst at work (presenteeism) is also a significant contributor to RA-related costs.[59] Caregivers, who may often be informal caregivers (e.g. relatives) also experience a considerable burden due to RA. On average, caregivers assist patients for 33 hours per week,[60] and many experience impacts on their own health as a result of their caregiving.[61] In addition to this impact on quality of life, informal caregivers may also suffer financially in terms of reduced working hours or lost earning potential as a result of the time commitment of their contributions to caregiving. Taking into account both the direct and indirect costs of RA, the economic burden to UK society is estimated at between £3.8 and £4.75 billion per year.[36]

3.3 *Current Treatment Pathway in Rheumatoid Arthritis and the Positioning of Baricitinib*

There are two classes of treatments for RA; symptom-treating drugs including NSAIDs, selective COX-2 inhibitors and corticosteroids; and DMARDs, which aim to slow disease progression and reduce joint damage.[38]

Two internationally renowned rheumatology organisations who represent patients with arthritis, health professional and scientific societies of rheumatology are the European League against Rheumatism (EULAR) and the American College of Rheumatology (ACR).[62,63] Each of these bodies has issued a set of treatment guidelines for rheumatoid arthritis, the key points of which are presented in Section 3.6. NICE guidance on the RA treatment pathway is available in the form of Clinical Guideline 79 and a number of Technology Appraisals, and was deemed the most relevant treatment pathway to present here.[14,36,64-66] A schematic of the current treatment pathway in RA, as recommended in the NICE Pathway for rheumatoid arthritis,[37] is presented in Figure 3. For patients with newly diagnosed active RA, NICE recommendations for first-line treatment consist of a combination of cDMARDs, including methotrexate and at least one other DMARD, within three months of the onset of symptoms.[36] Short-term glucocorticoids may also be offered to rapidly improve symptoms.[36] In patients for whom combination cDMARD therapy is not appropriate, cDMARD monotherapy should be started, placing greater emphasis on rapid dose escalation to achieve a clinically-effective dose.[36] For patients that do not respond adequately to therapy and have moderately-active RA (DAS28: 3.2–5.1), NICE recommends continuing cDMARDs with best supportive care.

Figure 3. NICE recommended treatment pathway for RA



Footnotes: Positions in the treatment pathway where baricitinib might be considered, in accordance with the NICE scope/final license,[3] are demonstrated by thick, solid border lines. Broken border lines indicate positions where baricitinib is not considered in the decision problem.
Abbreviations: RA = rheumatoid arthritis, DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, MTX = methotrexate, DAS28 = Disease Activity Score, bDMARD = biologic DMARD, ADA = adalimumab, CTZ = certolizumab pegol, ETN = etanercept, TCZ = tocilizumab, GOL = golimumab, IFX = infliximab, ABA = abatacept, RTX = rituximab.
Sources: [1]NICE CG79,[36] [2]NICE TA375,[67] [3]NICE TA195,[68] [4]NICE TA225,[64] [5]NICE TA247,[66] [6]TA415[65]

For patients that have not responded to intensive combination therapy with cDMARDs and have severe RA (DAS28 >5.1), NICE recommends the use of methotrexate in combination with bDMARDs (adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept).[14] For those who cannot take methotrexate because it is contraindicated or because of intolerance, adalimumab, etanercept, certolizumab pegol or tocilizumab monotherapy can be used.[14] For patients with severe disease who have had an inadequate response to, or are intolerant of other DMARDs (including at least one TNFi therapy), rituximab in combination with methotrexate should be used.[64,69,70] If rituximab is contraindicated or withdrawn due to an adverse event, adalimumab, etanercept, infliximab, certolizumab pegol, abatacept, golimumab and tocilizumab each in combination with methotrexate may be used.[65,69] Where rituximab is unable to be used because methotrexate is contraindicated (rituximab must be given in combination with methotrexate in line with the marketing authorisation for rituximab) or has been withdrawn due to an adverse event, adalimumab, etanercept or certolizumab pegol may be used as a monotherapy.[65,69] It should be noted, however, that a systematic literature review found that the likelihood of responding to a subsequent bDMARD treatment decreased as the number of previous treatments with TNFi agents increased.[71]

In accordance with the NICE scope[40] and the SmPC,[2] the potential positions of baricitinib in the recommended treatment pathway for RA are listed in Table 8.

Table 8. Potential positions for baricitinib in the RA treatment pathway

Population	Comparators
Moderately active RA that has not responded adequately to therapy with cDMARDs	• Combination therapy with conventional DMARDs (cDMARDs) (including methotrexate and at least one other DMARD, such as sulfasalazine and leflunomide) • cDMARD monotherapy with dose escalation • Best supportive care (only where cDMARDs are not appropriate due to intolerance)
Severely active RA that has not responded adequately to therapy with cDMARDs	• Adalimumab, etanercept, certolizumab pegol or tocilizumab only (each as monotherapy) • Biologic DMARDs in combination with methotrexate (adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab, abatacept)
Severely active RA that has not responded adequately to therapy with bDMARDs, including at least one TNFi agent	• Rituximab in combination with methotrexate

When rituximab is contraindicated or withdrawn due to adverse events	• Adalimumab, etanercept and certolizumab pegol (each as monotherapy) • Adalimumab, etanercept, infliximab, abatacept, tocilizumab or certolizumab pegol each in combination with methotrexate

Abbreviations: RA = rheumatoid arthritis, DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, bDMARD = biologic DMARD, TNFi = tumour necrosis factor inhibitor, NSAID = non-steroidal anti-inflammatory drug.

In reference to these positions in the RA treatment pathway, three phase III RCTs have been performed to assess the safety and efficacy of baricitinib in three different RA populations: inadequate responders to methotrexate,[9] cDMARDs[11] and TNFi bDMARDs.[12]

3.4 *Life Expectancy of Patients with Moderate-to-Severe Disease*

In addition to the impact of the condition on patient quality of life (Section 3.2), RA is also associated with a negative impact on patient life expectancy. The life expectancy of people with RA was found to be approximately 3.5 years less than the general population in an observational study of a large number of patients (8,613) treated with bDMARDs in Germany.[35] For those with the most severe forms of RA, life expectancy may be reduced by as much as 10–15 years.[32] Therefore, achieving low disease activity is critical to the quality and length of life in patients with RA.

As described in Section 3.2, comorbidities associated with RA have a significant impact on patient life expectancy. A number of conditions, including cardiovascular, respiratory and gastrological disorders, as well as infections and cancers, are observed more frequently in RA patients; these may also contribute to the impact of RA on life expectancy.[32] RA patients have an approximately two-fold higher risk of cardiovascular disease than the general population[55] and respiratory conditions account for approximately 10–20% of deaths in RA patients.[72]

3.5 *Relevant NICE Guidance and Pathways*

NICE clinical guidelines and published technology appraisals of relevance to this submission are listed below. Recommendations from each were summarised in Section 3.3:

NICE Clinical Guidelines 79 [CG79]: Rheumatoid arthritis in adults: management[36]

NICE Technology Appraisal Guidance 415 [TA415]: Certolizumab pegol for treating rheumatoid arthritis after inadequate response to a TNF-alpha inhibitor[65]

NICE Technology Appraisal Guidance 375 [TA375]: Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for rheumatoid arthritis not previously treated with DMARDs or after conventional DMARDs only have failed[14]

NICE Technology Appraisal Guidance 247 [TA247]: Tocilizumab for the treatment of rheumatoid arthritis[70]

NICE Technology Appraisal Guidance 225 [TA225]: Golimumab for the treatment of rheumatoid arthritis after the failure of previous disease-modifying antirheumatic drugs[64]

NICE Technology Appraisal Guidance 195 [TA195]: Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor[69]

It should be noted that TA375 has partially replaced TA225 and TA247 in relation to the first-line use of golimumab and tocilizumab only if the disease is severe and has not responded to intensive therapy with a combination of conventional DMARDs. However, second-line recommendations on the use of these therapies remain valid.

As highlighted in Section 1.1, subgroups considered in this submission are cDMARD-IR patients with moderate disease who are currently not eligible for biologic treatment, cDMARD-IR patients with severe disease for whom biologic treatments are recommended, and patients who have failed a prior TNFi treatment who again are eligible for biologic treatment.

3.6 *Relevant Clinical Guidelines*

As noted in Section 3.3, two internationally renowned rheumatology organisations are EULAR and ACR, each of which have produced a set of treatment guidelines for rheumatoid arthritis, the key details of which are provided below.

Following a clinical diagnosis of RA, the EULAR guidelines recommend commencing treatment with methotrexate or a combination of cDMARDs, including low-dose glucocorticoids. If clinical remission or low disease activity is not achieved within six months, switching to a second cDMARD strategy, or adding a bDMARD when poor prognostic factors are present is recommended. In patients responding insufficiently to methotrexate and/or other cDMARD strategies, bDMARDs should be commenced in combination with methotrexate. If a first bDMARD has failed, the guidelines advise to switch to another bDMARD. Tofacitinib (not currently licensed by the EMA) may be considered after biologic treatment has failed.[38] A 2016 update of the EULAR guidelines is expected to be published shortly. A draft version of the guidelines differs marginally from the 2013 guidelines, recommending that bDMARDs be used in combination with cDMARDs (MTX) rather than as a monotherapy, and that RA patients should not be distinguished by their disease duration but rather by treatment phase (cDMARD-naïve, cDMARD-experienced, bDMARD-experienced). The update also recommends that prognostic factors be used for patient stratification, and makes several recommendations about glucocorticoid treatments.[73]

Similarly, ACR guidelines recommend commencing treatment with cDMARD monotherapy, ideally methotrexate. If clinical remission or low disease activity is not achieved, switching to a combination of cDMARDs or a bDMARD (TNFi or other), with or without methotrexate and in no order of preference is recommended. In the case of insufficient response, another bDMARD may be used with or without methotrexate. Tofacitinib (not currently licensed by the EMA) may be considered after biologic treatment has failed.[39]

The EULAR and ACR guidelines are broadly consistent with NICE recommendations for the treatment of RA, with the use of one or more cDMARDs in newly diagnosed patients followed by one or more bDMARDs if sufficient treatment response is not achieved. However, the NICE recommendations differ from the EULAR and ACR guidelines with regards to dependence of bDMARD treatment on disease severity. In patients for whom cDMARDs have failed to produce an adequate response, EULAR guidelines state that bDMARDs should be used upon failure of cDMARDs, irrespective of disease severity,[38] whilst ACR guidance recommends initiating bDMARD therapy in moderate or severe cases.[39] In contrast, NICE guidance TA375 specifies that treatment of patients with moderate RA with bDMARDs is not deemed to be a cost-effective use of NHS resources, and prescription of bDMARDs is recommended in patients who have failed cDMARDs only if they possess severe disease (DAS28 >5.1).[14] Moreover, TA375 specifies that treatment with bDMARDs should only be continued if there is at least a moderate response using the EULAR criteria (which are based on the degree of change in DAS28 score) at 6 months after starting therapy. After initial response, treatment should be withdrawn if at least a moderate EULAR response is not maintained.[14] This is in accordance with EULAR and ACR guidelines, although the target is clinical remission according to the ACR-EULAR definition.[38,39]

3.7 *Issues Relating to Current Clinical Practice*

As outlined in the current treatment guidelines for RA, under current NICE guidance patients with moderate disease activity who respond inadequately to cDMARD therapy have no further treatment options beyond best supportive care, unless their disease becomes severe. The matter was raised in an appeal against this decision by the British Society of Rheumatology (BSR) and the NRAS, on the basis that a group of patients with moderate disease may exhibit rapid disease progression, and as such may warrant the use of bDMARDs. Whilst the appeal was dismissed, the issue of patients with moderate disease who have responded inadequately to cDMARDs remains relevant as this is a patient population that is lacking in treatment options. This submission includes evidence from baricitinib clinical trials in the cDMARD-IR moderate population (see Section 4.7) and presents the results of an economic analysis in this population in Section 5.7.

3.7.1 Further research into the moderate patient population

In order to further understand the moderate patient group following the NICE committee discussions of TA375, two exploratory projects were undertaken for this submission.

3.7.1.1 Guys and Thomas's NHS Trust moderate disease activity analysis

Firstly, a patient cohort managed by Guy's and St. Thomas's NHS Trust. This cohort commenced in 2004 and captured data relevant to 'treat to target' therapy (achievement of disease remission). A number of initial objectives were considered:

- To identify factors associated with disease progression (as measured by HAQ-DI) at 12 months in patients with persistent moderate disease.
- To identify factors associated with the rate of disease progression (as measured by HAQ-DI) over 12 months in patients with persistent moderate disease.
- To identify factors associated with a clinically relevant worsening of function (defined as an increase in HAQ-DI ≥ 0.22 over 12 months) in patients with moderate disease

Persistent moderate disease patients were defined as having at least two consecutive DAS28 scores ranging between 3.2 and 5.1 within a 12 month period, had a recorded HAQ-DI score at the time of the second DAS28 score and had at least one further HAQ-DI score. Baseline or T0 was the point of the second DAS28 score confirming moderate disease and 12 month point was taken as the closest time point to T0+12months at which a HAQ-DI score was recorded. Patients had to be biologic-naïve.

Preliminary results are described below.

[redacted] patients aligned to these criteria were identified in the cohort. The mean age of the cohort was [redacted] years with a mean disease duration of seven years. Other baseline characteristics are shown in Table 9. The mean endpoint time from baseline was [redacted] months.

Table 9: Baseline characteristics of GSTT moderate cohort

Characteristic		No. Patients (%)
Female Gender		[redacted]
Ethnicity	White	[redacted]
	Black	[redacted]
	Asian	[redacted]
	Mixed	[redacted]
	Other	[redacted]
RF-Positive		[redacted]
Anti-CCP Positive		[redacted]
DMARDs	DMARD Monotherapy	[redacted]
	DMARD Combination Therapy	[redacted]
	No DMARDs	[redacted]
Steroids		[redacted]

Abbreviations: GSTT = Guy's and St. Thomas's NHS Trust, RF = rheumatoid factor, DMARD = disease-modifying antirheumatic drug, anti-CCP = anti-cyclic citrullinated peptide

[redacted]

The distribution of change in HAQ-DI score is shown in Table 10 and Figure 4.

Table 10: Summary of Change in HAQ-DI Scores between Baseline and 12-month HAQ-DI Scores

Change in HAQ-DI Score	Number of Patients (%)
-3 to ≤-2	[redacted]
>-2 to ≤-1	[redacted]
>-1 to ≤0.75	[redacted]
>-0.75 to ≤0.5	[redacted]
>-0.5 to ≤0.25	[redacted]
>0.25 to ≤0	[redacted]
>0 to ≤0.25	[redacted]
>0.25 to ≤0.5	[redacted]
>0.5 to ≤0.75	[redacted]
>0.75 to ≤1	[redacted]
>1 to ≤2	[redacted]
>2 to 3	[redacted]

Abbreviations: HAQ-DI = Health Assessment Questionnaire-Disability Index

Figure 4: Histograms of Change in HAQ-DI Scores between End-Point and Baseline

[redacted]

Abbreviations: HAQ-DI = Health Assessment Questionnaire-Disability Index

The early analysis shows that there is a subset of patients with moderate disease who experienced a 0.22 or greater increase in HAQ-DI which, as expected, grows smaller when the magnitude of the change in HAQ-DI is greater. This suggests that the 'rapid-progressor' group discussed in TA375 that might benefit from more aggressive treatment is a small minority of the overall moderate population. These patients would require their disease to reach the severe state before treatment with options other than cDMARDs under current NICE guidance.

Preliminary regression modelling investigating the objectives above indicate that baseline HAQ is the dominant predictor but further work is hoped to be undertaken.

3.7.1.2 Systematic review of evidence regarding potential prognostic criteria for disease progression and outcomes in patients with moderate rheumatoid arthritis

The second exploratory undertaking was a systematic literature review (SLR) of the evidence regarding prognostic factors for rapid progression in RA patients with moderate disease. The review searched Medline®, Medline® In-process, Embase, the Cochrane Library as well as recent abstracts from five relevant congresses. The search was conducted in October 2016.

A large number of publications were identified ([redacted]), therefore a subset were prioritised for data extraction and full review. These [redacted] studies were in a confirmed moderate disease activity population and were conducted in settings likely more reflective of UK practise (e.g. Europe, North America). A number of progression outcomes were identified, including HAQ-DI change.

Results of the SLR demonstrated that whilst there is variation in the reported results and outcome measures, a number of factors are potentially predictive of disease progression such as disease duration, DAS28 score, HAQ-DI, ultrasound and the presence of anti-cyclic citrullinated peptide antibody. Two UK studies focussed on DAS-28. Kiely et al[74] found that patients who had not achieved a target DAS28 of less than 3.2 after the first year of cDMARD therapy were unlikely to do so with continued cDMARD therapy. Along similar lines, Nikiphorou et al[75] found that risk of joint failure and surgery was similar in patients with persistent low or high moderate disease activity despite cDMARD therapy compared to patients with persistently high DAS28 scores (i.e. severe disease).

Both of these pieces of ongoing research indicate that there is a sub-group of patients with moderate disease activity who would benefit from treatment beyond conventional DMARDs.

3.8 *Equality Considerations*

There are no identified equality issues arising in relation to this technology.

4 Clinical effectiveness

Summary of Clinical Effectiveness

- Oral baricitinib 4 mg (QD), compared with an active comparator or placebo, was associated with significant improvements in signs and symptoms, physical function, and patient reported outcomes in patients across a wide-spectrum of the RA treatment pathway, including:
    - MTX-IR patients (RA-BEAM)[9]
    - cDMARD-IR patients (RA-BUILD)[11]
    - anti-TNF-IR patients (RA-BEACON)[12]
- Two additional studies of relevance are RA-BEYOND (long-term extension study) presented in section 4.7.4 and RA-BEGIN (DMARD-naïve population- unlicensed indication) presented in Appendix 1.
- Baricitinib 4 mg (QD) is the first RA drug to demonstrate superiority in a head-to-head trial versus adalimumab with background MTX (RA-BEAM) in its core phase III programme.
- The primary endpoint of ACR20 at week 12 or 24 was met in all studies. In addition, baricitinib 4 mg (QD) showed statistically significant improvements to all comparators for ACR20/50/70 response rates at the primary timepoints.
- In studies measuring radiographic outcomes (RA-BEAM and RA-BUILD), progression of joint damage, as measured by mTSS, was significantly reduced in patients treated with baricitinib 4 mg with methotrexate compared to placebo and similar compared to adalimumab with methotrexate.
- Baricitinib, in combination with MTX, demonstrated consistent improvements in EULAR response across trials.
- Patient reported symptoms were statistically significantly improved versus placebo across all three trials, as determined by HAQ-DI, duration/severity of morning joint stiffness and FACIT-F.
- In addition, the results obtained in the HAQ-DI MCID and FACIT-F MCID response rates demonstrate the statistical superiority of baricitinib compared to placebo.

Summary of Safety

- Treatment with baricitinib was well-tolerated and a small proportion of patients discontinued from the baricitinib studies because of AEs.
- The most commonly reported adverse drug reactions (ADRs) in ≥2% of patients treated with baricitinib monotherapy or in combination with csDMARDs included increased LDL cholesterol, URTI and nausea. However, the majority of all ADRs were mild to moderate in severity.
- The proportion of patients with SAEs (including serious infections) was similar across treatment groups in the phase III studies and integrated placebo-controlled analysis sets, except for RA-BEAM, where a higher proportion of SAEs were reported with placebo and baricitinib versus adalimumab.
- Although baricitinib was associated with a higher incidence of SAEs compared with adalimumab through 52 weeks in RA-BEAM, their AE profiles were similar across clinically

significant categories of risk including MACE, malignancies, hypercholesterolemia, serious infections and herpes zoster.

- Despite a higher risk of cardiovascular disease, infection, and malignancy in the RA population, treatment with baricitinib did not result in increased risk of malignancy, serious or opportunistic infections, or MACE.
- Non-serious herpes simplex and herpes zoster infections were more frequent in patients treated with baricitinib than placebo, yet rates were not significantly higher than those seen with MTX or adalimumab.
- The majority of herpes zoster cases were mild to moderate in severity and complicated cases were uncommon.
- Increases in LDL cholesterol were one of the most commonly reported ADRs, yet increases in HDL-C were also seen with baricitinib so that the mean HDL/LDL ratio was unchanged. Furthermore, there was a significant decrease in the amount of small and very small LDL particles in RA-BEAM, which are considered the most atherogenic. Few major adverse cardiovascular events (MACE) events were also observed in the baricitinib clinical programme and no relationship was seen between MACE and increased LDL.
- Treatment with baricitinib also resulted in changes to haematology and clinical chemistry analytes. These included mean changes of greater magnitude for some analytes than seen with the active comparators, which are, therefore, likely to be related to the pharmacology of JAK inhibition (such as increases in lipids [including total cholesterol, low-density lipoprotein cholesterol and high-density lipoprotein cholesterol] and creatine phosphokinase).

Summary of the Clinical Effectiveness SLR

- An SLR was conducted to identify relevant evidence on the efficacy and safety of baricitinib for the treatment of moderate to severe RA.
- Searches of major electronic databases were conducted on 17 June 2015, and updated on 9–10 June 2016, whilst conference proceedings were searched 3–4 August 2015, and updated on 3–31 August 2016.
- In the original review 7,316 records were identified, with an additional 1,415 in the updated review. After level one screening 479 records (+180 from the update) progressed to level two screening. In total, 257 records were ultimately included, 138 primary and 118 secondary.
- Four primary studies were identified for baricitinib: RA-BEAM, RA-BEAM, RA-BEACON and RA-BEGIN.

Summary of the RA-BEAM clinical effectiveness results

- In MTX-IR patients with moderately to severely active RA despite stable background MTX, oral baricitinib 4 mg (QD) was associated with significant improvements in signs and symptoms, low disease activity and remission rates, physical function and HRQOL outcomes compared to placebo and to adalimumab.
- Patients treated with baricitinib 4 mg (QD) had statistically significantly less progression of structural joint damage compared to placebo at Week 24; the progression of structural joint damage was similar to the rates observed with adalimumab at Weeks 24 and 52.
- The beneficial treatment effect of baricitinib 4 mg (QD) compared with placebo was rapid, with improvements observed as early as Week 1 or 2 for ACR20/50/70 and Week 4 for SDAI remission.

- Baricitinib 4 mg was superior to adalimumab in terms of improvements in physical function (measured by HAQ-DI) and other PROs including worst tiredness and duration/severity of morning joint stiffness at Week 12; these differences were statistically significant. Improvements were seen as early Week 4 for severity of morning joint stiffness and Week 8 for worst tiredeness
- The safety and tolerability profile of baricitinib remained consistent with other baricitinib Phase 2 and 3 studies.

Summary of the RA-BUILD Clinical Effectiveness Results

- In patients with moderately to severely active RA despite prior treatment with cDMARDs, 2 mg (QD) and 4 mg (QD) baricitinib was associated with significant improvements in signs and symptoms, low disease activity and remission rates, physical function, PROs.
- Baricitinib 4 mg (QD) was superior to placebo in inhibiting progression of structural joint damage for both components of joint narrowing and erosion at Week 24 and this difference was statistically significant.
- Baricitinib 4 mg and 2 mg were superior to placebo for improvements in physical function at Weeks 12 and 24, as measured by HAQ-DI
- Overall, a more rapid and consistently larger treatment effect was seen for the 4 mg (QD) dose compared with 2 mg (QD) across different analyses including SDAI, CDAI and in components of the composite scores.
- A more rapid and consistently larger treatment effect was also seen for the 4 mg (QD) dose compared with 2 mg (QD) for PROs, including improvements in duration/severity of morning joint stiffness.
- The safety and tolerability profile of baricitinib remained consistent with other baricitinib Phase 2 and 3 studies.

Summary of the RA-BEACON clinical effectiveness results

- In patients with moderately to severely active RA despite previous treatment with bDMARDs, including ≥1 anti-TNF, 2 mg (QD) and 4 mg (QD) baricitinib produced clinical improvements in signs and symptoms, low disease activity and remission rates, physical function and patient reported outcomes that were sustained through 24 weeks of treatment.
- Baricitinib 4 mg and 2 mg were superior to placebo in terms of ACR20 response rates at Week 12 (primary endpoint) and this difference was statistically significant
- Baricitinib 4 mg (QD) was superior to placebo at Week 24 in achieving remission or LDA as defined by DAS28-hsCRP, DAS28-ESR, SDAI and CDAI; these differences were statistically significant. The difference was significant for baricitinib 2 mg (QD) at Week 24 only for LDA response as defined by DAS hsCRP ≤3.2 and SDAI ≤11.
- The effect of baricitinib 4 mg (QD) was generally greater in magnitude, more rapid, durable and consistent across different efficacy measures, particularly for the most clinically meaningful endpoints (ACR50/70, remission and LDA) compared with the baricitinib 2 mg (QD) dose.
- Patients receiving both doses of baricitinib experienced significant improvements in physical function (HAQ-DI) at both Week 12 and Week 24 compared with placebo
- The safety and tolerability profile of baricitinib remained consistent with other baricitinib

Phase 2 and 3 studies.

Summary of the RA-BEYOND clinical effectiveness results

- Results to date from RA-BEYOND have demonstrated sustained efficacy for patients continuing with the baricitinib treatment allocation from their originating study over an additional 48 weeks of treatment, even with a population of patients with a variety of previous therapies and disease durations.
- High patient retention rates (90%) indicate a favourable risk/benefit profile following prolonged treatment with baricitinib.
- Patients originating from the three Phase III trials who met or exceeded HAQ-DI MCID (≥0.22 and ≥0.3) at baseline in RA-BEYOND were shown to maintain their responses for a further 48 weeks of treatment with baricitinib 4 mg (QD).
- Step-down titration of patients who achieved sustained LDA or remission, as measured by CDAI, from 4 mg (QD) to 2 mg (QD) was found to result in modest but statistically significant increases in disease activity after 12 weeks, but the majority of patients retained their state of LDA or remission.

Summary of Network Meta-Analysis

- Network meta-analyses (NMA) were performed in order to assess the relative efficacy of baricitinib compared with the relevant comparators in cDMARD- (including MTX)-IR or anti-TNF-IR patients with moderate-to-severe RA.
- In the cDMARD-IR population base case analysis at week 24, baricitinib 4 mg was found to be associated with statistically significantly higher odds of ACR50 response compared to cDMARD, adalimumab 40 mg, placebo, etanercept and sulfasalazine. No statistically significant differences were found versus any other comparator for the ACR50 outcome, with the exception of certolizumab pegol, in which odds of ACR50 response was found to be significantly in favour of the comparator. Although statistically significant differences were not found versus most biologic comparators, in the majority of cases the point estimate of relative treatment effect was favourable to baricitinib 4 mg.
- In the anti-TNF-IR population base case analysis at week 24, baricitinib demonstrated significantly higher ACR50 response rates than the cDMARD comparator. No statistically significant differences were seen versus biologic comparators, with the exception of the comparison of baricitinib 4 mg and baricitinib 2 mg to tocilizumab 8 mg, and the comparison of baricitinib 2 mg to rituximab 1000 mg, in which statistically significant treatment effects in favour of the comparator were observed. Versus the other comparators, point estimates in some cases favoured baricitinib 4 mg and in other cases favoured the comparator treatment.

4.1 *Identification and selection of relevant studies*

A SLR was conducted to identify relevant evidence on the efficacy and safety of baricitinib for the treatment of moderate to severe RA. The SLR also included relevant comparators for the treatment of moderate to severe RA for the purposes of allowing a potential indirect treatment comparison with baricitinib. The original SLR was conducted on 17 June 2015 and was subsequently updated on 9–10 June 2016.

4.1.1 Search Strategy

A search strategy was developed to identify relevant clinical evidence related to baricitinib and relevant comparators for patients with RA, the details of which are described in the following sections.

4.1.2 Details of Search Strategy

The primary objective of the clinical SLR was to identify evidence from clinical and safety studies of the following current treatments for moderately to severely active RA (including early and established RA):

- Baricitinib
- Conventional DMARDs, including the following:
    - MTX (Trexall, Maxtrex, Rheumatrex, amethopterin, Rasuvo, Otrexup)
    - Sulfasalazine (Azulfidine, Salazopyrin, Sulazine, sulfazine)
    - Leflunomide (Arabloc, Arava, Lunava, Respo)
    - Hydroxychloroquine (Plaquenil, Axemal, Dolquine, Quensyl, Quineprox)
    - Relevant bDMARD comparators for this submission (in combination with cDMARDs or as monotherapy):
    - Infliximab (Remicade)
    - Adalimumab (Humira, Trudexa)
    - Certolizumab pegol (Cimzia)
    - Golimumab (Simponi)
    - Etanercept (Enbrel)
    - Abatacept (Orencia)
    - Rituximab (Rituxan)
    - Tocilizumab (Actemra, RoActemra)
- Other bDMARDs, cDMARDs and traditional DMARDs (tDMARDs) that are not relevant comparators for this submission, but which may inform the NMA

The review was conducted in accordance with the methodological principles for conducting SLRs as recommended by the Centre for Reviews and Dissemination's *Guidance for Undertaking Reviews in Health Care,* and the results are reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting checklist.[76,77]

The following electronic literature databases were searched:

- MEDLINE and MEDLINE In-Process (using PubMed platform)
- Embase (using Elsevier Platform)
- Biosciences Information Services (using Dialog platform)
- The Cochrane Library, including the following:
    - The Cochrane Central Register of Controlled Trials
    - The Cochrane Database of Systematic Reviews
    - Database of Abstracts of Reviews of Effectiveness

In the original search, conducted on 17 June 2015, results from electronic databases were limited to those later than 1999. For the update of this review, conducted 9–10 June 2016, electronic database searches were limited to articles published from 1 March 2015 to present in order to capture any relevant trials published since the original review.

In addition to the online literature database searches, in the original review conference proceedings were searched between 3–4 August 2015 and limited to results from 2013 to 2015. For the update of the review conference proceedings were searched between 3–31 August 2016 and limited to results from September 2015 to August 2016. The following conference abstracts were published in journal supplements or were indexed in Embase, so separate searches of these conferences were not necessary because these abstracts were retrieved during the electronic medical database searches:

- EULAR meetings held in 2013 and 2014
- ACR's annual meetings held in 2013 and 2014
- British Society for Rheumatology 2013 meeting
- Some of the conference abstracts were in the process of being indexed; such abstracts were identified through the following Internet sites:
- EULAR's 2015 meeting:
    - http://www.abstracts2view.com/eular/sessionindex.php
- ACR's annual meeting, 2015:
    - http://www.rheumatology.org/Publications/MeetingPublications/ACR/ARHP_Annual_Meeting_Publications/
- British Society for Rheumatology, meetings held in 2014 and 2015:
    - http://www.rheumatology.org.uk/

For the review update, the conference abstracts which were in the process of being indexed were identified through the following internet sites:

- EULAR meeting held in 2016:
    - http://www.abstracts2view.com/eular/sessionindex.php?day=2016&session=2016197&#session_2016197

To identify ongoing, discontinued, or completed clinical trials of baricitinib and its comparators, the following websites were searched between 3–4 August 2015 in the original review and between 3–31 August 2016 in the updated review:

- ClinicalTrials.gov:
    - http://clinicaltrials.gov/
- International Clinical Trials Registry Platform:
    - http://www.who.int/ictrp/en/
- European Union's Clinical Trials Register:
    - http://www.clinicaltrialsregister.eu/
- Klinische Prüfungen PharmNet.Bund:
    - http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/

Reference lists of any relevant studies, recent (published in the last 2 years) systematic reviews, and meta-analyses were searched for further studies of interest. In addition, reference lists of relevant articles identified from the following sources were searched between 3–4 August 2015 in the original review and between 3–31 August 2016 in the updated review:

- Scottish Medicine Consortium advice
- NICE's multiple technology appraisal and single technology appraisal documents
- United States Food and Drug Administration register
- European public assessment reports for human medicines, published by the European Medicines Agency

Supplemental to the formal SLR, additional manufacturer data on file was available for baricitinib. Full details of the search strategies employed for both the original SLR and the SLR update are presented in Appendix 2.

4.1.3 Study Selection

The inclusion and exclusion criteria for both the original search and the update were based upon the PICOS criteria displayed in Table 11. The study selection process was performed in the following two phases:

- Level 1 screening: Titles and abstracts of studies identified from the electronic databases and the Internet searches were double-screened by two independent researchers to determine eligibility according to the inclusion and exclusion criteria described in Table 11. However, due to the large number of studies included after the level 1 screen, the studies identified as being relevant for the review were re-screened using more stringent criteria (criteria marked in bold and with an asterisk in Table 11). Any discrepancies were resolved; when a consensus was not reached, a third researcher was consulted.

- Level 2 screening: Full texts of studies selected at level 1 were obtained and double-screened by two independent researchers to determine eligibility according to the inclusion and exclusion criteria shown in Table 11, where any additional inclusion or exclusion criteria not present in the level 1 screening are marked in bold and with a †. Any discrepancies were resolved; when a consensus was not reached, a third researcher was consulted.

Table 11. Eligibility criteria used in search strategy

Clinical effectiveness	Inclusion criteria	Exclusion criteria
Population	• Adult (≥ 18 years) patients with moderately to severely active RA (including patients with early and established RA) • Treatment-naïve patients • Patients who had intolerance or inadequate response to prior conventional DMARDs • Patients who had intolerance or inadequate response to previous bDMARDs	• Juvenile idiopathic arthritis • Studies that include only juveniles • Patients with mild RA[a]; if the study population is mixed (i.e., mild to severe), exclude those studies in which data are not reported separately for moderate or severely active RA

Intervention	• Baricitinib
	  ○ †Licensed treatments: at the labelled doses
	  ○ †Treatments not yet licensed: in any form or dose
	• Methotrexate (Trexall, Rheumatrex, amethopterin, Rasuvo, Otrexup)
	• Sulfasalazine (Azulfidine, Salazopyrin, Sulazine, sulfazine)
	• Leflunomide (Arabloc, Arava, Lunava, Respo)
	• Hydroxychloroquine (Plaquenil, Axemal, Dolquine, Quensyl, Quineprox)
	• Azathioprine (Azasan, Imuran, Azamun, Imurel)
	• Infliximab (Remicade)
	• Adalimumab (Humira, Trudexa,ABP 501, BI695501, CHS-1420, GP2017, M923, PF-06410293)
	• Certolizumab pegol (Cimzia)
	• Golimumab (Simponi)
	• Etanercept (Enbrel, Avent, BX2922, CHS-0214, ENIA11, Etacept, Etanar, GP2013, GP2015, HD203, LBEC0101, M923, PRX-106, SB4, TuNEX, Yisaipu)
	• Abatacept (Orencia)
	• Anakinra (Kineret)
	• Rituximab (Rituxan,Mabthera, Zytux, Reditux)
		• Studies that do not have an intervention of interest in at least 1 arm
		• Non-pharmacological studies, e.g., exercise, Chinese medicine, etc.
		• *Biosimilars
		• *Azathioprine (Azasan, Imuran)
		• *Studies comparing conventional DMARDs to non-DMARD treatments, such as NSAIDs or glucocorticoids

	• Tocilizumab (Actemra, RoActemra, atlizumab) • Sarilumab • Sirukumab • Tofacitinib (Xeljanz, Jakvinus, tasocitinib) ○ At the level 1 screening, all therapy versions (i.e., any dose or combination) of the interventions listed above will be included	
Comparators	• Any comparison between any of the listed interventions and each other or placebo	• Studies not reporting on at least one of the interventions of interest
Outcomes	• *Studies reporting efficacy and safety data, HRQOL, WPAI-RA, or health care resource utilisation • *MRI studies that specifically mention the Sharp/Van der Heijde bone erosion score †To be included in the review, a study must report at least 1 of the following outcomes of interest: • †Efficacy measurements: • †ACR criteria • †ACR score • †Proportion of patients achieving an ACR20 response • †Proportion of patients achieving an ACR50 response • †Proportion of patients achieving an ACR70 response • †ACR remission • †Proportion of patients achieving an ACR50 response in the subgroup of patients who are TNF inhibitor naïve, have inadequate response to TNF or other biologics, or who are intolerant to TNF or other biologics (if reported) • †Proportion of patients achieving an ACR20 response in the subgroup of patients who are TNF inhibitor naïve, have inadequate response to TNF or other biologics or who are intolerant to TNF or other biologics (if reported) • †Individual components of the ACR:	• *Studies that report only MRI outcomes and do not specifically mention the Sharp/Van der Heijde bone erosion score • *Studies that report only bone mineral density • *Studies that investigate ultrasound and radiography in assessing bone damage

	○ HAQ-DI ○ Pain VAS ○ Tender joint count ○ Swollen joint count ○ Physician's Global Assessment of Disease Activity ○ Patient's Global Assessment of Disease Activity ○ Modified Total Sharp score ○ Erosion score ○ Joint space narrowing score ○ DAS-28 ESR for RA ○ DAS-28 CRP for RA ○ SDAI ○ CDAI ○ Physical function assessed by HAQ or HAQ-DI • †Endpoints measuring the following: ○ Morning joint stiffness (severity and duration) and/or joint pain (may be assessed by different instruments) ○ Tiredness or fatigue (may be assessed by different instruments) • †EULAR or ACR remission defined as: ○ CDAI score ≤ 2.8 ○ SDAI score ≤ 3.3 ○ DAS-28 < 2.6 ○ RAPID3 ≤ 1 ○ DAS-44 < 1.6 ○ Boolean definition of remission (EULAR or ACR where all measures must be < 1) • WPAI-RA • Health care resource utilisation • †HRQOL outcomes from the following:	

	○ EQ-5D ○ SF-36 • †Safety outcomes reported at study endpoint: ○ Overall rate of AEs ○ Overall rate of serious AEs ○ Discontinuations due to ▪ Lack of efficacy ▪ AEs ○ Individual AEs, such as the following: ▪ Specific myelosuppressive events, e.g., anaemia, leukopaenia, neutropaenia, or thrombocytopaenia or lymphopaenia or lymphocytopaenia ▪ Thrombocytosis ▪ Serious infections ▪ Opportunistic infections ▪ Malignancies ▪ Cardiovascular events ▪ Elevations in ALT or AST (> 3 times upper limit of normal) with total bilirubin (> 2 times upper limit of normal) ▪ Injection-related combinations ▪ Intravenous reactions ○ Death ○ Initial or prolonged inpatient hospitalisation	

Study design	• Randomised, controlled, prospective clinical trials • Long-term follow-up studies (e.g. open-label follow-up studies with continuation of treatments in their respective randomised group) • Systematic reviews (including meta-analyses)[b]	• *Phase 2, randomised, controlled, prospective clinical trials • Non-randomised clinical trials • Single-arm studies • *Long-term follow-up or extension studies of RCTs in which patients do not remain in their respective randomised group • *Maintenance studies and step-down treatment studies • Preclinical studies • Phase 1 studies • Prognostic studies • Retrospective studies • Prospective observational studies • Case reports • Commentaries and letters (publication type) • Consensus reports • Pooled analyses • *Post hoc analyses • Non-systematic reviews • *Systematic reviews (including meta-analyses) published prior to 2014 • Secondary analyses • Animal models
Language restrictions	• *English-language publications	• *Non–English-language publications
Date restrictions	• 1999 to present	• None

Footnotes: [a]If the disease severity of included patients was not clearly stated in the article, the following approach was used and validated by Lilly: if DAS-28 scores were reported, then DAS-28 scores of > 3.2 were considered to be moderate RA; DAS-28 scores of > 5.1 were considered to be severe RA. If DAS-28 scores were not reported, then swollen and tender joint counts both > 6 was considered to be a good proxy for moderate to severe RA.
[b]Systematic reviews and meta-analyses will be used only for identification of primary studies that may have been missed in the electronic searches.
*Due to the high number of included studies from the abstract/title review, a secondary set of more stringent criteria were used to re-screen included studies.
†Additional criteria used during the full text review process.

Abbreviations: DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, bDMARD = biologic DMARD, OD = once daily, TNF = tumour necrosis factor inhibitor, ACR = American College of Rheumatology, ACR20/50/70 = 20/50/70% improvement in ACR criteria, EULAR = European League Against Rheumatism, EULAR = EULAR response index, DAS28 = Disease Activity Score modified to include the 28 diarthrodial joint count, hsCRP = high-sensitivity C-reactive protein, ESR = erythrocyte sedimentation rate, SDAI = Simplified Disease Activity Index, CDAI = Clinical Disease Activity Index, MJS = morning joint stiffness, WJP = worst joint pain, FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, mTSS = modified Total Sharp Score, WPAI-RA = Work Productivity and Activity Index-Rheumatoid Arthritis, EQ-5D-5L = EuroQoL 5 dimensions–5 levels, HAQ-DI = Health Assessment Questionnaire-Disability Index, SF-36v2 = Medical Outcomes Study 36-Item Short Form Health Survey Version 2 Acute, VAS = visual analogue scale, AE = adverse event

4.1.4 PRISMA Diagram

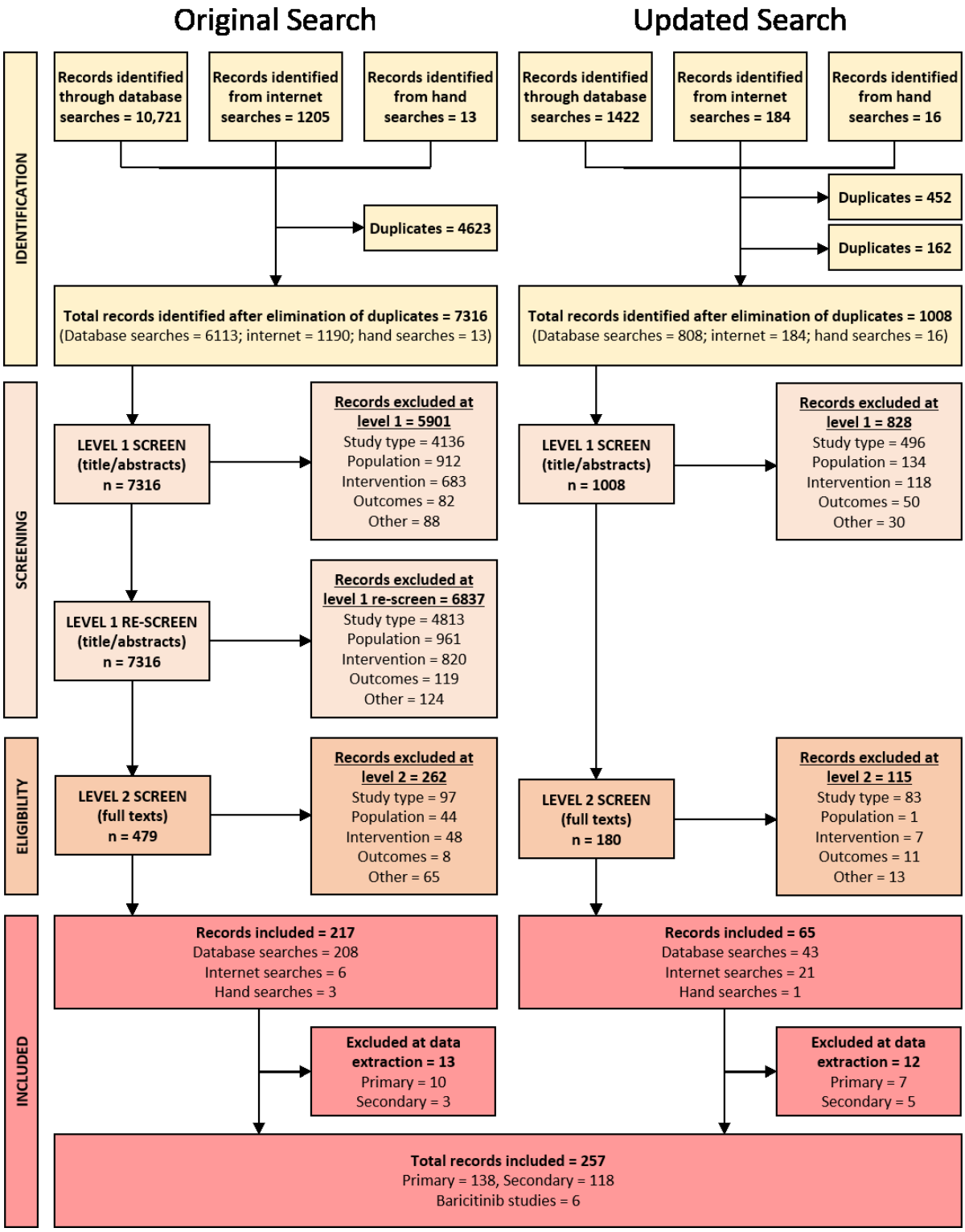
The PRISMA flow diagram of the evidence identified in the original and updated SLRs is presented in Figure 5.

In the original systematic review, a total of 7,316 records (title and abstracts) were identified for manual screening to identify all relevant studies that met the predefined inclusion and exclusion criteria (Table 11). In the review update, 1,008 records were identified for manual screening. After the initial screening of titles and abstracts (level 1 screening) for the original review, 1,415 publications were identified as potentially relevant. From the level 1 re-screen, 479 publications were progressed for further screening (level 2), whilst for the update 180 studies progressed to level 2 screening.

At the level 2 screening in the original review, 217 publications were selected for inclusion in the review. For the update, 65 publications were selected for inclusion in the review. In addition, during the data extraction stage of the original review a further 13 papers were identified as being unsuitable for inclusion in this review and therefore were excluded. In the update of the review 12 studies were excluded during the data extraction phase.

Therefore, over both the original review and its update, 257 articles were ultimately included in the review; 138 are primary publications and 118 are secondary publications. A complete list of primary publications with the corresponding secondary publications is provided in Appendix 3. All publications excluded at level 2, with reasons for exclusion, are listed in Appendix 4.

Figure 5. PRISMA diagram for study inclusion and exclusion



Abbreviation: PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

4.1.5 Study Linking

Study linking was performed as part of the clinical SLR, with 138 primary publications and 118 secondary publications identified. A complete list of all linked primary and secondary studies is provided in Appendix 3.

4.2 *List of relevant randomised controlled trials*

4.2.1 Relevant Randomised Controlled Trials

Four phase III RCTs were identified in the SLR that evaluated baricitinib at progressive stages of the treatment pathway for moderately to severely active RA, along with an additional long-term extension study, as follows:

- RA-BEAM (JADV) compared baricitinib (4 mg QD) to adalimumab (40 mg Q2W) and placebo with background MTX in MTX inadequate responders and bDMARD-naïve patients
- RA-BUILD (JADX) compared 2 mg or 4 mg (QD) baricitinib to placebo in cDMARD inadequate responders and bDMARD-naïve patients
- RA-BEACON (JADW) compared 2 mg or 4 mg (QD) baricitinib to placebo in patients TNFi inadequate responders to bDMARDs
- RA-BEGIN (JADZ) investigated the use of baricitinib (4 mg QD) monotherapy or baricitinib (4 mg QD) in combination with MTX (QW) compared to MTX (QW) monotherapy in DMARD-naïve patients *(unlicensed treatment position, therefore this study is presented in Appendix 1 only)*

The long-term extension trial RA-BEYOND (JADY) investigates the safety and efficacy of up to four years of additional treatment with baricitinib

An overview of the baricitinib phase III clinical development programme is presented in Table 12.

Table 12. List of relevant RCTs and long-term extension studies

Trial number (acronym)	Population	Intervention	Comparators	Primary study reference	Secondary study references
RA-BEAM (JADV)[9]	MTX-inadequate responders, bDMARD-naïve adult patients with moderately to severely RA	Baricitinib 4 mg, oral, QD (with background MTX)	Adalimumab 40 mg, SC injection, Q2W (with background MTX) Placebo (with background MTX)	RA-BEAM Clinical Study Report[9]	RA-BEAM Clinical Study Report[9]
RA-BUILD (JADX)[11]	cDMARD-inadequate responders, bDMARD-naïve adult patients with moderately to severely active RA	Baricitinib (2 mg, oral, QD) Baricitinib (4 mg, oral, QD). Patients on ≥1 cDMARDs (with or without MTX) continued to take background therapy during study.	Placebo (Patients on ≥1 cDMARDs (with or without MTX) continued to take background therapy during study)	Dougados *et al.* (2016)[8]	Emery P, Gaich CL, DeLozier AM, de Bono S, Liu J, Chang C, Dougados M. Patient-Reported Outcomes from a Phase 3 Study of Baricitinib in Patients with Rheumatoid Arthritis with Inadequate Response to Conventional Synthetic Disease-Modifying Antirheumatic Drugs. Arthritis Rheumatol. 2015; 67 (suppl 10).
RA-BEACON (JADW)[12]	bDMARD inadequate responders adult patients with moderately to severely active RA	Baricitinib 2 mg, oral, QD (with background cDMARDs) Baricitinib 4 mg, oral, QD (with background cDMARDs)	Placebo (with background cDMARDs)	Genovese *et al.* (2016)[10]	Genovese MC, Kremer J, Zamani O, Ludicico C, Krogulec M, Xie L *et al.* Baricitinib, an oral janus kinase (JAK)1/JAK2 inhibitor, in patients with active rheumatoid arthritis (RA) and an inadequate response to TNF inhibitors: results of the phase 3 RA-BEACON study. Ann Rheum Dis. 2015;74(suppl2):75. Smolen JS, Kremer J, Gaich C, DeLozier AM, Schlichting D, Xie L, and Genovese MC. Patient-reported outcomes from a phase 3 study of baricitinib in patients with rheumatoid arthritis (RA) and an inadequate response to tumor necrosis factor inhibitors. Annals of the Rheumatic Diseases. 2015;74:785-786.
RA-BEGIN (JADZ)[78]	DMARD-naïve adult patients with moderately to severely RA (unlicensed)	Baricitinib (4 mg, oral, QD) Baricitinib (4 mg, QD) + MTX* (oral, QW)	MTX* (oral, QW)	RA-BEGIN Clinical Study Report[78]	RA-BEGIN Clinical Study Report[78]
RA-BEYOND[15]	Patients with moderate to severe RA who completed Phase 2b study JADA or Phase 3 studies JADZ, JADV, JADX or JADW	Baricitinib (2 mg, oral, QD) Baricitinib (4 mg, oral, QD)	N/A	RA-BEYOND Clinical Study Report[15]	RA-BEYOND Clinical Study Report[15]

Footnotes: *Patients received MTX starting at 10 mg per week and escalated by 5 mg every 4 weeks to a maximum of 20 mg per week.
Abbreviations: DMARD = disease-modifying antirheumatic drug, RA = rheumatoid arthritis, bDMARD = biologic DMARD, cDMARD = conventional DMARD, MTX = methotrexate, QD = once daily, QW = once weekly.

4.3 *Summary of methodology of the relevant randomised controlled trials*

Details of each randomised controlled trial are presented in this section. A comparative summary of the methodology of the three phase III baricitinib trials included in this submission is provided in Table 16. The methods and results of RA-BEGIN, a phase III study of an unlicensed indication of baricitinib in cDMARD-naïve patients with moderate to severe RA, are presented in Appendix 1.

4.3.1 Trial designs

All of the baricitinib trials presented here were multicentre, double-blind, double-dummy, outpatient, phase III randomised controlled trials. RA-BEAM was a 52-week study, whilst RA-BUILD and RA-BEACON were 24 weeks in length.[9,11,12] Additionally, RA-BEAM included non-inferiority and pre-specified superiority testing against adalimumab as an active control. Upon completion of one of these originating studies, patients either continued into the post-treatment follow-up period of the study, which comprised a follow-up visit approximately 28 days after the last dose of study drug, or entered the long-term extension study RA-BEYOND to assess the long-term safety and efficacy of baricitinib.

Across the phase III studies, patients who were non-responders to therapy were eligible for rescue therapy at Week 16 or 24. No response was defined as a lack of improvement of ≥ 20% in tender joint count (TJC) and swollen joint count (SJC) at Week 14 and Week 16 compared with baseline for RA-BEAM, RA-BUILD and RA-BEACON. In RA-BEAM only, all patients in the placebo arm were switched to baricitinib + MTX treatment at Week 24 regardless of response status. To ensure investigational product blinding, patients with renal impairment who received 2 mg QD baricitinib during a study, continued to receive 2 mg QD baricitinib. Patients not experiencing improvement in signs and symptoms following at least four weeks of rescue treatment were discontinued from a study. After a patient was rescued, they were classified as a non-responder in subsequent efficacy analyses.[9,11,12]

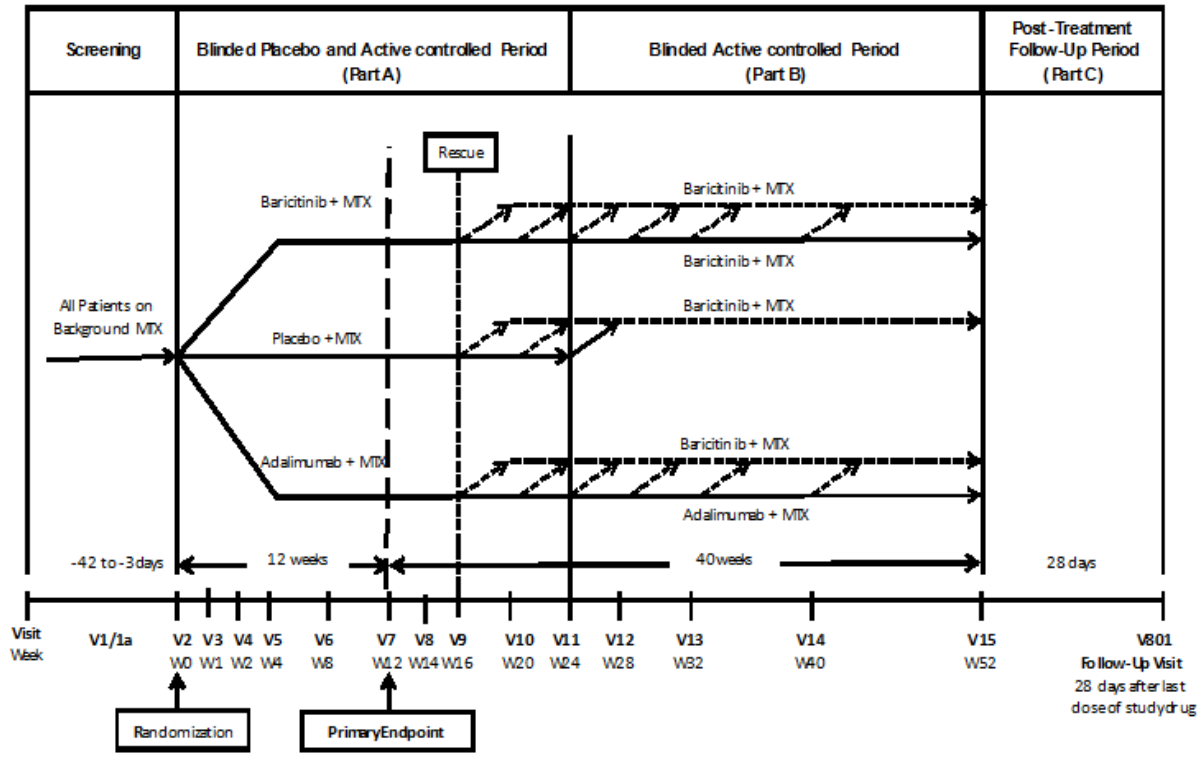
A summary of each phase III trial design is included below.

RA-BEAM

RA-BEAM was a Phase III, 52-week, double-blind RCT conducted to evaluate the efficacy and safety of baricitinib 4 mg (QD) compared to placebo and adalimumab (all patients on background MTX) in patients with moderately to severely active RA who had experienced an inadequate response to MTX and had never been treated with a bDMARD. Adalimumab was selected as an appropriate active comparator to baricitinib due to the fact that TNF inhibitors are the most commonly used biologic class of therapies for the treatment of moderately to severely active RA in the UK.[9]

The RA-BEAM trial design is illustrated in Figure 6. Adalimumab was administered from Week 0 to Week 50, in order to include a 2 week wash-out period in Week 50 to Week 52, prior to initiation of baricitinib therapy in patients proceeding to the long-term safety study RA-BEYOND.[9] All patients in the placebo arm were switched to baricitinib + MTX treatment at 24 weeks regardsless of response at 24 weeks.

Figure 6. Schematic of study design for RA-BEAM

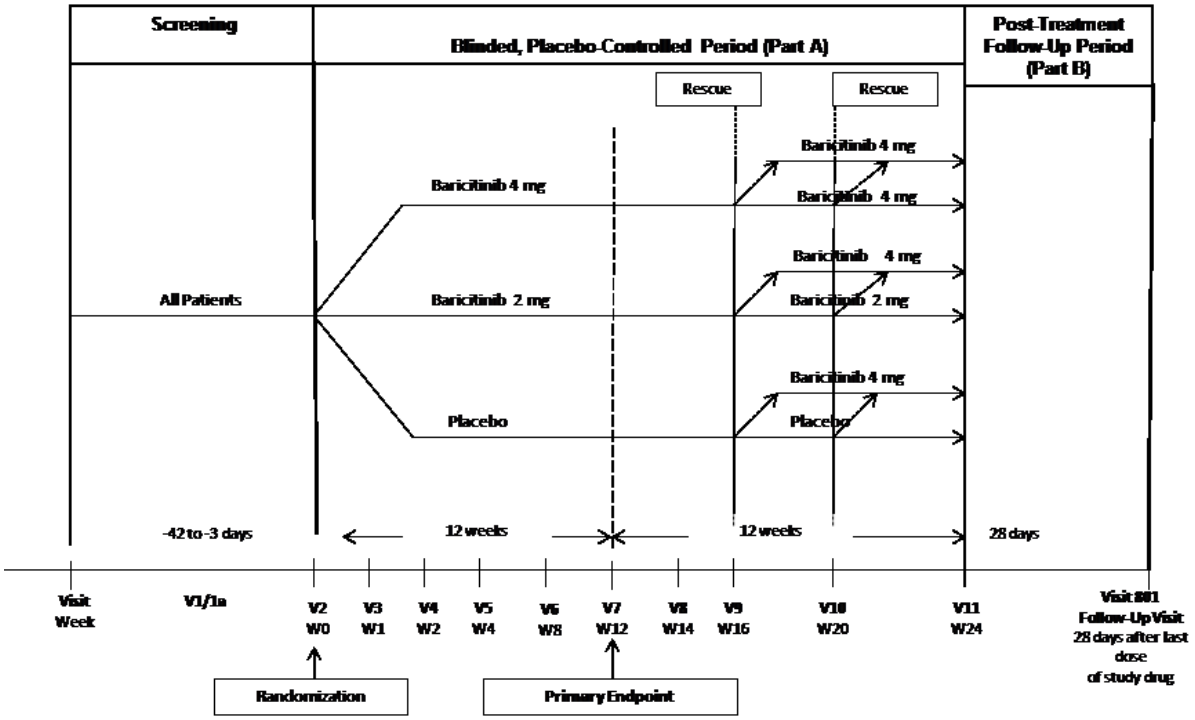


Footnotes: Diagonal dashed arrows indicate an option for rescue therapy. The diagonal solid arrow indicates a mandatory change to baricitinib treatment at Week 24 for placebo treated patients. Rescued patients reassigned to BAR4, regardless of original treatment assignment, and classified as non-responder in subsequent efficacy analyses of categorical variables. Patients with an eGFR of <60 mL/min/1.73 m^2 received baricitinib 2 mg QD, irrespective of original treatment allocation.
Abbreviations: V = study visit; W = study week, MTX = methotrexate, eGFR = estimated glomerular filtration rate.

RA-BUILD

RA-BUILD[11] was a Phase III, 24-week, double-blind RCT conducted to evaluate the efficacy and safety of baricitinib (2 mg and 4 mg, QD) compared to placebo in patients with moderately to severely active RA who had experienced an inadequate response or were intolerant to ≥1 cDMARD and had not received a bDMARD. All patients continued on their background cDMARD(s). The RA-BUILD trial design is illustrated in Figure 7.

Figure 7. Schematic of study design for RA-BUILD

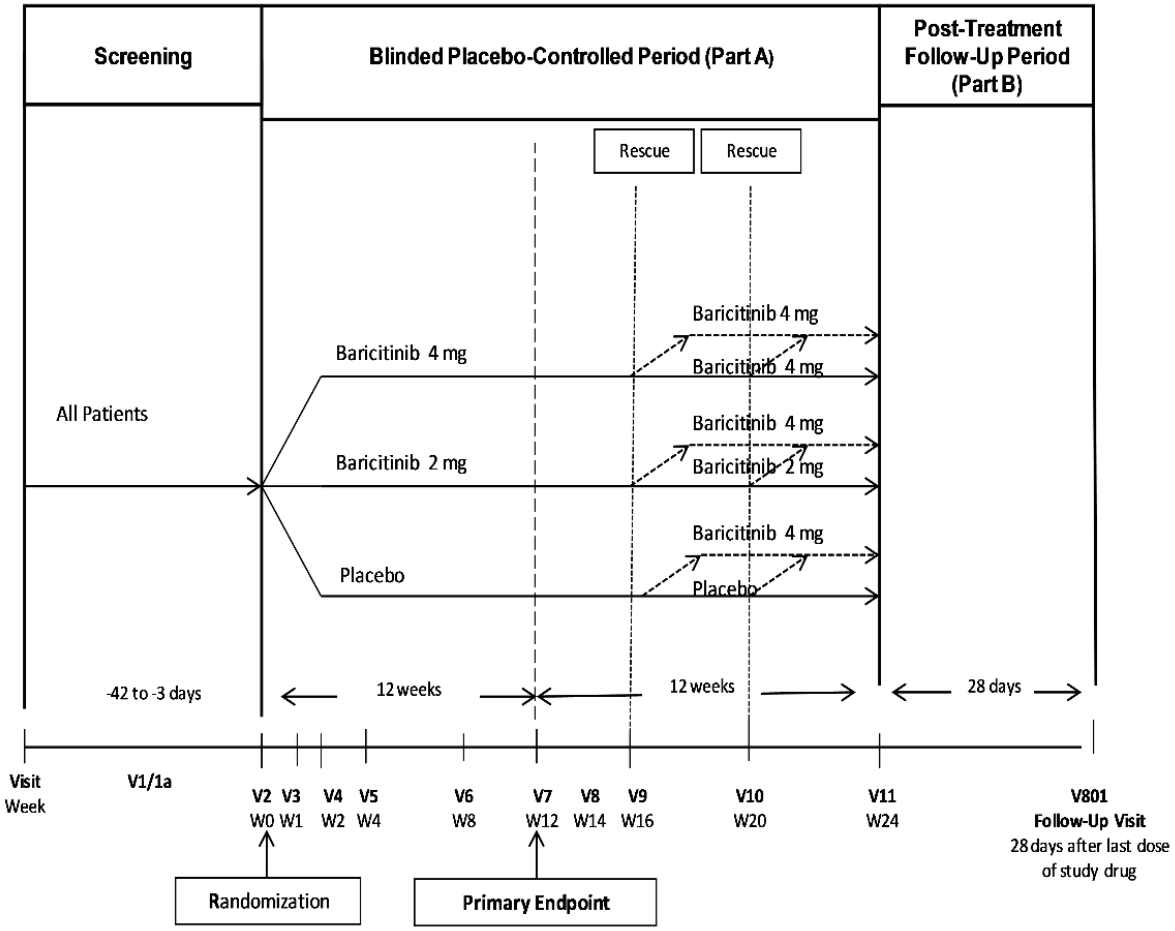


Footnotes: Diagonal dashed arrows indicate an option for rescue therapy. Rescued patients reassigned to BAR4, regardless of original treatment assignment, and classified as non-responder in subsequent efficacy analyses of categorical variables. Patients with an eGFR of <60 mL/min/1.73 m^2 received baricitinib 2 mg QD, irrespective of original treatment allocation.
Abbreviations: V = study visit; W = study week, MTX = methotrexate, eGFR = estimated glomerular filtration rate.

RA-BEACON

RA-BEACON[12] was a Phase III, 24-week, double-blind RCT conducted to evaluate the efficacy and safety of baricitinib (2 mg and 4 mg, QD) compared to placebo in patients with moderately to severely active RA who had experienced an inadequate response or intolerance to at least 1 TNF-inhibitor. All patients continued on background cDMARD(s) but any biologic treatment was discontinued prior to study entry. The RA-BEACON trial design is illustrated in Figure 8.

Figure 8. Schematic of study design for RA-BEACON



Footnotes: Diagonal dashed arrows indicate an option for rescue therapy. Rescued patients reassigned to BAR4, regardless of original treatment assignment, and classified as non-responder in subsequent efficacy analyses of categorical variables. Patients with a eGFR of <60 mL/min/1.73 m^2 received baricitinib 2 mg QD, irrespective of original treatment allocation.
Abbreviations: V = study visit; W = study week, MTX = methotrexate, eGFR = estimated glomerular filtration rate.

4.3.2 Randomisation and blinding

Randomisation schemes for the completed phase III studies were driven by the sample sizes needed to provide acceptable statistical power for the respective comparisons between treatment arms.

In RA-BEAM, patients were randomised (3:3:2) to receive baricitinib (4 mg QD), placebo or adalimumab (40 mg Q2W), all with background MTX. The 3:3:2 ratio was driven by the sample size needed (approximately 1,280 patients [480 baricitinib, 480 placebo, and 320 adalimumab]) to provide:

- >95% power to detect a difference between the baricitinib and placebo treatment groups in ACR20 response rate at Week 12
- Approximately 94% power to detect a difference in mTSS between the baricitinib and placebo treatment groups
- Approximately 93% power for the non-inferiority analysis based on a margin of 12% of ACR20 response rate at Week 12 between the baricitinib and adalimumab treatment groups.
- Randomisation was stratified by region and joint erosion status, and conducted by a computer-generated random sequence using an IVRS.

- In RA-BUILD, patients were randomised (1:1:1) to receive baricitinib (4 mg QD), baricitinib (2 mg QD) or placebo. The 1:1:1 ratio was driven by the sample size needed (approximately 660 patients [220 baricitinib 4 mg, 220 baricitinib 2 mg, and 220 placebo]) to provide:
- >95% power to detect a difference between baricitinib 4 mg (QD) and placebo in ACR20 response rate Week 12
- >90% power to detect a difference between baricitinib 2 mg (QD) and placebo in ACR20 response rate at Week 12

Randomisation was stratified by region and joint erosion status, and conducted by a computer-generated random sequence using an IVRS.

Finally, in RA-BEACON, patients were randomised (1:1:1) to receive baricitinib (4 mg QD), baricitinib (2 mg QD) or placebo. The 1:1:1 ratio was driven by the sample size needed (approximately 525 patients [175 baricitinib 4 mg, 175 baricitinib 2 mg, and 175 placebo]) to provide:

- 97% power to detect a difference between baricitinib 4 mg (QD) and placebo groups in ACR20 response rate at Week 12
- 80% power to detect a difference between baricitinib 2 mg (QD) and placebo groups in ACR20 response rate at Week 12

Randomisation was stratified by region and history of bDMARD use, and conducted by a computer-generated random sequence using an IVRS.

All of the baricitinib trials were double-blind, double-dummy studies in order to maintain the blind among 3 different treatment arms (baricitinib, placebo and adalimumab for RA-BEAM and baricitinib 4 mg [QD], baricitinib 2 mg [QD] and placebo for RA-BUILD and RA-BEACON). In particular, for RA-BEAM it was important that patients not assigned to adalimumab received a matching placebo subcutaneous (SC) injection biweekly. All patients with an eGFR <60 mL/min/1.73m^2 randomised to baricitinib 4 mg received a 2 mg dose.

All study personnel remained blinded through the completion of RA-BUILD and RA-BEACON. To preserve the blinding, the number of Lilly personnel who had access to the randomisation table and treatment assignments (as required for un-blinded safety reporting to regulatory agencies) was kept to a minimum. However, in RA-BEAM a number of sponsor personnel and external contractors were un-blinded to the 24-week data for the purpose of preparing for regulatory interactions. Robust safeguards were prospectively established to ensure that this un-blinded group did not have direct contact on patient-level matters with investigators or study site personnel, to ensure continued blinding of treatment assignments and to prevent bias in study conduct until the final database lock. As such, patients, investigators, and all other personnel involved in the conduct of the study remained blinded to the 24-week RA-BEAM data and to individual treatment assignments, until the contribution of their sites to the studies were completed. For the long-term extension study RA-BEYOND, both patients and their investigators remained blinded to the initial treatment assignment in the originator study.

4.3.3 Eligibility criteria

Key eligibility criteria for the baricitinib trials can be seen in Table 13. A complete list of all inclusion and exclusion criteria is given in Appendix 5.

Table 13. Eligibility criteria for baricitinib phase III trials

	RA-BEAM[9]	RA-BUILD[11]	RA-BEACON[12]
Inclusion criteria	• Adults with a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA[79] • The presence of at least 6/68 tender joints and at least 6/66 swollen joints • HsCRP measurement of ≥6 mg/L • At least 12 weeks of MTX therapy prior to study entry with 8 weeks at a stable dose (7.5–25 mg/week, but if <15 mg/week, documentation of clinical rationale should have been provided) • ≥3 joint erosions in hand, wrist or foot joints based on radiographs or ≥1 joint erosion and be RF or ACPA antibody positive	• Adults with a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA[79] • The presence of at least 6/68 tender joints and at least 6/66 swollen joints • HsCRP measurement of ≥1.2 x ULN • Had failed treatment at an approved dose with more than one cDMARD (experienced insufficient efficacy or were intolerant to treatment)	• Adults with a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA[79] • The presence of at least 6/68 tender joints and at least 6/66 swollen joints • HsCRP measurement of ≥ 1x ULN • Receiving stable doses of background cDMARD therapy • Had failed treatment at an approved dose with at least one biologic TNFi DMARD (experienced insufficient efficacy or were intolerant to treatment) • Patients who had received other bDMARDs could also participate
Exclusion criteria	• Receiving/had previously received prohibited RA therapies (bDMARDs) • Recent history of infection or tested positive for TB or other serious infections • Immunocompromised or had specific abnormal laboratory tests • Comorbidities that put patients at risk of adverse events when taking study drug A complete list of inclusion/exclusion criteria is presented in Appendix 5.	• Receiving/had previously received any bDMARD • Recent history of infection or tested positive for TB or other serious infections • Immunocompromised or had specific abnormal laboratory tests • Comorbidities that put patients at risk of adverse events when taking study drug A complete list of inclusion/exclusion criteria is presented in Appendix 5.	• Received bDMARDs within 28 days before randomisation (6 months before for rituximab) • Recent history of infection or tested positive for TB or other serious infections • Immunocompromised or had specific abnormal laboratory tests • Comorbidities that put patients at risk of adverse events when taking study drug A complete list of inclusion/exclusion criteria is presented in Appendix 5.

Abbreviations: RA = rheumatoid arthritis, DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, bDMARD = biological DMARD, MTX = methotrexate, IR = inadequate response, TNFi = tumour necrosis factor inhibitor, ACR = American College of Rheumatology, EULAR = European League Against Rheumatism, ACPA = anti-citrullinated protein antibody, hsCRP = high-sensitivity C-reactive protein, ULN = upper limit of normal, TB = tuberculosis, OD = once daily, Q2W = twice weekly.

4.3.4 Settings and locations where the data were collected

RA-BEAM, RA-BUILD and RA-BEACON were international, multicentre trials conducted in outpatient settings in 26 countries across North America, South America, Europe and Asia. Across the three studies there were a total of [redacted] study sites based in the UK which enrolled a total of [redacted] patients.

A summary of trial setting and data collection locations is provided in Table 16.

4.3.5 Study drugs and concomitant medications

Based on efficacy, safety, and PK data from the Phase II studies (JADC and JADA), a single dose of baricitinib 4 mg (QD) was selected for evaluation in RA-BEAM, and as one of the treatment arms in RA-BUILD and RA-BEACON. As baricitinib is excreted mostly by the kidneys, patients' renal function was assessed during screening for all studies. The dose of baricitinib for patients with an eGFR <60 mL/min/1.73 m^2 was baricitinib 2 mg (QD) if assigned to a baricitinib treatment. Patients with renal impairment were analysed according to the treatment group to which they were assigned at randomisation.

Permitted and prohibited concomitant medications are listed in Table 14. For RA-BEAM, prior treatment with MTX was required for study participation (at least 12 weeks with treatment at a stable dose for at least 8 weeks prior to study entry). For RA-BUILD and RA-BEACON, prior treatment with cDMARDs was required for study participation (at least 12 weeks with treatment at a stable dose for at least 8 weeks prior to study). Finally, for RA-BEACON, patients had failed treatment at an approved dose with at least one biologic TNF inhibitor (experienced insufficient efficacy or were intolerant to treatment).

Table 14. Permitted and prohibited concomitant medications

	RA-BEAM[9]	RA-BUILD[8]	RA-BEACON[10]
Permitted concomitant medications	• Hydroxychloroquine or sulfasalazine at a stable dose for at least 8 weeks prior to study entry • NSAIDs and/or prednisone (≤10 mg) (or equivalent) were permitted only if the patient was on a stable dose for at least 6 weeks prior to randomisation • Analgesics, but dose increases and/or introduction of a new analgesic were not permitted	• Hydroxychloroquine, sulfasalazine, leflunomide or azathioprine at a stable dose for at least 8 weeks prior to study entry • NSAIDs and/or prednisone (≤10 mg) (or equivalent) were permitted only if the patient was on a stable dose for at least 6 weeks prior to randomisation • Analgesics, but dose increases and/or introduction of a new analgesic were not permitted	• Patients taking MTX must have been using MTX for at least 12 weeks with treatment at a stable dose for at least 8 weeks • Patients taking hydroxychloroquine, sulfasalazine, leflunomide or azathioprine must have been taking a stable dose for at least 8 weeks prior to study entry • NSAIDs and/or prednisone (≤10 mg) (or equivalent) were permitted only if the patient was on a stable dose for at least 6 weeks prior to randomisation • Analgesics, but dose increases and/or

			introduction of a new analgesic were not permitted
Prohibited concomitant medications	• Live vaccines (non-live seasonal vaccinations and/or emergency vaccination were allowed) • Any DMARDs, other than stable doses of background DMARDs being used at the time of study entry • Any biologic therapy for any indication • Any interferon therapy • Any parenteral corticosteroid administered by IM or IV injection		

Abbreviations: RA = rheumatoid arthritis, NSAID = non-steroidal anti-inflammatory drug, MTX = methotrexate, DMARD = disease-modifying antirheumatic drug, IM = intramuscular, IV = intravenous.

A summary of study drugs and concomitant medications is provided in Table 16.

4.3.6 Primary, secondary and tertiary outcomes

4.3.6.1 Primary outcome

The primary efficacy outcome in each of the phase III studies was 20% improvement in American College of Rheumatology criteria (ACR20), defined as at least 20% improvement in the following ACR Core Set values:

- Tender joint count (68 joint count)
- Swollen joint count (66 joint count)
- An improvement of at least 20% in at least 3 of the following 5 assessments:
    - Patient's assessment of pain (VAS)
    - Patient's Global Assessment of Disease Activity (VAS)
    - Physician's Global Assessment of Disease Activity (VAS)
    - Patient's assessment of physical function as measured by the HAQ-DI
    - Acute phase reactant as measured by hsCRP

Using this measure, the primary objectives of the studies were:

- To determine whether baricitinib 4 mg (QD) was superior to placebo as assessed by the proportion of patients achieving an ACR20 response at Week 12 (RA-BEAM, RA-BUILD, RA-BEACON)

4.3.6.2 Secondary outcomes

A summary of selected secondary objectives of the three baricitinib phase III trials is presented in Table 15. A description of the secondary outcomes is provided in Appendix 6.

Table 15. Secondary objectives for baricitinib phase III trials

Secondary objectives	RA-BEAM[9]	RA-BUILD[8]	RA-BEACON[10]
	To evaluate the efficacy of baricitinib 4 mg (QD) as assessed by: • Proportion of patients who achieved an ACR50 response • Proportion of patients who achieved an ACR70 response • Change from baseline to Week 24 by mTSS (van der Heijde method) compared to placebo • Change from baseline to Week 12 in HAQ-DI score compared to placebo • Change from baseline to Week 12 in DAS28-hsCRP compared to placebo • Proportion of patients achieving ACR20 response at Week 12 compared to adalimumab • Proportion of patients achieving an SDAI score ≤3.3 at Week 12 compared to placebo • Median duration of morning joint stiffness in the 7 days prior to Week 12 compared to placebo as collected in electronic diaries • Change from baseline to Week 12 in DAS28-hsCRP compared to adalimumab • Mean severity of morning joint stiffness in the 7 days prior to Week 12 compared to placebo as collected in electronic diaries • Mean Worst Tiredness NRS in the 7 days prior to Week 12 compared	To evaluate the efficacy of baricitinib 4 mg (QD) versus placebo as assessed by: • Proportion of patients who achieved an ACR50 response • Proportion of patients who achieved an ACR70 response • Change from baseline to Week 12 in Health Assessment Questionnaire–Disability Index (HAQ-DI) score • Change from baseline to Week 12 in DAS28-hsCRP • Proportion of patients achieving an SDAI score ≤3.3 at Week 12 • Median duration of morning joint stiffness in the 7 days prior to Week 12 as collected in electronic diaries • Mean severity of morning joint stiffness in the 7 days prior to Week 12 as collected in electronic diaries • Mean Worst Tiredness NRS in the 7 days prior to Week 12 as collected in electronic diaries To evaluate the efficacy of baricitinib 2 mg (QD) versus placebo as assessed by: • Proportion of patients achieving ACR20 at Week 12 • Change from baseline to Week 12 in HAQ-DI score • Change from baseline to Week 12 in DAS28-hsCRP	To evaluate the efficacy of baricitinib 4 mg (QD) as assessed by: • Proportion of patients who achieved an ACR50 response • Proportion of patients who achieved an ACR70 response • Change from baseline to Week 12 in HAQ-DI score compared to placebo • Change from baseline to Week 12 in DAS28-hsCRP compared to placebo • Proportion of patients who achieved a SDAI score ≤3.3 at Week 12 compared to placebo To evaluate the efficacy of baricitinib 2 mg (QD) versus placebo as assessed by: • Proportion of patients who achieved ACR20 at Week 12 compared to placebo • Change from baseline to Week 12 in HAQ-DI score compared to placebo • Change from baseline to Week 12 in DAS28-hsCRP compared to placebo • Proportion of patients who achieved an SDAI score ≤3.3 at Week 12 compared to placebo • Proportion of patients achieving EULAR good+moderate and good responses at Week 12 compared to placebo

	to placebo as collected in electronic diaries • Proportion of patients achieving EULAR good+moderate and good responses at Week 12 compared to placebo	• Proportion of patients achieving SDAI score ≤3.3 at Week 12 • Proportion of patients achieving EULAR good+moderate and good responses at Week 12 compared to placebo.	

Abbreviations: RA = rheumatoid arthritis, DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, bDMARD = biological DMARD, MTX = methotrexate, TNFi = tumour necrosis factor inhibitor, ACR = American College of Rheumatology, EULAR = European League Against Rheumatism, ACPA = anti-citrullinated protein antibody, hsCRP = high-sensitivity C-reactive protein, ULN = upper limit of normal, TB = tuberculosis, OD = once daily, Q2W = twice weekly, HAQ-DI = Health Assessment Questionnaire-Disability Index, mTSS = modified Total Sharp Score, SDAI = Simplified Disease Activity Index, NRS = Numeric Rating Scale.

4.3.6.3 Other secondary outcomes

Other secondary objectives are listed in Appendix 7.

Table 16. Comparative summary of trial methodology

	RA-BEAM (MTX-IR)[9]	RA-BUILD (cDMARD-IR)[8]	RA-BEACON (TNFi-IR)[10]
Trial design	Multicentre, 52-week, double-blind, double-dummy, placebo- and active-controlled, parallel-group, outpatient, phase III randomised controlled trial.	Multicentre, 24-week, double-blind, double-dummy placebo-controlled, outpatient, phase III randomised controlled trial.	Multicentre, 24-week, double-blind, double-dummy, placebo-controlled, outpatient phase III randomised controlled trial.
Method of randomisation	Patients were randomised (3:3:2) to receive baricitinib (4 mg QD), placebo or adalimumab (40 mg Q2W), all with background MTX. Randomisation was stratified by region and joint erosion status, and conducted by a computer-generated random sequence using an IVRS.	Patients were randomised (1:1:1) to receive baricitinib (4 mg QD), baricitinib (2 mg QD) or placebo. Patients on ≥1 cDMARDs (with or without MTX) continued to take background therapy during study. Randomisation was stratified by region and joint erosion status, and conducted by a computer-generated random sequence using an IVRS.	Patients were randomised (1:1:1) to receive baricitinib (4 mg QD), baricitinib (2 mg QD) or placebo. Randomisation was stratified by region and history of bDMARD use, and conducted by a computer-generated random sequence using an IVRS.
Key eligibility criteria for participants	Adults with a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA[79]	Adults with a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA[79]	Adults with a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA[79]
Settings and locations where the data were collected	International: 281 study sites across 26 countries, including North America, South America, Europe and Asia. Seven study sites were included in the UK, [redacted]	International: 182 study sites across 22 countries, including North America, South America, Europe and Asia. Four study sites were included in the UK, [redacted]	International: 178 study centres across 24 countries, including North America, South America, Europe and Asia. Three study sites were included in the UK, [redacted]

Trial drugs	**Intervention:** Baricitinib* (oral, 4 mg, QD) (n=487) (with background MTX) **Comparators:** Adalimumab (SC, 40 mg, Q2W) (n=330) (with background MTX) Placebo (n=488) (with background MTX)	**Interventions:** Baricitinib (oral, 2 mg, QD) (n=229) (on background cDMARDs) Baricitinib‡ (oral, 4 mg, QD) (n=227) (on background cDMARDs) **Comparator:**[8] Placebo (n=228) (99% of patients were on ≥1 cDMARDs (with or without MTX), and continued to take background therapy during study in both the intervention and comparator arms)	**Interventions:** Baricitinib (oral, 2 mg, QD, on background cDMARDs) (n=174) Baricitinib‡ (oral, 4 mg, QD on background cDMARDs) (n=177) **Comparator:**[10] Placebo (n=176)
Primary outcomes	**ACR20** To determine whether baricitinib was superior to placebo by the proportion of patients achieving an ACR20 response at Week 12	**ACR20**[11] To determine whether baricitinib was superior to placebo by the proportion of patients achieving an ACR20 response at Week 12	**ACR20** To determine whether baricitinib 4 mg (QD) was superior to placebo by the proportion of patients who achieved an ACR20 response at Week 12

Secondary outcomes	**Baricitinib vs placebo (at Week 12):** • Change from baseline in HAQ-DI score • Change from baseline in DAS28-hsCRP • SDAI ≤3.3 response rates • Median duration of morning joint stiffness (7 days prior to week 12) • Mean severity of morning joint stiffness (7 days prior to week 12) • Mean worst tiredness (7 days prior to week 12) • Radiographic progression of structural joint damage (mTSS) **Baricitinib vs adalimumab (at Week 12):** • ACR20 response rates • Change from baseline in DAS28-hsCRP	**Baricitinib 4 mg (QD) vs placebo (at Week 12):** • Change from baseline in HAQ-DI score • Change from baseline in DAS28-hsCRP • SDAI ≤3.3 response rates • Median duration of morning joint stiffness (7 days prior to week 12) • Mean severity of morning joint stiffness (7 days prior to week 12) • Mean worst tiredness (7 days prior to week 12) **Baricitinib 2 mg (QD) vs placebo (at Week 12):** • ACR20 response rates • Change from baseline in HAQ-DI score • Change from baseline in DAS28-hsCRP • SDAI ≤3.3 response rates	**Baricitinib 4 mg (QD) vs placebo (at Week 12):** • Change from baseline in HAQ-DI score • Change from baseline in DAS28-hsCRP • SDAI ≤3.3 response rates **Baricitinib 2 mg (QD) vs placebo (at Week 12):** • ACR20 response rates • Change from baseline in HAQ-DI score • Change from baseline in DAS28-hsCRP • SDAI ≤3.3 response rates
Pre-planned subgroups	Subgroups analyses comparing baricitinib to placebo were performed using ACR20, ACR50, HAQ-DI and DAS28(hsCRP) at Weeks 12 and 24 and mTSS at Week 24. Subgroups evaluated included region, renal function, background therapy, joint erosion status, country, gender, age, race etc.	Subgroups analyses comparing baricitinib 4 mg (QD) to placebo and baricitinib 2 mg (QD) to placebo were performed using ACR20, ACR50, HAQ-DI and DAS28(hsCRP) at Weeks 12 and 24, and mTSS at Week 24. Subgroups evaluated included region, renal function, baseline joint erosion status, country, gender, age, race, etc.	Subgroup analyses comparing baricitinib 4 mg (QD) to placebo and baricitinib 2 mg (QD) to placebo were performed using ACR20, ACR50, HAQ-DI and DAS28-hsCRP at Weeks 12 and 24. Subgroups evaluated included region, renal function, history of bDMARD use at screening, country, gender, age, race, etc.

Footnotes: *The dose for patients with renal impairment, defined as estimated glomerular filtration rate <60 mL/min/1.73 m^2, was 2 mg baricitinib OD, ‡Patients with renal impairment received baricitinib 2 mg irrespective of active treatment group assignment.
Abbreviations: RA = rheumatoid arthritis, DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, bDMARD = biological DMARD, MTX = methotrexate, IR = inadequate response, TNFi = tumour necrosis factor inhibitor, DMC = data monitoring committee, IVRS = Interactive Voice Response System, ACR = American College of Rheumatology, EULAR = European League Against Rheumatism, ACPA = anti-citrullinated protein antibody, hsCRP = high-sensitivity C-reactive protein, ULN = upper limit of normal, TB = tuberculosis, OD = once daily, Q2W = twice

weekly, NSAID = non-steroidal anti-inflammatory drug, IM = intramuscular, IV = intravenous, HAQ-DI = Health Assessment Questionnaire-Disability Index, mTSS = modified Total Sharp Score, SDAI = Simplified Disease Activity Index, NRS = Numeric Rating Scale.

4.4 *Statistical analysis and definition of study groups in the relevant randomised controlled trials*

Analysis populations

A summary of the analysis populations for efficacy and safety outcomes for each of the four trials is presented Table 17.

Table 17. Summary of analysis populations

	RA-BEAM[9]	RA-BUILD[11]	RA-BEACON[12]
Efficacy analysis population	**Modified intention to treat:** All randomised patients who received at least 1 dose of study drug.	**Modified intention to treat:** All randomised patients who received at least 1 dose of study drug.[8]	**Modified intention to treat:** All randomised patients who received at least 1 dose of study drug.[10]
Safety analysis population	**Safety:** All randomised patients who received at least 1 dose of study drug and who did not discontinue from the study for the reason 'lost to follow-up' at the first post-baseline visit.	**Safety:** All randomised patients who received at least 1 dose of study drug and who did not discontinue from the study for the reason 'lost to follow-up' at the first post-baseline visit.	**Safety:** All randomised patients who received at least 1 dose of study drug and who did not discontinue from the study for the reason 'lost to follow-up' at the first post-baseline visit.

Table 18. Summary of statistical analyses for the primary efficacy analysis in the RCTs

Trial number (acronym)	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
RA-BEAM[9]	To determine whether baricitinib was superior to placebo as assessed by the proportion of patients achieving an ACR20 response at Week 12.[9]	ACR20 response at Week 12 was evaluated in mITT population. Treatment comparisons between the baricitinib 4 mg/day group and placebo group were made using a logistic regression analysis, with region, joint erosion status (1 to 2 joint erosions plus seropositivity vs at least 3 joint erosions) and treatment group as explanatory factors in the model. A P-value of 0.05 or less (two-sided) was considered to indicate statistical significance. Treatment comparisons between baricitinib and adalimumab were made using a pre-defined logistic regression model with treatment, region and joint erosion for ACR20 response at Week 12. If the lower bound of the 95% CI was >-12%, it was concluded that baricitinib is noninferior to adalimumab. Similarly, if the lower bound of the 95% CI was > 0%, Lilly concluded that baricitinib was superior to adalimumab.	A total of approximately 1,280 patients (480 in the baricitinib 4 mg/day group, 480 in the placebo group and 320 in the adalimumab group) was estimated to provide >95% power to detect a difference between the baricitinib and placebo treatment groups in ACR20 response rate (assuming 60% vs 35%) at Week 12.	Patients who received rescue treatment or discontinued from the study underwent non-responder imputation for the primary endpoint.

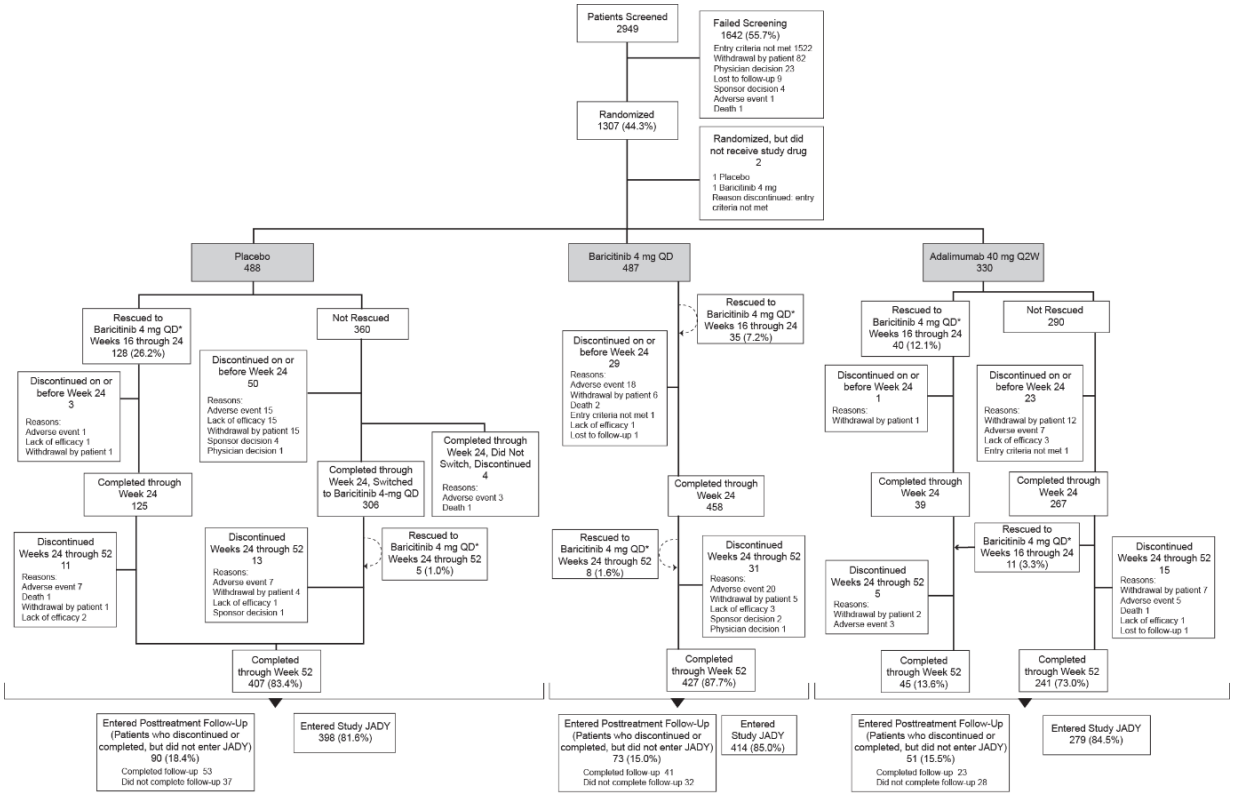
RA-BUILD[8]	To determine whether baricitinib 4 mg/day was superior to placebo as assessed by the proportion of patients who achieved an ACR20 response at Week 12.	ACR20 response at Week 12 was evaluated in the mITT population. Treatment comparisons between the baricitinib (4 mg QD) group and placebo were made using a logistic regression analysis, with treatment, region and presence of baseline joint erosions as explanatory factors in the model.	It was estimated that a sample size of approximately 220 patients per study group would provide >95% power for the comparison of the ACR20 response rate between the baricitinib (4 mg QD) and placebo groups (with assumed response rates of 60% and 35%, respectively) at Week 12.	Patients who received rescue treatment or discontinued from the study underwent non-responder imputation for the primary endpoint.
RA-BEACON[10]	To determine whether baricitinib 4 mg/day was superior to placebo as assessed by the proportion of patients who achieved an ACR20 response at Week 12.	ACR20 response at Week 12 was evaluated in the mITT population. Treatment comparisons between the baricitinib (4 mg QD) group and placebo were made using a logistic regression analysis, with region and history of bDMARD use at screening (<3, ≥3) as explanatory factors in the model. A P value of 0.05 or less (two-sided) was considered to indicate statistical significance.	It was estimated that a sample size of approximately 175 patients per study group would provide >90% power for the comparison of the ACR20 response rate between the 4 mg/day baricitinib group and the placebo group (with assumed response rates of 45% and 25%, respectively) at Week 12.	Patients who received rescue treatment or discontinued from the study underwent non-responder imputation for the primary endpoint.

Abbreviations: ACR20 = American College of Rheumatology 20% response, bDMARD = biologic disease-modifying antirheumatic drug; CI = confidence interval; mITT = modified intention to treat, MTX = methotrexate, QD = once daily.

4.5 *Participant flow in the relevant randomised controlled trials*

4.5.1 Participant Flow

Figure 9. CONSORT diagram showing patient flow in RA-BEAM[9]

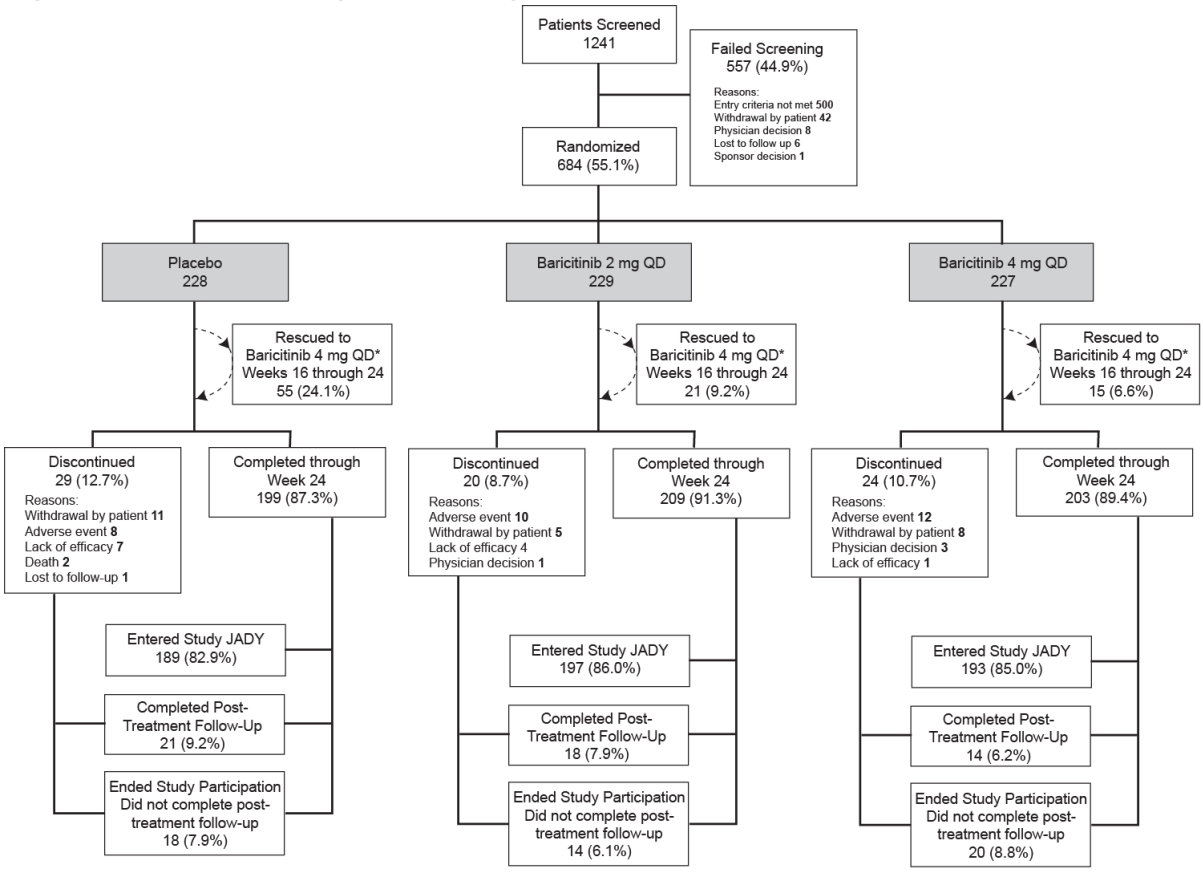


Footnotes: *Patients who met rescue criteria received baricitinib 4 mg (QD) and may also have received higher doses of ongoing concomitant non-steroidal anti-inflammatory drugs (NSAIDs), corticosteroids, or analgesics or new NSAIDs, corticosteroids, or analgesics could have been added.
Abbreviations: QD = once daily, Q2W = twice weekly, NSAID = non-steroidal anti-inflammatory drug.

RA-BEAM

A total of 1,307 patients were randomised in RA-BEAM. In total, 488, 487 and 330 patients were randomised to (and received at least one dose of) placebo, baricitinib and adalimumab, respectively. Of those patients not rescued to baricitinib, 50/360 (13.9%), 29/487 (6.0%) and 23/290 (7.9%) patients discontinued treatment on or before Week 24 in the placebo, baricitinib and adalimumab arms, respectively. The number of patients entering the long-term extension study was similar across all arms, with 398 (81.6%), 414 (85.0%) and 279 (84.5%) for the placebo, baricitinib and adalimumab arms, respectively. Full details of participant flow are presented in the CONSORT diagram in Figure 9.[9]

Figure 10. CONSORT diagram showing patient flow in RA-BUILD[8]

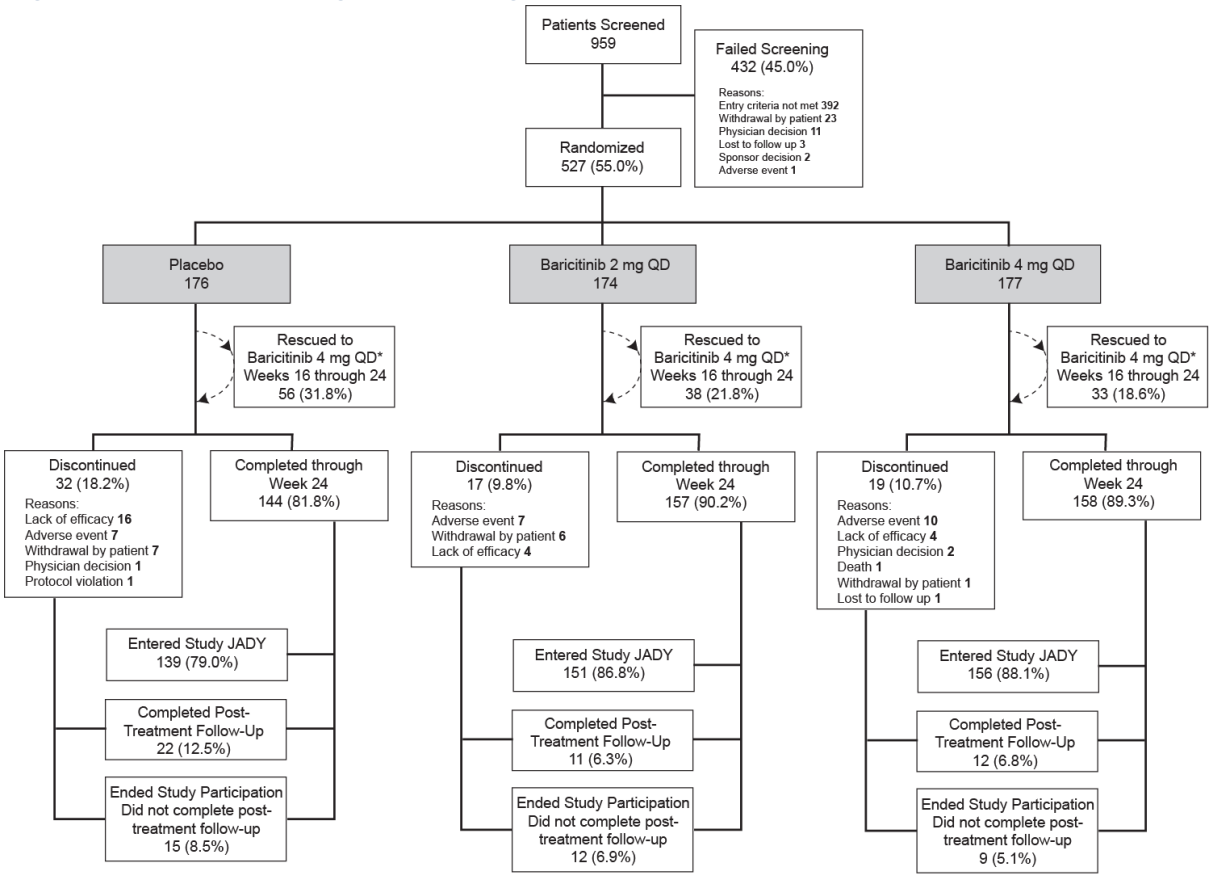


Footnotes: *Patients who met rescue criteria continued to receive baricitinib 4 mg (QD), but may have received a higher dose of or an additional allowed concomitant medication.
Abbreviation: QD = once daily.

RA-BUILD

A total of 684 patients were randomised in RA-BUILD. In total, 228, 229 and 227 patients were randomised to placebo, baricitinib 2 mg QD and baricitinib 4 mg QD, respectively. Of these, 29/228 (12.7%), 20/229 (8.7%) and 24/227 (10.7%) patients discontinued treatment on or before Week 24 in the placebo, baricitinib 2 mg QD and baricitinib 4 mg QD arms, respectively. The number of patients entering the long-term extension study was similar across all arms, with 189 (82.9%), 197 (86.0%) and 193 (85.0%) for the placebo, baricitinib 2 mg QD and baricitinib 4 mg QD arms, respectively. Full details of participant flow are presented in the CONSORT diagram in Figure 10.[11]

Figure 11. CONSORT diagram showing patient flow in RA-BEACON[10]



Footnotes: *Patients who met rescue criteria continued to receive baricitinib 4 mg (QD), but may have received a higher dose of or an additional allowed concomitant medication.
Abbreviations: QD = once daily, Q2W = twice weekly.

RA-BEACON

A total of 527 patients were randomised in RA-BEACON. In total, 176, 174 and 177 patients were randomised to placebo, baricitinib 2 mg (QD) and baricitinib 4 mg (QD), respectively. Of these, 32/176 (18.2%), 17/174 (9.8%) and 19/177 (10.7%) patients discontinued treatment on or before Week 24 in the placebo, baricitinib 2 mg QD and baricitinib 4 mg QD arms, respectively. The number of patients entering the long-term extension study was similar across all arms, with 139 (79.0%), 151 (86.8%) and 156 (88.1%) for the placebo, baricitinib 2 mg QD and baricitinib 4 mg QD arms, respectively. Full details of participant flow are presented in the CONSORT diagram in Figure 11.[12]

4.5.2 Baseline Characteristics

In the three Phase III trials presented here, baseline characteristics of the patient populations were generalisable to the UK RA population, as reported in Mercer et al (2016).[80] Furthermore, UK study centres and patients were included in all four Phase III trials.

RA-BEAM

The randomisation process resulted in well-balanced treatment arms with similar values reported across the recorded baseline characteristics, as summarised in Table 19.[9] As noted in Table 16, [redacted] patients in RA-BEAM were located in the UK, across seven study sites. At baseline, nearly 100% of patients had prior cDMARD therapy, with 22% of patients having received ≥3 prior cDMARDs, representing a treatment-refractory patient population. The trial population is similar to the UK RA population as reported in Mercer et al (2016).[80] For the full baseline characteristics of RA-BEAM participants, see Appendix 8.[9]

Table 19. Baseline characteristics of participants of RA-BEAM[9]

RA-BEAM[9,81]		Placebo (n=488)	Baricitinib 4 mg QD (n=487)	Adalimumab 40 mg Q2W (n=330)
Gender, n (%)[a]	Male	106 (21.7)	112 (23.0)	79 (23.9)
	Female	382 (78.3)	375 (77.0)	251 (76.1)
Age (years) [a]	Mean	53.4	53.5	52.9
	SD	11.8	12.2	12.3
	Median	54.5	55.0	54.5
	Range	19–83	23–80	20–86
Time from diagnosis of rheumatoid arthritis (years) [b]	Mean	8.9	8.7	8.3
	SD	8.0	8.6	7.9
	Median	6.6	6.2	6.0
	Range	0.05–39.91	0.03–56.42	0.25–34.50
DAS-28(CRP)	Mean	5.69	5.76	5.76
	SD	0.95	0.92	0.94
	Median	5.61	5.75	5.75
	Range	2.91–8.43	3.06–8.04	3.48–7.97
DAS-28(ESR)	Mean	6.40	6.45	6.43
	SD	1.01	0.94	0.96
	Median	6.35	6.45	6.41
	Range	3.29–9.07	3.51–8.81	3.84–8.99
HAQ-DI	Mean	1.55	1.57	1.59
	SD	0.67	0.68	0.70
	Median	1.50	1.56	1.63
	Range	0.0–3.0	0.0–3.0	0.0–3.0
ACPA-positive, n (%)		424 (86.9)	427 (87.7)	295 (89.4)

mTSS	Mean	45.05	42.46	44.36
	SD	50.24	50.11	50.93
	Median	23.25	21.50	25.50
	Range	0.0–300.5	0.0–284.5	0.5–309.5
RF-positive, n (%)		451 (92.4)	439 (90.1)	301 (91.2)

Abbreviations: QD = once daily, Q2W = twice weekly, DAS28 = Disease Activity Score, CRP = C-reactive protein, ESR = erythrocyte sedimentation rate, HAQ-DI = Health Assessment Questionnaire–Disability Index, ACPA = Anti-citrullinated protein antibody, mTSS = modified Total Sharp Score, RF = Rheumatoid factor.
Sources: a = JADV.11.1, b = JADV.11.2.

RA-BUILD

The randomisation process resulted in well-balanced treatment arms with similar values reported across the recorded baseline characteristics, as summarised in Table 20.[8] As noted in Table 16, [redacted] patients in RA-BUILD were located in the UK, across four study sites. The majority of patients (99%) had prior cDMARD therapy, with 25% of patients having received ≥3 cDMARDs, representing a cDMARD treatment-refractory patient population. The proportion of females is slightly higher than the UK cDMARD-treated population as reported in Mercer et al (2016) (81.9% versus 74%), but in all other characteristics the trial population is similar to the UK RA population.[80] For the full baseline characteristics of RA-BUILD participants, see Appendix 8.

Table 20. Baseline characteristics of participants of RA-BUILD[8,11]

RA-BUILD		Placebo QD (n=228)	Baricitinib 2 mg QD (n=229)	Baricitinib 4 mg QD (n=227)
Gender, n (%)	Male	39 (17.1)	45 (19.7)	40 (17.6)
	Female	189 (82.9)	184 (80.3)	187 (82.4)
Age (years)	Mean	51.4	52.2	51.8
	SD	12.5	12.3	12.1
	Median	53.0	52.0	53.0
	Range	21–79	22–82	20–80
Time from diagnosis of rheumatoid arthritis (years)	Mean	5.9	6.5	6.4
	SD	6.8	7.6	7.5
	Median	3.4	3.6	3.7
	Range	0.07–37.44	0.28–52.76	0.11–41.40
DAS-28(CRP)	Mean	5.53	5.57	5.55
	SD	0.91	0.96	0.87
	Median	5.50	5.49	5.53
	Range	2.27–7.50	3.05–8.03	3.30–7.91
DAS-28(ESR)	Mean	6.19	6.28	6.20
	SD	1.00	0.99	0.91
	Median	6.18	6.25	6.26
	Range	2.90–8.63	3.31–8.52	3.96–8.44
HAQ-DI	Mean	1.5	1.51	1.55

	SD	0.60	0.62	1.60
	Median	1.50	1.50	1.50
	Range	0.0–2.8	0.0–2.9	0.0–3.0
ACPA-positive, n (%)		172 (75.4)	169 (73.8)	163 (71.8)
mTSS	Mean	18.54	25.78	23.71
	SD	31.47	40.26	40.01
	Median	6.00	8.50	6.25
	Range	0.0–241.5	0.0–218.0	0.0–231.0
RF-positive, n (%)		171 (75.0)	177 (77.3)	173 (76.2)

Abbreviations: QD = once daily, DAS28 = Disease Activity Score, CRP = C-reactive protein, ESR = erythrocyte sedimentation rate, HAQ-DI = Health Assessment Questionnaire–Disability Index, ACPA = Anti-citrullinated protein antibody, mTSS = modified Total Sharp Score, RF = Rheumatoid factor.

RA-BEACON

The randomisation process resulted in well-balanced treatment arms with similar values reported across the recorded baseline characteristics, as summarised in Table 21.[9] As noted in Table 16, [redacted] patients in RA-BEACON were located in the UK, across three study sites. Almost 100% of patients had received prior cDMARD and/or bDMARD therapy at baseline and many were refractory to treatment with multiple conventional and biologic DMARDs. The proportion of females is slightly higher than the UK TNFi-treated population as reported in Mercer et al (2016) (81.8% versus 76%), and the trial population is slightly older than the other trials, reflecting the severe, treatment-refractory nature of the disease.[80] In all other characteristics, the trial population is similar to the UK TNFi-treated population.[80] For the full baseline characteristics of RA-BEACON participants, see Appendix 8.

Table 21. Baseline characteristics of participants of RA-BEACON[10,12]

RA-BEACON		Placebo (n=176)	Baricitinib 2 mg QD (n=174)	Baricitinib 4 mg QD (n=177)
Gender, n (%)	Male	31 (17.6)	37 (21.3)	28 (15.8)
	Female	145 (82.4)	137 (78.7)	149 (84.2)
Age (years)	Mean	56.0	55.1	55.9
	SD	10.7	11.1	11.3
	Median	57.0	55.0	58.0
	Range	24–77	21–82	24–82
Time from diagnosis of rheumatoid arthritis (years)	Mean	12.8	12.3	12.5
	SD	9.4	7.5	8.7
	Median	10.4	11.1	9.8
	Range	0.62–50.70	1.03–38.04	0.64–37.53
DAS-28(CRP)	Mean	5.89	6.03	5.87
	SD	0.94	0.89	1.00
	Median	5.80	5.99	5.83
	Range	3.64-8.24	3.94-8.07	3.31-8.06

DAS-28(ESR)	Mean	6.59	6.70	6.58
	SD	0.93	0.98	1.06
	Median	6.55	6.74	6.67
	Range	4.58-8.82	4.19-8.74	3.81-8.86
HAQ-DI	Mean	1.78	1.71	1.74
	SD	0.57	0.55	0.59
	Median	1.88	1.75	1.75
	Range	0.48-3.0	0.0-3.0	0.0-3.0
ACPA-positive, n (%)		125 (71.4)	124 (71.3)	119 (67.2)
mTSS	Mean	NR	NR	NR
	SD	NR	NR	NR
	Median	NR	NR	NR
	Range	NR	NR	NR
RF-positive, n (%)		130 (73.9)	128 (73.6)	128 (72.3)

Abbreviations: OD = once daily, DAS28 = Disease Activity Score, CRP = C-reactive protein, ESR = erythrocyte sedimentation rate, HAQ-DI = Health Assessment Questionnaire–Disability Index, ACPA = Anti-citrullinated protein antibody, mTSS = modified Total Sharp Score, RF = Rheumatoid factor.

4.6 *Quality assessment of the relevant randomised controlled trials*

A quality assessment of the three baricitinib phase III trials was carried out, and a summary is presented in Table 22. All baricitinib studies were of high quality and full quality assessments are provided in Appendix 9.

Table 22. Quality assessment results for RA-BEAM, RA-BUILD and RA- BEACON

Trial number (acronym)	RA-BEAM	RA-BUILD	RA-BEACON
Was randomisation carried out appropriately?	Yes	Yes	Yes
Was the concealment of treatment allocation adequate?	Yes	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes	Yes
Were there any unexpected imbalances in drop-outs between groups?	No	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	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
Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination)			

4.7 *Clinical effectiveness results of the relevant randomised controlled trials*

4.7.1 RA-BEAM

Summary of the RA-BEAM Clinical Effectiveness Results

- In patients receiving background MTX, baricitinib 4 mg (QD) was superior to adalimumab based on ACR20 response rate and DAS-28 CRP change from baseline at Week 12 and this difference was statistically significant. Efficacy was also durable with beneficial treatment effects sustained through to 52 weeks
- Treatment with baricitinib 4 mg (QD) was superior to placebo based on ACR20 response rate (primary endpoint) among MTX-IR patients at Week 12 and this difference was statistically significant
- Baricitinib showed statisticially significant improvements to adalimumab and placebo in terms of ACR50/70 response rates at Week 12
- Baricitinib 4 mg (QD) was superior to placebo in terms of remission criteria (DAS28(hsCRP), DAS(ESR), SDAI and CDAI) at Week 12 and Week 24 and this difference was statistically significant
- Remission rates were similar between baricitinib and adalimumab across the study
- The beneficial treatment effect of baricitinib 4 mg (QD) compared with placebo was rapid, with improvements observed as early as Week 1 or 2 for ACR20/50/70 and Week 4 for SDAI remission
- Patients treated with baricitinib 4 mg (QD) had significantly less progression of structural joint damage compared with placebo at Week 24; the progression of structural joint damage was similar to the rates observed with adalimumab at Weeks 24 and 52
- Baricitinib 4 mg (QD) demonstrated statistically significant improvments compared to adalimumab in physical function (measured by HAQ-DI) and other PROs including worst tiredness, and morning joint stifness severity and duration at Week 12. Improvements were seen as early as Week 4 for severity of morning joint stiffness and Week 8 for worst tiredness
- Improvements in physical function following treatment were also clinically meaningful as demonstrated by baricitinib 4 mg (QD) with the proportion of patients who met or exceeded the minimum clinically important difference (MCID) in HAQ-DI for the 2 cut-off values of ≥0.22 and ≥0.3 from baseline at Week 24 compared with adalimumab; these differences were statistically significant
- Baricitinib 4 mg (QD) showed statistically significant improvments to adalimumab at each visit up to Week 52 in the pain component of the ACR response criteria

The clinical effectiveness results presented in this section include the primary outcome, i.e. the superiority of baricitinib vs placebo, as assessed by the proportion of patients who achieved an ACR20 response at Week 12, as well as non-inferiority and superiority analyses for baricitinib (4 mg QD) compared to adalimumab using ACR20 response rates at Week 12. The following secondary outcomes are also presented: ACR50/70 response, EULAR response, DAS28(hsCRP) and DAS28(ESR), HAQ-DI, mTSS, SDAI,CDAI and FACIT-F. A summary of the outcomes is presented in Table 23, and any secondary outcomes not presented here are provided in Appendix 10. On a number of analyses presented, baricitinib (4 mg QD) demonstrated clinically meaningful improvements when compared to adalimumab, and consistently across all analyses compared to placebo across a range of relevant outcomes in methotrexate-IR patients.

Table 23. Summary of clinical efficacy results for RA-BEAM

Outcome measure	12 weeks			24 weeks			52 weeks		
	PBO (N=488)	BAR4 (N=487)	ADA (N=330)	PBO (N=488)	BAR4 (N=487)	ADA (N=330)	PBO (n=452)	BAR4 (N=487)	ADA (N=330)
ACR20 (%)[a,b,c]	40.2	69.6***+	61.2***	36.7	73.9***++	66.4***	N/A	71.3++	61.5
ACR50 (%)[d]	16.8	45.0***++	34.8***	19.3	50.5***	45.5***	N/A	55.9+	47.0
ACR70 (%)[e]	4.7	18.9***+	12.7***	8.0	29.8***+	21.8***	N/A	37.2	30.6
EULAR (good + moderate) response rate (%)[f]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EULAR (good) response rate (%)[f]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DAS28-hsCRP (≤3.2) response rate (%)[g]	13.7	43.9***+	35.2***	19.1	52.4***	47.9***	N/A	55.6+	48.2
DAS28-hsCRP (<2.6) response rate (%)[g]	4.3	24.4***	19.1***	7.8	34.5***	31.8***	N/A	39.6	39.1
HAQ-DI CFB LSM (SE)[h]	-0.34 (0.026)	-0.66***+ (0.026)	-0.56*** (0.030)	-0.35 (0.028)	-0.75***++ (0.028)	-0.63*** (0.033)	N/A	-0.77***++ (0.031)	-0.66 (0.036)
ΔmTSS CFB LSM (SE)[i]	N/A	N/A	N/A	0.90 (0.10)	0.41*** (0.10)	0.33*** (0.11)	1.80 (0.19)	0.71*** (0.18)	0.60*** (0.22)
SDAI LDA (≤11.0) response rate (%)[j]	15.8	42.1***+	34.8***	19.7	50.9***	48.5***	N/A	57.1+	49.4
SDAI remission	1.8	8.4***	7.3***	3.1	16.0***	13.6***	N/A	22.6	17.9

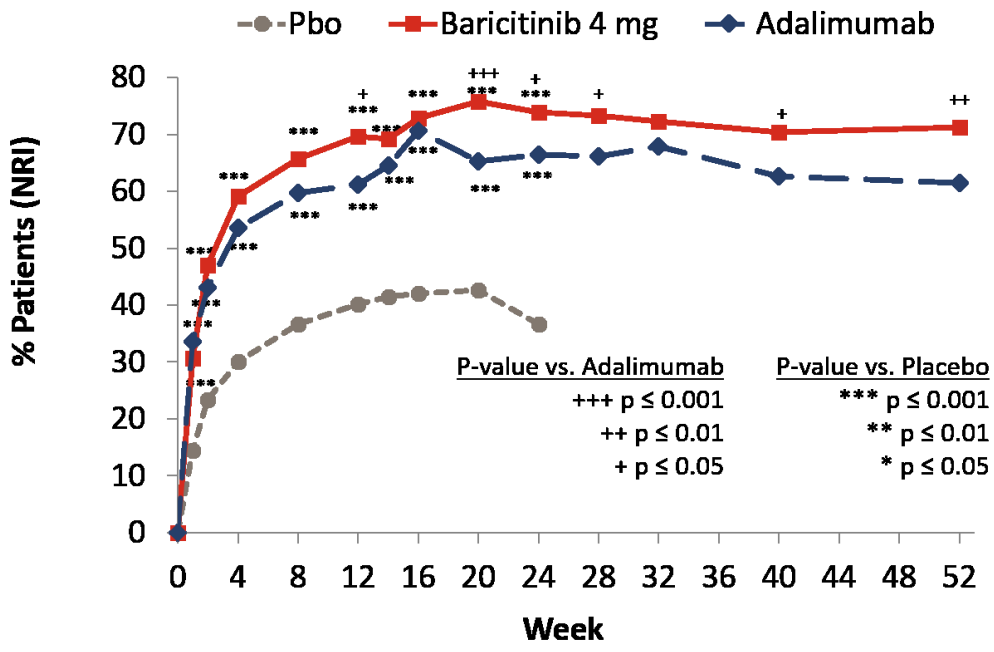
Outcome measure	12 weeks			24 weeks			52 weeks		
	PBO (N=488)	BAR4 (N=487)	ADA (N=330)	PBO (N=488)	BAR4 (N=487)	ADA (N=330)	PBO (n=452)	BAR4 (N=487)	ADA (N=330)
(≤3.3) response rate (%)[j]									
CDAI LDA (≤10.0) response rate (%)[k]	17.0	40.2***+	32.7***	19.7	49.9***	47.6***	N/A	56.9+	49.4
CDAI remission (≤2.8) response rate (%)[k]	2.3	8.4***	6.7**	3.9	16.0***	11.8***	N/A	21.6	17.6
FACIT-F (MCID) improvement ≥ 3.56 (%)[m]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
FACIT-F CFB LSM (SE)[m]	6.7 (0.42)	9.1*** (0.42)	8.7*** (0.49)	6.5 (0.46)	10.0*** (0.45)	9.3*** (0.54)	N/A	10.7+ (0.46)	9.3 (0.54)
MJS Duration (min)[n]	60.0	27.1***+	36.6***	N/A	N/A	N/A	N/A	N/A	N/A
MJS Severity LSM (SE)[o]	4.1 (0.10)	3.0***++ (0.10)	3.5*** (0.12)	N/A	N/A	N/A	N/A	N/A	N/A
EQ-5D-5L CFB LSM (SE)[q]	0.102 (0.009)	0.184*** (0.009)	0.167*** (0.011)	0.088 (0.010)	0.199*** (0.010)	0.175*** (0.012)	N/A	0.217+ (0.010)	0.182 (0.012)

Footnotes: Significance level definitions: * p ≤ 0.05; ** p ≤ 0.01; *** p ≤ 0.001 vs placebo; + p ≤ 0.05; ++ p ≤ 0.01; +++ p ≤0.001 vs adalimumab.
Abbreviations: PBO = placebo, ADA = adalimumab, BAR4 = baricitinib 4 mg QD, ACR = American College of Rheumatology, ACR20/50/70 = 20/50/70% improvement in ACR core set outcomes, DAS28 = Disease Activity Score, hsCRP = high-sensitivity C-reactive protein, ESR = erythrocyte sedimentation rate, SDAI = Simplified Disease Activity Index, LDA = Low Disease Activity, CDAI = Clinical Disease Activity Index, mTSS = modified Total Sharp Score, FACIT-F = Functional Assessment of Chronic Illness Therapy-Fatigue, LSM = least squares mean, SE = standard error, EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels, MJS = Morning Joint Stiffness, CFB = change from baseline.

4.7.1.1 Primary outcome

The primary outcome presented for RA-BEAM is the proportion (or percentage) of patients who achieved an ACR20 response rate at Week 12 in the baricitinib (4 mg QD) arm compared to the placebo arm. RA-BEAM met the primary endpoint, with a statistically significantly greater proportion of patients achieving an ACR20 response at Week 12 in the baricitinib (4 mg QD) arm (339/487, 69.6%) compared to the placebo arm (196/488, 40.2%, p=0.001), as shown in Figure 12 and Table 24. This statistically significant improvement was observed as early as Week 1 and was maintained until Week 24 for the baricitinib (4 mg QD) arm. Baricitinib (4 mg QD) is therefore superior to placebo with regards to ACR20 response at Week 12.[9]

Figure 12. ACR20 response rate over 52 weeks



Footnotes: Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo, +p≤0.05, ++p≤0.01, +++p≤0.001 vs adalimumab.
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, N = number of mITT patients, NRI = non-responder imputation, Pbo = placebo.

Table 24. ACR20 response rate at Week 12 using NRI*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
N	488	487	330	N/A	N/A
Response rate, n (%)	196 (40.2)	339 (69.6)	202 (61.2)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	29.4	8.4
95% CI	N/A	N/A	N/A	(23.5, 35.4)	(1.7, 15.1)
Odds ratio	N/A	N/A	N/A	3.6	1.5
95% CI	N/A	N/A	N/A	(2.7, 4.7)	(1.1, 2.0)
P-value	N/A	N/A	N/A	0.001	0.014

Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, NRI = non-responder imputation, PBO = placebo, BAR4 = baricitinib 4 mg QD, ADA = adalimumab, CI = confidence interval.
Source: *JADV CSR Table JADV.11.62.

4.7.1.2 Secondary outcomes

The secondary outcomes presented in this submission are: ACR20 (non-inferiority testing: baricitinib versus adalimumab, and superiorty testing when appropriate) ACR50/70 response, EULAR response, DAS28(hsCRP) and DAS28(ESR), HAQ-DI, mTSS, SDAI, CDAI, FACIT-F, duration/severity of MJS, WPAI-RA and EQ-5D-5L. Other secondary outcomes are presented in Appendix 10.

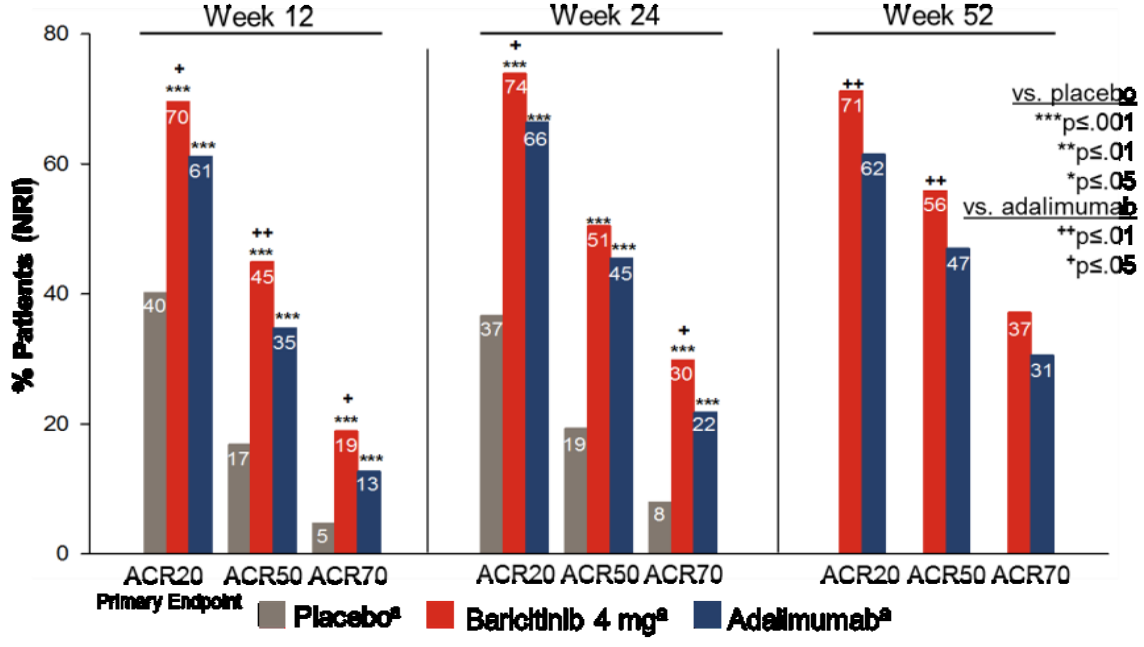
Non-inferiority and superiority of baricitinib vs adalimumab

Non-inferiority of baricitinib to adalimumab in the proportion of patients who achieved an ACR20 response rate at Week 12 was tested using a pre-specified lower bound of the 95% CI of −12%. The difference in the response rate was 8.4%, and the lower bound of the 95% CI was greater than −12% at 1.7%. Therefore, it was concluded that baricitinib was non-inferior to adalimumab with respect to ACR20 response rate at Week 12. Additionally, superiority testing was also pre-specified in the Statistical Analysis Plan (SAP). As such, due to the lower bound of the 95% CI for the response rate difference between baricitinib (4 mg QD) and adalimumab being >0%, it was also concluded that baricitinib (4 mg QD) was superior to adalimumab based on ACR20 response at Week 12 (Figure 12).

ACR50 and ACR70

In addition to results achieved for ACR20, baricitinib (4 mg QD) consistently demonstrated efficacy across the increasingly stringent ACR50 and ACR70 outcomes at Week 12 (Table 25). A significantly greater proportion of patients achieved ACR50 response (p=0.001) and ACR70 response (p=0.001) for baricitinib (4 mg QD) compared to placebo. Additionally, ACR50 and ACR70 response rates were also compared between the baricitinib (4 mg QD) and adalimumab arms. At Week 12, a significantly greater proportion of patients achieved both ACR50 (p=0.005) and ACR70 (p=0.020) responses in baricitinib (4 mg QD) compared to adalimumab. Significant responses were also maintained for baricitinib compared to placebo at Weeks 24 and 52, and for baricitinib compared to adalimumab at Week 24 (ACR50 and ACR70) and Week 52 (ACR50) (Figure 13).[9]

Figure 13. ACR20/50/70 responses at Weeks 12, 24 and 52



Footnotes: [a]All patients on background MTX. Primary endpoint=ACR20 for baricitinib 4 mg QD vs placebo at Week 12. Patients who were rescued or permanently discontinued were imputed thereafter as non-responders. Non-rescued placebo patients were switched to BAR4 mg QD at Week 24. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo, +p≤0.05, ++p≤0.01, +++p≤0.001 vs adalimumab.
Abbreviations: ACR50 = 50% improvement in American College of Rheumatology criteria, NRI = Non-responder imputation.

Table 25. ACR50/ACR70 response rate at Week 12 using NRI*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs

					ADA
ACR50					
N	488	487	330	N/A	N/A
Response rate, n (%)	82 (16.8)	219 (45.0)	115 (34.8)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	28.2	10.1
95% CI	N/A	N/A	N/A	(22.6, 33.7)	(3.3, 16.9)
Odds ratio	N/A	N/A	N/A	4.2	1.5
95% CI	N/A	N/A	N/A	(3.1, 5.7)	(1.1, 2.1)
P-value	N/A	N/A	N/A	0.001	0.005
ACR70					
N	460	473	313	N/A	N/A
Response rate, n (%)	23 (5.0)	92 (19.5)	42 (13.4)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	14.5	6.0
95% CI	N/A	N/A	N/A	(10.2, 18.1)	(0.8, 11.2)
Odds ratio	N/A	N/A	N/A	4.8	1.6
95% CI	N/A	N/A	N/A	(3.0, 7.8)	(1.1, 2.4)
P-value	N/A	N/A	N/A	0.001	0.026

Abbreviations: ACR50/70 = 50%/70% improvement in American College of Rheumatology criteria, NRI = non-responder imputation, PBO = placebo, BAR4 = baricitinib 4 mg (QD), ADA = adalimumab, CI = confidence interval.
Source: *JADV CSR Table JADV.11.63, JADV CSR Table JADV.11.64.

EULAR and DAS28 responses

As shown in Table 26, the proportion of patients achieving a EULAR good + moderate response was significantly greater with baricitinib 4 mg QD, compared to placebo, at Weeks 12 and 24 (both p=0.001). Additionally, the proportion of patients achieving a EULAR good response was significantly greater in baricitinib 4 mg QD, compared to placebo, at Weeks 12 and 24 (both p=0.001).

These findings were also observed for baricitinib 4 mg QD compared to adalimumab, with significantly greater proportions of patients achieving a EULAR good + moderate response at Weeks 12, 24, and 52 (p=0.002, p=0.011 and p=0.020, respectively), as shown in Table 26. Furthermore, when considering EULAR good response, there was a significantly greater response with baricitinib 4 mg QD compared to adalimumab at a number of weeks (8, 12, 14, 16, 20, 24, 28 32, and 52, all p-values ≤0.035). A summary of these results is presented in Table 26.[9]

Table 26. EULAR responses at Week 12 using NRI*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
Good and Moderate Response					
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Good Response					
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

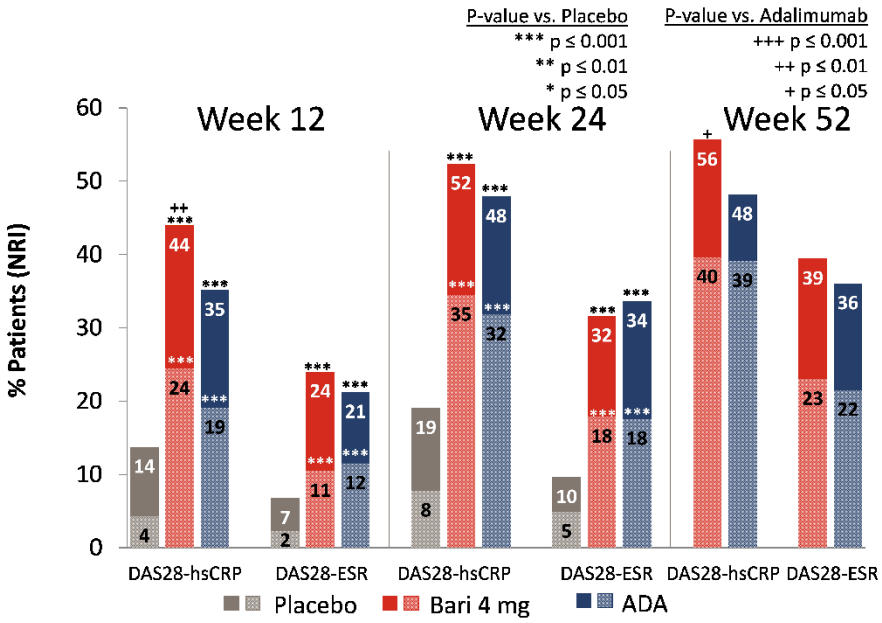
Abbreviations: EULAR = EULAR response criteria score, NRI = non-responder imputation, PBO = placebo, BAR4 = baricitinib 4 mg (QD), ADA = adalimumab, CI = confidence interval.
Source: *JADV CSR Table JADV.14.28.

Similarly, as shown in Figure 14, compared to placebo, a significantly greater proportion of patients treated with baricitinib 4 mg (QD) achieved LDA and remission at Week 12, as defined by DAS28(hsCRP) and DAS28(ESR) (≤3.2 [LDA] and <2.6 [remission]) compared to placebo. These significant improvements in disease activity versus placebo were maintained through to Week 24.

Additionally, a significantly greater proportion of patients in the baricitinib 4 mg (QD) group achieved LDA, as defined by DAS28(hsCRP) ≤3.2, at Week 12 compared with adalimumab. This significant improvement in disease activity compared with adalimumab was also observed at Week 52.

A summary of patients achieving DAS28(hsCRP) LDA and remission at Week 12 for baricitinib 2 mg (QD) and 4 mg (QD) is presented in Table 27.

Figure 14. DAS28(hsCRP) and DAS28(ESR) results at Weeks 12, 24 and 52



Footnotes: Total height of each bar indicates % of patients reaching DAS28 ≤3.2. Lower (shaded) portion of each bar indicates % of patients reaching DAS28 <2.6. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo, +p≤0.05, ++p≤0.01, +++p≤0.001 vs adalimumab.
Abbreviations: DAS28 = Disease Activity Score 28 joints, hsCRP = high-sensitivity C-reactive protein, ESR = erythrocyte sedimentation rate, BAR4 = baricitinib 4 mg (QD), ADA = adalimumab, NRI = non-responder imputation.

Table 27. DAS28(hsCRP) response rate at Week 12 using NRI*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
≤ 3.2*					
N	488	487	330	N/A	N/A
Response rate, n (%)	67 (13.7)	214 (43.9)	116 (35.2)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	30.2	8.8
95% CI	N/A	N/A	N/A	(24.9, 35.6)	(2.0, 15.6)
Odds ratio	N/A	N/A	N/A	5.2	1.5
95% CI	N/A	N/A	N/A	(3.79, 7.21)	(1.11, 2.01)
P-value	N/A	N/A	N/A	0.001	0.008
< 2.6*					
N	488	487	330	N/A	N/A
Response rate, n (%)	21 (4.3)	119 (24.4)	63 (19.1)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	20.1	5.3
95% CI	N/A	N/A	N/A	(15.9, 24.4)	(-0.4, 11.0)
Odds ratio	N/A	N/A	N/A	7.6	1.4
95% CI	N/A	N/A	N/A	(4.66, 12.43)	(0.97, 1.96)
P-value	N/A	N/A	N/A	0.001	0.078

Abbreviations: DAS28 = Disease Activity Score, hsCRP = high-sensitivity C-reactive protein, NRI = non-responder imputation, PBO = placebo, BAR4 = baricitinib 4 mg (QD), ADA = adalimumab, CI = confidence interval.

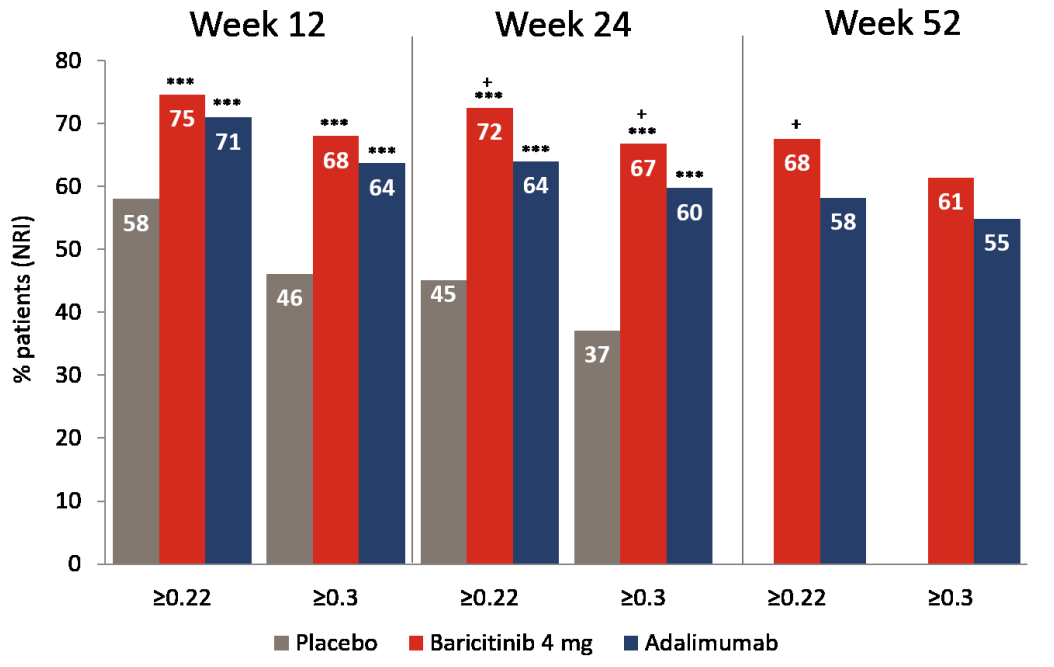
Source: *JADV CSR Table JADV.14.22.

Health Assessment Questionnaire-Disability Index

The HAQ-DI is a patient-reported measure of physical function. A statistically significantly greater change from baseline in HAQ-DI score was observed at Week 12 for baricitinib (4 mg QD) compared to placebo (p=0.001, Table 28). Significance was observed as early as Week 1 and was maintained through Week 24 for baricitinib (4 mg QD). Compared to adalimumab, a statistically significant improvement in HAQ-DI score was observed at Weeks 12, 24, and 52 for the baricitinib (4 mg QD) arm (p=0.005, p=0.003 and p=0.007, respectively), observed as early as Week 4 and maintained through Week 52.

Improvements in physical function following treatment with baricitinib were also shown to be clinically meaningful. A statistically significantly greater proportion of patients treated with baricitinib compared to placebo met or exceeded the minimum clinically important difference (MCID) in HAQ-DI for the two cut-off values of ≥0.22 and ≥0.3 (Figure 15), and this effect remained statistically significant until Week 24. Similarly, a significantly greater proportion of patients met or exceeded the MCID at Week 24 compared with adalimumab, as shown in Figure 15.[9]

Figure 15. HAQ-DI percentage of patients who met or exceeded the MCID at Weeks 12, 24 and 52



Footnotes: Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo, +p≤0.05, ++p≤0.01, +++p≤0.001 vs adalimumab.
Abbreviations: HAQ-DI = Health Assessment Questionnaire–Disability Index, NRI = non-responder imputation.

Table 28. HAQ-DI change from baseline at Week 12 using NRI*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR vs ADA
N	484	482	327	N/A	N/A
LSM (SE)	-0.34 (0.026)	-0.66 (0.026)	-0.56 (0.030)	N/A	N/A
LSM	N/A	N/A	N/A	-0.32 (0.032)	-0.10 (0.036)

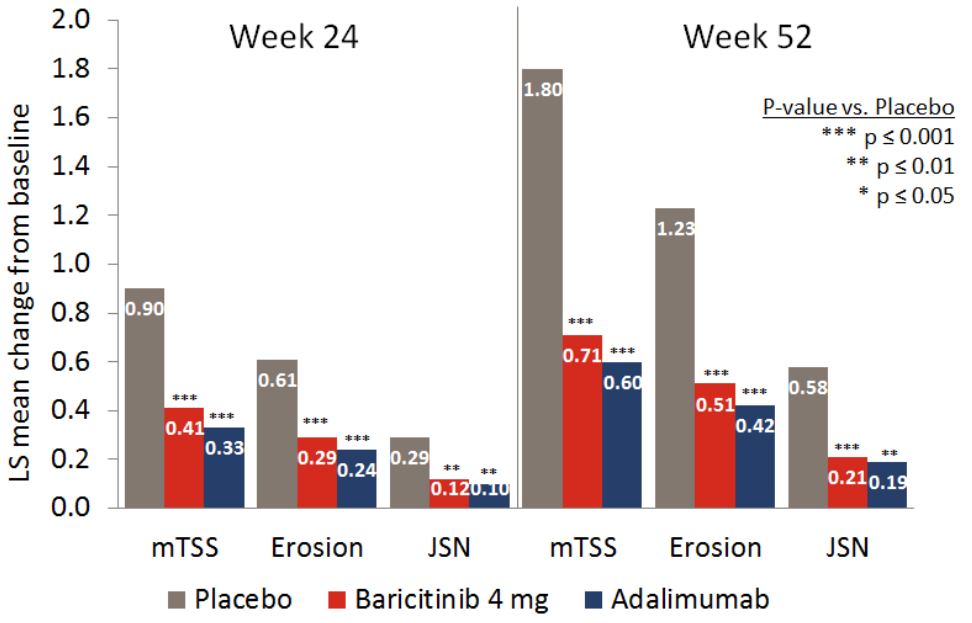
difference (SE)					
95% CI	(-0.39, -0.29)	(-0.71, -0.61)	(-0.62, -0.50)	(-0.38, -0.26)	(-0.17, -0.03)
P-value	0.001	0.001	0.001	0.001	0.005

Abbreviations: HAQ-DI = Health Assessment Questionnaire–Disability Index, NRI = non-responder imputation, PBO = placebo, BAR= baricitinib, ADA = adalimumab, LSM = least squares mean, SE = standard error, CI = confidence interval.
Source: *JADV CSR Table JADV.14.23.

Modified total Sharp score

Radiographic progression of structural joint damage was measured using the modified Total Sharp Score (mTSS). As shown in Figure 16 and Table 29, a statistically significant decrease in progression of mTSS was observed at Weeks 24 and 52, for the baricitinib (4 mg QD) arm compared to placebo (p=0.001).[9] Similar results were seen for the individual components of bone erosion score and joint space narrowing, with significant differences at Weeks 24 and 52 for baricitinib 4 mg (QD) vs placebo. Baricitinib was comparable to adalimumab in inhibiting the progression of joint damage at both Week 24 and Week 52.[9]

Figure 16. mTSS, bone erosion and joint space narrowing results at Weeks 24 and 52



Footnotes: Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo, +p≤0.05, ++p≤0.01, +++p≤0.001 vs adalimumab. Linear extrapolation used to impute scores following rescue or discontinuation
Abbreviations: mTSS = modified Total Sharp Score, Erosion = bone erosion score, JSN = joint space narrowing, LS = least squares.

Table 29. Modified Total Sharp Score change from baseline at Week 24 using LE*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
N	452	470	312	N/A	N/A
LSM (SE)	0.90 (0.10)	0.41 (0.10)	0.33 (0.11)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	-0.49 (0.12)	0.07 (0.14)

95% CI	(0.70, 1.09)	(0.22, 0.60)	(0.11, 0.56)	(-0.73, -0.25)	(-0.19, 0.34)
P-value	0.001	0.001	0.004	0.001	0.594

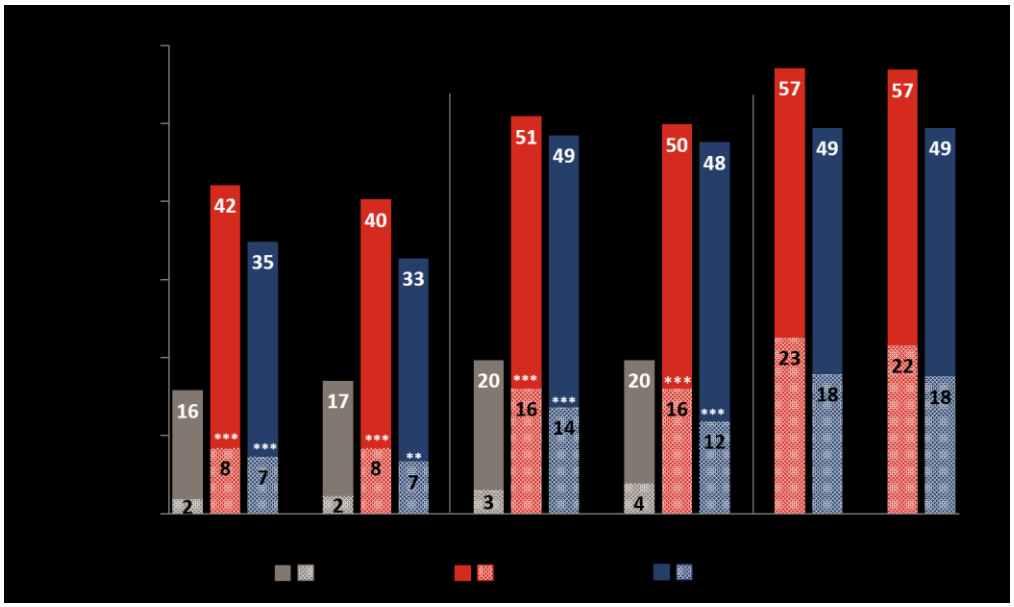
Abbreviations: mTSS = modified Total Sharp Score, LSM = least squares mean, SE = standard error, CI = confidence interval, PBO = placebo, BAR = baricitinib, ADA = adalimumab, LE = linear extrapolation.
Source: *JADV CSR Table JADV.11.10.

SDAI and CDAI

The Simplified Disease Activity Index (SDAI) is a tool for measurement of disease activity in RA that integrates measures of physical examination, acute phase reactant, patient self-assessment, and physician assessment. As shown in Figure 17 and Table 30, statistically significantly greater proportions of patients achieved SDAI scores of ≤3.3 (i.e. remission) and ≤11 (i.e. low disease activity) with baricitinib compared to placebo at Weeks 12 and 24, and an SDAI score of ≤11 with baricitinib compared to adalimumab at Weeks 12 and 52.

The Clinical Disease Activity Index (CDAI) is similar to the SDAI, but it allows for immediate scoring of patient disease activity because it does not use a laboratory result. Similar to the SDAI results, statistically significantly greater proportions of patients achieved CDAI scores of ≤2.8 (i.e. remission) and ≤10 (i.e. low disease activity) with baricitinib compared to placebo at Weeks 12 and 24, and a CDAI score of ≤11 with baricitinib compared to adalimumab at Weeks 12 and 52 (Figure 17 and Table 31).

Figure 17. SDAI and CDAI at Weeks 12, 24 and 52



Footnotes: Total height of each bar = SDAI ≤ 11, CDAI ≤ 10. Lower (shaded) portion of each bar = SDAI ≤ 3.3, CDAI ≤ 2.8. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo, +p≤0.05, ++p≤0.01, +++p≤0.001 vs adalimumab
Abbreviations: SDAI = Simplified Disease Activity Index, CDAI = Clinical Disease Activity Index, NRI = non-responder imputation, BAR4 = baricitinib 4 mg (QD), ADA = adalimumab.

Table 30. SDAI response rate at Week 12 using NRI*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
≤ 11 (LDA)					
N	488	487	330	N/A	N/A
Response rate, n (%)	77 (15.8)	205 (42.1)	115 (34.8)	N/A	N/A
Difference in	N/A	N/A	N/A	26.3	7.2

response rate					
95% CI	N/A	N/A	N/A	(20.9, 31.8)	(0.5, 14.0)
Odds ratio	N/A	N/A	N/A	4.06	1.40
95% CI	N/A	N/A	N/A	(2.98, 5.53)	(1.04, 1.88)
P-value	N/A	N/A	N/A	0.001	0.028
≤ 3.3 (Remission)					
N	488	487	330	N/A	N/A
Response rate, n (%)	9 (1.8)	41 (8.4)	24 (7.3)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	6.6	1.1
95% CI	N/A	N/A	N/A	(3.8, 9.3)	(-2.6, 4.9)
Odds ratio	N/A	N/A	N/A	N/A	N/A
95% CI	N/A	N/A	N/A	N/A	N/A
P-value	N/A	N/A	N/A	0.001	0.600

Abbreviations: SDAI = Simplified Disease Activity Index, NRI = non-responder imputation, PBO = placebo, BAR = baricitinib, ADA = adalimumab, CI = confidence interval, LDA = low disease activity
Source: *JADV CSR Table JADV.14.25.

Table 31. CDAI response rate at Week 12 using NRI*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
≤ 10 (LDA)					
N	488	487	330	N/A	N/A
Response rate, n (%)	83 (17.0)	196 (40.2)	108 (32.7)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	23.2	7.5
95% CI	N/A	N/A	N/A	(17.8, 28.7)	(0.8, 14.2)
Odds ratio	N/A	N/A	N/A	3.42	1.39
95% CI	N/A	N/A	N/A	(2.52, 4.63)	(1.03, 1.87)
P-value	N/A	N/A	N/A	0.001	0.034
≤ 2.8 (Remission)					
N	488	487	330	N/A	N/A
Response rate, n (%)	11 (2.3)	41 (8.4)	22 (6.7)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	6.2	1.8
95% CI	N/A	N/A	N/A	(3.4, 9.0)	(-1.9, 5.4)
Odds ratio	N/A	N/A	N/A	N/A	N/A
95% CI	N/A	N/A	N/A	N/A	N/A
P-value	N/A	N/A	N/A	0.001	0.424

Abbreviations: CDAI = Clinical Disease Activity Index, NRI = non-responder imputation, PBO = placebo, BAR4 = baricitinib 4 mg (QD), ADA = adalimumab, CI = confidence interval, LDA = low disease activity
Source: *JADV CSR Table JADV.11.44.

Functional Assessment of Chronic Illness Therapy–Fatigue

The FACIT–F is a brief, 13-item, symptom-specific questionnaire that specifically assesses the self-reported severity of fatigue and its impact upon daily activities and functioning. As shown in Table 32, compared to placebo, a statistically significant improvement in change from baseline FACIT–F score was observed at Week 12 (p=0.001)for the baricitinib (4 mg QD) arm. In addition, a statistically significant improvement was observed as early as Week 4 (the first post-baseline observation) and was maintained through Week 24 for the baricitinib (4 mg QD) arm compared to placebo. Additionally, the proportion of patients who met or exceeded the MCID ((≥3.56) in FACIT-F was statistically significantly greater at Weeks 12 (p≤0.05) and 24 (p≤0.001) for the baricitinib (4 mg QD) arm, observed as early as Week 4 (the first post-baseline observation) and maintained through to Week 24, compared to placebo (Figure 18 and Table 33).

Compared to adalimumab, a statistically significant improvement in FACIT-F score was observed at Weeks 20 (p=0.046), 28 (p=0.032), and 52 (p=0.033) for the baricitinib (4 mg QD) arm. Additionally, the proportion of patients who met or exceeded the MCID in FACIT-F was significantly greater at Weeks 28 (p=0.022) and 40 (p=0.041) for the baricitinib (4 mg QD) arm compared to adalimumab.[9]

Figure 18. FACIT-F percentage of patients achieving or exceeding MCID at 52 weeks

[redacted]

Footnotes: Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo, +p≤0.05, ++p≤0.01, +++p≤0.001 vs adalimumab.
Abbreviations: FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue , MCID = minimum clinically important difference.

Table 32. FACIT-F change from baseline at Week 12 using mLOCF and NRI*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
N	475	479	320	N/A	N/A
LSM (SE)	6.7 (0.42)	9.1 (0.42)	8.7 (0.49)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	2.4 (0.53)	0.4 (0.59)
95% CI	(5.9, 7.5)	(8.3, 9.9)	(7.8, 9.7)	(1.4, 3.5)	(-0.8, 1.5)
P-value	0.001	0.001	0.001	0.001	0.521

Abbreviations: FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, PBO = placebo, BAR = baricitinib, ADA = adalimumab, LSM = least squares mean, SE = standard error, CI = confidence interval, NRI = non-responder imputation, mLOCF = modified last observation carried forward.
Source: *JADV CSR Table JADV.14.52.

Table 33. Percentage of patients achieving or exceeding the MCID at Week 12 using mLOCF and NRI*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

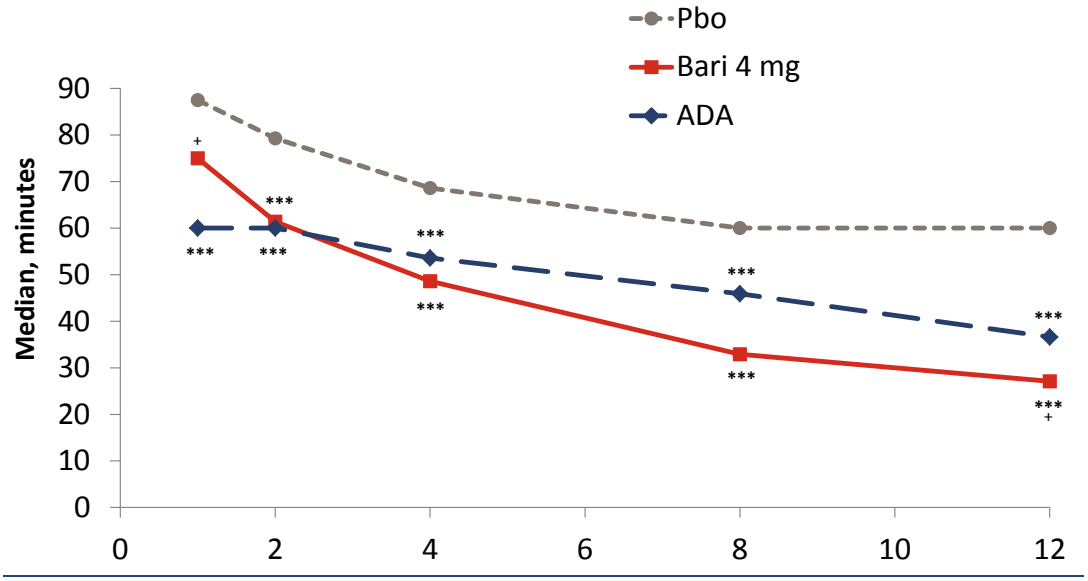
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, MCID = minimum clinically important difference, PBO = placebo, BAR = baricitinib, ADA = adalimumab, LSM = least squares mean, SE = standard error, CI = confidence interval, NRI = non-responder imputation, mLOCF = modified last observation carried forward.
Source: *JADV CSR Table JADV.14.52.

Duration and severity of morning joint stiffness (MJS)

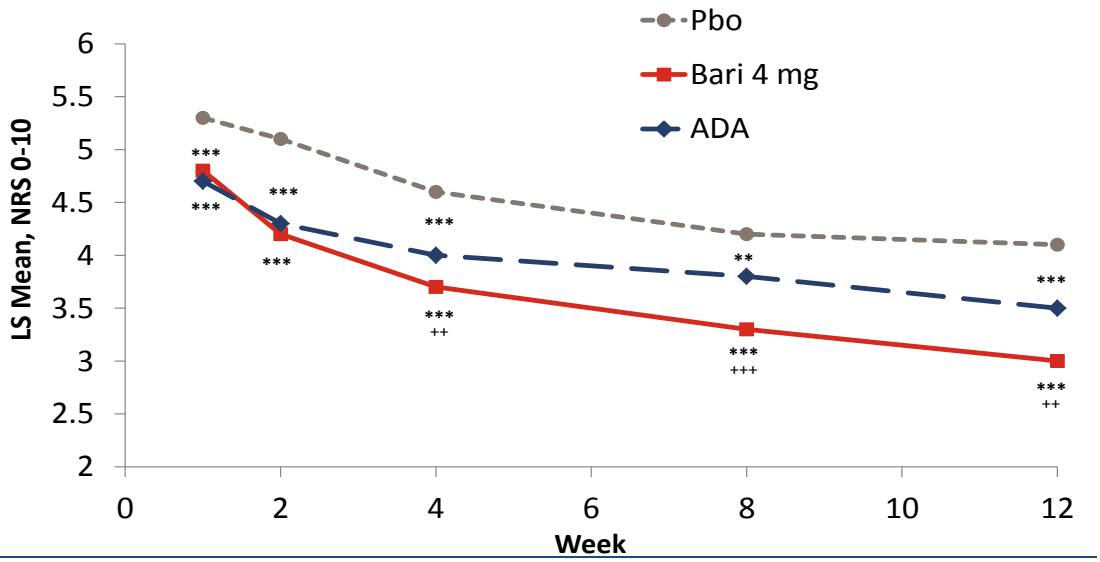
Baricitinib 4 mg (QD) showed a statistically significant improvement to placebo in the reduction of the median duration and mean severity of MJS in the 7 days prior to Week 12.[9] Barictinib also demonstrated a statistically significant diference compared to adalimumab in the reduction of the median duration and mean severity of MJS and th in the 7 days prior to Week 12 (Figure 19 and Figure 20).

Figure 19: Duration of morning stiffness in RA-BEAM



Abbreviations: ADA = adalimumab; bari = baricitinib; Pbo = placebo
***p≤0.001; + p≤0.05 vs adalimumab
Data are median durations of morning joint stiffness in minutes, based on the preceding week of daily diary entries. Daily question: “Please indicate how long your morning joint stiffness lasted today”

Figure 20: Severity of morning joint stiffness in RA-BEAM



Abbreviations: ADA = adalimumab; bari = baricitinib; Pbo = placebo
ADA = adalimumab; Bari = baricitinib; LS = least squares; NRS = numeric rating scale; Pbo = placebo
** p≤0.01; ***p≤0.001; +++ p≤0.001; ++ p≤0.01 (vs adalimumab)
Data are LS mean scores for severity of morning joint stiffness, based on the preceding week of daily diary entries. Higher values indicate greater severity. Daily question: "Please rate the overall level of morning joint stiffness you had from the time you woke up today."

Work Productivity and Activity Impairment-Rheumatoid Arthritis (WPAI-RA)

Baricitinib 4 mg (QD) was statistically better than placebo in the absenteeism, presenteeism, work productivity loss and activity impairment scores of the WPAI-RA at Week 12 (p=0.04 for absenteeism, p≤0.001 for all others). Baricitinib 4 mg (QD) also demonstrated statistically significant improvements compared to adalimumab in the work productivity loss and activity impairment scored of the WPAI-RA instrument at Week 12 (p=0.013 and p=0.003, respectively).

European Quality of Life-5 Dimensions-5 levels (EQ-5D-5L)

As shown in Table 34 baricitinib 4 mg (QD) was statistically significantly superior to placebo at Week 24 using the EQ-5D-5L UK index score.[9]

Table 34. EQ-5D-5L Index Score change from baseline at Week 24 using mLOCF*

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
N	475	479	320	N/A	N/A
LSM (SE)	0.088 (0.010)	0.199 (0.010)	0.175 (0.012)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	0.110 (0.013)	0.023 (0.014)
95% CI	(0.068, 0.108)	(0.179, 0.219)	(0.152, 0.199)	(0.085, 0.136)	(-0.005, 0.052)
P-value	0.001	0.001	0.001	0.001	0.108

Abbreviations: PBO = placebo, BAR4 = baricitinib 4 mg (QD), ADA = adalimumab, LSM = least squares mean, SE = standard error, CI = confidence interval, EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels, mLOCF = modified last observation carried forward.
Source: *JADV CSR Table JADV.14.55.

4.7.2 RA-BUILD

Summary of the RA-BUILD Clinical Effectiveness Results

- Baricitinib 4 mg and 2 mg were superior to placebo in terms of ACR20/50/70, remission and LDA as measured by DAS-hCRP, DAS-ESR, CDAI and SDAI at 12 and 24 weeks and these differences were all statistically significant
- Baricitinib 4 mg was superior to placebo in inhibiting progression of structural joint damage for both components of joint narrowing and erosion at Week 24 and this difference was statistically significant
- Overall, a more rapid and consistently larger treatment effect was seen for the 4 mg dose compared with 2 mg across different analyses including SDAI, CDAI and in components of the composite scores
- A more rapid and consistently larger treatment effect was also seen for the 4 mg dose compared with 2 mg for patient reported outcomes, including improvements in duration/severity of morning joint stiffness.
- Results from RA-BUILD for baricitinib 4 mg reinforce the efficacy results observed in RA-BEAM conducted in patients with RA and who were MTX-IR
- Baricitinib 4 mg and 2 mg were superior to placebo in improving worst tiredness in the 7 days prior to Week 12 and this difference was statistically significant. Improvements were seen as early as Week 2 with the 4 mg dose
- Baricitinib 4 mg and 2 mg were superior to placebo in improving MJS duration and severity the 7 days prior to Week 12; these improvements were seen as early as Week 1 for baricitinib 4 mg for severity and as early as Week 2 for duration (4 mg) with improvements maintained across 12 weeks
- Baricitinib 4 mg and 2 mg were superior to placebo for improvements in physical function at Weeks 12 and 24, as measured by HAQ-DI
- Improvements in physical function from baseline were also clinically meaningful as measured by the 2 cut-off values of ≥0.22 and ≥0.3 to define the MCID in HAQ-DI. The proportion of patients who met or exceeded the MCID was higher with baricitinib 4 mg and 2 mg compared with placebo at both Weeks 12 and 24 and these differences were statistically significant

The clinical effectiveness results presented in this section include the primary outcome, i.e. the proportion of patients who achieved an ACR20 response at Week 12, and the following secondary outcomes: ACR50/70 response, DAS28(hsCRP) and DAS28(ESR), EULAR response, HAQ-DI, mTSS, SDAI, CDAI and FACIT-F. A summary of the outcomes is presented in Table 35, and any secondary outcomes not presented here are provided in Appendix 10. Throughout all analyses presented, 4 mg QD baricitinib and 2 mg QD baricitinib demonstrated consistent, clinically meaningful efficacy when compared to placebo across a range of relevant outcomes in cDMARD-refractory patients.

Table 35. Summary of clinical effectiveness results for RA-BUILD

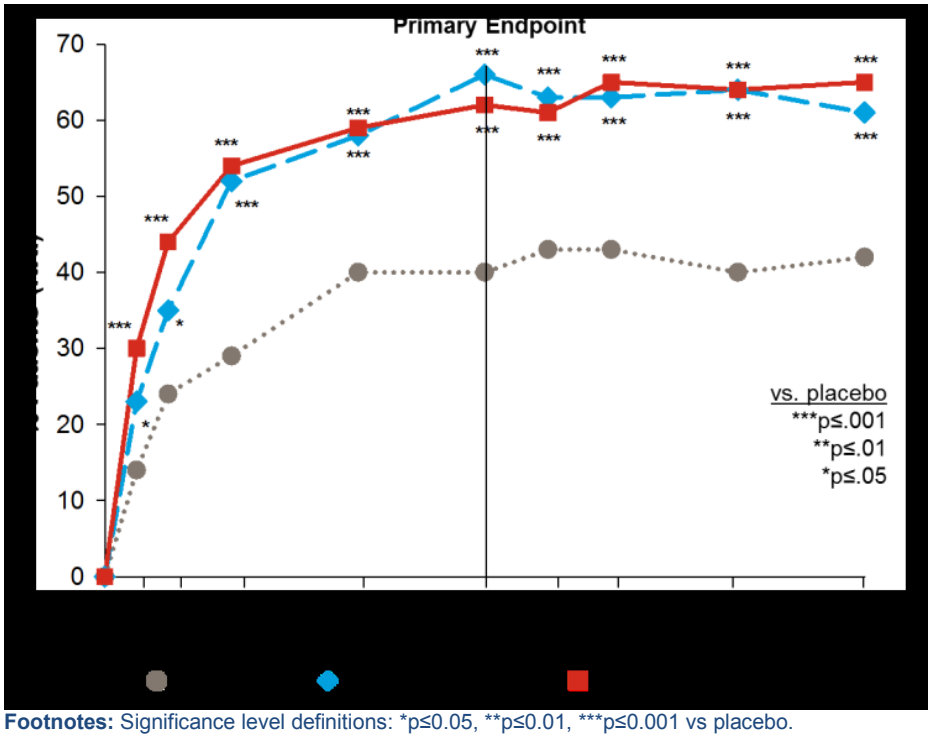
Outcome measures	12 weeks			24 weeks		
	PBO (N=228)	BAR2 (N=229)	BAR4 (N=227)	PBO (N=228)	BAR2 (N=229)	BAR4 (N=227)
ACR20 (%)[a]	39.5	65.9***	61.7***	42.1	61.1***	65.2***
ACR50 (%)[b]	12.7	33.6***	33.5***	21.5	41.5***	44.1***
ACR70 (%)[c]	3.1	17.9***	18.1***	7.9	25.3***	24.2***
EULAR (good + moderate) response rate (%)[d]	53.5	79.0***	79.3***	53.5	72.1***	78.0***
EULAR (good) response rate (%)[d]	15.4	34.1***	38.3***	21.9	45.4***	50.7***
DAS28-hsCRP (≤3.2) response rate (%)[e]	17.1	35.8***	39.2***	23.7	46.3***	51.5***
DAS28-hsCRP (<2.6) response rate (%)[e]	8.8	25.8***	25.6***	10.5	30.6***	33.0***
HAQ-DI CFB LSM (SE)[f]	-0.34 (0.037)	-0.54*** (0.036)	-0.53*** (0.037)	-0.35 (0.040)	-0.58*** (0.039)	-0.58*** (0.040)
ΔmTSS CFB LSM (SE)[g]	N/A	N/A	N/A	0.70 (0.14)	0.33* (0.14)	0.15** (0.14)
SDAI LDA (≤11.0) response rate (%)[h]	19.7	33.2**	34.8***	28.5	48.0***	52.4***
SDAI remission (≤3.3) response rate (%)[h]	0.9	9.2***	8.8***	3.9	16.6***	15.0***
CDAI LDA (≤10.0) response rate (%)[i]	20.6	34.5**	34.8***	27.6	45.4***	52.0***
CDAI remission (≤2.8) response rate (%)[i]	1.8	10.0***	9.3***	3.9	15.3***	15.4***
FACIT-F (MCID) improvement ≥ 3.56 (%)[k]	58.8	63.3	64.8	42.5	59.0***	59.9***
FACIT-F CFB LSM (SE)[k]	7.5 (0.64)	8.5 (0.61)	9.1 (0.64)	7.9 (0.67)	9.2 (0.64)	10.1* (0.67)
MJS Duration (min)[l]	60.0	44.4**	34.6***	N/A	N/A	N/A
MJS Severity LSM (SE)[m]	4.1 (0.15)	3.5** (0.15)	3.4*** (0.16)	N/A	N/A	N/A
EQ-5D-5L CFB LSM (SE)[o]	0.092 (0.014)	0.165*** (0.013)	0.162***(0.014)	0.091 (0.014)	0.157*** (0.014)	0.186*** (0.014)

Footnotes: Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs. placebo.
Abbreviations: PBO = placebo, BAR= baricitinib, ACR = American College of Rheumatology, ACR20/50/70 = 20/50/70% improvement in ACR core set outcomes, DAS28 = Disease Activity Score, hsCRP = high-sensitivity C-reactive protein, ESR = erythrocyte sedimentation rate, SDAI = Simplified Disease Activity Index, LDA = Low Disease Activity, CDAI = Clinical Disease Activity Index, mTSS = modified Total Sharp Score, FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, LSM = least squares mean, SE = standard error, EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels, MJS = Morning Joint Stiffness.

4.7.2.1 Primary outcome

The primary outcome presented for RA-BUILD is the proportion (or percentage) of patients who achieved an ACR20 response rate at Week 12 for 4 mg QD baricitinib compared to placebo. RA-BUILD met the primary endpoint, with a significantly greater proportion of patients achieving an ACR20 response at Week 12 in the 4 mg QD baricitinib arm (140/227, 61.7%) compared to placebo (90/228, 39.5%, p=0.001). Although a secondary analysis, the comparison between placebo and the 2 mg QD baricitinib treatment arm also demonstrated a statistically significant improvement in ACR20 response rate at Week 12. Compared to placebo, a statistically significant improvement was observed as early as Week 1 for both 2 mg (QD) and 4 mg (QD) baricitinib (Figure 21 and Table 36).[11]

Figure 21. ACR20 response rate over 24 weeks



Footnotes: Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, N = number of mITT patients, NRI = non-responder imputation.

Table 36. ACR20 response rate at Week 12 using NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	228	229	227	N/A	N/A
Response rate, n (%)	90 (39.5)	151 (65.9)	140 (61.7)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	26.5	22.2
95% CI	N/A	N/A	N/A	(17.6, 35.3)	(13.2, 31.2)
Odds ratio	N/A	N/A	N/A	3.0	2.5
95% CI	N/A	N/A	N/A	(2.0, 4.4)	(1.7, 3.7)
P-value	N/A	N/A	N/A	0.001	0.001

Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, NRI = non-responder imputation, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.

Source: *JADX CSR Table JADX.14.21.

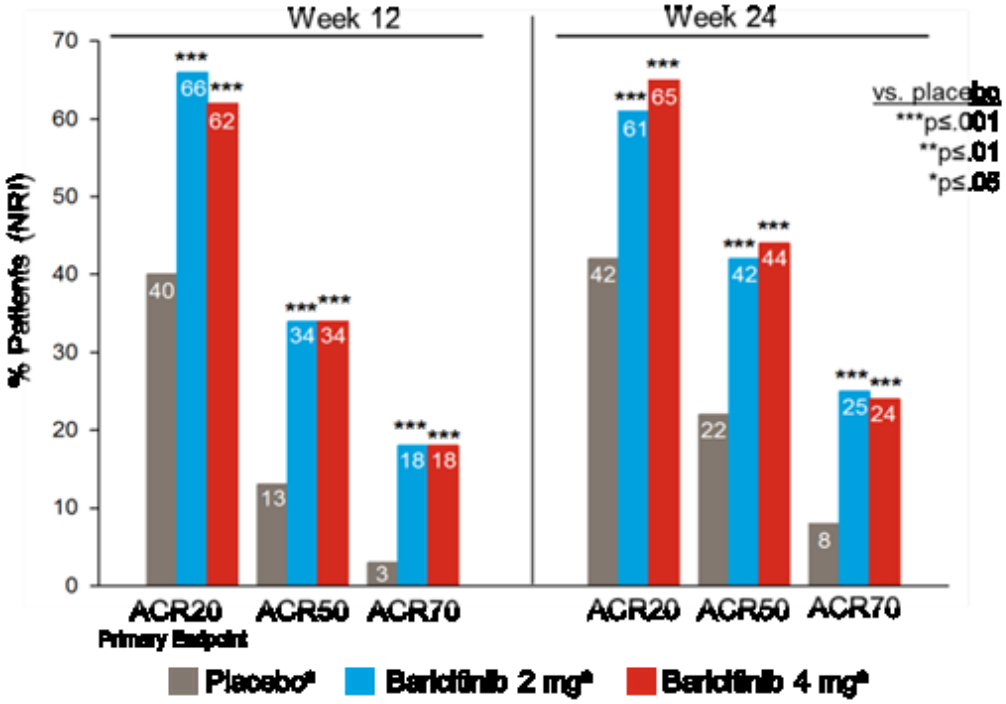
4.7.2.2 Secondary outcomes

The secondary outcomes presented in this submission are: ACR50/70 response, DAS28(hsCRP) and DAS28(ESR), EULAR response, HAQ-DI, mTSS, SDAI and CDAI and FACIT-F. Other secondary outcomes are presented in Appendix 10.

ACR50 and ACR70

In addition to results achieved for ACR20, baricitinib 2 mg QD and baricitinib 4 mg QD consistently demonstrated efficacy across the increasingly stringent ACR50 and ACR70 outcomes at Week 12. A significantly greater proportions of patients achieved an ACR50 response (p=0.001) and an ACR70 response (p=0.001) with 4 mg QD baricitinib compared to placebo, and a significantly greater proportion of patients achieved an ACR50 response (p=0.001) and an ACR70 response (p=0.001) with 2 mg QD baricitinib compared to placebo. Significant responses were maintained up to Week 24 (Figure 22, Table 37).[11]

Figure 22. ACR20/50/70 responses at Weeks 12 and 24



Footnotes: Primary endpoint=ACR20 for baricitinib 4 mg QD vs placebo at Week 12. Patients who were rescued or permanently discontinued were imputed thereafter as non-responders. [a]Patients were on background cDMARD unless documented intolerance or contraindication. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: ACR = American College of Rheumatology, NRI = Non-responder imputation.

Table 37. ACR50/70 response rate at Week 12 using NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4vs PBO
ACR50					
N	228	229	227	N/A	N/A
Response rate, n (%)	29 (12.7)	77 (33.6)	76 (33.5)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	20.9	20.8
95% CI	N/A	N/A	N/A	(13.4, 28.4)	(13.3, 28.3)
Odds ratio	N/A	N/A	N/A	3.5	3.5
95% CI	N/A	N/A	N/A	(2.2, 5.6)	(2.2, 5.7)
P-value	N/A	N/A	N/A	0.001	0.001
ACR70					
N	228	229	227	N/A	N/A
Response rate, n (%)	7 (3.1)	41 (17.9)	41 (18.1)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	14.8	15.0
95% CI	N/A	N/A	N/A	(9.4, 20.3)	(9.5, 20.5)
Odds ratio	N/A	N/A	N/A	6.9	7.2
95% CI	N/A	N/A	N/A	(3.0, 15.9)	(3.2, 16.6)
P-value	N/A	N/A	N/A	0.001	0.001

Abbreviations: ACR50/70 = 50/70% improvement in American College of Rheumatology criteria, NRI = non-responder imputation, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.
Source: *JADX CSR Table JADX.11.19, JADX CSR Table JADX.11.20.

EULAR and DAS28 responses

Compared to placebo, the proportion of patients achieving a EULAR good + moderate response was significantly greater at Weeks 12 and 24 for both the baricitinib arms (p=0.001 for all comparisons). A statistically significant improvement was observed as early as Week 1 and was maintained through Week 24 for both the baricitinib arms compared to placebo.[11]

Additionally, the proportion of patients achieving a EULAR good response was significantly greater, compared to placebo at Weeks 12 and 24, for both the baricitinib arms (p=0.001 for all comparisons). A statistically significant improvement was observed as early as Week 1 and was maintained through Week 24 for the 4 mg QD baricitinib arm compared to placebo, whilst a statistically significant improvement was observed as early as Week 2 and was maintained through Week 24 for the 2 mg QD baricitinib arm compared to placebo. A summary of results for Week 12 is presented in Table 38.[11]

Table 38. EULAR Response Rate at Week 12 using NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
Good and Moderate Response					
N	228	229	227	N/A	N/A
Response rate, n (%)	122 (53.5)	181 (79.0)	180 (79.3)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	25.5	25.8
95% CI	N/A	N/A	N/A	(17.2, 33.9)	(17.4, 34.1)
Odds ratio	N/A	N/A	N/A	3.3	3.5
95% CI	N/A	N/A	N/A	(2.2, 5.0)	(2.3, 5.4)
P-value	N/A	N/A	N/A	0.001	0.001
Good Response					
N	228	229	227	N/A	N/A
Response rate, n (%)	35 (15.4)	78 (34.1)	87 (38.3)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	18.7	23.0
95% CI	N/A	N/A	N/A	(11.0, 26.4)	(15.1, 30.8)
Odds ratio	N/A	N/A	N/A	2.9	3.6
95% CI	N/A	N/A	N/A	(1.8, 4.6)	(2.3, 5.7)
P-value	N/A	N/A	N/A	0.001	0.001

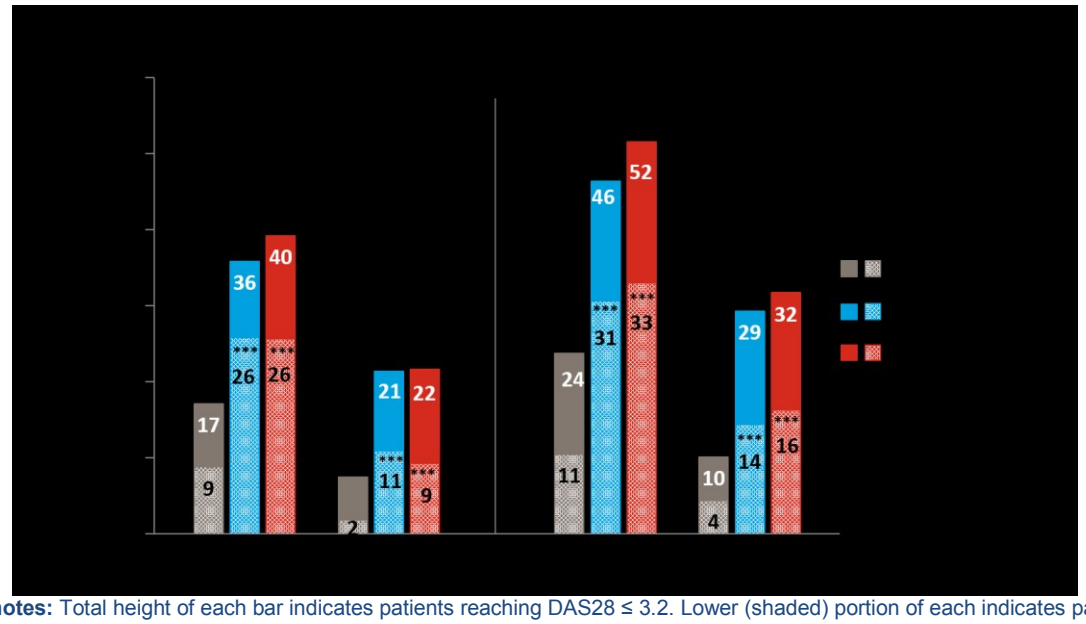
Abbreviations: EULAR = EULAR response criteria score, NRI = non-responder imputation, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.
Source: *JADX CSR Table JADX.11.39.

Similarly, as shown in Figure 23, compared to placebo, a significantly greater proportion of patients treated with baricitinib 4 mg (QD) achieved LDA and remission at Week 12, as defined by DAS28(hsCRP) and DAS28(ESR) (≤3.2 [LDA] and <2.6 [remission]) compared to placebo. These significant improvements in disease activity versus placebo were maintained through to Week 24.

Additionally, a significantly greater proportion of patients in the baricitinib 2 mg (QD) group achieved LDA and remission, as defined by DAS28(hsCRP) and DAS28(ESR) (≤3.2 and <2.6), at Week 12 compared to placebo. This significant improvement in disease activity compared with placebo was also maintained through Week 24.

A summary of patients achieving DAS28(hsCRP) LDA and remission at Week 12 for baricitinib 2 mg (QD) and 4 mg (QD) is presented in Table 39.[11]

Figure 23. DAS28(hsCRP) and DAS28(ESR) results at Weeks 12 and 24



Footnotes: Total height of each bar indicates patients reaching DAS28 ≤ 3.2. Lower (shaded) portion of each indicates patients reaching DAS28 < 2.6. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: DAS28 = Disease Activity Score 28 joints, hsCRP = high-sensitivity C-reactive protein, ESR = erythrocyte sedimentation rate, NRI = non-responder imputation, PBO = placebo.

Table 39. DAS28(hsCRP) Response Rate at Week 12 using NRI*

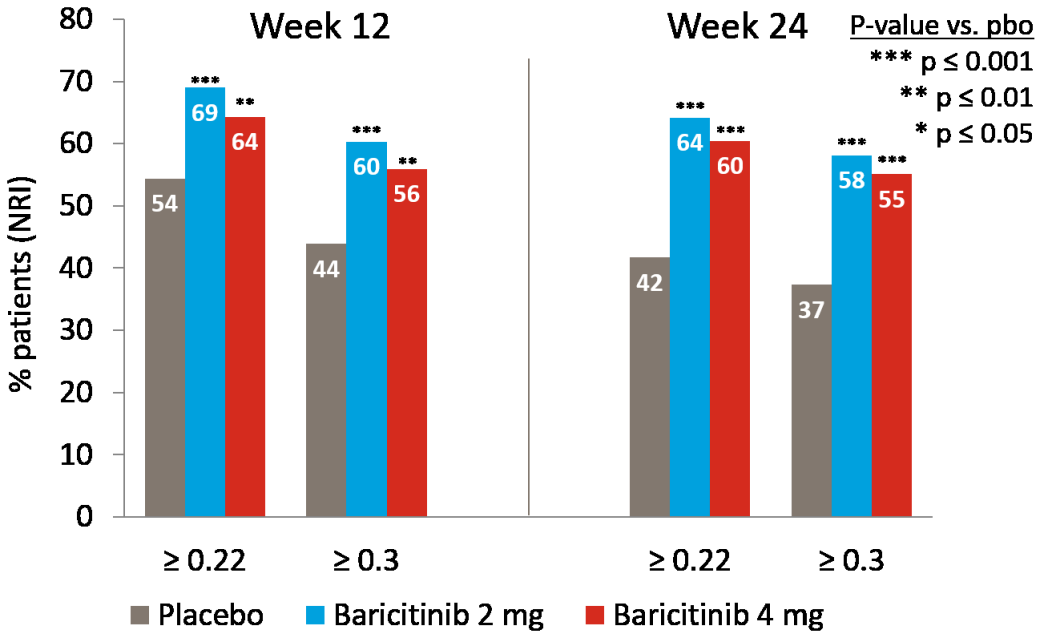
Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
≤ 3.2					
N	228	229	227	N/A	N/A
Response rate, n (%)	39 (17.1)	82 (35.8)	89 (39.2)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	18.7	22.1
95% CI	N/A	N/A	N/A	(10.8, 26.6)	(14.1, 30.1)
Odds ratio	N/A	N/A	N/A	2.76	3.29
95% CI	N/A	N/A	N/A	(1.77, 4.31)	(2.11, 5.13)
P-value	N/A	N/A	N/A	0.001	0.001
< 2.6					
N	228	229	227	N/A	N/A
Response rate, n (%)	20 (8.8)	59 (25.8)	58 (25.6)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	17.0	16.8
95% CI	N/A	N/A	N/A	(10.2, 23.7)	(10.0, 23.5)
Odds ratio	N/A	N/A	N/A	3.69	3.73
95% CI	N/A	N/A	N/A	(2.12, 6.42)	(2.14, 6.51)
P-value	N/A	N/A	N/A	0.001	0.001

Abbreviations: DAS28 = Disease Activity Score, hsCRP = high-sensitivity C-reactive protein, NRI = non-responder imputation, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.
Source: *JADX CSR Table JADX.14.15.

Health Assessment Questionnaire-Disability Index

The HAQ-DI is a patient-reported physical function measure. As shown in Table 40, compared to placebo, a statistically significantly greater change from baseline in HAQ-DI score was observed for baricitinib 4 mg and 2 mg (QD) at Week 12. This superiority was maintained up to Week 24. Moreover, baricitinib 4 mg and 2 mg (QD) were statistically superior to placebo in terms of the proportion of patients who met or exceeded the MCID in HAQ-DI from baseline for the ≥0.22 and ≥0.3 cut-off values at both Week 12 and Week 24 (Figure 24 and Table 40).[11]

Figure 24: HAQ-DI percentage of patients who met or exceeded the MCID at Weeks 12 and 24



Footnotes: Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: HAQ-DI = Health Assessment Questionnaire–Disability Index, NRI = non-responder imputation, PBO = placebo.

Table 40. HAQ-DI change from baseline at Week 12 using NRI*

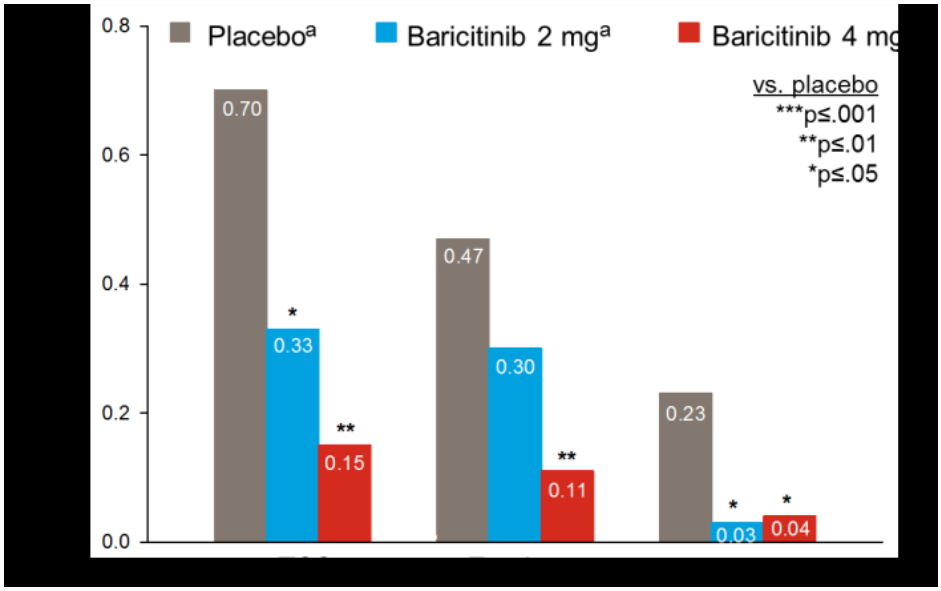
Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	221	228	222	N/A	N/A
LSM (SE)	-0.34 (0.037)	-0.54 (0.036)	-0.53 (0.037)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	-0.21 (0.049)	-0.20 (0.049)
95% CI	(-0.41, -0.26)	(-0.62, -0.47)	(-0.61, -0.46)	(-0.30, -0.11)	(-0.30, -0.10)
P-value	0.001	0.001	0.001	0.001	0.001

Abbreviations: HAQ-DI = Health Assessment Questionnaire–Disability Index, NRI = non-responder imputation, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), LSM = least squares mean, SE = standard error, CI = confidence interval.
Source: *JADX CSR Table JADX.11.10.

Modified Total Sharp Score

Radiographic progression of structural joint damage was measured using the mTSS, although this was an exploratory endpoint in RA-BUILD, and the study was not powered to show statistical meaningfulness for this measure.[11] As shown in Figure 25, using linear extrapolation for missing or post-rescue data, a statistically significant decrease in progression of mTSS was observed at Weeks 24, for the baricitinib 4 mg (QD) and 2 mg (QD) arms compared to placebo. Similar results were seen for the individual components of bone erosion score and joint space narrowing, with significant improvements in both erosion and joint space narrowing for baricitinib 4 mg (QD) and in joint space narrowing for baricitinib 2 mg (QD). Overall, the largest and most consistent treatment effect for the inhibition of radiographic progression of structural joint damage compared with placebo was observed with the 4 mg baricitinib dose, with robust results across different methods of missing data imputation. A summary of results at Week 24 is presented in Table 41.

Figure 25. mTSS bone erosion and joint space narrowing at Week 24



Footnotes: [a]Patients were on background cDMARD unless documented intolerance or contraindication. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: LS = least squares, mTSS = modified Total Sharp Score.

Table 41. Modified Total Sharp Score change from baseline at Week 24 using LE*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	190	208	198	N/A	N/A
LSM (SE)	0.70 (0.14)	0.33 (0.14)	0.15 (0.14)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	-0.38 (0.18)	-0.55 (0.19)
95% CI	(0.42, 0.98)	(0.06, 0.59)	(-0.13, 0.43)	(-0.74, -0.01)	(-0.92, -0.19)
P-value	0.001	0.017	0.300	0.043	0.004

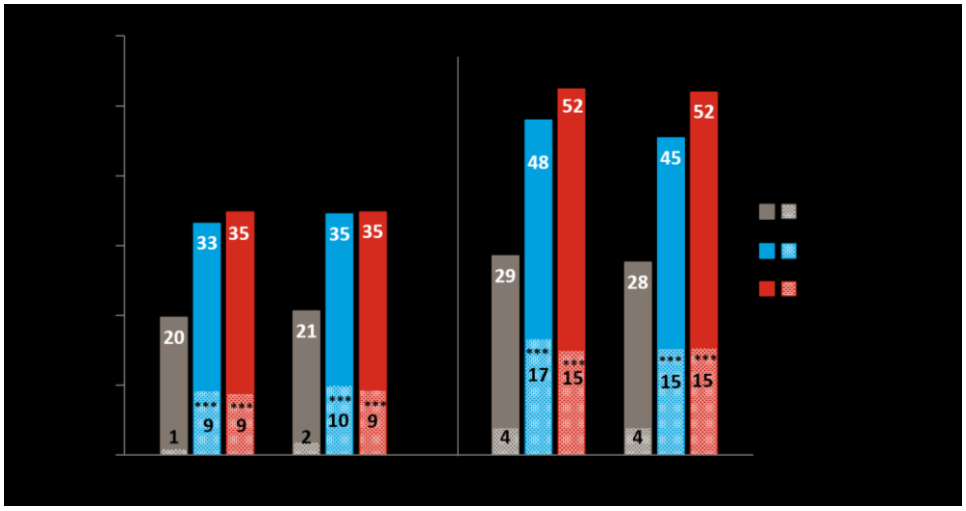
Abbreviations: mTSS = modified Total Sharp Score, LSM = least squares mean, SE = standard error, CI = confidence interval, PBO = placebo, BAR = baricitinib, LE = linear extrapolation.
Source: *JADX CSR Table JADX.11.45.

SDAI and CDAI

The SDAI is a tool for measurement of disease activity in RA that integrates measures of physical examination, acute phase reactant, patient self-assessment, and physician assessment. As shown in Figure 26, statistically significant improvements in the proportion of patients achieving SDAI of ≤3.3 (i.e. remission) and ≤11 (i.e. low disease activity) were observed for baricitinib 4 mg (QD) and 2 mg (QD) vs placebo at Weeks 12 and 24 (Table 42).

The CDAI is similar to the SDAI, but it allows for immediate scoring because it does not use a laboratory result. Similar to the SDAI results, statistically significant improvements in the proportion of patients achieving CDAI of ≤2.8 (i.e. remission) and ≤10 (i.e. low disease activity) were observed for baricitinib 4 mg (QD) and 2 mg (QD) vs placebo at Weeks 12 and 24 (Table 43).

Figure 26. SDAI and CDAI at Weeks 12 and 24



Footnotes: Total height of each bar = SDAI ≤ 11, CDAI ≤ 10. Lower (shaded) portion of each bar = SDAI ≤ 3.3, CDAI ≤ 2.8. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: SDAI = Simplified Disease Activity Index, CDAI = Clinical Disease Activity Index, NRI = non-responder imputation, PBO = placebo BAR = baricitinib.

Table 42. SDAI Response Rate at Week 12 NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
≤ 11 (LDA)					
N	228	229	227	N/A	N/A
Response rate, n (%)	45 (19.7)	76 (33.2)	79 (34.8)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	13.5	15.1
95% CI	N/A	N/A	N/A	(5.5, 21.4)	(7.0, 23.1)
Odds ratio	N/A	N/A	N/A	2.03	2.25
95% CI	N/A	N/A	N/A	(1.31, 3.13)	(1.46, 3.48)
P-value	N/A	N/A	N/A	0.002	0.001
≤ 3.3 (Remission)					
N	228	229	227	N/A	N/A
Response rate, n (%)	2 (0.9)	21 (9.2)	20 (8.8)	N/A	N/A

Difference in response rate	N/A	N/A	N/A	8.3	7.9
95% CI	N/A	N/A	N/A	(4.4, 12.2)	(4.1, 11.8)
Odds ratio	N/A	N/A	N/A	N/A	N/A
95% CI	N/A	N/A	N/A	N/A	N/A
P-value	N/A	N/A	N/A	0.001	0.001

Abbreviations: SDAI = Simplified Disease Activity Index, NRI = non-responder imputation, BAR = baricitinib, PBO = placebo, CI = confidence interval, LDA = low disease activity
Source: *JADX CSR Table JADX.14.16.

Table 43. CDAI Response Rate at Week 12 using NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
≤ 10					
N	228	229	227	N/A	N/A
Response rate, n (%)	47 (20.6)	79 (34.5)	79 (34.8)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	13.9	14.2
95% CI	N/A	N/A	N/A	(5.8, 22.0)	(6.1, 22.3)
Odds ratio	N/A	N/A	N/A	2.05	2.13
95% CI	N/A	N/A	N/A	(1.33, 3.14)	(1.39, 3.28)
P-value	N/A	N/A	N/A	0.002	0.001
≤ 2.8					
N	228	229	227	N/A	N/A
Response rate, n (%)	4 (1.8)	23 (10.0)	21 (9.3)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	8.3	7.5
95% CI	N/A	N/A	N/A	(4.0, 12.5)	(3.4, 11.6)
Odds ratio	N/A	N/A	N/A	N/A	N/A
95% CI	N/A	N/A	N/A	N/A	N/A
P-value	N/A	N/A	N/A	0.001	0.001

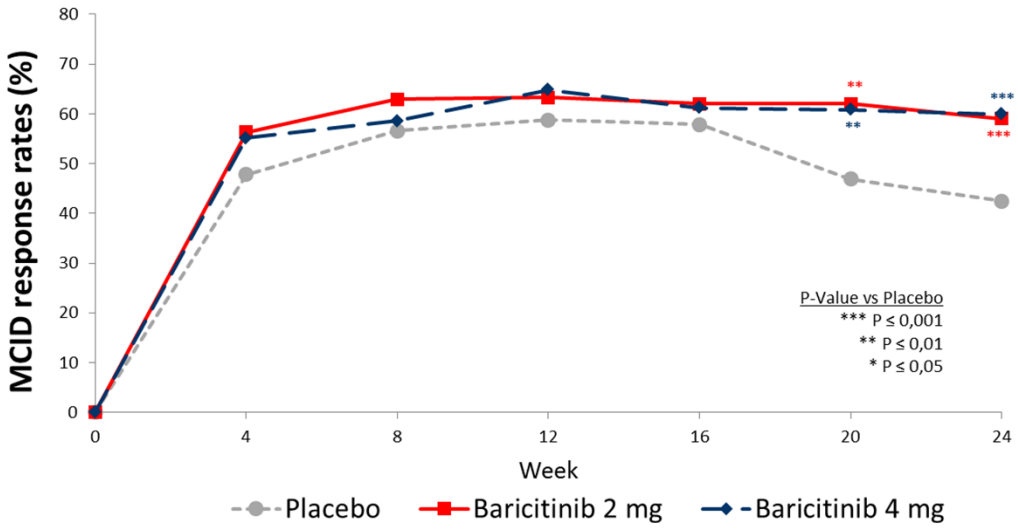
Abbreviations: CDAI = Clinical Disease Activity Index, NRI = non-responder imputation, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), PBO = placebo, CI = confidence interval.
Source: *JADX CSR Table JADX.14.34.

Functional Assessment of Chronic Illness Therapy–Fatigue

The FACIT-F is a brief, 13-item, symptom-specific questionnaire that specifically assesses the self-reported severity of fatigue and its impact upon daily activities and functioning. Compared to placebo, a statistically significant improvement in change from baseline FACIT-F score was observed at Week 4 and from Weeks 16 through 24 for the 4 mg QD baricitinib arm (p≤0.049 for all timepoints), but not for Week 12 (Table 44). A statistically significant improvement was observed only at Week 20 for the 2 mg QD baricitinib arm (p=0.017) compared to placebo.

Additionally, compared to placebo the proportion of patients who met or exceeded the MCID was statistically significantly greater at Weeks 20 and 24 for the both the baricitinib arms (p≤0.004 for all comparisons, Figure 27), but not for Week 12 (Table 45).

Figure 27. FACIT-F percentage of patients achieving or exceeding MCID at 24 weeks



Footnotes: Data are LS mean scores for FACIT-F. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, PBO = placebo, BAR = baricitinib.

Table 44. FACIT-F Change from Baseline at Week 12 using mLOCF and NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	216	227	216	N/A	N/A
LSM (SE)	7.5 (0.64)	8.5 (0.61)	9.1 (0.64)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	1.0 (0.83)	1.6 (0.84)
95% CI	(6.2, 8.7)	(7.3, 9.7)	(7.8, 10.3)	(-0.7, 2.6)	(-0.1, 3.2)
P-value	0.001	0.001	0.001	0.247	0.063

Abbreviations: FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, MCID = minimum clinically important difference, PBO = placebo, BAR = baricitinib, LSM = least squares mean, SE = standard error, CI = confidence interval, mLOCF = modified last observation carried forward, NRI = non-responder imputation.
Source: *JADX CSR Table JADX.11.41.

Table 45. FACIT-F proportion of patients who met or exceeded the MCID at Week 12 using mLOCF and NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	228	229	227	N/A	N/A
Response rate, n (%)	134 (58.8)	145 (63.3)	147 (64.8)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	4.5	6.0
95% CI	N/A	N/A	N/A	(-4.4, 13.5)	(-2.9, 14.9)
Odds ratio	N/A	N/A	N/A	1.21	1.28
95% CI	N/A	N/A	N/A	(0.83, 1.77)	(0.87, 1.88)
P-value	N/A	N/A	N/A	0.323	0.209

Abbreviations: FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, MCID = minimum clinically important difference, PBO = placebo, BAR = baricitinib, LSM = least squares mean, SE = standard error, CI = confidence interval, mLOCF = modified last observation carried forward, NRI = non-responder imputation.
Source: *JADX CSR Table JADX.11.41.

Duration and severity of morning joint stiffness

Baricitinib 4 mg (QD) and 2 mg (QD) were statistically superior to placebo in reducing the median duration and mean severity (respectively) of MJS in the 7 days prior to Week 12.[11]

Work Productivity and Activity Impairment-Rheumatoid Arthritis (WPAI-RA)

Patients treated with baricitinib 4 mg (QD) demonstrated significant improvements from baseline in the presenteeism (p=0.021), work productivity loss (p=0.049) and activity impairment (p=0.003) scores of the WPAI-RA compared with placebo at Week 12. Likewise, significant improvements from baseline among patients treated with baricitinib 2 mg (QD) were observed at Week 12 for the work productivity loss (p=0.014) and activity impairment scores (p=0.004).[11]

European Quality of Life-5 Dimensions-5 levels (EQ-5D-5L)

As shown in Table 46, baricitinib 4 mg (QD) and 2 mg (QD) were statistically superior to placebo at Week 24 for the EQ-5D-5L UK algorithm index score in RA-BUILD.[11]

Table 46. EQ-5D-5L Index Score change from baseline at Week 24 using mLOCF*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	216	227	216	N/A	N/A
LSM (SE)	0.091 (0.014)	0.157 (0.014)	0.186 (0.014)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	0.066 (0.019)	0.095 (0.019)
95% CI	(0.063, 0.119)	(0.130, 0.184)	(0.158, 0.215)	(0.030, 0.103)	(0.058, 0.132)
P-value	0.001	0.001	0.001	0.001	0.001

Abbreviations: PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), LSM = least squares mean, SE = standard error, CI = confidence interval, EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels, mLOCF = modified last observation carried forward.
Source: *JADX CSR Table JADX.11.43.

4.7.3 RA-BEACON

Summary of the RA-BEACON Clinical Effectiveness Results

- Baricitinib 4 mg and 2 mg were superior to placebo in terms of ACR20 response rates at Week 12 (primary endpoint) and this difference was statistically significant
- The efficacy of baricitinib was not influenced by the number of prior bDMARD (<3 and ≥3) (no treatment interaction, p=0.328 and p=0.066 for baricitinib 2 mg and 4 mg, respectively)
- Baricitinib 4 mg and 2 mg were superior to placebo in terms of ACR50/70 response rates at Week 12, which was maintained through to Week 24, and these differences were statistically significant
- Baricitinib 4 mg was superior to placebo at Week 24 in achieving remission or LDA as defined by DAS28-hsCRP, DAS28-ESR, SDAI and CDAI; these differences were statistically significant. The difference was significant for baricitinib 2 mg at Week 24 only for LDA response as defined by DAS hsCRP ≤3.2 and SDAI ≤11
- The effect of baricitinib 4 mg was generally greater in magnitude, more rapid, durable and consistent across different efficacy measures, particularly for the most clinically meaningful endpoints (ACR50/70, remission and LDA) compared with the baricitinib 2 mg dose
- Patients receiving both doses of baricitinib experienced significant improvements in physical function (HAQ-DI) at both Week 12 and Week 24 compared with placebo
- These improvements from baseline were also clinically meaningful as measured by the 2 cut-off values of ≥0.22 and ≥0.3 to define the MCID in HAQ-DI (HRQOL outcome). The proportion of patients who met or exceeded the MCID was higher in baricitinib 4 mg and 2 mg compared with placebo at both Week 12 and Week 24 and these differences were statistically significant
- More than one third of patients enrolled in RA-BEACON had an inadequate response to both TNFis and non-TNFi bDMARDs. Given the highly refractory nature of the patient population, achieving LDA in >30% of the total patient population at Week 12 and Week 24, as measured by DAS28-hsCRP ≤3.2, is notable as patients with inadequate disease control following treatment with biologics is an unmet need in RA

The clinical effectiveness results presented in this section include the primary outcome, i.e. the proportion patients who achieved an ACR20 response at Week 12, and the following secondary outcomes: ACR50/70 response, DAS28(hsCRP) and DAS28(ESR), EULAR response, HAQ-DI, SDAI and CDAI and FACIT-F. A summary of the major outcomes is presented in Table 47, and any secondary outcomes not presented here are provided in Appendix 10. Throughout all analyses presented, baricitinib (2 mg QD or 4 mg QD) demonstrated consistent, clinically-meaningful efficacy when compared to placebo across a range of relevant outcomes in TNFi-refractory patients.

Table 47. Summary of clinical effectiveness results for RA-BEACON

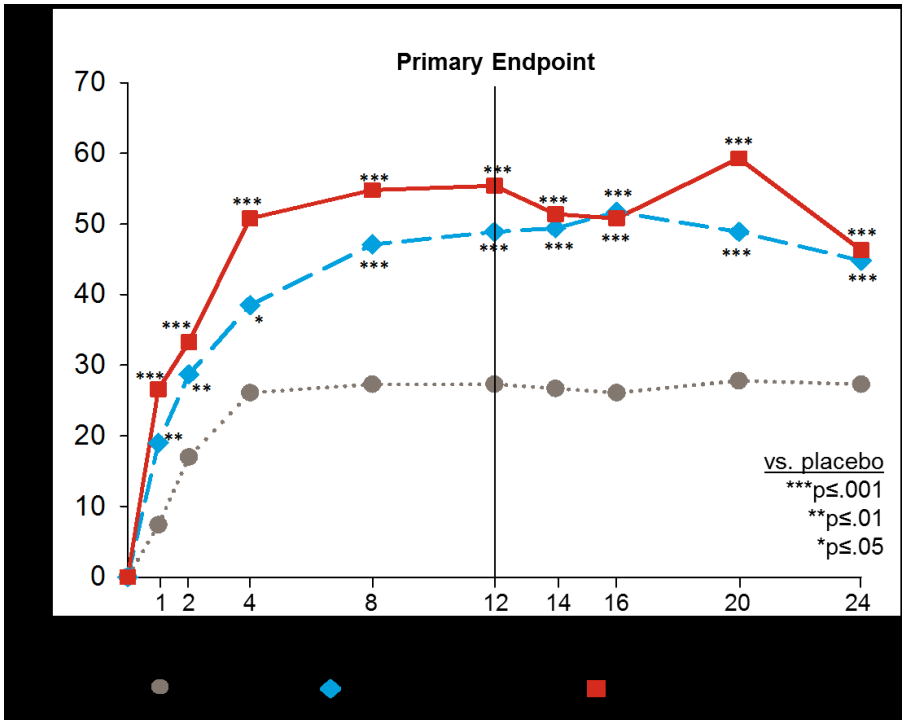
Outcome measures	12 weeks			24 weeks		
	PBO N=176	BAR2 N=174	BAR4 N=177	PBO N=176	BAR2 N=174	BAR4 N=177
ACR20 (%)[a,b]	27.3	48.9***	55.4***	27.3	44.8***	46.3***
ACR50 (%)[c]	8.0	20.1**	28.2***	13.1	23.0*	29.4***
ACR70 (%)[d]	2.3	12.6***	11.3**	3.4	13.2***	16.9***
EULAR (good + moderate) response rate (%)[e]	42.6	66.1***	72.3***	37.5	54.0**	59.9***
EULAR (good) response rate (%)[e]	8.5	24.1***	29.9***	11.4	20.1*	31.6***
DAS28-hsCRP (≤3.2) response rate (%)[f]	9.1	24.1***	31.6***	11.4	20.1*	33.3***
DAS28-hsCRP (<2.6) response rate (%)[f]	4.0	10.9*	16.4***	6.3	10.9	21.5***
HAQ-DI CFB LSM (SE)[g]	-0.17 (0.04)	-0.37*** (0.04)	-0.40*** (0.04)	-0.15 (0.05)	-0.37*** (0.05)	-0.42*** (0.05)
SDAI LDA (≤11.0) response rate (%)[h]	9.1	22.4***	28.2***	14.2	22.4*	31.1***
SDAI remission (≤3.3) response rate (%)[h]	1.7	2.3	5.1	2.3	4.6	9.0**
CDAI LDA (≤10.0) response rate (%)[i]	10.8	23.6**	27.7***	15.3	23.0	31.1***
CDAI remission (≤2.8) response rate (%)[i]	1.7	2.9	5.6	3.4	4.6	9.0*
FACIT-F (MCID) improvement ≥ 3.56 (%)[k]	48.3	63.8**	62.7**	37.5	50.0*	52.5**
FACIT-F CFB LSM (SE)[k]	5.2 (0.9)	8.3 (0.9)**	8.1 (0.9)**	5.7 (0.9)	8.1 (0.9)*	9.2 (0.9)**
MJS Duration (min)[l]	-3.5	-21.0**	-24.0***	-8.0	-25.5**	-27.0**
EQ-5D-5L CFB LSM (SE)[n]	0.036 (0.019)	0.114*** (0.019)	0.169*** (0.018)	0.038 (0.019)	0.111** (0.019)	0.159*** (0.019)

Footnotes: Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), ACR = American College of Rheumatology, ACR20/50/70 = 20/50/70% improvement in ACR core set outcomes, DAS28 = Disease Activity Score, hsCRP = high-sensitivity C-reactive protein, ESR = erythrocyte sedimentation rate, SDAI = Simplified Disease Activity Index, LDA = Low Disease Activity, CDAI = Clinical Disease Activity Index, mTSS = modified Total Sharp Score, FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, LSM = least squares mean, SE = standard error, EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels, MJS = Morning Joint Stiffness.

4.7.3.1 Primary outcome

The primary outcome presented for RA-BEACON is the proportion (or percentage) or patients who achieved an ACR20 response at Week 12 in the 4 mg QD baricitinib arm compared to placebo. RA-BEACON met the primary endpoint, with a significantly greater proportion of patients achieving an ACR20 response at Week 12 for the 4 mg QD baricitinib arm (98/177, 55.4%) compared to the placebo (48/176, 27.3%, p=0.001), as shown in Figure 28 and Table 48. This statistically significant improvement was observed as early as Week 1 and was maintained until Week 24. An additional comparison that compared the 2 mg QD baricitinib arm (85/174, 48.9%) to the placebo arm (48/176, 27.3%, p=0.001) also demonstrated a significant improvement in response rate from Week 1 to Week 24 (Figure 28).[12]

Figure 28. ACR20 response rate over 24 weeks



Footnotes: [a]Patients were on background methotrexate. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, N = number of mITT patients, NRI = non-responder imputation.

Table 48. ACR20 response rate at Week 12 using NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	176	174	177	N/A	N/A
Response rate, n (%)	48 (27.3)	85 (48.9)	98 (55.4)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	21.6	28.1
95% CI	N/A	N/A	N/A	(11.7, 31.5)	(18.2, 37.9)
Odds ratio	N/A	N/A	N/A	2.7	3.4
95% CI	N/A	N/A	N/A	(1.7, 4.2)	(2.2, 5.4)

P-value	N/A	N/A	N/A	0.001	0.001

Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval, NRI = non-responder imputation.
Source: *JADW CSR Table JADW.11.9.

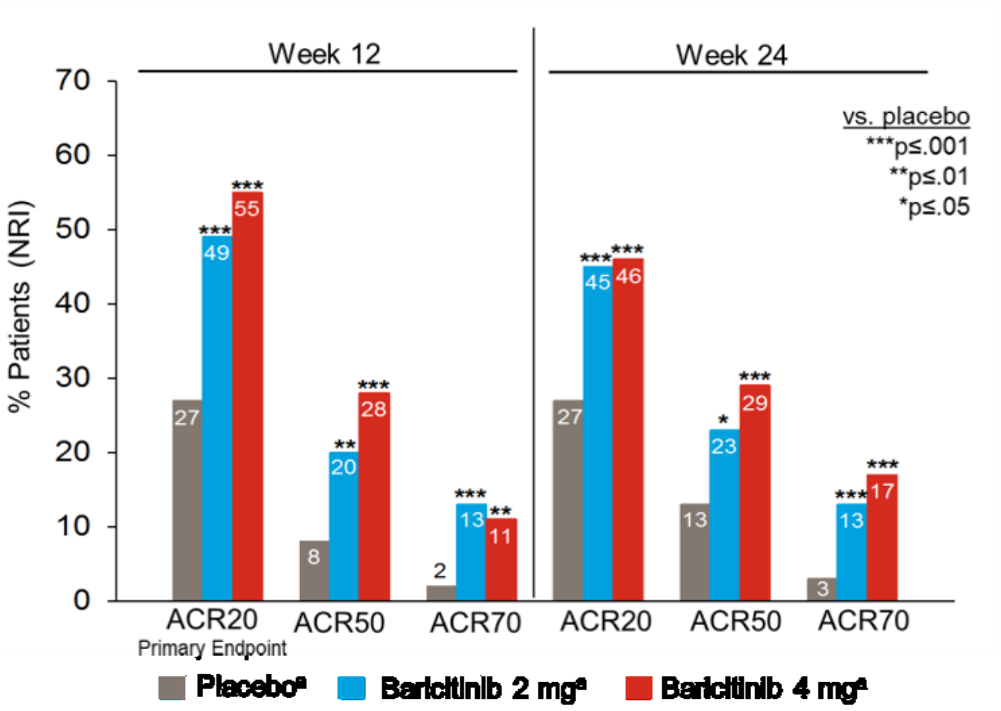
4.7.3.2 Secondary outcomes

The secondary outcomes presented in this submission are: ACR50/70 response, DAS28(hsCRP) and DAS28(ESR), EULAR response, HAQ-DI, SDAI and CDAI and FACIT-F. Other secondary outcomes are presented in Appendix 10.

ACR50 and ACR70

In addition to results achieved for ACR20, baricitinib 2 mg QD and baricitinib 4 mg QD consistently demonstrated efficacy across the increasingly stringent ACR50 and ACR70 outcomes at Week 12 (Figure 29 and Table 49). A significantly greater proportion of patients achieved an ACR50 response (p=0.001) and ACR70 response (p=0.002) for 4 mg QD baricitinib compared to placebo, and a significantly greater proportion of patients achieved an ACR50 response (p=0.002) and an ACR70 response (p=0.001) for 2 mg QD baricitinib compared to placebo. Significant responses were maintained up to Week 24 (Figure 29).[12]

Figure 29. ACR20/50/70 responses at Weeks 12 and 24



Footnotes: Primary endpoint=ACR20 for baricitinib 4 mg QD vs placebo at Week 12. Patients who were rescued or permanently discontinued were imputed thereafter as non-responders. [a]Patients were on background cDMARD. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: ACR = American College of Rheumatology, NRI = Non-responder imputation.

Table 49. ACR50/70 response rate at Week 12 using NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
ACR50					
N	176	174	177	N/A	N/A
Response rate, n (%)	14 (8.0)	35 (20.1)	50 (28.2)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	12.2	20.3
95% CI	N/A	N/A	N/A	(5.0, 19.3)	(12.5, 28.0)
Odds ratio	N/A	N/A	N/A	3.0	4.7
95% CI	N/A	N/A	N/A	(1.6, 5.9)	(2.5, 8.9)
P-value	N/A	N/A	N/A	0.002	0.001
ACR70					
N	176	174	177	N/A	N/A
Response rate, n (%)	4 (2.3)	22 (12.6)	20 (11.3)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	10.4	9.0
95% CI	N/A	N/A	N/A	(5.0, 15.8)	(3.9, 14.2)
Odds ratio	N/A	N/A	N/A	N/A	N/A
95% CI	N/A	N/A	N/A	N/A	N/A
P-value	N/A	N/A	N/A	0.001	0.002

Abbreviations: ACR50/70 = 50/70% improvement in American College of Rheumatology criteria, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval, NRI = non-responder imputation.
Source: *JADW CSR Table JADW.11.15, JADW CSR Table JADW.11.16.

EULAR and DAS responses

The proportion of patients achieving a EULAR good + moderate response was significantly greater compared to placebo at Week 12 for both the baricitinib arms (p=0.001 in both comparisons). A statistically significant improvement was observed as early as Week 1 and was maintained through Week 24 for the baricitinib arms compared to placebo.[12]

Additionally, the proportion of patients achieving a EULAR good response was significantly greater, compared to placebo at Week 12, for the baricitinib arms (p=0.001 for both comparisons). A statistically significant improvement was observed as early as Week 1 and was maintained through Week 24 for the 4 mg QD baricitinib arm compared to placebo, whilst a statistically significant improvement was observed as early as Week 4 and was maintained through Week 24 for the 2 mg QD baricitinib arm compared to placebo. A summary of EULAR responses at Week 12 is shown in Table 50.[12]

Table 50. EULAR Response Rate at Week 12 using NRI*

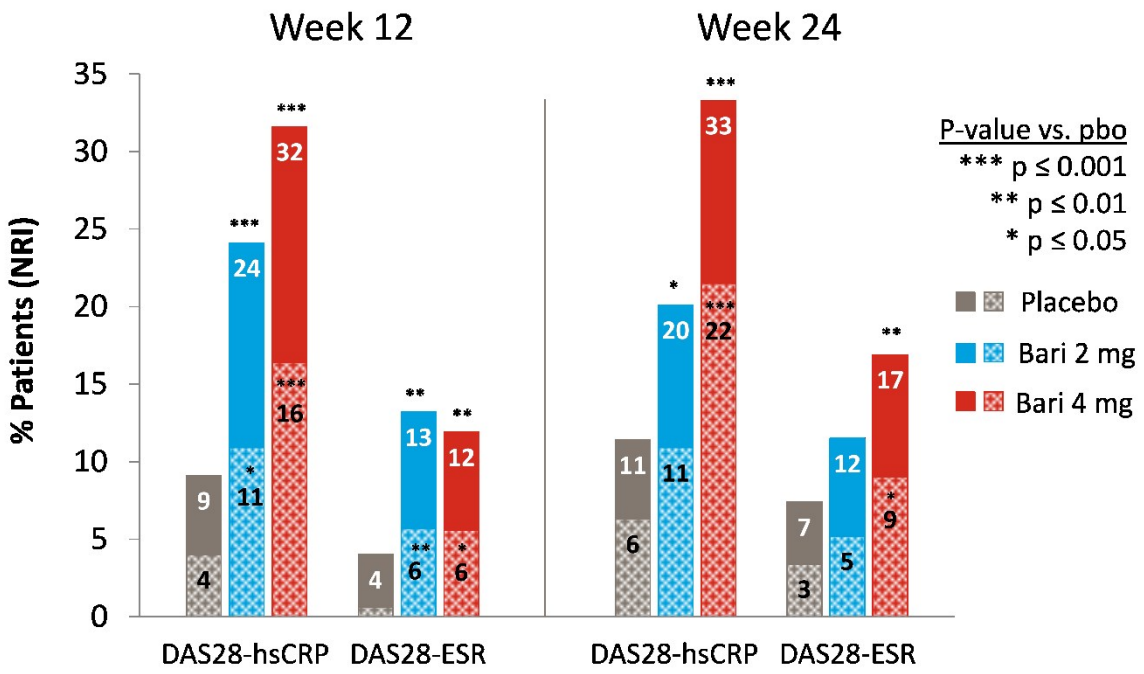
Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
Good and Moderate Response					
N	176	174	177	N/A	N/A
Response rate, n (%)	75 (42.6)	115 (66.1)	128 (72.3)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	23.5	29.7
95% CI	N/A	N/A	N/A	(13.3, 33.6)	(19.9, 39.5)
Odds ratio	N/A	N/A	N/A	2.7	3.6
95% CI	N/A	N/A	N/A	(1.8, 4.2)	(2.3, 5.7)
P-value	N/A	N/A	N/A	0.001	0.001
Good Response					
N	176	174	177	N/A	N/A
Response rate, n (%)	15 (8.5)	42 (24.1)	53 (29.9)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	15.6	21.4
95% CI	N/A	N/A	N/A	(8.0, 23.2)	(13.5, 29.3)
Odds ratio	N/A	N/A	N/A	3.6	4.8
95% CI	N/A	N/A	N/A	(1.9, 6.8)	(2.6, 9.0)
P-value	N/A	N/A	N/A	0.001	0.001

Abbreviations: EULAR = EULAR response criteria score, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval, NRI = non-responder imputation.
Source: *JADW CSR Table JADW.14.33.

Similarly, Figure 30 and Table 51 demonstrate the statistically significantly greater proportion of patients achieving LDA or remission for baricitinib 2 mg (QD) and 4 mg (QD) in DAS28(hsCRP) and DAS(ESR) at Weeks 12 and 24 compared to placebo.[12]

More than one third of patients enrolled in RA-BEACON had an inadequate response to both TNFi's and non-TNFi bDMARDs. Given the highly refractory nature of the patient population, that LDA, as measured by DAS28(hsCRP) ≤3.2, was achieved by >30% of the total patient population receiving baricitinib 4 mg QD at Week 12 and Week 24, is a notable result (Figure 30 and Table 51). Patients with inadequate disease control following treatment with biologics represent an unmet need in RA.[12]

Figure 30. DAS28(hsCRP) and DAS28(ESR) results over 12 and 24 weeks



Footnotes: Total height of each bar indicates patients reaching DAS28 ≤ 3.2. Lower (shaded) portion of each indicates patients reaching DAS28 < 2.6. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: DAS28 = Disease Activity Score 28 joints, hsCRP = high-sensitivity C-reactive protein, ESR = erythrocyte sedimentation rate, NRI = non-responder imputation, PBO = placebo.

Table 51. DAS28(hsCRP) Response Rate at Week 12 using NRI*

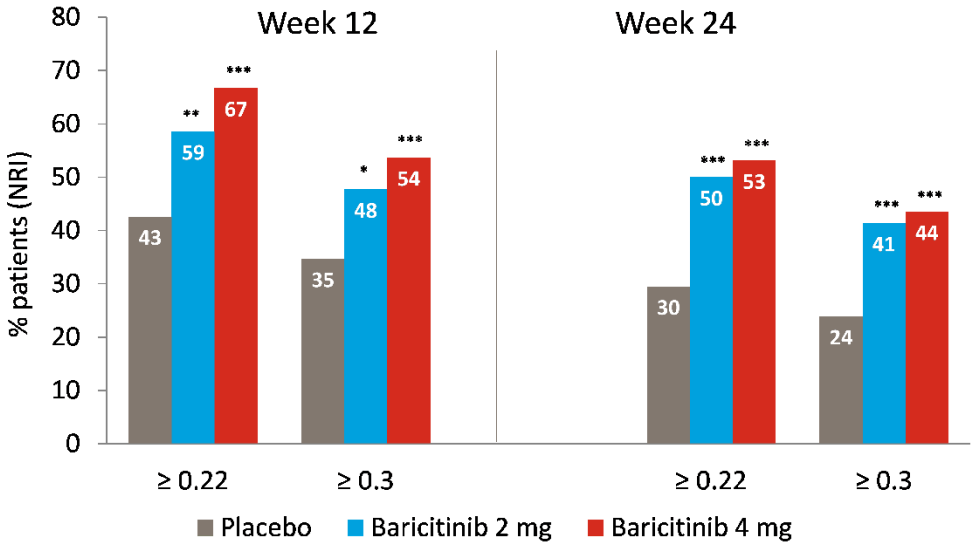
Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
≤ 3.2					
N	176	174	177	N/A	N/A
Response rate, n (%)	16 (9.1)	42 (24.1)	56 (31.6)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	15.0	22.5
95% CI	N/A	N/A	N/A	(7.4, 22.7)	(14.5, 30.6)
Odds ratio	N/A	N/A	N/A	3.3	4.8
95% CI	N/A	N/A	N/A	(1.8, 6.2)	(2.6, 8.9)
P-value	N/A	N/A	N/A	0.001	0.001
< 2.6					
N	176	174	177	N/A	N/A
Response rate, n (%)	7 (4.0)	19 (10.9)	29 (16.4)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	6.9	12.4
95% CI	N/A	N/A	N/A	(1.5, 12.4)	(6.2, 18.6)
Odds ratio	N/A	N/A	N/A	N/A	N/A
95% CI	N/A	N/A	N/A	N/A	N/A
P-value	N/A	N/A	N/A	0.015	0.001

Abbreviations: DAS28 = Disease Activity Score, hsCRP = high-sensitivity C-reactive protein, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval, NRI = non-responder imputation.
Source: *JADW CSR Table JADW.11.25.

Health Assessment Questionnaire-Disability Index

The HAQ-DI is a patient-reported measure of physical function. As shown in Table 52 statistically significant improvements in change from baseline HAQ-DI scores were observed at Week 12 for the 2 mg QD baricitinib (p=0.001) and 4 mg QD baricitinib (p=0.001) arms compared to placebo. Compared to placebo, a statistically significant improvement was observed as early as Week 1 and was maintained through Week 12 for both baricitinib arms. Moreover, baricitinib 4 mg and 2 mg (QD) were statistically superior to placebo in the proportion of patients who met or exceeded the MCID in HAQ-DI for the two cut-off values of ≥0.22 and ≥0.3 at Week 12 and Week 24 (Figure 31).[12]

Figure 31. HAQ-DI percentage of patients who met or exceeded the MCID at Weeks 12 and 24



Footnotes: Data indicate patients meeting or exceeding MCID HAQ-DI (≥0.22 and ≥0.3). *p≤0.05, **p≤0.01, ***p≤0.001 vs. placebo. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs. placebo.
Abbreviations: HAQ-DI = Health Assessment Questionnaire–Disability Index, NRI = non-responder imputation.

Table 52. HAQ-DI Change from Baseline at Week 12 using NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	176	174	177	N/A	N/A
LSM (SE)	-0.17 (0.04)	-0.37 (0.04)	-0.40 (0.04)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	-0.20 (0.05)	-0.23 (0.05)
95% CI	(-0.26, -0.09)	(-0.46, -0.29)	(-0.48, -0.31)	(-0.30, -0.10)	(-0.33, -0.13)
P-value	0.001	0.001	0.001	0.001	0.001

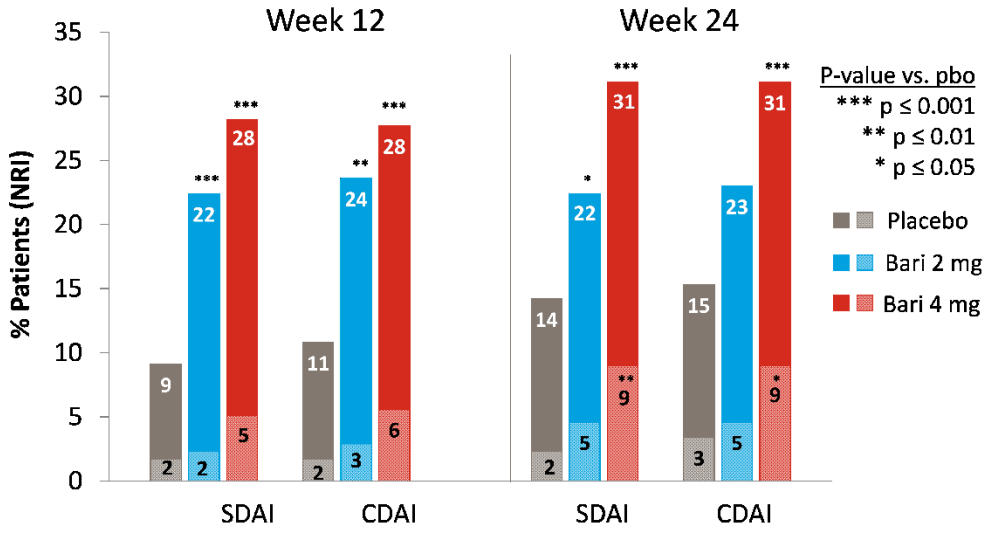
Abbreviations: HAQ-DI = Health Assessment Questionnaire–Disability Index, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), LSM = least squares mean, SE = standard error, CI = confidence interval, NRI = non-responder imputation.
Source: *JADW CSR Table JADW.11.10.

SDAI and CDAI

The SDAI is a tool for measurement of disease activity in RA that integrates measures of physical examination, acute phase reactant, patient self-assessment, and physician assessment. The CDAI is similar to the SDAI, but it allows for immediate scoring because it does not use a laboratory result. As shown in Figure 32, baricitinib 4 mg (QD) was superior to placebo in achieving LDA, as defined by the proportion of patients achieving SDAI ≤ 11 and CDAI ≤ 10, at Weeks 12 and 24. Baricitinib 4 mg (QD) was also superior to placebo in achieving remission, as defined by SDAI ≤ 3.3 and CDAI ≤ 2.8, at Week 24.

Similar to the baricitinib 4 mg (QD) results, baricitinib 2 mg (QD) was superior to placebo in achieving LDA, as defined by the proportion of patients achieving SDAI ≤ 3.3 and CDAI ≤ 2.8, at Week 12 and was superior for LDA, as defined by SDAI, at Week 24. Summaries of SDAI and CDAI results are shown in Table 53 and Table 54, respectively.

Figure 32. SDAI and CDAI at Weeks 12 and 24



Footnotes: Total height of each bar = SDAI ≤ 11, CDAI ≤ 10. Lower (shaded) portion of each bar = SDAI ≤ 3.3, CDAI ≤ 2.8. Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs placebo.
Abbreviations: SDAI = Simplified Disease Activity Index, CDAI = Clinical Disease Activity Index, NRI = non-responder imputation, BAR = baricitinib.

Table 53. SDAI Response Rate at Week 12 using NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
≤ 11 (LDA)					
N	176	174	177	N/A	N/A
Response rate, n (%)	16 (9.1)	39 (22.4)	50 (28.2)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	13.3	19.2
95% CI	N/A	N/A	N/A	(5.8, 20.8)	(11.3, 27.0)
Odds ratio	N/A	N/A	N/A	3.0	4.1
95% CI	N/A	N/A	N/A	(1.6, 5.7)	(2.2, 7.6)
P-value	N/A	N/A	N/A	0.001	0.001

≤ 3.3 (Remission)					
N	176	174	177	N/A	N/A
Response rate, n (%)	3 (1.7)	4 (2.3)	9 (5.1)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	0.6	3.4
95% CI	N/A	N/A	N/A	(-2.3, 3.5)	(-0.4, 7.1)
Odds ratio	N/A	N/A	N/A	N/A	N/A
95% CI	N/A	N/A	N/A	N/A	N/A
P-value	N/A	N/A	N/A	0.723	0.140

Abbreviations: SDAI = Simplified Disease Activity Index, NRI = non-responder imputation, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval, NRI = non-responder imputation.
Source: *JADW CSR Table JADW.14.16.

Table 54. CDAI Response Rate at Week 12 using NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
≤ 10 (LDA)					
N	176	174	177	N/A	N/A
Response rate, n (%)	19 (10.8)	41 (23.6)	49 (27.7)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	12.8	16.9
95% CI	N/A	N/A	N/A	(5.0, 20.6)	(8.9, 24.9)
Odds ratio	N/A	N/A	N/A	(5.0, 20.6)	(8.9, 24.9)
95% CI	N/A	N/A	N/A	2.7	3.3
P-value	N/A	N/A	N/A	(1.5, 4.9)	(1.8, 5.9)
≤ 2.8 (Remission)					
N	176	174	177	N/A	N/A
Response rate, n (%)	3 (1.7)	5 (2.9)	10 (5.6)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	1.2	3.9
95% CI	N/A	N/A	N/A	(-2.0, 4.3)	(0.0, 7.8)
Odds ratio	N/A	N/A	N/A	N/A	N/A
95% CI	N/A	N/A	N/A	N/A	N/A
P-value	N/A	N/A	N/A	0.501	0.087

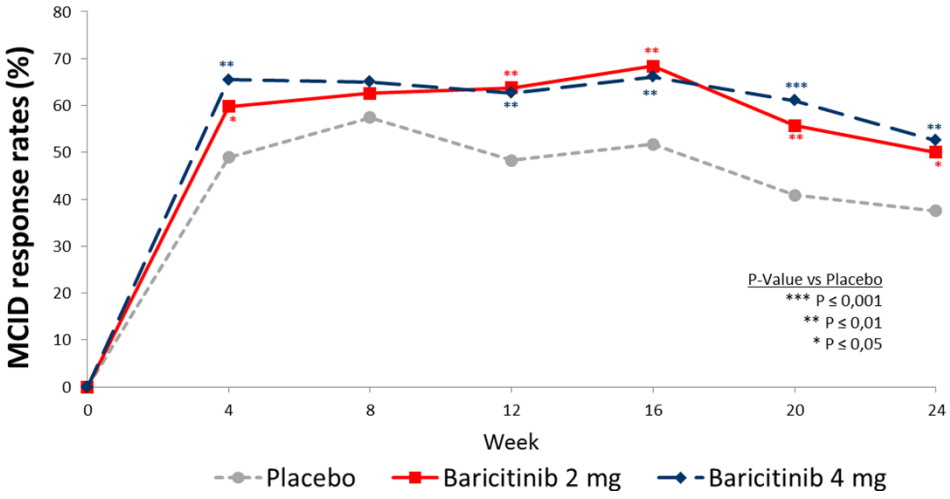
Abbreviations: CDAI = Clinical Disease Activity Index, NRI = non-responder imputation, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval, NRI = non-responder imputation.
Source: *JADW CSR Table JADW.14.30.

Functional Assessment of Chronic Illness Therapy–Fatigue

The FACIT-F is a brief, 13-item, symptom-specific questionnaire that specifically assesses the self-reported severity of fatigue and its impact upon daily activities and functioning. Compared to placebo, a statistically significant improvement in change from baseline FACIT-F score was observed at Weeks 12 (Table 55) and 24 for 4 mg QD baricitinib (p≤0.001 for both timepoints) and 2 mg QD baricitinib (p≤0.001 for both timepoints).

Additionally, compared to placebo, the proportion of patients who met or exceeded the MCID was statistically significantly greater at Weeks 12 and 24 for the baricitinib arms (p≤0.015 for all arms and timepoints, Figure 33).[12]

Figure 33. FACIT-F percentage of patients who met or exceeded the MCID over 24 weeks



Footnotes: Significance level definitions: *p≤0.05, **p≤0.01, ***p≤0.001 vs. placebo.
Abbreviations: FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, MCID = minimum clinically iimportant difference

Table 55. FACIT-F Change from Baseline at Week 12 using mLOCF and NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	170	170	174	N/A	N/A
LSM (SE)	5.2 (0.9)	8.3 (0.9)	8.1 (0.9)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	3.1 (1.0)	2.9 (1.0)
95% CI	(3.5, 6.9)	(6.6, 9.9)	(6.5, 9.8)	(1.0, 5.1)	(0.9, 5.0)
P-value	0.001	0.001	0.001	0.004	0.005

Abbreviations: FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, PBO = placebo, BAR = baricitinib, LSM = least squares mean, SE = standard error, CI = confidence interval, mLOCF = modified last observation carried forward, NRI = non-responder imputation.
Source: *JADW CSR Table JADW.11.40.

Table 56. FACIT-F percentage of patients who met or exceeded the MCID at Week 12 using mLOCF and NRI*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	176	174	177	N/A	N/A
Response rate, n (%)	85 (48.3)	111 (63.8)	111 (62.7)	N/A	N/A
Difference in response rate	N/A	N/A	N/A	15.5	14.4
95% CI	N/A	N/A	N/A	(5.2, 25.8)	(4.2, 24.7)
Odds ratio	N/A	N/A	N/A	1.9	1.8

95% CI	N/A	N/A	N/A	(1.2, 2.9)	(1.2, 2.8)
P-value	N/A	N/A	N/A	0.004	0.007

Abbreviations: FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, MCID = minimum clinically important difference, PBO = placebo, BAR = baricitinib, CI = confidence interval, mLOCF = modified last observation carried forward, NRI = non-responder imputation.
Source: *JADW CSR Table JADW.11.40.

Duration of morning joint stiffness

Compared to placebo, a statistically significant decrease in the duration of morning joint stiffness was observed at Weeks 12 and 24 for the baricitinib groups. Compared to placebo, a statistically significant decrease was observed as early as Week 1 and was maintained through Week 24 for the baricitinib 4 mg group. Compared to placebo, a statistically significant decrease was observed as early as Week 4 and was maintained through Week 24 for the baricitinib 2 mg group.

Work Productivity and Activity Impairment-Rheumatoid Arthritis (WPAI-RA)

Baricitinib 4 mg (QD) and 2 mg (QD) were superior to placebo in improving the activity impairment score of the WPAI-RA at Weeks 12 and 24 (4 mg: p≤0.001 for both time points; 2 mg: p=0.005 and p=0.030 for Weeks 12 and 24, respectively).[12]

European Quality of Life-5 Dimensions-5 levels (EQ-5D-5L)

As shown in Table 57, both baricitinib 4 mg (QD) and 2 mg (QD) doses were superior to placebo in improving the EQ-5D-5L (UK algorithm) index score at Week 24.[12]

Table 57. EQ-5D-5L Index Score change from baseline at Week 24 mLOCF*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	167	168	173	N/A	N/A
LSM (SE)	0.038 (0.019)	0.111 (0.019)	0.159 (0.019)	N/A	N/A
LSM difference (SE)	N/A	N/A	N/A	0.073 (0.023)	0.121 (0.023)
95% CI	(-0.000, 0.076)	(0.073, 0.149)	(0.122, 0.197)	(0.027, 0.119)	(0.076, 0.167)
P-value	0.051	0.001	0.000	0.002	0.001

Abbreviations: EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), LSM = least squares mean, SE = standard error, CI = confidence interval, mLOCF = modified last observation carried forward.
Source: *JADW CSR Table JADW.11.42.

4.7.4 RA-BEYOND

RA-BEYOND is an ongoing, phase III, long-term extension study to assess the safety and efficacy of baricitinib in patients with RA.[202] Patients who participated and completed the treatment period in the following studies were eligible for enrolment into RA-BEYOND:

- Study JADA (phase II)
- RA-BEAM
- RA-BUILD
- RA-BEACON
- RA-BEGIN

The primary objective of RA-BEYOND is to assess the long-term safety and tolerability of baricitinib through evaluation of:

- Treatment-emergent adverse events (TEAEs), adverse events of special interest (AESIs) and serious adverse events (SAEs)
- Temporary investigational product interruptions and permanent investigational product discontinuations
- Vital signs and laboratory evaluations (including chemistry and haematology).

The secondary objective of RA-BEYOND is to evaluate the long-term maintenance of treatment response to baricitinib, as assessed by:

- The proportion of patients:
    - who maintain an ACR20/50/70 response
    - who maintain a DAS28(hsCRP)/DAS28(ESR) ≤3.2, CDAI ≤10, and HAQ-DI improvement of ≥0.22 and ≥0.3
    - with mTSS change ≤0 from baseline
- The change from baseline in:
    - structural joint damage as measured by mTSS
    - joint space narrowing and bone erosion score
    - duration of MJS
    - European Quality of Life-5 Dimensions-5 Level (EQ-5D-5L) scores
- Evaluation of healthcare resource utilisation

Results from the August 2015 data cut-off have demonstrated sustained efficacy for patients continuing with the baricitinib treatment allocation from their originating study (RA-BEAM, RA-BUILD or RA-BEACON) over an additional 48 weeks of treatment. At the time of the cut-off, this comprised 2,290 patients.[15] These results indicate continued efficacy with baricitinib treatment, especially as patients in RA-BEYOND have been exposed to a variety of previous therapies and a range of disease durations. Furthermore, the high patient retention rates (90%) seen to date in RA-BEYOND support a favourable risk/benefit profile following prolonged treatment with baricitinib.[15]

Table 58 shows the categorical response rates of patients at Week 12 of their originating studies and 48 weeks after entry into RA-BEYOND who were:

- Originally randomised to baricitinib
- Who were not rescued in an originating study
- Who entered RA-BEYOND at least 48 weeks before the August 2015 submission data cutoff.

Table 58. Durability of efficacy: responses at Week 12 of originating study and after 48 weeks in RA-BEYOND

Outcome measure*		RA-BEAM BAR4 N=[redacted]	RA-BUILD		RA-BEACON	
			BAR2 N=[redacted]	BAR4 N=[redacted]	BAR2 N=[redacted]	BAR2 N=[redacted]
ACR20 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ACR50 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ACR70 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DAS28(hsCRP)<2.6 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DAS28(hsCRP) ≤3.2 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CDAI ≤2.8 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CDAI ≤10.0 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SDAI ≤3.3 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

SDAI ≤11 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ-DI improvement ≥0.3 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: JADY CSR.[15]
Footnotes: [a]NRI (rescue not available at week 12), [b]NRI without considering rescue status, *Baseline in the originating study is used in the response rate calculation. The time points are weeks since randomisation in the originating study. Analyses exclude patients who were rescued or switched in the originating studies. RA-BEYOND populations analysed here only include patients who have completed 48 weeks of RA-BEYOND or would have completed 48 weeks if not discontinued. As such, not all patients from the originating study were included in analyses, leading to a different sample size than the patient population from the original study. Therefore, Week 12 results presented in the table above may differ to those presented earlier for the respective originating study due to the difference in sample sizes, which affects the proportion of patients achieving an outcome measure. Data after patients step down to baricitinib 2 mg are imputed based on the model predicted values using data from baricitinib treatment period in the originating and RA-BEYOND studies. NRI without considering rescue is used to impute missing data. Note: Baseline in the originating study is used in the response rate calculation. The time points are weeks since randomisation in the originating study. Data after patients' step-down to baricitinib 2 mg are imputed based on the model predicted values using data from baricitinib treatment period in the originating and RA-BEYOND studies. NRI without considering rescue is used to impute missing data. Note: One year after entry in RA-BEYOND is Week 100 for the 52-week studies (RA-BEGIN and RA-BEAM), Week 72 for the 24-week studies (RA-BUILD and RA-BEACON)
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology Criteria; ACR50 = 50% improvement in American College of Rheumatology Criteria; ACR70 = 70% improvement in American College of Rheumatology Criteria; BAR = baricitinib; CDAI = Clinical Disease Activity Index; DAS28–hsCRP = Disease Activity Score in 28 joints high-sensitivity C-reactive protein; DAS28-ESR = Disease Activity Score in 28 joints-erythrocyte sedimentation rate; HAQ-DI = Health Assessment Questionnaire-Disability Index; NRI = non-responder imputation; PGA = Physician's Global Assessment of Disease Activity; PtGA = Patient's Global Assessment of Disease Activity; SDAI = Simplified Disease Activity Index

Maintenance of response from RA-BEYOND entry up to 48 weeks for HAQ-DI

Patients originating from the RA-BEAM, RA-BUILD and RA-BEACON studies who met or exceeded HAQ-DI MCID (≥0.22 and ≥0.3) at baseline in RA-BEYOND, were shown to maintain their response for a further 48 weeks of treatment with baricitinib 4 mg (QD), as shown in Figure 34, Figure 35 and Figure 36, respectively. These results support the hypothesis that the clinically meaningful effect of baricitinib on physical function is maintained long-term across different RA patient populations.

Figure 34. Maintenance of HAQ-DI MCID across 48 weeks among RA-BEAM patients entering RA-BEYOND (n=78)

Abbreviations: HAQ-DI = Health Assessment Questionnaire-Disability Index; JADY = RA-BEYOND; MCID = minimum clinically important difference.
Source: JADY CSR - Section 7.1.2.3., HAQ-DI Improvement >=0.22 and >=0.3; table JADY 11.141

Figure 35. Maintenance of HAQ-DI MCID across 48 weeks among RA-BUILD patients entering RA-BEYOND (n=135)

Abbreviations: HAQ-DI = Health Assessment Questionnaire-Disability Index; MCID = minimum clinically important difference.
Source: JADY CSR - Section 7.1.2.3., HAQ-DI Improvement >=0.22 and >=0.3 - Table JADY 11.142

Figure 36. Maintenance of HAQ-DI MCID across 48 weeks among RA-BEACON patients entering RA-BEYOND (n=121)

Abbreviations: HAQ-DI = Health Assessment Questionnaire-Disability Index; MCID = minimum clinically important difference
Source: JADY CSR - Section 7.1.2.3. HAQ-DI Improvement >=0.22 and >=0.3 - Table JADY 11.143

Baricitinib 4 mg to 2 mg step-down titration

RA-BEYOND includes a dose step-down experiment where patients who received baricitinib 4 mg (QD) for at least 15 months, achieved sustained LDA or remission (as measured by CDAI) and had not previously been rescued; are re-randomised in a double-blind manner to continue baricitinib 4 mg (QD) or step-down to a reduced dose of baricitinib 2 mg (QD).[15]

Interim results showed that dose reduction from baricitinib 4 mg to 2 mg (QD) was associated with modest but statistically significant increases in disease activity across measures after 12 weeks, and higher rescue rates compared to patients who continued to receive baricitinib 4 mg (QD). Despite the modest increases in disease activity the majority of patients in both the 4 mg and 2 mg (QD) groups retained their state of LDA or remission that led to their step-down across the study sets. Table 59 summarises the CDAI response rates observed in patients at Week 12 after stepping down to a reduced dose of baricitinib 2 mg (QD). Of note, patients who did not maintain disease control when stepping down to the 2 mg dose were able to recapture disease control when they were stepped back up to the 4 mg dose.

Table 59. Summary of CDAI response rates at Week 12 after step-down in RA-BEYOND

	Combined Studies RA-BEAM/ BUILD/BEACON	Combined Studies RA-BEAM/BUILD	RA-BEGIN	RA-BEACON

	BAR2 [redacted]	BAR4 [redacted]	BAR2 [redacted]	BAR4 [redacted]	BAR2 [redacted]	BAR4 [redacted]	BAR2 [redacted]	BAR4 [redacted]
CDAI Week 12 Disease Activity (using NRI)								
CDAI ≤10 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CDAI ≤2.8 (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CDAI BAR4 vs BAR2 % difference in response rate at 12 weeks' post-randomisation (using NRI)								
CDAI ≤10 (%)	[redacted]		[redacted]		[redacted]		[redacted]	
p-value	[redacted]		[redacted]		[redacted]		[redacted]	
CDAI ≤2.8 (%)	[redacted]		[redacted]		[redacted]		[redacted]	
p-value	[redacted]		[redacted]		[redacted]		[redacted]	

Abbreviations: BAR = baricitinib; CDAI = Clinical Disease Activity Index; NRI = non-responder imputation; N = number of modified intent-to-treat patients

4.8 *Subgroup analysis*

Results for the primary endpoint of each of the three trials (ACR20 response at Week 12) are presented for the following subgroups:

RA-BEAM (Section 4.8.2):

- Disease activity (DAS-hsCRP ≤5.1) and disease activity (DAS28-hsCRP >5.1)

RA-BUILD (Section 4.8.1):

- Disease activity (DAS-hsCRP ≤5.1) and Disease activity (DAS28-hsCRP >5.1)

RA-BEACON (Section 4.8.3):

- Disease activity (DAS-hsCRP ≤5.1) and disease activity (DAS28-hsCRP >5.1)
- <3 previous bDMARDs used and ≥3 previous bDMARDs used

The subgroup analyses demonstrate the efficacy of baricitinib in patients with either moderate or severe disease activity at baseline.[82-84] Efficacy is also demonstrated regardless of whether <3 or ≥3 bDMARDs have been received in the TNFi-IR population.[84] A full list of all subgroup analyses performed is presented in Appendix 11.

4.8.1 RA-BEAM

ACR20 response data at Week 12 for patients who had either moderate or severe disease activity at baseline in RA-BEAM are presented in Table 60 and Table 61, respectively. Similar results to the mITT population were observed for the moderate population, with a statistically significantly greater proportion of patients achieved an ACR20 response at Week 12 with baricitinib (4 mg QD) compared to placebo in both subgroups. Baricitinib (4 mg QD) was statistically significantly superior to adalimumab in the moderate disease activity subgroup only.

Table 60. ACR20 response rate at Week 12 – moderate disease activity subgroup (baseline DAS28-hsCRP ≤5.1)

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: *P value from interaction of subgroup with treatment. Green cell denotes a statistically significant difference in favour of baricitinib.
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, PBO = placebo, BAR= baricitinib, ADA = adalimumab, CI = confidence interval.
Source: JADV CSR Table JADV.14.63.[82]

Table 61. ACR20 response rate at Week 12 – severe disease activity subgroup (baseline DAS28-hsCRP >5.1)

Outcomes	PBO	BAR4	ADA	BAR4 vs PBO	BAR4 vs ADA
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: *P value from interaction of subgroup with treatment. Green cell denotes a statistically significant difference in favour of baricitinib.
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), ADA = adalimumab, CI = confidence interval.
Source: JADV CSR Table JADV.14.63.[82]

4.8.2 RA-BUILD

ACR20 response data at Week 12 for patients who had either moderate or severe disease activity at baseline in RA-BUILD are presented in Table 62 and Table 63, respectively. Similar to the mITT population, a statistically significantly greater proportion of patients achieved an ACR20 response at Week 12 with either baricitinib 4 mg or 2 mg (QD) compared to placebo in both subgroups.

Table 62. ACR20 response at Week 12 – moderate disease activity subgroup (baseline DAS28-hsCRP ≤5.1)

Outcomes	PBO	BAR4	BAR2	BAR4 vs PBO	BAR2 vs PBO
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: *P value from interaction of subgroup with treatment. Green cell denotes a statistically significant difference in favour of baricitinib.
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria; DAS28 = Disease Activity Score 28 joints; hsCRP = high-sensitivity C-reactive protein; PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.
Source: JADX CSR Table JADX.14.61.[83]

Table 63. ACR20 response at Week 12 – severe disease activity subgroup (baseline DAS28-hsCRP >5.1)

Outcomes	PBO	BAR4	BAR2	BAR4 vs PBO	BAR2 vs PBO
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: *P value from interaction of subgroup with treatment. Green cell denotes a statistically significant difference in favour of baricitinib.
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria; DAS28 = Disease Activity Score 28 joints; hsCRP = high-sensitivity C-reactive protein; PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.
Source: JADX CSR Table JADX.14.61.[83]

4.8.3 RA-BEACON

ACR20 response data at Week 12 for patients who had either moderate or severe disease activity at baseline in RA-BEACON are presented in Table 64 and Table 65, respectively. A numerically greater proportion of patients achieved an ACR20 response at Week 12 in the baricitinib 4 mg or 2 mg (QD) groups compared to placebo in both subgroups (statistical analysis for interaction was not possible due to the small number of responders in the placebo arm for patients with moderate disease activity at baseline).

Table 64. ACR20 response rate at Week 12 – moderate disease activity subgroup (baseline DAS28-hsCRP ≤5.1)*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: *P value from interaction of subgroup with treatment in the logistic regression model. When logistic regression sample size requirements (> 5 responders in any category for any subgroup) are not met, odds ratios, p-values and 95% CIs are not produced.
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.
Source: JADW CSR Table JADW.14.46.[84]

Table 65. ACR20 response rate at Week 12 – severe disease activity subgroup (baseline DAS28-hsCRP >5.1)*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: *P value from interaction of subgroup with treatment in the logistic regression model. When logistic regression sample size requirements (>5 responders in any category for any subgroup) are not met, odds ratios, p-values and 95% CIs are not produced.

Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.

Source: JADW CSR Table JADW.14.46.[84]

ACR20 response data at Week 12 for patients who had previously used either <3 or ≥3 bDMARDs at baseline in RA-BEACON are presented in Table 66 and Table 67, respectively. Similar to the mITT population, a statistically significantly greater proportion of patients achieved an ACR20 response at Week 12 with baricitinib 4 mg (QD) compared to placebo.

Table 66. ACR20 response rate at Week 12 – for subgroup with number of previous bDMARDs used <3*

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: *P value from interaction of subgroup with treatment in the logistic regression model. Green cell denotes a statistically significant difference in favour of baricitinib.

Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.

Source: JADW CSR Table JADW.14.46.[84]

Table 67. ACR20 response rate at Week 12 – for subgroup with number of previous bDMARDs used ≥3

Outcomes	PBO	BAR2	BAR4	BAR2 vs PBO	BAR4 vs PBO
N	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Response rate, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Difference in response rate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Odds ratio	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
P-value*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: *P value from interaction of subgroup with treatment in the logistic regression model. Green cell denotes a statistically significant difference in favour of baricitinib.

Abbreviations: ACR20 = 20% improvement in American College of Rheumatology criteria, PBO = placebo, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), CI = confidence interval.

Source: JADW CSR Table JADW.14.46.[84]

4.9 *Meta-analysis*

RA-BEGIN, RA-BEAM, RA-BUILD and RA-BEACON were the only RCTs identified in the systematic literature review, and evaluate the efficacy and safety of baricitinib in separate patient populations, namely DMARD-naïve, methotrexate-inadequate response, cDMARD-inadequate response and TNF inhibitor-inadequate response populations. Whilst a meta-analysis of RA-BUILD and BEACON was theoretically feasible, the fact that a comprehensive network meta-analysis of all relevant comparators was conducted and was more informative meant that this approach was favoured instead of a meta-analysis of these two studies.

4.10 *Indirect and mixed treatment comparisons*

4.10.1 Search strategy

Network meta-analyses (NMA) were performed in order to assess the relative efficacy of baricitinib compared with the relevant comparators in cDMARD (including MTX)-IR or anti-TNF-IR patients with moderate-to-severe RA.

The methodology of the SLR that identified studies to potentially inform the NMAs is described in Section 4.1.

4.10.2 Study selection for the NMA

As reported in Section 4.1, a total of 257 records were ultimately included in the systematic literature review, reporting 138 unique RCTs to be considered for inclusion in the NMAs.

For inclusion in the NMA, studies needed to meet the following criteria:

- The study needed to enable treatment(s) to be connected to at least one other treatment in the network via a common comparator
- Treatment comparisons must have been made within at least one of the two populations of interest (cDMARD-IR or TNF-IR)
- The study needed to report the outcomes of interest for the NMA, which were ACR response (20%, 50% and 70% improvement) or EULAR response (moderate and good response). As described further in Section 4.10.7, a probit model was used for the analysis of outcomes in the NMA, in line with the approach taken by the AG in TA375. It should be noted that it was not necessary for studies to report EULAR response to be included in the NMA if ACR responses only were reported; as described in Section 4.10.5 a conversion was applied to generate EULAR response data from ACR responses
- Another factor considered in defining the base-case network for the cDMARD-IR population was the use of prior biologics. The clinical SLR did not identify any studies evaluating certolizumab (one of the comparators defined in the final scope) in which cDMARD-IR patients were exclusively naïve to biologic treatments. As such, in order to allow inclusion of CTZ as a comparator in the cDMARD-IR NMA, studies were permitted to be included in the analysis if <20% of patients in the study had received prior biologic treatment. This relaxation of criteria allowed certolizumab to be included in the base-case network

Comparators

The treatments included in the NMAs for the cDMARD-IR population and TNF-IR population are shown in Table 68.

Table 68. Summary table of treatments included in each NMA

cDMARD-IR population	Anti-TNF-IR population
• ABA + cDMARD • ADA • ADA + cDMARD • BAR (2 mg QD) + cDMARD • BAR (4 mg QD) + cDMARD • cDMARD • CTZ + cDMARD • ETN • ETN + cDMARD • ETN + SSZ • GOL + cDMARD • IFX + cDMARD • MTX • PBO • RTX[a] • RTX + cDMARD[a] • TCZ • TCZ + cDMARD • TOFA + cDMARD • SSZ • SSZ + HCQ + cDMARD	• ABA + cDMARD • BAR (2 mg QD) + cDMARD • BAR (4 mg QD) + cDMARD • cDMARD • GOL + cDMARD • RTX + MTX • TCZ + cDMARD • TCZ + MTX

Footnote: [a]The cDMARD-IR NMA was performed from a global perspective, and, as a result, RTX, which is not normally considered for the treatment of the cDMARD-IR population in the UK, was included in the analysis. However, the RTX trials included in the network only provide information on RTX and MTX (see Figure 37) and therefore only increase the amount of evidence available in the network for the estimation of treatment effect for the RTX and MTX nodes. Therefore, the inclusion of the RTX studies is not expected to impact the validity of the treatment effect estimates for baricitinib versus its relevant comparators.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN-b etanercept biosimilar, GOL = golimumab, IFX-b = infliximab biosimilar, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine

Analysis time point

For the majority of studies identified in the SLR, the primary outcome was measured at 24 weeks. This time point (defined as study visits scheduled at 24 weeks +/- 4 weeks, allowing study visits from Week 20 to Week 28) was therefore used in the base case NMA analyses. A sensitivity analysis with endpoints measured at 12 weeks (defined as study visits scheduled at 12 weeks [+/-3 weeks]) was also performed and is presented in Section 4.10.8.2, as 12 weeks represents the primary endpoint for the three baricitinib trials presented in Section 4.

4.10.3 Summary of trials included in the NMA

Table 69 and Table 70 present a summary of the trials considered for inclusion in the base case NMA at 24 weeks and the sensitivity analysis at 12 weeks based on the criteria described above, for the cDMARD-IR and anti-TNF-IR populations, respectively. The reporting of ACR and EULAR outcomes from each study considered for inclusion is also detailed in this table. Ultimate inclusion or not of each study in the 24 week or 12 week analyses is indicated by green (included) or red (not included) shading in the 24 week endpoint and 12 week endpoint columns, respectively. Where studies listed in this table were not ultimately included in these analyses at a given timepoint, the reason is provided in the table and accompanying footnotes.

Table 69. Summary of trials considered for inclusion in the NMA for the cDMARD-IR population

Study name	Treatment 1	Treatments 2 / 3	Control	Study design	Endpoints analysed 24 weeks (base case)	Endpoints analysed 12 weeks (sensitivity analysis)
Abe (2006)	IFX 3 mg + MTX (n=49)	---	MTX (n=47)	DB	NA (no MTX arm)	ACR20/50/70
ATTRACT	IFX 3 mg (n=86)		MTX (n=88)	DB	ACR20	ACR20
De Filippis (2006)	IFX 3 mg + MTX (n=16)	ETN 25 mg + MTX (n=16)	---	OL	ACR20/50/70	ACR20/50/70
START	IFX 3 mg + MTX (n=360)	---	MTX (n=363)	DB	ACR20/50/70; EULAR	*No outcomes reported*
ARMADA	ADA 40 mg + MTX (n=67)	---	MTX (n=62)	DB	ACR20/50/70	ACR20/50/70
CHANGE	ADA 40 mg (n=91)	---	Placebo (n=87)	DB	ACR20/50/70	ACR20/50/70
Keystone (2004)	ADA 40 mg + MTX (n=207)	---	MTX (n=200)	DB	ACR20/50/70	ACR20/50/70
Kim (2007)	ADA 40 mg + MTX (n=65)	---	MTX (n=63)	DB	ACR20/50/70	ACR20/50/70
STAR	ADA 40 mg + cDMARD (n=318)	---	cDMARD (n=318)	DB	ACR20/50/70	ACR20/50/70
Van de Putte (2004)	ADA 40 mg (n=113)	---	Placebo (n=110)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR
CNTO 148[a]	GOL 50 mg + MTX (n=35)	---	MTX (n=35)	DB	*[ACR20/50/70][a]*	ACR20/50/70
GO-FORTH	GOL 50 mg + MTX (n=89)	---	MTX (n=90)	DB	ACR20/50/70 EULAR	ACR20/50/70; EULAR
GO-FORWARD	GOL 50 mg + MTX (n=89)	---	MTX (n=133)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR
Li (2013)	GOL 50 mg + MTX (n=132)	---	MTX (n=132)	DB	ACR20/50/70	ACR20/50/70
J-RAPID[b]	CTZ + MTX (n=82)	---	MTX (n=77)	DB	ACR20/50/70	ACR20/50/70; EULAR
Kang (2013)[b]	CTZ + MTX (n=81)	---	MTX (n=40)	DB	ACR20/50/70	*No outcomes reported*
RAPID1[b]	CTZ + MTX (n=393)	---	MTX (n=199)	DB	ACR20/50/70; EULAR	ACR20/50/70
RAPID2[b]	CTZ + MTX (n=246)	---	MTX (n=127)	DB	ACR20/50/70	ACR20/50/70
REALISTIC[c]	CTZ + cDMARD (n=531/851)	---	cDMARD (n=132 /212)	DB	*NA (12-wk study)*	ACR20/50/70
AIM	ABA 10 mg + MTX (n=433)	---	MTX (n=219)	DB	ACR20/50/70	ACR20/50/70
AMPLE	ABA 125 mg + MTX (n=318)	ADA 40 mg + MTX (n=328)	---	SB	ACR20/50/70	ACR20/50/70
ATTEST	ABA 10 mg + MTX (n=156)	IFX 3 mg + MTX (n=165)	MTX (n=110)	DB	ACR20/50/70	*No outcomes reported*
ACT-RAY	TCZ 8 mg + MTX (n=279)	TCZ 8 mg (n=277)	---	OL & DB	ACR20/50/70; EULAR	ACR20/50/70

ADACTA	TCZ 8 mg (n=163)	ADA 40 mg (n=162)	---	DB	ACR20/50/70; EULAR	*No outcomes reported*
AMBITION[b]	TCZ 8 mg (N=286)	---	MTX (N=284)	DB	ACR20/50/70 EULAR	ACR20/50/70
BREVACTA[d]	TCZ 162 mg + cDMARD (n=348/437)	---	cDMARD (n=172 /219)	DB	ACR20/50/70	*No outcomes reported*
LITHE[b]	TCZ 8 mg + MTX (n=398)	---	MTX (n=393)	DB	ACR20/50/70	ACR20/50/70
Nishimoto (2004)	TCZ 8 mg (n=55)	---	Placebo (n=54)	DB	*NA (12-wk study)*	ACR20/50/70; EULAR
OPTION[b]	TCZ 8 mg + MTX (N=205)	---	MTX (N=204)	DB	ACR20/50/70; EULAR	ACR20/50/70
SATORI	TCZ 8 mg + MTX (n=61)	---	MTX (n=66)	DB	ACR20/50/70; EULAR	ACR20/50/70
SUMMACTA	TCZ 162 mg (n=631)	TCZ 8 mg (n=631)	---	DB	ACR20/50/70	ACR20/50/70
TOWARD[b]	TCZ 8 mg + cDMARD (n=805)	---	cDMARD (n=415)	DB	ACR20/50/70 EULAR	ACR20/50/70
ORAL SCAN[b]	TOFA 5 mg + MTX (n=321)	TOFA 10 mg + MTX (n=316)	MTX (n=160)	DB	ACR20/50/70	ACR 50/70
ORAL STANDARD[b]	TOFA 5 mg + MTX (n=204)	TOFA 10 mg + MTX (n=201) ---------- ADA 40 mg + MTX (n=204)	MTX (n=108)	DB	ACR20/50/70	ACR20/50/70
APPEAL	ETN 25 mg + MTX (n=197)	---	cDMARD + MTX (n=103)	OL	*NA (16-wk study)*	ACR20/50/70
Combe (2006)	ETN 25 mg + SSZ (n=101)	SSZ (n=50)	ETN 25 mg (n=103)	DB	ACR20/50/70	*[ACR20/50/70[e]*
JESMR	ETN 25 mg + MTX (n=77)	ETN 25 mg (n=74)	---	OL	ACR20/50/70; EULAR	*[EULAR][f]*
Lan (2004)	ETN 25 mg + MTX (n=29)	---	MTX (n=29)	DB	*NA (12-wk study)*	ACR20/50/70
Machado (2014)	ETN 50 mg + MTX (n=284)	---	cDMARD + MTX (n=145)	OL	ACR20/50/70; EULAR	ACR20/50/70
Moreland 1999/ Mathias 2000	ETN 25 mg (n=78)	---	Placebo (n=80)	DB	ACR20/50/70	ACR20/50/70
RACAT[b]	ETN 50 mg + MTX (n=175)	SSZ +HCQ + MTX (n=178)	---	DB	ACR20/50/70	*No outcomes reported*
TEMPO	ETN 25 mg + MTX (n=231)	ETN 25 mg (n=223)	MTX (n=228)	DB	ACR20/50/70	ACR20/50/70
Weinblatt (1999)	ETN 25 mg + MTX (n=59)	---	MTX (n=30)	DB	ACR20/50/70	ACR20/50/70
Edwards (2004)	RTX 1000 mg (n=40)	RTX 1000 mg + MTX (n=40)	MTX (n=40)	DB	ACR20/50/70; EULAR	*No outcomes reported*
RA-SCORE	RTX 1000 mg (n=63)	---	MTX (n=60)	DB	ACR20/50/70; EULAR	*No outcomes reported*
SERENE	RTX 1000 mg + MTX (n=168)	RTX 2000 mg + MTX (n=172)	MTX (n=172)	DB	ACR20/50/70; EULAR	*No outcomes reported*

RA-BEAM	BAR 4 mg + MTX (n=487)	ADA 40 mg + MTX (n=330)	MTX (n=488)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR
RA-BUILD	BAR 2 mg + cDMARD (n=229)	BAR 4 mg + cDMARD (n=227)	cDMARD (n=228)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR

Footnotes: Studies in green cells indicate allowance of prior bDMARD treatment up to 20%. [a]CNTO 148: 24-week results were excluded from the analysis due to switch of the placebo group to IFX at week 20. [b]Study includes prior bDMARD use up to 20%. [c]REALISTIC: only results from the subgroup of REALISTIC patients that were cDMARD-IR are used in the analysis. [d]BREVACTA: only results from the subgroup of BREVACTA patients that were cDMARD-IR are used in the analysis. [e]Data not analysed for consistency with the approach taken in the NICE MTA (TA375) [f]No ACR data were available for this time point, thus it was not possible to calculate the EULAR response.
Abbreviations: NMA = Network meta-analysis, ACR20/50/70 = 20/50/70% improvement in ACR disease activity index, EULAR = European League Against Rheumatism, ABTS = subcutaneous abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX = infliximab, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab, NA = not available, DB = double-blind, OL = open-label, cDMARD = conventional disease-modifying anti-rheumatoic drug, TOFA = tofacitinib

Table 70. Summary of trials included in the NMA for the anti-TNFi-IR population

Study name	Treatment 1	Treatment 2	Control	Study design	Endpoints 24 weeks	Endpoints 12 weeks
ATTAIN	ABA 10 mg + cMDARD (n=258)	---	cDMARD (n=133)	DB	ACR20/50/70	ACR20
REALISTIC[a]	CTZ + cDMARD (n=320/851)	---	cDMARD (n=80/212)	DB	NA (12-week study)	ACR20/50/70; *[EULAR]*[b]
GO-AFTER[c]	GOL 50 mg +/- cDMARD (n=153)	---	cDMARD (n=155)	DB	ACR20/50/70	ACR20/50/70
RADIATE	TCZ 8 mg + MTX (n=175)	---	MTX (n=160)	DB	ACR20/50/70	ACR20/50/70
BREVACTA[d]	TCZ 162 mg + cDMARD (n=89/437)	---	cDMARD (n=47/219)	DB	ACR20/50/70	No outcomes reported
ORAL STEP	TOFA 5 mg + MTX (n=133)	TOFA 10 mg + MTX (n=134)	MTX (n=132)	DB	*[ACR20/50/70]*[e]	ACR20/50/70
REFLEX	RTX 1000 mg + MTX (n=311)	---	MTX (n=209)	DB	ACR20/50/70; EULAR	ACR20/50/70
BEACON	BAR 2 mg + cDMARD (n=174)	BAR 4 mg + cDMARD (n=177)	cDMARD (n=176)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR

Footnotes: [a]REALISTIC: only results from the subgroup of REALISTIC patients that were anti-TNF-IR are used in the analysis. [c]GO-AFTER: approx. 30% of patients did not have concomitant cDMARD. [b]Insufficient ACR response data for the anti-TNF-IR subgroup were available from the REALISTIC study in order to perform the conversion to EULAR response. [d]BREVACTA: only results from the subgroup of BREVACTA patients that were anti-TNF-IR are used in the analysis. [e]ORAL STEP: results at Week 24 were excluded from the analysis due to a disconnect in the network.
Abbreviations: NMA = Network meta-analysis, ACR20/50/70 = 20/50/70% improvement in ACR disease activity index, EULAR = European League Against Rheumatism, ABTS = subcutaneous abatacept, BAR = baricitinib, CTZ = certolizumab pegol, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX = infliximab, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab, NA = not available, DB = double-blind, OL = open-label, cDMARD = conventional disease-modifying anti-rheumatoic drug, TOFA = tofacitinib

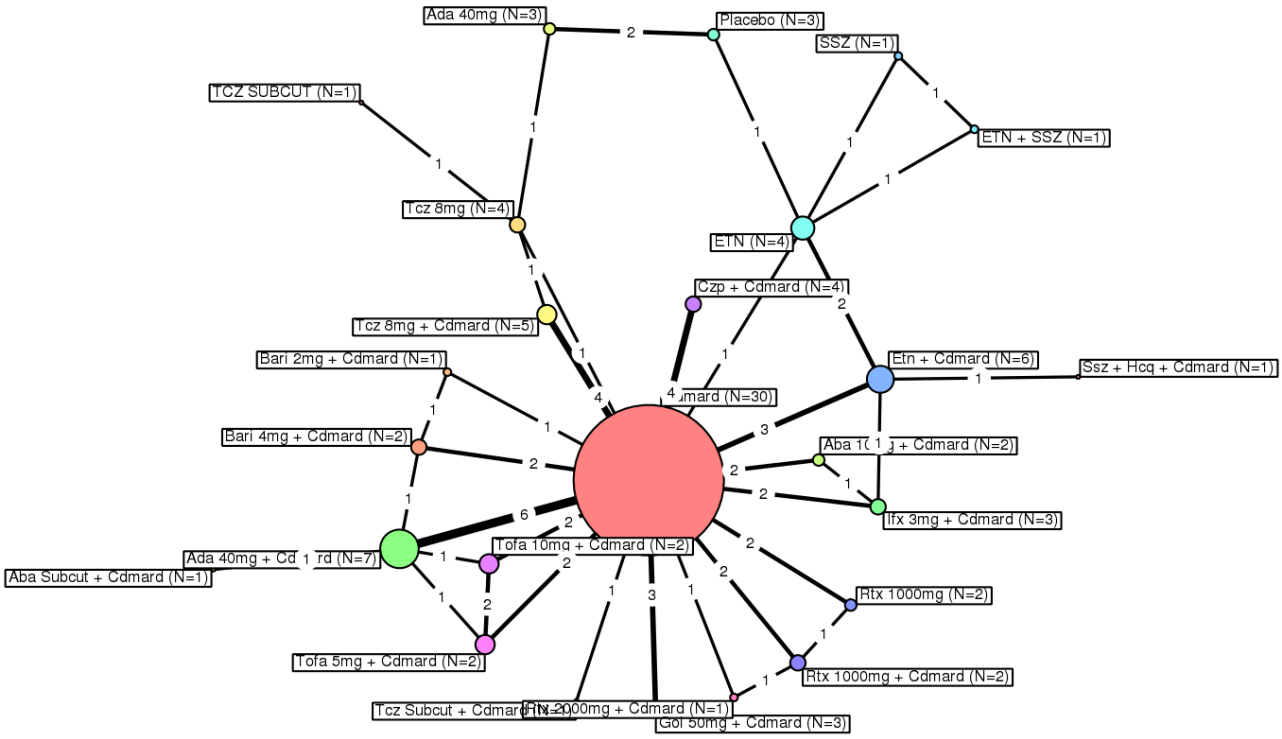
Network diagrams

The treatment networks for the RCTs included in the base case analyses (i.e. the Week 24 time point) for the cDMARD-IR and anti-TNF-IR populations are presented below. The network diagrams for the Week 12 time point are presented in Appendix 12.

cDMARD-IR population

The treatment network for ACR response at Week 24 in the cDMARD-IR population is presented in Figure 37.

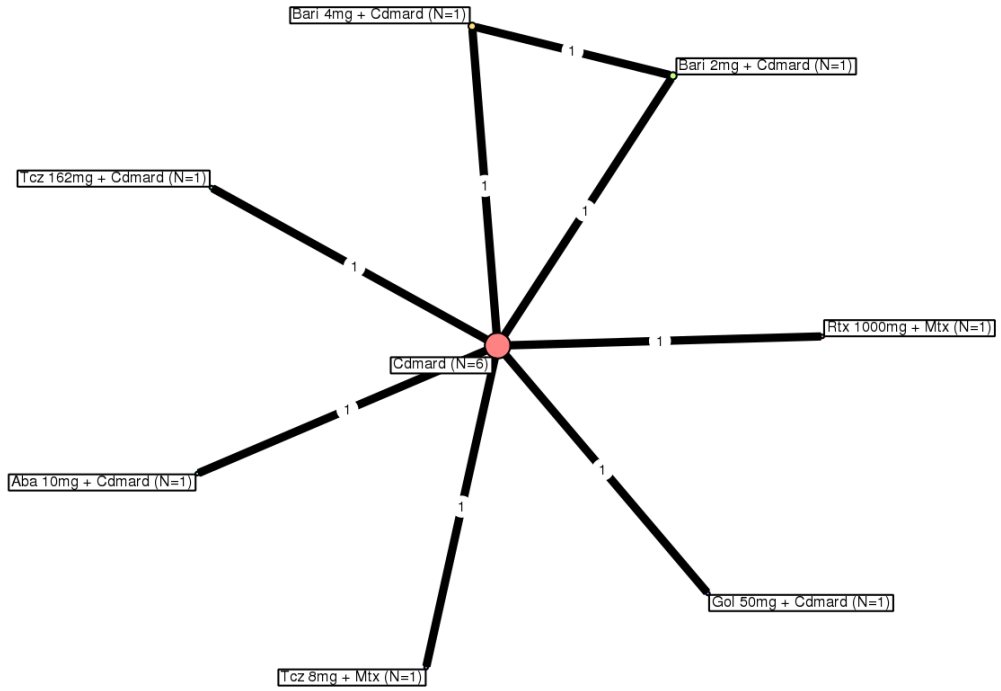
Figure 37. Network of studies contributing to ACR outcomes at Week 24 in the cDMARD-IR population



Anti-TNF-IR Population

The treatment network for ACR response at Week 24 in the anti-TNF-IR population is presented in Figure 38.

Figure 38. Network of studies contributing to ACR outcomes at Week 24 in the anti-TNF-IR population



4.10.4 Trials identified in the SLR that were not included in the NMA

Trials identified in the clinical SLR that were not included in the NMA analyses are listed in Appendix 13.

4.10.5 Methods and outcomes of included studies

Populations

Network meta-analyses were performed separately for the cDMARD-IR and anti-TNF-IR populations. However, similar to the NMA performed by the AG in TA 375,[85] separate analyses were not performed for patients with either moderate or severe disease activity. This was a result of insufficient data being available from the SLR for the moderate and severe patient subgroups.

The patient characteristics of the studies included in the cDMARD-IR and anti-TNF-IR analyses are presented in Table 71 and Table 72, respectively. These were generally similar across all studies. As noted in section 4.10.2, in order to allow the inclusion of certolizumab in the base-case network, relaxation of prior-biologic use criteria had to be applied. An analysis with strict no prior-biologic criteria was also conducted and presented as a sensitivity analysis.

The input ACR response data for each of the populations of interest can be found in Appendix 14. The input EULAR response data for each of the populations of interest can be provided upon request.

Table 71. Baseline Characteristics - cDMARD (including MTX)-Inadequate Response

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
Abe (2006) Japan 36 weeks	Patients had active disease despite treatment with MTX for more than 3 months. The MTX dosage must have been stable 6 mg per week or more during the last 4 weeks.	Active RA: ≥6 tender joints (of 68) and ≥6 swollen joints (of 66), plus ≥2 of the following: morning stiffness ≥45 minutes, ESR ≥28 mm per hour, or CRP ≥2 mg/dL.	IFX 3 mg/kg + MTX (n = 49)	No	Yes	Steinbrocker disease: class I = 5 (10.2%), II = 36 (73.5), and III = 8 (16.3)	18.40	55.20 (10.90)	NR	NR	NR	NR
			MTX QW (n = 47)	No	Yes	Steinbrocker disease: class I = 2 (4.3), II = 33 (70.2), and III = 12 (25.5)	25.50	55.10 (7.60)	NR	NR	NR	NR
ACT-RAY Dougados *et al.* (2014) Secondary report: Dougados *et al.* (2013) Multinational 52 weeks	Active RA despite MTX treatment (had been receiving MTX for ≥12 weeks, with a stable dose of ≥15 mg per week for ≥6 weeks). Exclusion criteria included previous use of biologics and any cDMARD treatment other than MTX during the preceding month.	Active moderate-to-severe RA: DAS28-ESR > 4.4 at baseline and ≥4.0 at screening; radiographic evidence of RA-related joint erosions.	TCZ 8 mg/kg + MTX (n = 279)	No	Yes	NR	18.10	53.00 (13.50)	55/83 (66.30)	NR	NR	DAS28-ESR = 6.33 (0.98)
			TCZ 8 mg/kg + PBO (n = 277)	No	Yes	NR	21.40	53.60 (11.90)	41/64 (64.10)	NR	NR	DAS28-ESR = 6.36 (1.00)

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
ADACTA Gabay *et al.* (2013) Secondary report: Gabay *et al.* (2013) Multinational (15 countries in North and South America, Europe, and Austral-Asia) 24 weeks	Patients had to be taking MTX or have done previously, be unable to tolerate MTX, or be inappropriate candidates for continued MTX treatment in the judgment of the investigator. No prior treatment with a bDMARD.	Severe RA for ≥ 6 months: DAS-28 > 5.1 at baseline, SJC (66 joints) ≥ 6 and TJC (68 joints) ≥ 8, and hs-CRP ≥ 1 mg/dL or ESR ≥ 28 mm per hour.	ADA 40 mg QOW (n = 163)	No	Yes	NR	17.90	53.30 (12.40)	119/162 (73.00)	43.10 (12.60)	45.60 (13.90)	DAS-28 ESR = 6.76 (0.90)
			TCZ 8 mg/kg (n = 163)	No	Yes	NR	20.90	54.40 (13.00)	122/163 (75.00)	40.80 (12.30)	43.40 (13.20)	DAS-28 ESR = 6.72 (0.90)
AIM Kremer *et al.* (2006) Secondary report: Russell *et al.* (2007) Worldwide (116 centres) 52 weeks	RA was persistent and active despite MTX treatment. All patients were treated with MTX (≥ 15 mg per week) for ≥ 3 months, with a stable dose for 28 days before enrolment. 72 patients had been treated with other DMARDs, 1 with biologics, 462 with CCS, and 551 with NSAIDs.	Active RA: ≥ 10 swollen joints, ≥ 12 tender joints, and CRP levels ≥ 10.0 mg/L while receiving MTX.	ABA 10 mg/kg + MTX (n = 433)	Nod	Mixed (only 1 patient)	NR	22.20	51.50 (12.90)	354/433 (81.80)	NR	NR	DAS-28 CRP = 6.40 (0.80)
			MTX 15 mg QW (n = 219)	Nod	Yes	NR	18.30	50.40 (12.40)	172/219 (78.50)	NR	NR	DAS-28 CRP = 6.40 (0.11)
AMPLE Schiff *et al.* (2014) Secondary reports: Weinblatt *et al.*	Patients had an inadequate response to MTX, and had not received previous bDMARD therapy.	Moderate-to-severe, active disease: DAS28-CRP ≥3.2, and either a history of seropositivity for	ABA 125 mg + MTX (n = 318)	No	Yes	NR	18.60	51.40 (12.60)	240/318 (75.50)	NR	NR	DAS28-CRP = 5.50 (1.10)
			ADA 40 mg QOW +	No	Yes	NR	17.70	51.00 (12.80)	254/328 (77.40)	NR	NR	DAS28-CRP

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
(2013a); Fleischmann *et al.* (2014b); Weinblatt *et al.* (2014); Fleischmann *et al.* (2013b); Maldonado *et al.* (2013); Fleischmann *et al.* (2013d); Weinblatt *et al.* (2013b) North and South America 104 weeks		anti-CCP antibodies or RF, and/or an elevated ESR or CRP level.	MTX (n = 328)									= 5.50 (1.10)
APPEAL Kim *et al.* (2012) Secondary report: Bae *et al.* (2013) Asia Pacific (Hong Kong, India, Korea, Malaysia, Philippines, Taiwan, and Thailand) 16 weeks	In the opinion of the investigator, the patients were currently receiving an adequate dose of oral MTX 1 day per week (≥7.5 mg per week but ≤25 mg per week) at a stable dose for a minimum of 3 months. No previous or current treatment with ETN, other TNF inhibitors or other biologic agents, or concurrent treatment with a DMARD other than MTX within 3 months prior to screening.	Active, moderate-to-severe RA: DAS28 ≥3.2 with either ESR ≥28 mm per hour or morning stiffness ≥45 minutes; and class I, II, or III functional status as defined by ACR criteria.	ETN 25 mg BIW + MTX (n = 197)	No	Yes	NR	8.60	48.40 (12.00)	NR	NR	NR	DAS28-ESR = 6.10 (1.10) DAS28-CRP = 5.23 (1.10)
			DMARD + MTX (n = 103)	No	Yes	NR	11.70	48.50 (11.30)	NR	NR	NR	DAS28-ESR = 6.10 (1.10) DAS28-CRP = 5.34 (1.10)
ARMADA Weinblatt *et al.*	Participants were treated with MTX for ≥6 months and must have been	Active disease: presence of ≥9 tender joints (of	ADA 40 mg QOW + MTX (n = 67)	No	Yes	NR	25.40	57.20 (11.40)	NR	NR	NR	NR

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
(2003) US and Canada 24 weeks	taking a stable weekly dose (12.5-25 mg, or 10 mg if intolerant to higher doses) for ≥4 weeks. All participants had failed treatment with ≥1 DMARD besides MTX, but ≤4 DMARDs. No prior treatment with anti-CD4 therapy or TNF antagonists.	68) and 6 swollen joints (of 66).	MTX QW (n = 62)	No	Yes	NR	17.70	56.00 (10.8)0	NR	NR	NR	NR
ATTEST Schiff *et al.* (2008), Multinational 52 weeks	Inadequate response to MTX, as demonstrated by ongoing active disease. All had received MTX ≥15 mg per week for ≥3 months prior and washed out all DMARDs except for MTX. No prior experience of ABA or anti-TNF therapy. Most patients were receiving NSAIDs and/or CCS.	Active RA: at randomisation ≥10 swollen joints, ≥ 12 tender joints, and CRP levels ≥1 mg/dL using a high-sensitivity assay.	ABA 10 mg/kg + MTX (n = 156)[a]	No	Yes	NR	16.70	49.00 (12.50)	136/156 (87.20)	NR	NR	DAS28-ESR = 6.90 (1.00)
			IFX 3 mg/kg + MTX (n = 165)[a]	No	Yes	NR	17.60	49.10 (12.00)	140/165 (84.80)	NR	NR	DAS28-ESR = 6.80 (0.90)
			MTX QW (n = 110)[a]	No	Yes	NR	12.70	49.40 (11.50)	85/110 (77.30)	NR	NR	DAS28-ESR = 6.80 (1.00)
ATTRACT Maini *et al.* (1999) Secondary report: Lipsky *et al.* (2000) North America and Europe 54 weeks	Active disease despite treatment with MTX (MTX for ≥3 months; stable dose at ≥12·5 mg per week, for ≥4 weeks). Over 70% were receiving NSAIDs and >60% were receiving CCS. Mean number (SD) of previous DMARDs (excluding MTX): IFX = 2.8 (1.5); PBO = 2.5 (1.4).	Active RA: ≥6 swollen and tender joints plus 2 of: morning stiffness ≥45 minutes, ESR >28 mm per hour, CRP >2 mg/dL.	IFX 3 mg/kg	No	Yes	NR	19.00	54.00 (11.00)	72/86 (84.00)	NR	NR	NR
			PBO + MTX QW	No	Yes	NR	20.00	51.00 (12.00)	67/88 (77.00)	NR	NR	NR
BREVACTA	Inadequate response to	Moderate-to-	TCZ	No	Mixed[b]	NR	14.20	52.10 (11.45)	349/432	NR	NR	DAS28-

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
Kivitz *et al.* (2014) Worldwide 24 weeks	≥1 DMARD that, in up to 20% of patients, could include ≥1 anti-TNF agent. Patients had received ≥1 traditional DMARDs at a stable dose for ≥8 weeks. Previous DMARDs, mean (SD): TCZ = 1.3 (0.7); PBO = 1.4 (0.8). Previously failed anti-TNF treatment: TCZ = 20.4%; PBO = 21.5%.	severe: SJC ≥6 (66-joint count), TJC ≥8 (68-joint count), radiographic evidence of ≥1 joints with a definite erosion attributable to RA, and a CRP level ≥10 mg/L and/or ESR ≥28 mm per hour.	162 mg Q2W (n = 437)						(80.80)			ESR = 6.70 (0.92)
			PBO (n = 219)	No	Mixed[b]	NR	17.40	52.00 (11.71)	178/218 (81.70)	NR	NR	DAS28-ESR = 6.60 (0.94)
CHANGE Miyasaka *et al.* (2008) Japan 24 weeks	Patients had failed treatment with ≥ 1 prior DMARD. The most common previous DMARDs were MTX (87.2%), SSZ (73.9%), and bucillamine (67.6%).	ACR criteria for active RA, had ≥ 10 swollen joints and ≥ 12 tender joints (excluding distal interphalangeal joints).	ADA 40 mg QOW (n = 91)	No	Yes	NR	20.90	56.90 (10.30)	81/91 (89.00)	NR	NR	NR
			PBO (n = 87)	No	Yes	NR	23.00	53.40 (12.80)	75/87 (86.20)	NR	NR	NR
CNTO 148 Kay *et al.* (2008) NR 52 weeks	Persistent disease activity despite receiving MTX at a stable dosage of ≥10 mg per week. Patients had to have been treated with MTX for ≥ 3 months.	Active RA defined by the ACR 1987 revised criteria. Persistent disease activity was defined as ≥6 swollen joints and ≥6 tender joints and ≥2 of the following criteria: CRP level ≥1.5 mg/dL, ESR ≥28 mm in first	GOL 50 mg Q4W + MTX (n = 35)	No	Yes	NR	14.30	Median = 57.00 (7.41)	NR	NR	NR	DAS28-ESR median = 6.40 (1.26) DAS28-CRP median = 5.30 (1.26)
			MTX ≥10 mg QW (n = 35)	No	Yes	NR	25.70	Median = 52.00 (14.81)	NR	NR	NR	DAS28-ESR median = 6.40 (1.26)

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
		hour, and morning stiffness ≥30 minutes.										DAS28-CRP median = 5.30 (0.89)
Combe (2006) Secondary report: Combe *et al.* (2009) NR 104 weeks	Patients had received stable doses of SSZ (2-3 g daily) for ≥4 months and demonstrated an inadequate response. Patients had not received ETN or other TNF antagonists, and they must not have received a DMARD other than SSZ within 3 months before baseline.	Active RA: ≥ 6 swollen and ≥ 10 tender joints and ≥ 1 of the following: ESR ≥ 28 mm at the end of first hour; serum CRP ≥ 20 mg/L and morning stiffness for ≥ 45 minutes.	ETN 25 mg BIW (n = 103)	No	Yes	NR	21.40	51.30 (13.50)	NR	NR	NR	DAS44-ESR = 5.10 (1.10)
			ETN 25 mg BIW + SSZ 2-3 g QD (n = 101)	No	Yes	NR	19.80	50.60 (12.30)	NR	NR	NR	DAS44-ESR = 5.20 (1.20)
			SSZ 2/2.5/3 g QD (n = 50)	No	Yes	NR	18.00	53.30 (12.80)	NR	NR	NR	DAS44-ESR = 5.00 (1.10)
De Filippis (2006) Italy 54 weeks	Patients not responding to DMARDs for >6 months, including a stable dose of MTX (between 10-15 mg per week) in the 3 months before entering the study.	Active disease: presence of >3 swollen joints and 3 of the following: ESR >28 mm per hour, CRP >1.9 mg/dL, morning stiffness >45 minutes, >5 swollen joints, and >10 tender joints.	ETN 25 mg BIW + MTX (n = 16)	No	Yes	NR	NR	44.70 (14.17)	NR	NR	NR	NR
			IFX 3 mg/kg + MTX (n = 16)		Yes	NR	NR	46.79 (10.90)	NR	NR	NR	NR
Edwards (2004) Secondary report: Strand *et*	Active disease despite treatment with ≥10 mg of MTX per week.	Active disease: ≥8 swollen and 8 tender joints and ≥2 of the	RTX 1000 mg (n = 40)	No	Yes	NR	27.00	54.00 (10.00)	40/40 (100)	NR	NR	DAS28-ESR = 6.80 (1.00)

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
al. (2006) Australia, Canada, Israel, and 8 European countries 48 weeks		following: a serum CRP level ≥15 mg/L; an ESR ≥ 28 mm per hour, or morning stiffness >45 minutes.	RTX 1000 mg + MTX (n = 40)	No	Yes	NR	25.00	54.00 (12.00)	40/40 (100)	NR	NR	DAS28-ESR = 6.80 (0.90)
			MTX ≥ 10 mg QW (n = 40)	No	Yes	NR	20.00	54.00 (11.00)	40/40 (100)	NR	NR	DAS28-ESR = 6.90 (0.70)
GO-FORTH Tanaka *et al.* (2012) Secondary report: Tanaka *et al.* (2016) Japan 156 weeks	MTX inadequate responders. Patients had received ≥ 6 mg per week oral MTX for RA for ≥3 months. Patients had not received anti-TNF biologic therapy, alkylating agents (cyclophosphamide), or any investigational agents within the previous 4 months.	Active RA: ≥4/66 swollen joints and ≥4/68 tender joints and met ≥2 of the following: CRP >1.5 mg/dL or ESR >28 mm per hour; morning stiffness ≥30 minutes; radiographic evidence of bone erosion; or anti-CCP antibody positive or RF positive.	GOL 50 mg Q4W + MTX (n = 89)	No	Yes	NR	15.10	50.40 (9.90)	NR	NR	NR	DAS28-ESR = 5.50 (1.18)
			MTX 6-8 mg QW (n = 90)	No	Yes	NR	17.00	51.10 (11.60)	NR	NR	NR	DAS28-ESR = 5.60 (0.99)
GO-FORWARD Keystone *et al.* (2009b) Secondary reports: Genovese *et al.* (2012); Keystone *et al.*	MTX inadequate responders; had been on a stable MTX dose of ≥15 mg per week but ≤25 mg per week during the 4-week period immediately preceding screening. Patients had tolerated ≥15 mg per	Active RA: ≥4 swollen joints (out of 66) and ≥4 tender joints (out of 68) and ≥2 of the following: CRP ≥1.5 mg/dL or ESR ≥28 mm	GOL 50 mg Q4W + MTX (n = 89)	No	Yes	NR	19.10	Median = 52.00 (10.37)	77/89 (86.50)	NR	NR	Median DAS28-ESR = 6.11 (1.17) Median DAS28-CRP = 5.77

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
(2010); Keystone *et al.* (2016) Argentina, Australia, Canada, Chile, Germany, Hungary, Mexico, New Zealand, Poland, South Korea, Taiwan and the US 52 weeks	week MTX for ≥3 months. Patients had not used any anti-TNF agent, RTX, natalizumab or cytotoxic agents. Over 70% had received cDMARDs other than MTX.	per hour; ≥30 minutes of morning stiffness; bone erosion determined by x-ray and/or MRI; or anti-CCP or RF positive.										(1.17)
			MTX 10-25 mg QW (n = 133)	No	Yes	NR	18.00	Median = 52.00 (11.85)	108/133 (81.20)	NR	NR	Median DAS28-ESR = 6.11 (0.97) Median DAS28-CRP = 5.46 (1.05)
J-RAPID Yamamoto *et al.* (2014) Japan 24 weeks	Inadequate response to MTX. 13.4% in the CTZ arm and 19.5% in the MTX arm had used anti-TNFs. Patients had not received any biologic therapy for RA within the 6 preceding months. Patients who had received previous treatment with ≥ 2 TNF inhibitors or who had not initially responded to previous TNF-inhibitor therapy were excluded.	Active RA: ≥ 9 tender and 9 swollen joints (among 68 and 66 joints of ACR definition, respectively), and satisfied ≥ 1 of the following: ESR ≥ 30 mm per hour or CRP ≥ 1.5 mg/dL.	CTZ 400-200 mg QOW + MTX (n = 82)	No	Mixed[d]	NR	15.90	50.60 (11.40)	71/82 (86.60)	NR	NR	DAS-28 ESR = 6.20 (0.80)
			MTX 6-8 mg QW (n = 77)	No	Mixed[d]	NR	14.30	51.90 (11.10)	66/77 (85.70)	NR	NR	DAS-28 ESR = 6.50 (0.90)
JESMR Kameda *et al.* (2010) Secondary report: Kameda *et al.* (2011)	Patients were receiving ≥6 mg per week of MTX for a minimum of 3 months, and were dose stable for ≥4 weeks. Patients were not undergoing antirheumatic therapy other than MTX	≥6 tender joints and 2 swollen joints, either a serum CRP level >2 mg/dL or ESR ≥ 28 mm at 1 hour.	ETN 25 mg BIW + MTX (n = 77)	No	Yes	NR	20.00	56.50 (11.10)	65/75 (86.70)	NR	NR	DAS28 -ESR = 6.00 (1.00)
			ETN 25 mg BIW (n = 74)	No	Yes	NR	12.70	58.10 (12.60)	65/71 (91.50)	NR	NR	DAS28-ESR = 6.10

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
Japan 104 weeks	and PSL equivalents, and had not received ETN or other biologic treatments.											(0.90)
Kang (2013) Korea (abstract only) 24 weeks	Patients had an inadequate response to MTX.	Active RA ≥ 9 tender or swollen joints, CRP > 15 mg/L or ESR ≥ 30 mm per hour. Mean DAS-28 = 7.40	MTX 10-20 mg QW (n = 40)	No	Unclear	NR	NR	NR	NR	NR	NR	NR
			CTZ 400-200 mg QOW + MTX (n = 81)	No	Unclear	NR	NR	NR	NR	NR	NR	NR
Keystone (2004) US and Canada 52 weeks	Patients had active RA despite being on MTX therapy for ≥3 months at a stable dose of 12.5-25 mg per week (or ≥ 10 mg per week in patients intolerant to MTX) for ≥4 weeks. Exclusion criteria included prior use of anti-CD4 antibody therapy or TNF antagonists.	Established, moderate-to-severe, active RA according to the 1987 revised ACR criteria.	ADA 40 mg QOW + MTX (n = 207)	No	Unclear	NR	23.70	56.10 (13.50)	169/207 (81.60)	NR	NR	NR
			MTX QW (n = 200)	No	Unclear	NR	27.00	56.10 (12.00)	179/200 (89.50)	NR	NR	NR
Kim (2007) Korea 24 weeks	Patients had insufficient response to MTX after treatment for ≥6 months with a stable dosage ≥4 weeks prior to screening. Patients received ≥1 prior DMARD other than MTX but could have had efficacy failures to ≤4 standard DMARDs other than MTX.	Patients met ACR criteria for active RA, and had ≥6 swollen joints and ≥9 tender joints at baseline.	PBO (n = 63)	No	Unclear	TB (skin test), abnormal chest radiograph, calcified granulomas, pleural scarring	14.30	49.80 (10.50)	52/63 (82.50)	NR	NR	NR
			ADA 40 mg QOW (n = 65)	No	Unclear	TB (skin test), abnormal chest radiograph, calcified granulomas,	4.60	48.50 (10.20)	50/65 (76.90)	NR	NR	NR

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
						pleural scarring						
Lan (2004) Taiwan 12 weeks	Patients had been receiving MTX at a stable dose of 12.5-20 mg per week.	Active RA defined as >6 swollen and tender joints for ≥3 months.	ETN 25 mg BIW (n = 29)	No	Unclear	NR	17.00	47.55 (NR)	NR	NR	NR	NR
			MTX 12.5-20 mg QW (n = 29)	No	Unclear	NR	10.00	50.79 (NR)	NR	NR	NR	NR
Li (2013) Secondary report: Li *et al.* (2015) China 56 weeks	Active RA despite MTX therapy. Patients had received & tolerated MTX 7.5-20mg/week for ≥ 4 weeks	≥ 4/66 swollen joints and ≥ 4/68 tender joints, CRP≥ 15 mg/L, ESR≥ 28 mm/h, and CCP and RF positive	PBO Q4W + MTX 7.5-20 mg QW (n = 132)	No	Yes	NR	21.20	46.70 (12.20)	122/132 (92.4)	NR	NR	DAS-28 CRP = 5.50 (1.10)
			GOL 50 mg Q4W + MTX 7.5-20 mg QW (n = 132)	No	Yes	NR	16.70	47.70 (11.50)	115/132 (87.1)	NR	NR	DAS-28 CRP = 5.40 (1.10)
Machado (2014) Secondary report: Machado *et al.* (2016) Latin America 24 weeks	Patients had active RA despite MTX monotherapy (≥7.5 and ≤25 mg per week) for ≥3 months. Exclusion criteria included previous treatment with ETN or other biologic agents. The majority (>73%) had been treated with CCS and NSAIDs.	Established, moderately to severely active RA indicated by ≥6 swollen joints, ≥8 tender joints, and ESR ≥28 mm per hour.	ETN 50 mg QW + MTX (n = 284)	No	Yes	NR	11.74	48.40 (12.00)	242/281 (86.10)	NR	NR	DAS28-ESR = 6.60 (0.70)
			DMARD + MTX (n = 145)	No	Yes	NR	9.86	48.60 (11.30)	119/142 (83.80)	NR	NR	DAS28-ESR = 6.70 (0.70)
Moreland (1999) North America 26 weeks	Patients were required to have had an inadequate response to 1-4 DMARDs (such as AZA, MTX, SSZ, penicillamine, HCQ, or oral or injectable gold); an inadequate response was defined as	Established, active RA defined as ≥ 12 tender joints, ≥ 10 swollen joints, and at least 1 of: ESR ≥ 28 mm per	PBO (n = 80)	No	Unclear	NR	24.00	51.00 (NR)	63/80 (79.00)	NR	NR	NR
			ETN 25 mg BIW (n = 78)	No	Unclear	NR	26.00	53.00 (NR)	62/78 (79.00)	NR	NR	NR

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
	discontinuation of therapy because of lack of effect.	hour, CRP > 20 mg/L, or morning stiffness ≥ 45 minutes.										
Nishimoto (2004) UK and Japan 12 weeks	All patients had been treated unsuccessfully (due to lack of efficacy) with at least 1 DMARD or immune-suppressant.	Established, active RA defined in terms of high counts of swollen and tender joints and increased ESRs and CRP levels.	TCZ 8 mg/kg (n = 55)	No	Unclear	NR	16.36	56.00 (NR)	NR	NR	NR	NR
			PBO (n = 54)	No	Unclear	NR	26.42	53.00 (NR)	NR	NR	NR	NR
ORAL SCAN Van Der Heijde *et al.* (2013) Multinational 104 weeks	Stable doses of MTX were required (15-25 mg weekly for ≥ 6 weeks; stable doses < 15 mg were allowed only if there were safety issues at higher doses). Established, active RA. Active disease defined by ≥ 6 tender or painful joints (68-joint count) and ≥ 6 swollen joints (66-joint count) and by an ESR of > 28 mm per hour or a CRP > 7 mg/L.		TOFA 5 mg BID (n = 321)	No	Mixed	NR	16.20	53.70 (11.60)	241/321 (75.20)	NR	NR	DAS-28 ESR = 6.34 (NR) DAS-28 CRP = 5.22 (NR)
			TOFA 10 mg BID (n = 316)	No	Mixed	NR	13.60	52.00 (11.40)	245/316 (77.60)	NR	NR	DAS-28 ESR = 6.25 (NR) DAS-28 CRP = 5.20 (NR)
			PBO to TOFA 5 mg BID (n = 81)	No	Mixed	NR	19.80	53.20 (11.50)	65/81 (79.70)	NR	NR	DAS-28 ESR = 6.25 (NR) DAS-28 CRP

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
												= 5.14 (NR)
			PBO to TOFA 10 mg BID (n = 79)	No	Mixed	NR	8.90	52.10 (11.80)	59/79 (75.30)	NR	NR	DAS-28 ESR = 6.29 (NR) DAS-28 CRP = 5.18 (NR)
ORAL STANDARD van Vollenhoven *et al.* (2012) Secondary report: van Vollenhoven *et al.* (2013); Strand *et al.* (2016) Worldwide 52 weeks	Patients were receiving 7.5-25 mg of MTX weekly and had an incomplete response (defined as sufficient residual disease activity to meet entry criteria). Exclusion criteria included prior treatment with ADA and lack of response to prior anti-TNF biologic treatment.	Established, active RA defined as the presence of ≥ 6 tender or painful joints (of 68 joints) and ≥ 6 swollen joints (of 66 joints) and either an ESR > 28 mm per hour or a CRP level > 7 mg/L.	TOFA 5 mg BID + MTX (n = 204)	No	Mixed	NR	14.70	53.00 (11.90)	136/204 (66.80)	NR	NR	DAS-28 ESR = 6.60 (NR) DAS-28 CRP = 5.40 (NR)
			TOFA 10 mg BID + MTX (n = 201)	No	Mixed	NR	16.40	52.90 (11.80)	133/ 201 (66.20)	NR	NR	DAS-28 ESR = 6.50 (NR) DAS-28 CRP = 5.40 (NR)
			ADA 40 mg QOW + MTX (n = 204)	No	Mixed	NR	20.60	52.50 (11.70)	139/204 (68.20)	NR	NR	DAS-28 ESR = 6.40 (NR) DAS-28 CRP = 5.30 (NR)
			MTX	No	Mixed	NR	23.20	55.50 (13.70)	40/56	NR	NR	DAS-28

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
			7.5-25 mg QW for 12 or 24 weeks then TOFA 5 mg BID + MTX (n = 56)						(71.40)			ESR = 6.60 (NR) DAS-28 CRP = 5.60 (NR)
			MTX 7.5-25 mg QW for 12 or 24 weeks then TOFA 10 mg BID + MTX (n = 52)	No	Mixed	NR	25.00	51.90 (13.70)	32/52 (60.80)	NR	NR	DAS-28 ESR = 6.30 (NR) DAS-28 CRP = 5.30 (NR)
RACAT O'Dell *et al.* (2013) Secondary report: Chang *et al.* (2014) NR 48 weeks	Patients had active disease despite MTX therapy at stable doses of 15-25 mg weekly for ≥ 12 weeks. 48% were taking oral GCS at baseline.	Established, active RA with DAS-28 ≥ 4.4	SSZ 1-2 g QD + HCQ 400 mg QD + MTX (n = 178)	No	Unclear	NR	56.70	57.80 (13.00)	117/178 (65.70)	36.00 (11.50)	NR	DAS-28 = 5.80 (0.90)
			ETN 50 mg QW + MTX (n = 175)	No	Unclear	NR	51.40	56.00 (13.20)	117/175 (66.90)	36.40 (11.20)	NR	DAS-28 = 5.90 (0.90)
			BAR 4 mg QD (n = 227)	No	Yes	NR	NR	NR	NR	NR	NR	NR
			PBO (n = 228)	No	Yes	NR	NR	NR	NR	NR	NR	NR
RAPID 1 Keystone *et al.* (2008) Secondary report: Strand *et al.* (2009)	Patients had active RA despite treatment with MTX for ≥ 6 months, with a stable dosage of ≥ 10 mg per week for ≥ 2 months prior to baseline. Exclusion criteria	Established, active RA defined as ≥ 9 tender and 9 swollen joints at screening and at baseline, with	MTX 10 mg QW (n = 199)	No	Unclear	NR	16.10	52.20 (11.20)	165/199 (82.80)	NR	NR	Median DAS-28 ESR = 7.00 (NR)
			CTZ 400-200 mg	No	Unclear	NR	17.60	51.40 (11.60)	313/393 (79.60)	NR	NR	Median DAS-28

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
Worldwide 52 weeks	included any previous biologic therapy that resulted in a severe hypersensitivity or anaphylactic reaction and patients who had previously failed to respond to treatment with an anti-TNF agent.	either an ESR Westergren ≥ 30 mm per hour or a CRP level > 15 mg/L.	QOW + MTX (n = 393)									ESR = 6.90 (NR)
RAPID-2 Smolen *et al.* (2009) Secondary report: Strand *et al.* (2011) Multinational 24 weeks	Patients had received prior MTX for ≥ 6 months (stable dose ≥ 10 mg per week for ≥ 2 months before baseline). Patients were excluded if they had received previous treatment with a biologic agent resulting in a severe hypersensitivity or anaphylactic reaction, or had not initially responded to previous anti-TNF therapy. 59% were taking steroids at baseline.	Established, active RA defined by: 9 tender joints, 9 swollen joints, fulfilment of 1 of the following: ≥ 30 mm per hour ESR (Westergren), or CRP > 15 mg/L.	MTX QW (n = 127)	No	Mixed[d]	NR	15.70	51.50 (11.80)	97/127 (78.20)	NR	NR	DAS-28 ESR = 6.83 (0.87)
			CTZ 400-200 mg QOW + MTX (n = 246)	No	Mixed[d]	NR	16.30	52.20 (11.10)	186/246 (77.50)	NR	NR	DAS-28 ESR = 6.85 (0.84)
RA-SCORE Peterfy *et al.* (2016) (abstract only) Worldwide 52 weeks	Patients had prior cDMARD & MTX exposure. Patients had an inadequate response to MTX. No previous treatment with biologics.	Active disease defined by DAS28CRP≥ 3.2	RTX 500 mg + MTX (n = 62)	No	Yes	NR	27.40	48.70 (11.10)	NR	NR	NR	DAS-28 ESR = 6.30 (1.20) DAS-28 CRP = 5.60 (1.10)

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
			RTX 1000 mg + MTX (n = 60)	No	Yes	NR	16.70	50.70 (11.70)	NR	NR	NR	DAS-28 ESR = 6.00 (1.10) DAS-28 CRP = 5.30 (1.00)
			PBO + MTX (n = 63)	No	Yes	NR	23.80	50.30 (11.90)	NR	NR	NR	DAS-28 ESR = 6.30 (1.10) DAS-28 CRP = 5.60 (1.10)
REALISTIC Weinblatt *et al.* (2012) Secondary report: Pope *et al.* (2015) US and Canada 12 weeks	Patients showed an unsatisfactory response or intolerance to DMARD (MTX, LEF, SSZ, chloroquine or HCQ, AZA, and/or gold). Patients were excluded who received treatment either with > 2 TNF inhibitors, RTX or ABA. 21% had discontinued previous anti-TNF inhibitors for efficacy reasons and 16% for non-efficacy reasons.	Active RA defined by ≥5 tender and ≥4 swollen joints (28-joint count) and either ≥10 mg/L CRP or ≥28 mm per hour ESR (Westergren method) at screening.	CTZ 400-200 mg QOW (n = 851)	No	Unclear[b]	NR	22.40	55.40 (12.40)	555/851 (73.90)	NR	NR	DAS28-ESR = 6.40 (0.90) DAS28-CRP = 5.70 (0.90)
			MTX QW (n = 212)	No	Unclear[b]	NR	20.30	53.90 (12.70)	137/212 (76.50)	NR	NR	DAS28-ESR = 6.40 (0.90) DAS28-CRP = 5.70 (0.90)
SATORI Nishimoto *et al.*	All patients had an inadequate response to MTX (8 mg per week for	Active RA defined as ≥ 6 tender joints (of	MTX 8 mg QW (n = 66)	No	Unclear	NR	25.00	50.80 (12.20)	NR	NR	NR	DAS28c = 6.20 (0.90)

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
(2009) Japan 24 weeks	≥8 weeks until enrolment). Inadequate response was defined as the presence of active disease.	49), ≥ 6 swollen joints (of 46), ESR ≥ 30 mm per hour, or CRP of ≥ 10 mg/L.	TCZ 8 mg/kg + MTX (n = 61)	No	Unclear	NR	9.84	52.60 (10.60)	NR	NR	NR	DAS28 c = 6.10 (0.90)
SERENE Emery *et al.* (2010) Secondary report: Emery *et al.* (2010) Worldwide 48 weeks	Active RA despite MTX (10-25 mg per week for ≥12 weeks). Patients had not previously received biologic treatment for RA. 45% of patients were receiving oral steroids at baseline.	Active disease was defined as SJC and TJC both ≥8, and either CRP ≥0.6 mg/dL or ESR ≥28 mm per hour.	MTX 10-25 mg QW (n = 172)	No	Yes	NR	14.50	52.16 (12.39)	129/172 (75.00)	NR	NR	DAS28-ESR = 6.54 (1.02) DAS28-CRP = 5.95 (0.97)
			RTX 1000 mg + MTX (n = 68)	No	Yes	NR	20.40	51.91 (12.93)	126/167 (75.40)	NR	NR	DAS28-ESR = 6.40 (0.95) DAS28-CRP = 5.81 (0.91)
			RTX 2000 mg+ MTX (n = 170)	No	Yes	NR	18.80	51.30 (12.64)	125/170 (73.50)	NR	NR	DAS28 -ESR = 6.49 (1.06) DAS28 -CRP = 5.86 (0.97)
STAR Furst *et al.* (2003) US and Canada 24 weeks	Patients did not respond to standard antirheumatic therapy including traditional DMARDs, low-dose CCS, NSAIDs, and/or analgesics.	Active RA defined by ≥6 swollen joints and ≥9 tender joints (excluding distal	ADA 40 mg QOW + DMARD (n = 318)	No	Yes	NR	20.40	55.00 (12.80)	201/318 (63.40)	NR	NR	NR
			DMARD	No	Yes	NR	20.80	55.80 (12.40)	198/318	NR	NR	NR

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
	Exclusion criteria included patients treated with anti-CD4 therapy or biologic DMARDs (eg, TNF antagonists, interleukin-1 receptor antagonists).	interphalangeal joints).	(n = 318)						(62.30)			
START Westhovens *et al.* (2006) Secondary report: West-hovens (2007) NR 54 weeks	Active disease despite receiving MTX; patients may or may not have been treated with other concomitant DMARDs. All had been receiving MTX for ≥3 months prior to randomisation (stable dose for ≥4 weeks). 59% were taking oral CCS, 41% NSAIDs, and 6.2% narcotics/ opioid analgesics at baseline.	Active RA was defined as the presence of 6 swollen joints and 6 tender joints.	MTX ≤25 mg QW (n = 363)	No	Unclear	Chronic renal infection, chronic sinusitis, osteomyelitis, chronic chest infection with bronchiectasis, diabetes, and chronic renal failure	16.80	Median = 52.00 (28.41)	284/363 (80.70)	NR	NR	NR
			IFX 3 mg/kg + MTX (n = 360)	No	Unclear	Chronic renal infection, chronic sinusitis, osteomyelitis, chronic chest infection with bronchiectasis, diabetes, and chronic renal failure	20.00	Median = 53.00 (27.67)	293/360 (82.80)	NR	NR	NR
TEMPO Klareskog *et al.* (2004) Secondary reports: van der Heijde *et al.* (2007); van der	Patients had a less than satisfactory response at the discretion of the investigator to ≥ 1 DMARD other than MTX. Individuals previously treated with MTX could	Established, active RA: ≥ 10 swollen and ≥ 12 painful joints and ≥ 1 of: ESR ≥ 28 mm per hour;	ETN 25 mg BIW (n = 223)[a]	No	Unclear	NR	23.00	53.20 (13.80)	167/223 (75.00)	NR	NR	DAS44 -ESR = 5.70 (1.10)
			MTX 7.5-20 mg QD (n = 228)[a]	No	Unclear	NR	21.00	53.00 (12.80)	163/228 (71.00)	NR	NR	DAS44 -ESR = 5.50

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
Heijde *et al.* (2005); van der Heijde *et al.* (2006) Worldwide 52 weeks	be enrolled provided they had not had clinically important toxic effects or lack of response, at the discretion of the investigator. Patients were ineligible if they had previously received ETN or other TNF antagonists.	plasma CRP ≥ 20 mg/L; or morning stiffness for 45 minutes or more.										(1.20)
			ETN 25 mg BIW + MTX (n = 231)[a]	No	Unclear	NR	26.00	52.50 (12.40)	176/231 (76.00)	NR	NR	DAS44 -ESR = 5.50 (1.20)
van de Putte (2004) Europe, Canada, and Australia 26 weeks	Treatment with at least 1 DMARD had previously failed.	Severe RA. Active disease defined as ≥12 tender joints (of 68), ≥10 swollen joints (of 66), and either an ESR ≥28 mm in the first hour or a serum CRP ≥20 mg/L.	ADA 40 mg QOW (n = 113)	No	Unclear	NR	20.40	52.70 (13.30)	90/113 (79.60)	NR	NR	DAS28c = 7.07 (0.86)
			PBO (n = 110)	No	Unclear	NR	22.70	53.50 (13.20)	90/110 (81.80)	NR	NR	DAS28c = 7.09 (0.87)
Weinblatt (1999) NR 24 weeks	Active RA despite taking MTX ≥ 6 months, and at a stable dose of 15-25 mg per week for the last 4 weeks (weekly doses as low as 10 mg were acceptable for patients who could not tolerate higher doses). At baseline, 78% were taking NSAIDs and 62% CCS.	Active RA, as manifested by at least 6 joints that were swollen and 6 that were tender at the time of enrolment.	MTX 15-25 mg QW (n = 30)	No	Unclear	NR	27.00	53.00 (NR)	27/30 (90.00)	NR	NR	NR
			ETN 25 mg BIW + MTX (n = 59)	No	Unclear	NR	10.00	48.00 (NR)	50/59 (84.00)	NR	NR	NR

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
AMBITION Jones *et al.* (2010) Worldwide 24 weeks	Oral GCS and NSAIDs permitted if stable dose for ≥ 6 weeks. Included patients who had temporarily discontinued MTX and those who discontinued anti-TNF treatment for reasons other than efficacy. Excluded patients if they had been unsuccessfully treated with an anti-TNF agent, had received MTX in the preceding 6 months or discontinued previous MTX treatment because of clinically important AEs or lack of efficacy.	Moderate to severe RA for ≥ 3 months. Active RA was defined by the presence of ≥ 6 SJC from a total of 66, ≥ 8 TJC from a total of 68, and a CRP ≥ 1 mg/dL or ESR ≥ 28 mm per hour.	TCZ 8 mg/kg (n = 288)	MTX naïve[c]	Mixed[d]	NR	17.40	51.10 (13.10)	NR	NR	NR	DAS-28 ESR = 6.80 (1.00)
			MTX 7.5-20 mg QW (n = 284)	MTX naïve[c]	Mixed[d]	NR	18.50	50.10 (12.80)	NR	NR	NR	DAS-28 ESR = 6.80 (0.90)
LITHE Kremer *et al.* (2011) Secondary report: Fleischmann *et al.* (2013); Kremer *et al.* (2016) Australia, Brazil, China, Denmark, Finland, France, Greece, Italy, Mexico, Norway, Poland, South Africa, Spain, and US	Patients had an inadequate response to MTX. Exclusion criteria included failure to respond to anti-TNF therapy. 75% had also taken DMARDs other than MTX.	Established RA, moderate to severe in the opinion of the investigator	TCZ 8 mg/kg (n = 398)	No	Mixed[d]	NR	18.00	53.40 (11.70)	330/398 (83.00)	NR	NR	DAS-28 = 6.60 (1.00)
			MTX 10-25 mg QW (n = 393)	No	Mixed[d]	NR	17.00	51.30 (12.40)	322/393 (82.00)	NR	NR	DAS-28 = 6.50 (1.00)

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
104 weeks												
OPTION Smolen *et al.* (2008) Secondary report: Ramos-Remus *et al.* (2008) Multinational 24 weeks	Inadequate response to MTX for 12 weeks of longer with a stable dose of 10-25 mg per week for 8 weeks or longer before the study. 55% of patients took steroids and 67% took NSAIDs at baseline.	Established, moderately to severely active RA defined by a SJC ≥ 6 plus a TJC ≥ 8 and CRP > 10 mg/L or ESR of 28 ≥ mm per hour.	TCZ 8 mg/kg + MTX (n = 205)	No	Mixed[d]	NR	15.00	50.80 (11.80)	171/205 (83.00)	NR	NR	DAS-28 = 6.80 (0.90)
			MTX 10-25 mg QW (n = 204)	No	Mixed[d]	NR	22.00	50.60 (12.10)	144/204 (71.00)	NR	NR	DAS-28 = 6.80 (0.90)
SUMMACTA Burmester *et al.* (2014) Secondary report: Burmester *et al.* (2013) Worldwide 96 weeks	Patients must have received ≥ 1 traditional DMARD at a stable dose for ≥ 8 weeks and had an inadequate response to DMARD (up to 20% of patients may have failed 1 or more anti-TNF). Exclusion criteria included allergies to biologic agents and previous treatment with TCZ. 54% were receiving GCS at baseline.	Patients had SJC of ≥ 4 (66-joint count) and TJC ≥ 4 (68-joint count) and CRP ≥ 10 mg/L and/or ESR ≥ 28 mm per hour at screening.	TCZ 162 mg QW (n = 631)	No	Mixed[a]	NR	17.40	52.40 (12.29)	456/620 (73.50)	NR	NR	DAS-28 ESR = 6.60 (1.00)
			TCZ 8 mg/kg (n = 631)	No	Mixed[a]	NR	17.30	52.50 (12.50)	465/625 (74.40)	NR	NR	DAS-28 ESR = 6.70 (1.01)
TOWARD Genovese *et al.* (2008) Multinational 24 weeks	Patients had moderate to severe RA despite receiving stable doses of permitted DMARDs (MTX, chloroquine, HCQ, parenteral gold, SSZ, AZA, and LEF) for ≥ 8 weeks prior to study entry. Exclusion criteria included patients who were unsuccessfully treated with an anti-TNF	Established, moderate to severe RA with SJC ≥ 6, a TJC ≥ 8, and a CRP ≥ 1 mg/dL or an ESR ≥ 28 mm per hour.	TCZ 8 mg/kg + DMARD (n = 805)	No	Mixed	NR	19.00	53.00 (13.00)	NR	NR	NR	DAS-28 ESR = 6.70 (1.00)
			DMARD (n = 413)	No	Mixed	NR	16.00	54.00 (13.00)	NR	NR	NR	DAS-28 ESR = 6.60 (1.00)

Trial ID; Country; Study Duration	Details of/ Response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (Number Randomised)	cDMA RD Naive/ MTX Naive	bDMARD Naive	Comorbidities	Males, %	Age (Yrs) Mean (SD)	RF +, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
	agent. At baseline 74% were taking NSAIDs and 53% oral steroids.											

Footnotes: Baseline characteristics for RA-BEAM and RA-BUILD are presented in Section 4.5.2.

a = Number treated; number randomised not reported. b = Some patients (>20%) are bDMARDs experienced. c = Does not report if ESR or CRP was used. d = Limited number of patients (≤ 20%) are cDMARDs experienced.

Abbreviations: ABA = abatacept; ACR = American College of Rheumatology; ADA = adalimumab; AZA = azathioprine; BAR = baricitinib; bDMARD = biologic disease-modifying antirheumatic drug; BIW = twice weekly; CCP = cyclic citrullinated peptide; CCS = corticosteroid; CDAI = Clinical Disease Activity Index; cDMARD = conventional disease-modifying antirheumatic drug; CRP = C-reactive protein; CTZ = certolizumab pegol; DAS = Disease Activity Score; DAS28 = Disease Activity Score modified to include the 28 diarthrodial joint count; DAS44 = Disease Activity Score in 44 Joints; DMARD = disease-modifying antirheumatic drug; ESR = erythrocyte sedimentation rate; ETN = etanercept; GCS = glucocorticoid; GOL = golimumab; HCQ = hydroxychloroquine; hs-CRP = high-sensitivity C-reactive protein; ID = identification; IFX = infliximab; LEF = leflunomide; MRI = magnetic resonance imaging; MTX = methotrexate; NR = not reported; NSAID = non-steroidal anti-inflammatory drug; PBO = placebo; PSL = prednisolone; Q2W = every 2 weeks; Q4W = every 4 weeks; QD = once daily; QOW = every other week; QW = once weekly; RA = rheumatoid arthritis; RF = rheumatoid factor; RTX = rituximab; SD = standard deviation; SDAI = Simplified Disease Activity Index; SJC = swollen joint count; SSZ = sulfasalazine; TB = tuberculosis; TCZ = tocilizumab; TJC = tender joint count; TNF = tumour necrosis factor; UK = United Kingdom; US = United States.

Table 72. Patient characteristics in studies included in the anti-TNF-IR analyses

Trial ID; Country; Study Duration	Details of/ response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (No. randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Co-morbidities	Males, %	Age (Years) Mean (SD)	RF Positive, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
ATTAIN Genovese *et al.* (2005) Secondary reports: West-hovens *et al.* (2006a) North America and Europe 24 weeks	Inadequate response to anti–TNF therapy with ETN, IFX, or both at the approved dose after ≥ 3 months of treatment. Enrolled patients currently receiving anti-TNF therapy and past users. Patients were previously treated with NSAIDs. Prior cDMARDs included MTX, AZA, gold, HCQ, SSZ and chloroquine.	Moderate-to-severe RA: ≥ 10 swollen joints and ≥ 12 tender joints.	ABA 10 mg/kg	No	No	NR	22.90	53.40 (12.40)	73.30	NR	NR	DAS28 - ESR = 6.50 (0.90)
			PBO	No	No	NR	20.30	52.70 (11.30)	72.90	NR	NR	DAS28 - ESR = 6.50 (0.80)
BREVACTA Kivitz *et al.* (2014) Worldwide 24 weeks	Inadequate response to ≥1 DMARD that, in up to 20% of patients, could include ≥1 anti-TNF agent. Patients had received ≥1 traditional DMARDs at a stable dose for ≥8 weeks. Previous DMARDs, mean (SD): TCZ = 1.3 (0.7); PBO = 1.4 (0.8). Previously failed anti-TNF treatment: TCZ = 20.4%; PBO = 21.5%.	Moderate to severe: SJC ≥6 (66-joint count), TJC ≥8 (68-joint count), radiographic evidence of ≥1 joints with a definite erosion attributable to RA, and a CRP level ≥10 mg/L and/or ESR ≥28 mm per hour.	TCZ 162 mg Q2W (n = 437)	No	Mixed[a]	NR	14.20	52.10 (11.45)	349/432 (80.80)	NR	NR	DAS28-ESR = 6.70 (0.92)
			PBO (n = 219)	No	Mixed[a]	NR	17.40	52.00 (11.71)	178/218 (81.70)	NR	NR	DAS28-ESR = 6.60 (0.94)
ORAL STEP Burmester *et al.* (2013b) Secondary	Patients had previous inadequate response or intolerance to 1 or more approved TNF inhibitors, as established by the investigator, administered in	Established, moderate-to-severe RA and active disease.	TOFA 5 mg BID (n = 133)[b]	No	No	Hypertension, hypercholesterolaemia, osteoporosis, diabetes mellitus, COPD	15.04	55.40 (11.50)	80/132 (60.61)	NR	NR	DAS28-ESR = 6.50 (1.10) DAS28-

Trial ID; Country; Study Duration	Details of/ response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (No. randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Co-morbidities	Males, %	Age (Years) Mean (SD)	RF Positive, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
report: Strand *et al.* (2015) Worldwide 24 weeks	accordance with its label. Patients must have taken oral or parenteral MTX continuously for 4 months or more before the first study dose and be on a stable dose of 7·5-25 mg per week (7·5-20 mg per week in Republic of Ireland) for 6 weeks or more before the first study dose.	Active disease was defined as 6 or more tender or painful joints (of 68-joint count) and 6 or more swollen joints (of 66-joint count) and either ESR (Wester-gren method) higher than 28 mm per hour or CRP of more than 66-67 nmol/L (7 mg/L).										CRP = 5.40 (1.00)
			TOFA 10 mg BID (n = 134)[b]	No	No	Hypertension, hypercholesterolaemia, osteoporosis, diabetes mellitus, COPD	13.43	55.10 (11.30)	83/134 (61.94)	NR	NR	DAS28-ESR = 6.40 (0.90) DAS28-CRP = 5.30 (0.90)
			MTX QW (n = 132)[b]	No	No	Hypertension, hypercholesterolaemia, osteoporosis, diabetes mellitus, COPD	19.70	54.40 (11.30)	86/131 (65.65)	NR	NR	DAS28-ESR = 6.40 (1.10) DAS28-CRP = 5.40 (1.00)
GO-AFTER Smolen *et al.* (2009b) Austria, Australia, Canada, Finland, Germany, Netherlands, New Zealand, Spain, UK, and US	Had been treated with ≥ 1 dose of a TNF inhibitor (ETN, IFX, or ADA). Previous treatment with TNF inhibitor could have been discontinued for any reason, and was categorised as lack of effectiveness, intolerance, or other. Patients had never received natalizumab or RTX.	Active RA: persistent disease activity with ≥ 4 swollen and 4 tender joints	GOL 50 mg Q4W (n = 153)	Unclear	No	NR	26.00	Median = 55.00 (12.59)	108/149 (72.00)	NR	NR	Median DAS28[c] = 6.30 (1.19)
			PBO (n = 155)	Unclear	No	NR	15.00	Median = 54.00 (13.33)	110/151 (73.00)	NR	NR	Median DAS28[c] = 6.30 (1.19)

Trial ID; Country; Study Duration	Details of/ response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (No. randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Co-morbidities	Males, %	Age (Years) Mean (SD)	RF Positive, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
24 weeks												
RADIATE Emery *et al.* (2008a) Secondary reports: Strand *et al.* (2012b); Emery *et al.* (2009c) North America and western Europe 24 weeks	Patients had to be treated with MTX for ≥12 weeks before baseline (stable dose ≥8 weeks) and had failed to respond or an intolerance to 1 or more TNF antagonists within the past year. At baseline, 56% were receiving oral steroids.	Moderate to severely active RA for ≥6 months, SJC ≥6, TJC ≥8, and CRP >1.0 mg/dL or ESR >28 mm per hour.	TCZ 8 mg/kg + MTX (n = 175)	No	No	NR	16.00	53.90 (12.70)	134/170 (79.00)	NR	NR	DAS28[c] = 6.79 (0.93)
			MTX 10-25 mg QW (n = 160)	No	No	NR	21.00	53.40 (13.30)	119/158 (75.00)	NR	NR	DAS28[c] = 6.80 (1.06)
REALISTIC Weinblatt *et al.* (2012) Secondary report: Pope *et al.* (2015) US and Canada 12 weeks	Patients showed an unsatisfactory response or intolerance to DMARD (MTX, LEF, SSZ, chloroquine or HCQ, AZA, and/or gold). Patients were excluded who received treatment either with >2 TNF inhibitors, RTX or ABA. 21% had discontinued previous anti-TNF inhibitors for efficacy reasons and 16% for non-efficacy reasons.	Active RA defined by ≥5 tender and ≥4 swollen joints (28-joint count) and either ≥10 mg/L CRP or ≥28 mm per hour ESR (Wester-gren method) at screening.	CTZ 400-200 mg QOW (n = 851)	No	Unclear[a]	NR	22.40	55.40 (12.40)	555/851 (73.90)	NR	NR	DAS28-ESR = 6.40 (0.90) DAS28-CRP = 5.70 (0.90)
			MTX QW (n = 212)	No	Unclear[a]	NR	20.30	53.90 (12.70)	137/212 (76.50)	NR	NR	DAS28-ESR = 6.40 (0.90) DAS28-CRP = 5.70 (0.90)
REFLEX Cohen *et al.* (2006) Secondary reports: Keystone *et al.* (2009a); Keystone *et*	Patients had an inadequate response to previous or current treatment with anti-TNF agents IFX (≥3 mg/kg; ≥4 infusions), ADA (40 mg every other week for ≥3 months), or ETN (25 mg twice weekly for ≥3	Active RA defined as ≥8 swollen joints (of 66) and ≥ 8 tender joints (of 68), CRP level	MTX 10-25 mg QW (n = 209)	No	No	NR	19.00	52.80 (12.60)	165/209 (79.00)	NR	NR	DAS28[c] = 6.80 (1.00)
			RTX 1000 mg + MTX (n = 311)	No	No	NR	19.00	52.20 (12.20)	242/308 (79.00)	NR	NR	DAS28[c] = 6.90 (1.00)

Trial ID; Country; Study Duration	Details of/ response to Prior Treatment	Details of Disease Severity/ Activity	Treatment (No. randomised)	cDMARD Naive/ MTX Naive	bDMARD Naive	Co-morbidities	Males, %	Age (Years) Mean (SD)	RF Positive, n/N (%)	Baseline Scores		
										CDAI, Mean (SD)	SDAI, Mean (SD)	DAS, Mean (SD)
al. (2008) US, Europe, Canada, and Israel 104 weeks	months), or were intolerant to ≥1 administration of these agents. Patients were taking MTX (10-25 mg per week) for ≥12 weeks prior to screening, with the last 4 weeks at a stable dosage. 63% were taking GCS at baseline.	≥1.5 mg/dL or ESR ≥28 mm per hour, radiographic evidence of ≥1 joint with a definite erosion attributable to RA, as determined by a central reading site.										

Footnote: [a]Some patients (>20%) are bDMARD experienced. [b]Number treated; number randomised not reported. [c]Does not report if ESR or CRP was used
Abbreviations: ABA = abatacept; ACR = American College of Rheumatology; ADA = adalimumab; AZA = azathioprine; BAR = baricitinib; bDMARD = biologic disease-modifying antirheumatic drug; BIW = twice weekly; CCP = cyclic citrullinated peptide; CCS = corticosteroid; CDAI = Clinical Disease Activity Index; cDMARD = conventional disease-modifying antirheumatic drug; CRP = C-reactive protein; CZP = certolizumab pegol; DAS = Disease Activity Score; DAS28 = Disease Activity Score modified to include the 28 diarthrodial joint count; DAS44 = Disease Activity Score in 44 Joints; DMARD = disease-modifying antirheumatic drug; ESR = erythrocyte sedimentation rate; ETN = etanercept; GCS = glucocorticoid; GOL = golimumab; HCQ = hydroxychloroquine; hs-CRP = high-sensitivity C-reactive protein; ID = identification; IFX = infliximab; LEF = leflunomide; MRI = magnetic resonance imaging; MTX = methotrexate; NR = not reported; NSAID = non-steroidal anti-inflammatory drug; PBO = placebo; PSL = prednisolone; Q2W = every 2 weeks; Q4W = every 4 weeks; OD = once daily; QOW = every other week; QW = once weekly; RA = rheumatoid arthritis; RF = rheumatoid factor; RTX = rituximab; SD = standard deviation; SDAI = Simplified Disease Activity Index; SJC = swollen joint count; SSZ = sulfasalazine; TB = tuberculosis; TCZ = tocilizumab; TJC = tender joint count; TNF = tumour necrosis factor; UK = United Kingdom; US = United States

Outcomes considered in the NMA

The following endpoints were selected for inclusion in the NMA:

- ACR response (20%, 50%, and 70% improvement in criteria)
- EULAR response (no, moderate, and good)

ACR response is a common efficacy and health-related quality of life outcome evaluated in clinical trials in RA and was thus selected for analysis in the NMA.

The category of EULAR response is specified in the bDMARD treatment continuation rules stipulated by NICE since TA375 (see Section 3.5) and is therefore the efficacy outcome used to determine short-term treatment continuation or discontinuation in the economic model, as presented in Section 5.

Many of the studies identified in the systematic literature review did not report EULAR response (see Section 4.10.3). To resolve this issue, in studies where EULAR response was not reported, ACR response rates were converted to EULAR based on a conversion algorithm that was used by the AG in TA375.[85] This algorithm converts ACR response to EULAR response based on data from the United States (US) veterans database (VARA).[85] This conversion algorithm is not universally validated. Therefore, as most of the EULAR data arise from the conversion of ACR by this algorithm, the EULAR response models should be viewed with caution, and clinical interpretation should focus on the ACR response data. However, it should be noted that this approach was necessary due to the importance of providing EULAR data for inclusion in the cost-effectiveness model to be consistent with TA375.

4.10.6 Risk of bias

Quality assessments of each trial included in the NMA were performed according to standards recommended by NICE.[86] Results of these quality assessments are presented in Appendix 15. No notable risk of bias was identified and therefore no adjustment was made to the analysis.

4.10.7 Methods of analysis and presentation of results

Methods of analysis

The NMAs were conducted using Bayesian mixed treatment comparison techniques as described in the National Institute for Health and Care Excellence (NICE) Decision Support Unit (DSU) Technical Support Documents (TSDs).[87-89]

The ACR responses [20%, 50% and 70%] and EULAR responses [moderate or good, and good] were analysed as ordinal outcomes, using probit models. The principles given in the NICE DSU technical support document 3 by Dias and colleagues for ordered categorical data were followed, the key details of which are reproduced below. The approach utilised uses a multinomial likelihood model with a probit link:

$$pikj = \Phi\ (\mu i + zij + \delta i,bkI\{k \neq 1\}$$

where j represents the different response (ACR or EULAR) thresholds, k is an arm of a trial i and therefore pijk is the probability that a patient in arm k of trial i belongs to category j. The pooled effect of the experimental treatment versus the control (in this case, the cDMARD arm of the included studies) is to change the probit (Z) score of the control by δi,bk standard deviations. The

term zij specifies the cut-offs at which the individual moves from one category to the next in trial i. This model allows inclusion of trials using different thresholds or trials reporting different numbers of thresholds- which is relevant here due to the reporting of multiple ACR and EULAR response categories across studies. The analysis also follows the guidance from TSD2 by Dias and colleagues to re-write the multinomial likelihood as a series of conditional binomials.

Fixed-effects (FE) and random-effects (RE) simultaneous Bayesian models were fitted for both the cDMARD- and TNFi-IR populations. Model fit was assessed using the deviance information criterion (DIC) and the posterior mean of the total residual deviance. When comparing two DIC values, a difference of 5 of more was regarded as a meaningful difference. With total residual deviance, a model was considered a good fit if the total residual deviance was approximately equal to the number of data points available.[90,91]

The RE models for the TNF-IR population were unstable and did not converge; hence results from the fixed effects model only are presented for this population. Based on model fit (DIC values for each model are presented in Appendix 16), the random-effects models were chosen as the primary approach for the cDMARD-IR population. Simultaneous models for baseline and treatment effects were used over separate models for several reasons: since the data for both baseline and treatment effects came from the same sources; there were some networks that had zero cells and fitting this type of model increased the stability of the relevant models; and the evidence for TNF-IR networks was sparse.

The analyses were performed in JAGS version 3.4 for Bayesian computation. For each model, the first 530,000 simulations were discarded to allow for model convergence, and an additional 1,060,000 simulations (thinning factor 53) were used to estimate the posterior probabilities from a sample of 40,000 using two chains. Convergence was verified by trace plots, monitoring the Monte Carlo error, and with Gelman-Rubin diagnostics.[92] The code used to run the NMA, can be found in Appendix 17.

Heterogeneity and inconsistency

Between-study heterogeneity for each pairwise comparison was assessed using Higgins I^2 (meta-analysis), a statistic which describes the percentage of variation between studies that is due to heterogeneity as opposed to chance. This analysis required at least two studies per comparison and significant heterogeneity was assumed if the corresponding p value from the analysis was <0.05. The results of the assessment of heterogeneity are presented in Section 4.10.9. A sensitivity analysis was conducted in which those studies identified to contribute significant heterogeneity were removed from the network; the results of this can be found in Section 4.10.8.2. It should be noted that in the anti-TNF-IR population all treatment comparisons were informed by a single study only; no analysis of heterogeneity was therefore possible in this population.

For the assessment of inconsistency, where closed loops of evidence were available in the network, node-splitting analyses were performed to evaluate consistency between direct and indirect evidence for the corresponding treatment comparisons. The node-splitting analyses for inconsistency are described further in Section 4.10.9.

Presentation of results

NMA allows comparisons to be made between each treatment in the network, and therefore allows estimation of the relative effectiveness of baricitinib versus all relevant comparators.

The following results are presented for the two bases case analyses (cDMARD-IR population and TNF-IR population) in the main body of the submission:

- For the relative results of analysis conducted on the ordinal endpoint of ACR50, odds ratios and 95% credible intervals (CrI) are presented.
- Median event rates and 2.5 and 97.5 percentiles are presented for ACR response (20%, 50% and 70%) and EULAR moderate or good, and good response.
- Posterior distribution treatment effect results for all interventions relative to cDMARDs
- Median ranking of treatments

ACR and EULAR ranking results for the base case analyses are presented in Appendix 18.

For each sensitivity analysis, median event rates and 2.5 and 97.5 credible intervals are presented for ACR response (20%, 50% and 70%) and EULAR moderate or good, and good response.

4.10.8 Results of the NMA

4.10.8.1 Base case analyses

cDMARD-IR population

ACR50 results at Week 24

The odds ratios and 95% credible intervals for the ACR50 outcome at Week 24 for baricitinib 4 mg and 2 mg (QD) versus each comparator in the cDMARD-IR population are presented in Table 73, Table 74 and Table 75. Both baricitinib 4 mg and 2 mg (QD) were statistically significantly superior to cDMARDs, placebo and sulfasalazine for the ACR50 outcome. In addition, baricitinib 4 mg QD was statistically significantly superior to etanercept. With the exception of CTZ + cDMARD, for which a statistically significantly greater response was achieved compared to both baricitinib 4 mg and 2 mg (QD), there were no statistically significant differences between either dose of baricitinib and the remaining comparators included in the analyses.

Table 73. ACR50 results at Week 24 for the cDMARD-IR population – Odds ratios and credible intervals (1/3)

Intervention	cDMARD	BAR4 + cDMARD	BAR2 + cDMARD	TCZ 8 mg	TCZ 8 mg + cDMARD	ADA 40 mg	ABA 10 mg + cDMARD	ABA SC + cDMARD	ADA 40 mg + cDMARD
BAR4 + cDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Results are presented from probit random effects model. Intervention on the left hand column is the numerator. Green cells indicate a statistically significant difference in favour of baricitinib. Blue cells indicate a statistically significant difference not in favour of baricitinib.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), cDMARD = conventional disease-modifying antirheumatic drugs, SC = subcutaneous, TCZ = tocilizumab

Table 74. ACR50 results at Week 24 for the cDMARD-IR population – Odds ratios and credible intervals (2/3)

Intervention	IFX 3mg + cDMARD	TCZ Subcut + cDMARD	Placebo	ETN	ETN + SSZ	SSZ	ETN + cDMARD	RTX 1000 mg	RTX 1000 mg + cDMARD
BAR4 + cDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Results are presented from probit random effects model. Intervention on the left hand column is the numerator. Green cells indicate a statistically significant difference in favour of baricitinib. Blue cells indicate a statistically significant difference not in favour of baricitinib.
Abbreviations: BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), cDMARD = conventional disease-modifying antirheumatic drugs, ETN =etanercept, IFX = infliximab, RTX = rituximab, SC = subcutaneous, SSZ = sulfasalazine, TCZ = tocilizumab

Table 75. ACR50 results at Week 24 for the cDMARD-IR population – Odds ratios and credible intervals (3/3)

Intervention	GOL 50 mg + cDMARD	CTZ + cDMARD	TOFA 10mg + cDMARD	TOFA 5mg + cDMARD	SSZ + HCQ + cDMARD	RTX 2000 mg + cDMARD	TCZ SC
BAR4 + cDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Results are presented from probit random effects model. Intervention on the left hand column is the numerator. Green cells indicate a statistically significant difference in favour of baricitinib. Blue cells indicate a statistically significant difference not in favour of baricitinib.
Abbreviations: BAR2 = baricitinib 2 mg (QD), BAR4 = baricitinib 4 mg (QD), cDMARD = conventional disease-modifying antirheumatic drugs, GOL = golimumab, HCQ = hydrochloroquine, RTX = rituximab, SSZ = sulfasalazine, TCZ = tocilizumab, TOFA = tofacitinib.

ACR response rates at Week 24

ACR 20%, 50% and 70% response rates at Week 24 for the cDMARD-IR population are presented in Table 76. The median ACR20 response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted], respectively. The median ACR50 response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted], respectively. The median ACR70 response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted], respectively.

Table 76. Median ACR response rates at Week 24 for the cDMARD-IR population

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
cDMARD	ACR20	[redacted]	[redacted]	[redacted]
cDMARD	ACR50	[redacted]	[redacted]	[redacted]
cDMARD	ACR70	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TCZ 8 mg	ACR20	[redacted]	[redacted]	[redacted]
TCZ 8 mg	ACR50	[redacted]	[redacted]	[redacted]
TCZ 8 mg	ACR70	[redacted]	[redacted]	[redacted]
TCZ 8 mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
TCZ 8 mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
TCZ 8 mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
ADA 40 mg	ACR20	[redacted]	[redacted]	[redacted]
ADA 40 mg	ACR50	[redacted]	[redacted]	[redacted]
ADA 40 mg	ACR70	[redacted]	[redacted]	[redacted]
ABA 10 mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
ABA 10 mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
ABA 10 mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
ABA SC + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
ABA SC + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
ABA SC + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
ADA 40 mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
ADA 40 mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
ADA 40 mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
IFX 3 mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
IFX 3 mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
IFX 3 mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TCZ SC + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
TCZ SC + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
TCZ SC + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
Placebo	ACR20	[redacted]	[redacted]	[redacted]
Placebo	ACR50	[redacted]	[redacted]	[redacted]
Placebo	ACR70	[redacted]	[redacted]	[redacted]

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
ETN	ACR20	[redacted]	[redacted]	[redacted]
ETN	ACR50	[redacted]	[redacted]	[redacted]
ETN	ACR70	[redacted]	[redacted]	[redacted]
ETN + SSZ	ACR20	[redacted]	[redacted]	[redacted]
ETN + SSZ	ACR50	[redacted]	[redacted]	[redacted]
ETN + SSZ	ACR70	[redacted]	[redacted]	[redacted]
SSZ	ACR20	[redacted]	[redacted]	[redacted]
SSZ	ACR50	[redacted]	[redacted]	[redacted]
SSZ	ACR70	[redacted]	[redacted]	[redacted]
ETN + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
ETN + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
ETN + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
RTX 1000 mg	ACR20	[redacted]	[redacted]	[redacted]
RTX 1000 mg	ACR50	[redacted]	[redacted]	[redacted]
RTX 1000 mg	ACR70	[redacted]	[redacted]	[redacted]
RTX 1000 mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
RTX 1000 mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
RTX 1000 mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TOFA 5mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
TOFA 5mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
TOFA 5mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
SSZ + HCQ + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
SSZ + HCQ + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
SSZ + HCQ + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	ACR20	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	ACR50	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	ACR70	[redacted]	[redacted]	[redacted]

Footnotes: Results are presented for the random effects analysis.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, ACR = ACR response criteria

EULAR response rates at Week 24

EULAR response rates at Week 24 for the cDMARD-IR population are presented in Table 77. The median EULAR moderate + good response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted], respectively. The median EULAR good response rates were [redacted] and [redacted], respectively.

Table 77. Median EULAR response rates at Week 24 for the cDMARD-IR population

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
cDMARD	Good	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	Good	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	Good	[redacted]	[redacted]	[redacted]
TCZ 8mg	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ 8mg	Good	[redacted]	[redacted]	[redacted]
TCZ 8mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ 8mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
ADA 40mg	Moderate+Good	[redacted]	[redacted]	[redacted]
ADA 40mg	Good	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
ABA Subcut + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
ABA Subcut + cDMARD	Good	[redacted]	[redacted]	[redacted]
ADA 40mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
ADA 40mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
IFX 3mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
IFX 3mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
TCZ Subcut + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ Subcut + cDMARD	Good	[redacted]	[redacted]	[redacted]
Placebo	Moderate+Good	[redacted]	[redacted]	[redacted]
Placebo	Good	[redacted]	[redacted]	[redacted]
ETN	Moderate+Good	[redacted]	[redacted]	[redacted]
ETN	Good	[redacted]	[redacted]	[redacted]
ETN + SSZ	Moderate+Good	[redacted]	[redacted]	[redacted]
ETN + SSZ	Good	[redacted]	[redacted]	[redacted]
SSZ	Moderate+Good	[redacted]	[redacted]	[redacted]
SSZ	Good	[redacted]	[redacted]	[redacted]
ETN + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
ETN + cDMARD	Good	[redacted]	[redacted]	[redacted]
RTX 1000mg	Moderate+Good	[redacted]	[redacted]	[redacted]
RTX 1000mg	Good	[redacted]	[redacted]	[redacted]
RTX 1000mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
RTX 1000mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
GOL 50mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	Good	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
TOFA 5mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
TOFA 5mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
SSZ + Hcq + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
SSZ + Hcq + cDMARD	Good	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	Good	[redacted]	[redacted]	[redacted]

Footnotes: Results are presented for the random effects analysis.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, EULAR = EULAR response criteria

Posterior distribution of treatment effect versus cDMARDs

Figure 39 and Figure 40 present the posterior distributions of relative treatment effect for each intervention of interest versus cDMARDs for the probit models for ACR response and EULAR response, respectively. In these posterior distribution plots, the *d* parameter is the measure of relative treatment effect from the random effects probit model, where a d < 0 represents a favourable treatment effect for the intervention in question versus the reference treatment of cDMARD. With regards to both ACR and EULAR response, for both baricitinib 4 and 2 mg (QD) the vast majority of the posterior probability distributions was seen to correspond to favourable relative treatment effect versus cDMARDs.

Figure 39. Posterior distribution of treatment effect relative to cDMARDs for ACR responses at Week 24 in the cDMARD-IR population

Footnotes: Results are presented from the EULAR response random effects probit model.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CZP = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, ACR = ACR response criteria

Figure 40. Posterior distribution of treatment effect relative to cDMARDs for EULAR responses at Week 24 in the cDMARD-IR population

Footnotes: Results are presented from the EULAR response random effects probit model.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CZP = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, EULAR = EULAR response criteria

Finally, median ranking results for each intervention based on the probit models for ACR and EULAR are presented in Appendix 18.

Anti-TNF-IR population

ACR50 results at Week 24

The odds ratios and 95% credible intervals for the ACR50 outcome at Week 24 for baricitinib 4 mg and 2 mg (QD) versus each comparator in the cDMARD-IR population are presented in Table 78. Both baricitinib 4 mg and 2 mg (QD) were statistically significantly superior to cDMARDs. Tocilizumab 8 mg + cDMARD was statistically significantly superior to both doses of baricitinib and rituximab 1000 mg + MTX was statistically significantly superior to baricitinib 2 mg (QD). For all other comparisons with baricitinib 4 mg and 2 mg (QD), there was no statistically significant difference.

Table 78. ACR50 results at Week 24 for the anti-TNF-IR population – odds ratios and credible intervals

Intervention	cDMARD	BAR4 + cDMARD	BAR2 + cDMARD	ABA 10mg + cDMARD	TCZ Subcut + cDMARD	GOL 50 mg + cDMARD	TCZ 8 mg + cDMARD	RTX 1000 mg + MTX
BAR4 + cDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: Results are presented from the ACR response probit fixed effects model. Intervention on the left hand column is the numerator. Green cells indicate a statistically significant difference in favour of baricitinib. Blue cells indicate a statistically significant difference not in favour of baricitinib.
Footnotes: Results are presented from the EULAR response random effects probit model.
Abbreviations: ABA = abatacept, BAR = baricitinib, ETN etanercept, GOL = golimumab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response, TNF = tumour necrosis factor inhibitor; HCQ = hydroxychloroquine, ACR = ACR response criteria

ACR 20%, 50% and 70% response rates at Week 24 for the anti-TNF-IR population are presented in Table 79. The median ACR20 response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted] respectively. The median ACR50 response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted], respectively. The median ACR70 response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted], respectively.

Table 79. Median ACR response rates at Week 24 for the anti-TNF-IR population

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
cDMARD	ACR20	[redacted]	[redacted]	[redacted]
cDMARD	ACR50	[redacted]	[redacted]	[redacted]
cDMARD	ACR70	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TCZ 162 mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
TCZ 162 mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
TCZ 162 mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	ACR20	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	ACR50	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	ACR70	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	ACR20	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	ACR50	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	ACR70	[redacted]	[redacted]	[redacted]

Footnotes: Results are presented from the ACR response probit fixed effects model.
Abbreviations: ABA = abatacept, BAR = baricitinib, ETN etanercept, GOL = golimumab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response, TNF = tumour necrosis factor, HCQ = hydroxychloroquine, ACR = ACR response criteria

EULAR response rates at Week 24 for the anti-TNF-IR population are presented in Table 80. The median EULAR moderate or good response rates for baricitinib 4 mg and 2 mg (QD) were [redacted] and [redacted], respectively. The median EULAR good response rates were [redacted] and [redacted], respectively.

Table 80. Median EULAR response rates at Week 24 for the anti-TNF-IR population

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
cDMARD	Good	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	Good	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	Good	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
TCZ 162 mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ 162 mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	Good	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	Moderate+Good	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	Good	[redacted]	[redacted]	[redacted]

Footnotes: Results are presented from the ACR response probit fixed effects model.
Abbreviations: ABA = abatacept, BAR = baricitinib, ETN = etanercept, GOL = golimumab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response, TNF = tumour necrosis factor, HCQ = hydroxychloroquine, EULAR = EULAR response criteria

Posterior distribution of treatment effect versus cDMARDs

Figure 41 and Figure 42 present the posterior distributions of relative treatment effect for each intervention of interest versus cDMARDs for the fixed effects probit models for ACR response and EULAR response, respectively. In these posterior distribution plots, the *d* parameter is the measure of relative treatment effect from the probit model, where a $d < 0$ represents a favourable treatment effect for the intervention in question versus the reference treatment of cDMARD. With regards to both ACR and EULAR response, for both baricitinib 4 and 2 mg (QD) the vast majority of the posterior probability distributions was seen to correspond to favourable relative treatment effect versus cDMARDs.

Figure 41. Posterior distribution of treatment effect relative to cDMARDs for ACR responses at Week 24 in the anti-TNF-IR population

[redacted]

Footnotes: Results are presented from the ACR response probit fixed effects probit model.
Abbreviations: ABA = abatacept, BAR = baricitinib, ETN = etanercept, GOL = golimumab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response, TNF = tumour necrosis factor, HCQ = hydroxychloroquine, ACR = ACR response criteria

Figure 42. Posterior distribution of treatment effect relative to cDMARDs for EULAR responses at Week 24 in the anti-TNF-IR population

[redacted]

Footnotes: Results are presented from the EULAR response probit fixed effects probit model.
Abbreviations: ABA = abatacept, BAR = baricitinib, ETN = etanercept, GOL = golimumab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, cDMARD = conventional disease-modifying antirheumatic drug, IR = insufficient response, TNF = tumour necrosis factor, HCQ = hydroxychloroquine, EULAR = EULAR response criteria

4.10.8.2 Sensitivity Analyses

In addition to the base case analyses, three sensitivity analyses were performed:

cDMARD-IR Population

1. Sensitivity analysis 1: Week 12 timepoint (random effects)
2. Sensitivity analysis 2: Exclusion of studies with prior biologic use (Week 24 timepoint; random effects)
3. Sensitivity analysis 3: Exclusion of studies with high heterogeneity (Week 24 timepoint; random effects) – see Section 4.10.7 and Section 4.10.9 for details of the assessment of heterogeneity. Studies identified to have high heterogeneity for the ACR and EULAR response outcomes and removed from the analysis are presented in Table 81.

Table 81. Studies excluded from sensitivity analysis 3

Outcome	Study
ACR20	TEMPO
	SATORI
	ARMADA
ACR50	Weinblatt (1999)
	ARMADA
ACR70	-
EULAR no response	TEMPO
	SATORI
	ARMADA
EULAR moderate response	-
EULAR good response	TEMPO
	OPTION, SATORI
	ARMADA, BEAM

Anti-TNF-IR Population

In this population, the only sensitivity analysis conducted was that of the week 12 timepoint.

Sensitivity analysis 1 (cDMARD-IR population: Week 12 timepoint [random effects])

The mean ACR and EULAR response rates for the cDMARD-IR population at the Week 12 timepoint are presented in Table 82 and Table 83.

Table 82. Median ACR response rates at Week 12 for the cDMARD-IR population (sensitivity analysis 1)

Intervention	Endpoint	Median		2.5 percentile		97.5 percentile	
cDMARD	ACR20	[redacted]		[redacted]		[redacted]	
cDMARD	ACR50						
cDMARD	ACR70						
BAR4 + cDMARD	ACR20						
BAR4 + cDMARD	ACR50						
BAR4 + cDMARD	ACR70						
BAR2 + cDMARD	ACR20						
BAR2 + cDMARD	ACR50						
BAR2 + cDMARD	ACR70						
IFX 3mg + cDMARD	ACR20						
IFX 3mg + cDMARD	ACR50						
IFX 3mg + cDMARD	ACR70						

Intervention	Endpoint	Median		2.5 percentile		97.5 percentile	
TCZ 8mg	ACR20						
TCZ 8mg	ACR50						
TCZ 8mg	ACR70						
TCZ 8mg + cDMARD	ACR20						
TCZ 8mg + cDMARD	ACR50						
TCZ 8mg + cDMARD	ACR70						
ABA 10mg + cDMARD	ACR20						
ABA 10mg + cDMARD	ACR50						
ABA 10mg + cDMARD	ACR70						
ABA Subcut + cDMARD	ACR20						
ABA Subcut + cDMARD	ACR50						
ABA Subcut + cDMARD	ACR70						
ADA 40mg + cDMARD	ACR20						
ADA 40mg + cDMARD	ACR50						
ADA 40mg + cDMARD	ACR70						
ETN + cDMARD	ACR20						
ETN + cDMARD	ACR50						
ETN + cDMARD	ACR70						
ADA 40mg	ACR20						
ADA 40mg	ACR50						
ADA 40mg	ACR70						
Placebo	ACR20						
Placebo	ACR50						
Placebo	ACR70						
GOL 50mg + cDMARD	ACR20						
GOL 50mg + cDMARD	ACR50						
GOL 50mg + cDMARD	ACR70						
CTZ + cDMARD	ACR20						
CTZ + cDMARD	ACR50						
CTZ + cDMARD	ACR70						
ETN	ACR20						
ETN	ACR50						
ETN	ACR70						
TOFA 10mg + cDMARD	ACR20						
TOFA 10mg + cDMARD	ACR50						
TOFA 10mg + cDMARD	ACR70						
TOFA 5mg + cDMARD	ACR20						
TOFA 5mg + cDMARD	ACR50						
TOFA 5mg + cDMARD	ACR70						
TCZ SUBCUT	ACR20						
TCZ SUBCUT	ACR50						
TCZ SUBCUT	ACR70						

Footnote: Results are presented from the EULAR response random effects probit model.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = convetional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, ACR = ACR response criteria

Table 83. Median EULAR responses rates at Week 12 for the cDMARD-IR population (sensitivity analysis 1)

Intervention	Endpoint	Median		2.5 percentile		97.5 percentile	
cDMARD	Moderate+Good						
cDMARD	Good						
BAR4 + cDMARD	Moderate+Good						
BAR4 + cDMARD	Good						
BAR2 + cDMARD	Moderate+Good						
BAR2 + cDMARD	Good						
IFX 3mg + cDMARD	Moderate+Good						
IFX 3mg + cDMARD	Good						
TCZ 8mg	Moderate+Good						
TCZ 8mg	Good						
TCZ 8mg + cDMARD	Moderate+Good						
TCZ 8mg + cDMARD	Good						
ABA 10mg + cDMARD	Moderate+Good						
ABA 10mg + cDMARD	Good						
ABA Subcut + cDMARD	Moderate+Good						
ABA Subcut + cDMARD	Good						
ADA 40mg + cDMARD	Moderate+Good						
ADA 40mg + cDMARD	Good						
ETN + cDMARD	Moderate+Good						
ETN + cDMARD	Good						
ADA 40mg	Moderate+Good						
ADA 40mg	Good						
Placebo	Moderate+Good						
Placebo	Good						
GOL 50mg + cDMARD	Moderate+Good						
GOL 50mg + cDMARD	Good						
CTZ + cDMARD	Moderate+Good						
CTZ + cDMARD	Good						
ETN	Moderate+Good						
ETN	Good						
TOFA 10mg + cDMARD	Moderate+Good						
TOFA 10mg + cDMARD	Good						
TOFA 5mg + cDMARD	Moderate+Good						
TOFA 5mg + cDMARD	Good						
TCZ SUBCUT	Moderate+Good						
TCZ SUBCUT	Good						

Footnote: Results are presented from the EULAR response random effects probit model.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = convetional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, EULAR = EULAR response criteria

Sensitivity analysis 2 (cDMARD-IR population: exclusion of studies with prior biologic use [Week 24 timepoint; random effects])

The mean ACR and EULAR response rates for the cDMARD-IR population with no prior biologic use at the Week 24 timepoint are presented in Table 84 and Table 85, respectively.

Table 84. Median ACR response rates at Week 24 for the cDMARD-IR population with no prior biologic use (sensitivity analysis 2)

Intervention	Endpoint	[redacted]		[redacted]		[redacted]	
cDMARD	ACR20						
cDMARD	ACR50						
cDMARD	ACR70						
BAR4 + cDMARD	ACR20						
BAR4 + cDMARD	ACR50						
BAR4 + cDMARD	ACR70						
BAR2 + cDMARD	ACR20						
BAR2 + cDMARD	ACR50						
BAR2 + cDMARD	ACR70						
TCZ 8mg	ACR20						
TCZ 8mg	ACR50						
TCZ 8mg	ACR70						
TCZ 8mg + cDMARD	ACR20						
TCZ 8mg + cDMARD	ACR50						
TCZ 8mg + cDMARD	ACR70						
ADA 40mg	ACR20						
ADA 40mg	ACR50						
ADA 40mg	ACR70						
ABA 10mg + cDMARD	ACR20						
ABA 10mg + cDMARD	ACR50						
ABA 10mg + cDMARD	ACR70						
ABA Subcut + cDMARD	ACR20						
ABA Subcut + cDMARD	ACR50						
ABA Subcut + cDMARD	ACR70						
ADA 40mg + cDMARD	ACR20						
ADA 40mg + cDMARD	ACR50						
ADA 40mg + cDMARD	ACR70						
IFX 3mg + cDMARD	ACR20						
IFX 3mg + cDMARD	ACR50						
IFX 3mg + cDMARD	ACR70						
TCZ Subcut + cDMARD	ACR20						
TCZ Subcut + cDMARD	ACR50						
TCZ Subcut + cDMARD	ACR70						
Placebo	ACR20						
Placebo	ACR50						
Placebo	ACR70						
ETN	ACR20						
ETN	ACR50						
ETN	ACR70						
ETN + SSZ	ACR20						
ETN + SSZ	ACR50						
ETN + SSZ	ACR70						
SSZ	ACR20						
SSZ	ACR50						
SSZ	ACR70						

Intervention	Endpoint	[redacted]		[redacted]		[redacted]	
ETN + cDMARD	ACR20						
ETN + cDMARD	ACR50						
ETN + cDMARD	ACR70						
RTX 1000mg	ACR20						
RTX 1000mg	ACR50						
RTX 1000mg	ACR70						
RTX 1000mg + cDMARD	ACR20						
RTX 1000mg + cDMARD	ACR50						
RTX 1000mg + cDMARD	ACR70						
GOL 50mg + cDMARD	ACR20						
GOL 50mg + cDMARD	ACR50						
GOL 50mg + cDMARD	ACR70						
RTX 2000mg + cDMARD	ACR20						
RTX 2000mg + cDMARD	ACR50						
RTX 2000mg + cDMARD	ACR70						

Footnote: Results are presented from the random effects probit model.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = convetional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, ACR = ACR response criteria

Table 85. Median EULAR response rates at Week 24 for the cDMARD-IR population with no prior biologic use (sensitivity analysis 2)

Intervention	Endpoint	Median		2.5 percentile		97.5 percentile	
cDMARD	Moderate+Good						
cDMARD	Good						
BAR4 + cDMARD	Moderate+Good						
BAR4 + cDMARD	Good						
BAR2 + cDMARD	Moderate+Good						
BAR2 + cDMARD	Good						
TCZ 8mg	Moderate+Good						
TCZ 8mg	Good						
TCZ 8mg + cDMARD	Moderate+Good						
TCZ 8mg + cDMARD	Good						
ADA 40mg	Moderate+Good						
ADA 40mg	Good						
ABA 10mg + cDMARD	Moderate+Good						
ABA 10mg + cDMARD	Good						
ABA Subcut + cDMARD	Moderate+Good						
ABA Subcut + cDMARD	Good						
ADA 40mg + cDMARD	Moderate+Good						
ADA 40mg + cDMARD	Good						
IFX 3mg + cDMARD	Moderate+Good						
IFX 3mg + cDMARD	Good						
TCZ Subcut + cDMARD	Moderate+Good						
TCZ Subcut + cDMARD	Good						
Placebo	Moderate+Good						
Placebo	Good						
ETN	Moderate+Good						

ETN	Good	[redacted]		[redacted]		[redacted]	
ETN + SSZ	Moderate+Good						
ETN + SSZ	Good						
SSZ	Moderate+Good						
SSZ	Good						
ETN + cDMARD	Moderate+Good						
ETN + cDMARD	Good						
RTX 1000mg	Moderate+Good						
RTX 1000mg	Good						
RTX 1000mg + cDMARD	Moderate+Good						
RTX 1000mg + cDMARD	Good						
GOL 50mg + cDMARD	Moderate+Good						
GOL 50mg + cDMARD	Good						
RTX 2000mg + cDMARD	Moderate+Good						
RTX 2000mg + cDMARD	Good						

Footnote: Results are presented from the random effects probit model.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = convetional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, EULAR = EULAR response criteria

Sensitivity analysis 3 (cDMARD-IR population: removal of heterogeneous studies [Week 24 time point])

The mean ACR and EULAR response rates for the cDMARD-IR population at the Week 24 time point with the exclusion of studies with high heterogeneity are presented in Table 88 and Table 89, respectively.

Table 86. Median ACR response rates at Week 24 for the cDMARD-IR population excluding studies with high heterogeneity (sensitivity analysis 3)

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
cDMARD	ACR20	[redacted]	[redacted]	[redacted]
cDMARD	ACR50	[redacted]	[redacted]	[redacted]
cDMARD	ACR70	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TCZ 8mg	ACR20	[redacted]	[redacted]	[redacted]
TCZ 8mg	ACR50	[redacted]	[redacted]	[redacted]
TCZ 8mg	ACR70	[redacted]	[redacted]	[redacted]
TCZ 8mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
TCZ 8mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
TCZ 8mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
ADA 40mg	ACR20	[redacted]	[redacted]	[redacted]
ADA 40mg	ACR50	[redacted]	[redacted]	[redacted]
ADA 40mg	ACR70	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
ABA Subcut + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
ABA Subcut + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
ABA Subcut + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
ADA 40mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
ADA 40mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
ADA 40mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
IFX 3mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
IFX 3mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
IFX 3mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TCZ Subcut + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
TCZ Subcut + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
TCZ Subcut + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
Placebo	ACR20	[redacted]	[redacted]	[redacted]
Placebo	ACR50	[redacted]	[redacted]	[redacted]
Placebo	ACR70	[redacted]	[redacted]	[redacted]
ETN	ACR20	[redacted]	[redacted]	[redacted]
ETN	ACR50	[redacted]	[redacted]	[redacted]
ETN	ACR70	[redacted]	[redacted]	[redacted]
ETN + SSZ	ACR20	[redacted]	[redacted]	[redacted]
ETN + SSZ	ACR50	[redacted]	[redacted]	[redacted]
ETN + SSZ	ACR70	[redacted]	[redacted]	[redacted]
SSZ	ACR20	[redacted]	[redacted]	[redacted]

SSZ	ACR50	[redacted]	[redacted]	[redacted]
SSZ	ACR70	[redacted]	[redacted]	[redacted]
ETN + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
ETN + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
ETN + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
RTX 1000mg	ACR20	[redacted]	[redacted]	[redacted]
RTX 1000mg	ACR50	[redacted]	[redacted]	[redacted]
RTX 1000mg	ACR70	[redacted]	[redacted]	[redacted]
RTX 1000mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
RTX 1000mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
RTX 1000mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TOFA 5mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
TOFA 5mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
TOFA 5mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
SSZ + Hcq + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
SSZ + Hcq + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
SSZ + Hcq + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	ACR20	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	ACR50	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	ACR70	[redacted]	[redacted]	[redacted]

Footnote: Results are presented from the random effects probit model.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = convetional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, ACR = ACR response criteria

Table 87. Median EULAR response rates at Week 24 for the cDMARD-IR population excluding studies with high heterogeneity (sensitivity analysis 3)

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
cDMARD	Good	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	Good	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	Good	[redacted]	[redacted]	[redacted]
TCZ 8mg	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ 8mg	Good	[redacted]	[redacted]	[redacted]
TCZ 8mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ 8mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
ADA 40mg	Moderate+Good	[redacted]	[redacted]	[redacted]
ADA 40mg	Good	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
ABA 10mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
ABA Subcut + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
ABA Subcut + cDMARD	Good	[redacted]	[redacted]	[redacted]
ADA 40mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
ADA 40mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
IFX 3mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
IFX 3mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
TCZ Subcut + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ Subcut + cDMARD	Good	[redacted]	[redacted]	[redacted]
Placebo	Moderate+Good	[redacted]	[redacted]	[redacted]
Placebo	Good	[redacted]	[redacted]	[redacted]
ETN	Moderate+Good	[redacted]	[redacted]	[redacted]
ETN	Good	[redacted]	[redacted]	[redacted]
ETN + SSZ	Moderate+Good	[redacted]	[redacted]	[redacted]
ETN + SSZ	Good	[redacted]	[redacted]	[redacted]
SSZ	Moderate+Good	[redacted]	[redacted]	[redacted]
SSZ	Good	[redacted]	[redacted]	[redacted]
ETN + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
ETN + cDMARD	Good	[redacted]	[redacted]	[redacted]
RTX 1000mg	Moderate+Good	[redacted]	[redacted]	[redacted]
RTX 1000mg	Good	[redacted]	[redacted]	[redacted]
RTX 1000mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
RTX 1000mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	Good	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
TOFA 10mg + cDMARD	Good	[redacted]	[redacted]	[redacted]

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
TOFA 5mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
TOFA 5mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
SSZ + Hcq + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
SSZ + Hcq + cDMARD	Good	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
RTX 2000mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ SUBCUT	Good	[redacted]	[redacted]	[redacted]

Footnote: Results are presented from the random effects probit model.
Abbreviations: ABA = abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, PBO = placebo, SSZ= sulfasalazine, cDMARD = convetional disease-modifying antirheumatic drug, IR = insufficient response; TOFA = tofacitinib, HCQ = hydroxychloroquine, EULAR = EULAR response criteria

Sensitivity analysis 4 (anti-TNF-IR population: Week 12 timepoint)

The mean ACR and EULAR response rates for the anti-TNF-IR population with no prior biologic use at the Week 24 timepoint are presented in Table 88 and Table 89, respectively. It should be noted that

Table 88. Median ACR response rates at Week 12 for the anti-TNF-IR population (sensitivity analysis 4)

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
cDMARD	ACR20	[redacted]	[redacted]	[redacted]
cDMARD	ACR50	[redacted]	[redacted]	[redacted]
cDMARD	ACR70	[redacted]	[redacted]	[redacted]
BAR 4 + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
TOFA 10mg + Mtx	ACR20	[redacted]	[redacted]	[redacted]
TOFA 10mg + Mtx	ACR50	[redacted]	[redacted]	[redacted]
TOFA 10mg + Mtx	ACR70	[redacted]	[redacted]	[redacted]
TOFA 5mg + Mtx	ACR20	[redacted]	[redacted]	[redacted]
TOFA 5mg + Mtx	ACR50	[redacted]	[redacted]	[redacted]
TOFA 5mg + Mtx	ACR70	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	ACR20	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	ACR50	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	ACR70	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	ACR20	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	ACR50	[redacted]	[redacted]	[redacted]
CTZ + cDMARD	ACR70	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	ACR20	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	ACR50	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	ACR70	[redacted]	[redacted]	[redacted]

Footnote: Results are presented from the ACR response fixed effects probit model.
Abbreviations: ABA = abatacept, BAR = baricitinib, ETN etanercept, GOL = golimumab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, cDMARD = convetional disease-modifying antirheumatic drug, IR = insufficient response, TNF = tumour necrosis factor, HCQ = hydroxychloroquine, ACR = ACR response criteria, TOFA = tofacitinib

Table 89. Median EULAR response rates at Week 12 for the anti-TNF-IR population (sensitivity analysis 4)

Intervention	Endpoint	Median	2.5 percentile	97.5 percentile
cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
cDMARD	Good	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
BAR4 + cDMARD	Good	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
BAR2 + cDMARD	Good	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	Moderate+Good	[redacted]	[redacted]	[redacted]
GOL 50mg + cDMARD	Good	[redacted]	[redacted]	[redacted]
TOFA 10mg + Mtx	Moderate+Good	[redacted]	[redacted]	[redacted]
TOFA 10mg + Mtx	Good	[redacted]	[redacted]	[redacted]
TOFA 5mg + Mtx	Moderate+Good	[redacted]	[redacted]	[redacted]
TOFA 5mg + Mtx	Good	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	Moderate+Good	[redacted]	[redacted]	[redacted]
TCZ 8mg + Mtx	Good	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	Moderate+Good	[redacted]	[redacted]	[redacted]
RTX 1000mg + Mtx	Good	[redacted]	[redacted]	[redacted]

Footnote: Results are presented from the EULAR response fixed effects probit model.
Abbreviations: ABA = abatacept, BAR = baricitinib, ETN etanercept, GOL = golimumab, MTX = methotrexate, RTX = rituximab, TCZ = tocilizumab, cDMARD = convetional disease-modifying antirheumatic drug, IR = insufficient response, TNF = tumour necrosis factor, HCQ = hydroxychloroquine, ACR = ACR response criteria, TOFA = tofacitinib

4.10.9 Heterogeneity Between Studies

Higgins' I^2

For each pairwise comparison, heterogeneity was assessed through the use of Higgins' I^2 statistic. Significant heterogeneity was assumed if the corresponding p value from the analysis was <0.05. The results of the assessment of heterogeneity are presented in Table 90 for the cDMARD-IR population. No analysis of heterogeneity was possible in the anti-TNF-IR population as all treatment comparisons were informed by a single study only.

Results from sensitivity analysis 3, in which studies identified with high heterogeneity were excluded, are presented in Section 4.10.8.2. Median values for the ACR and EULAR response outcomes where comparable to the base case analysis.

Table 90: Results of assessment of heterogeneity using Higgins' I^2

Time point	Endpoint	Study(ies)	Source
24 weeks	ACR20	TEMPO[a] SATORI ARMADA	I^2=86.0%; p=0.0008 I^2=66.6%; p=0.0295 I^2=77.1%, p=0.0016
24 weeks	ACR50	Weinblatt (1999) ARMADA[b]	I^2=84.7%, p=0.0015 I^2=53.1%, p=0.0583
24 weeks	EULAR no response	TEMPO[a] SATORI ARMADA	I^2=87.1%; p=0.0004 I^2=90.4%; p<.0001 I^2=66.7%; p=0.0102
24 weeks	EULAR good	TEMPO[a] OPTION, SATORI	I^2=92.3%; p<.0001 I^2=92.3%; p<.0001

	response	ARMADA, BEAM[c]	I^2=75.6%; p=0.0010

a- TEMPO: this only affects the comparison ETN+cDMARD vs cDMARD
b- ARMADA: as there are more studies in the network for this comparator, the heterogeneity diminishes a little. However, for reasons of consistency this study was still excluded from the sensitivity analysis on heterogeneity.
c- BEAM: this only affects the comparison ADA+cDMARD vs cDMARD

Node-splitting

Where closed loops of evidence were available in the network for the cDMARD-IR population at the Week 24 timepoint, node-splitting analyses were performed to evaluate consistency between direct and indirect evidence for the corresponding treatment comparisons. The node-splitting results for the comparison between baricitinib (4 mg QD) and adalimumab, as the only comparison involving baricitinib for which there was a closed loop, are presented in Table 91. As indicated by the non-significant p values, there is no indication of inconsistency between the direct and indirect evidence for this treatment comparison.

Table 91. Results of the node-splitting analysis for the comparison between baricitinib (4 mg QD) and adalimumab

Outcome	P value
ACR20	0.325
ACR50	0.959
ACR70	0.778
EULAR no response	0.0551
EULAR moderate response	0.278
EULAR good response	0.0675

Footnote: Results are presented for the random effects probit model.
Abbreviations: ACR20/50/70 = 20/50/70% improvement in ACR criteria, EULAR = European League Against Rheumatism

Node-splitting analyses for the anti-TNF-IR population were not considered to be sufficiently informative for inclusion in this submission, as the TNF-IR networks contained only one study per comparator and hence there were no cases where separate studies provided direct and indirect evidence for the comparison of baricitinib to a relevant comparator (see Section 4.10.3).

4.10.10 Justification of Fixed Effects or Random Effects Analyses

Fixed- and random-effect simultaneous Bayesian models were fitted for all populations. However, the random-effect model for the anti-TNF-IR populations was unstable and did not converge, and thus results for this analysis are not presented in the submission. Random-effect simultaneous Bayesian models could be fitted for the cDMARD-IR population. Based on model fit, random-effects models were chosen as the primary approach for the cDMARD-IR population; DIC values are presented in Appendix 16.

4.10.11 Overall Summary of NMA Evidence

cDMARD-IR Population

For the base case analysis at week 24, baricitinib 4 mg was found to be associated with a statistically significantly higher odds of an ACR50 response compared to cDMARD, adalimumab 40 mg, placebo, etanercept and sulfasalazine. No statistically significant differences were found versus any other comparator for the ACR50 outcome, with the exception of the comparison of baricitinib 4 mg to certolizumab pegol, in which odds of an ACR50 response was found to be significantly in favour of certolizumab. This pattern of results was also observed for baricitinib 2 mg. For baricitinib 4 mg, although statistically significant differences were not found versus most biologic comparators, in the majority of cases the point estimate of relative treatment effect was favourable to baricitinib 4 mg.

Posterior distributions of treatment effect found baricitinib 4 mg and baricitinib 2 mg to be associated with high probabilities of a favourable treatment effect versus cDMARDs for both ACR and EULAR responses at week 24.

Median ACR and EULAR response rates were seen to be relatively consistent for baricitinib across the base case 24 week analysis and the sensitivity analysis at 12 weeks. In the sensitivity analysis exploring the exclusion of studies including patients with prior biologic use, median rates of ACR and EULAR response for baricitinib 4 mg and 2 mg were similar, though slightly higher, than for the base case analysis.

Anti-TNF-IR Population

For the base case analysis at week 24 in the anti-TNF-IR population, baricitinib demonstrated significantly higher ACR50 response rates than the cDMARD comparator. No statistically significant differences were seen versus biologic comparators, with the exception of the comparison of baricitinib 4 mg and baricitinib 2 mg to tocilizumab 8 mg, and the comparison of baricitinib 2 mg to rituximab 1000 mg, in which statistically significant treatment effects in favour of the comparator were observed. Versus the other comparators, point estimates in some cases favoured baricitinib 4 mg and in other cases favoured the comparator treatment.

Posterior distributions from the probit model demonstrated very high probability of a favourable treatment effect for both baricitinib 4 mg and baricitinib 2 mg versus cDMARD.

In the sensitivity analysis exploring ACR and EULAR outcomes at 12 weeks, median ACR and EULAR response rates for both baricitinib 4 mg and baricitinib 2 mg were found to be slightly higher than for the base case analysis at week 24.

4.11 *Non-randomised and non-controlled evidence*

Given the availability of three well-designed prospective RCTs providing evidence for baricitinib in the populations of interest, non-randomised evidence was not searched for as part of the clinical SLR described in Section 4.1. Lilly is not aware of any currently available non-randomised or non-controlled studies providing evidence on the clinical effectiveness of baricitinib.

4.12 *Adverse reactions*

Summary of Baricitinib Safety Analysis

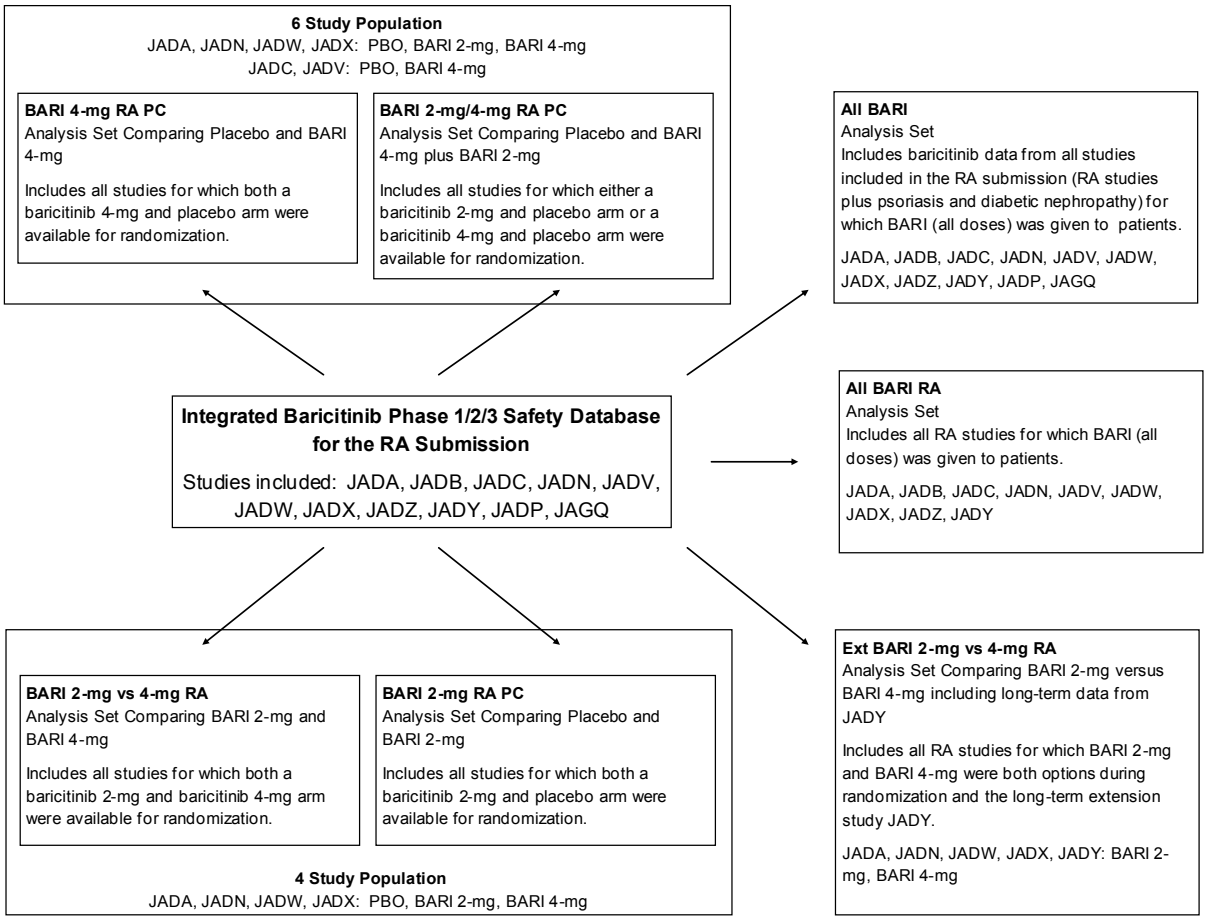
- Treatment with baricitinib was well-tolerated and a small proportion of patients discontinued from the baricitinib studies because of AEs.
- The most commonly reported adverse drug reactions (ADRs) in ≥2% of patients treated with baricitinib monotherapy or in combination with csDMARDs included increased LDL cholesterol, URTI and nausea. However, the majority of all ADRs were mild to moderate in severity.
- The proportion of patients with SAEs (including serious infections) was similar across treatment groups in the phase III studies and integrated placebo-controlled analysis sets, except for RA-BEAM, where a higher proportion of SAEs were reported with placebo and baricitinib versus adalimumab.
- Although baricitinib was associated with a higher incidence of SAEs compared with adalimumab through 52 weeks in RA-BEAM, their AE profiles were similar across clinically significant categories of risk including MACE, malignancies, hypercholesterolemia, serious infections and herpes zoster.
- Despite a higher risk of cardiovascular disease, infection, and malignancy in the RA population, treatment with baricitinib did not result in increased risk of malignancy, serious or opportunistic infections, or MACE.
- Non-serious herpes simplex and herpes zoster infections were more frequent in patients treated with baricitinib than placebo, yet rates were not significantly higher than those seen with MTX or adalimumab.
- The majority of herpes zoster cases were mild to moderate in severity and complicated cases were uncommon.
- Increases in LDL cholesterol were one of the most commonly reported ADRs, yet increases in HDL-C were also seen with baricitinib so that the mean HDL/LDL ratio was unchanged. Furthermore, there was a significant decrease in the amount of small and very small LDL particles in RA-BEAM, which are considered the most atherogenic. Few major adverse cardiovascular events (MACE) events were also observed in the baricitinib clinical programme and no relationship was seen between MACE and increased LDL.
- Treatment with baricitinib also resulted in changes to haematology and clinical chemistry analytes. These included mean changes of greater magnitude for some analytes than seen with the active comparators, which are, therefore, likely to be related to the pharmacology of JAK inhibition (such as increases in lipids [including total cholesterol, low-density lipoprotein cholesterol and high-density lipoprotein cholesterol] and creatine phosphokinase).

4.12.1 Integrated Safety Analysis

In addition to the four Phase III trials described above, the overall safety and tolerability of baricitinib was evaluated in a number of integrated safety analyses comprising data from baricitinib Phase I–III RA trials (including RA-BEGIN), as well as trials of baricitinib in the indications of psoriasis and diabetic nephropathy.[13] The integrated safety analyses were presented as part of the regulatory submission to the EMA. Details of each analysis set are presented in Figure 43.

The analysis set deemed most appropriate to present to provide a general overview of the safety and tolerability of baricitinib in RA was the BAR 4 mg RA PC analysis set, which evaluated the safety of baricitinib 4 mg QD compared to placebo, and comprised data from three Phase II studies (JADA, JADC and JADN; details of which are presented in Appendix 19) and three Phase III studies (JADV [RA-BEAM], JADX [RA-BUILD] and JADW [RA-BEACON].[13] It should be noted that JADZ (RA-BEGIN) was not included in the BAR 4 mg RA PC analysis set as it did not include a placebo arm.[13]

Figure 43. Schematic of integrated safety sets across the baricitinib phase I, II and III trials



Abbreviations: BAR = baricitinib, PC = placebo-controlled, RA = rheumatoid arthritis.
Source: Eli Lilly and Company. Data on File (Summary of Clinical Safety. Figure 2.7.4.3. Page 46). 2016[93]

4.12.1.1 Exposure (BAR 4 mg RA PC)

All analyses of the BAR 4 mg RA PC data set included the time period up to 16 weeks (before rescue became an option in the Phase III studies, see Section 4.3).[94] This time period includes the 12-week placebo-controlled treatment period from studies JADA, JADC, and JADN, and the first 16 weeks of the placebo-controlled period from studies JADV, JADW and JADX, allowing for analysis of data from the greatest number of patients in the baricitinib (4 mg QD) dose group compared to placebo.[94] Analyses were also performed with data up to 24 weeks (censoring data after the time of rescue) for safety topics for which longer controlled exposure was considered to be important.[94]

The analysis set included [redacted] patients in the baricitinib (4 mg QD) arm and [redacted] patients in the placebo arm. There were [redacted] and [redacted] patient-years of exposure to baricitinib (4 mg QD) and placebo up to Week 16,[95] respectively. The equivalent figures for Week 24 were [redacted] and [redacted] patient-years.[96]

4.12.1.2 Overall Safety/Tolerability Profile (BAR 4 mg RA PC)

Baricitinib was generally well-tolerated by patients included in the BAR 4 mg RA PC analysis set, with similar exposure-adjusted incidence rates (EAIR) of overall and severe treatment-emergent adverse events (TEAEs), serious adverse events, temporary interruptions due to adverse events and permanent discontinuations due to adverse events/death and death recorded up to Week 24 compared to placebo, see Table 92.[97]

Table 92. Overview of tolerability profile up to Week 24 (BAR 4 mg RA PC)

Adverse event, n (%) [EAIR]	Baricitinib (4 mg QD) (n=997)	Placebo (n=1070)
Overall treatment-emergent adverse events	695 (69.7) [169.8]	659 (61.6) [167.3]
Severe treatment-emergent adverse events	53 (5.3) [12.9]	43 (4.0) [10.9]
Serious adverse events*	53 (5.3) [12.9]	50 (4.7) [12.7]
Permanent discontinuation due to adverse events/death	47 (4.7) [11.5]	35 (3.3) [8.9]
Temporary interruption due to an adverse event	109 (10.9) [27.1]	89 (8.3) [23.0]
Death	3 [0.7]	2 [0.5]

Footnotes: Treatment adverse events were defined as adverse events that either first occurred or worsened in severity after the first dose of study treatment. Patients with multiple occurrences of the same event are counted under the highest severity. *Defined as any AE associated with a patient outcome that met the International Conference on Harmonisation E2A criteria for an SAE.[98]
Sources: Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.34. Page 197).[97] 2016.[97] Eli Lilly Data on File (Clinical Overview. Rheumatoid Arthritis. EMA Submission. Table 2.5.5.2. Page 64). 2016[99,100]
Abbreviation: EAIR = exposure-adjusted incidence rate, QD = once daily

Common Treatment-Emergent Adverse Events

Common TEAEs were defined as those reported in ≥2% (before rounding) of patients in either the baricitinib (4 mg QD) or placebo arm and these are presented in Table 93. There were 14 common TEAEs in the analysis up to 16 weeks.[101] Ten of the 14 events were reported by a numerically larger proportion of patients in the baricitinib (4 mg) arm; however, of these, a statistically significantly greater incidence in the baricitinib (4 mg) arm compared to placebo was

only reported for increased blood creatinine phosphokinase and hypercholesterolemia.[101] Further details on these TEAEs are discussed in Section 4.12.1.5.

Table 93. Treatment-emergent adverse events occurring in ≥2% in either treatment arm by MedDRA Preferred Term up to Week 16 (BAR 4 mg RA PC)

Treatment-emergent adverse event preferred term, n (%)[102]	Baricitinib (4 mg QD) (n=997)	Placebo (n=1070)	Baricitinib (4 mg QD) versus placebo	
			OR (95% CI)[a]	P value[b]
Nasopharyngitis	[redacted]	[redacted]	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Headache	[redacted]	[redacted]	[redacted]	[redacted]
Blood creatinine increased	[redacted]	[redacted]	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Bronchitis	[redacted]	[redacted]	[redacted]	[redacted]
Hypercholesterolaemia	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]
Pharyngitis	[redacted]	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]
Rheumatoid arthritis	[redacted]	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: [a]Mantel-Haenszel odds ratio and 95% CI (CI calculated if ≥4 events in treatment group and ≥1 in PBO). PBO is denominator. P value from Cochran-Mantel-Haenszel test stratified by study. Breslow-Day test p-value ≥0.10 denoted by 'c'; otherwise, the p value > 0.10. MedDRA version 18.0.
Source: Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.54. Page 903). 2016.[102]

Abbreviation: CI = confidence internal; MedDRA = Medical Dictionary for Regulatory Activities, OR = odds ratio.

Serious Adverse Events

Serious adverse events (SAEs) occurring in at least 2 patients in either the baricitinib (4 mg QD) group or the placebo arm are reported in Table 94. Eleven SAEs were found to occur at an incidence of ≥2 patients in either group, and no significant difference was found between baricitinib (4 mg QD) and placebo.[103]

Table 94. Serious adverse events occurring in at least 2 patients in any group by MedDRA Preferred Term - Serious by ICH - up to Week 16 (BAR 4 mg RA PC)

Serious adverse event preferred term, n (%)	Baricitinib (4 mg QD) (n=997)	Placebo (n=1070)	Baricitinib (4 mg QD) versus placebo	
			OR[a]	P value[b]
Herpes zoster	[redacted]	[redacted]	[redacted]	[redacted]
Cellulitis	[redacted]	[redacted]	[redacted]	[redacted]
Coronary artery disease	[redacted]	[redacted]	[redacted]	[redacted]
Cataract	[redacted]	[redacted]	[redacted]	[redacted]
Fall	[redacted]	[redacted]	[redacted]	[redacted]

Pneumonia	[redacted]	[redacted]	[redacted]	[redacted]
Rheumatoid arthritis	[redacted]	[redacted]	[redacted]	[redacted]
Asthma	[redacted]	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]	[redacted]
Bronchitis	[redacted]	[redacted]	[redacted]	[redacted]
Hyperglycaemia	[redacted]	[redacted]	[redacted]	[redacted]

Footnotes: [a]Mantel-Haenszel odds ratio and 95% CI (CI calculated if ≥4 events in treatment group and ≥1 in PBO). PBO is denominator. [b]P value from Cochran-Mantel-Haenszel test stratified by study. Breslow-Day test p-value ≥0.10 denoted by 'c'; otherwise, the p value > 0.10. MedDRA version 18.0.
Source: Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.65. Page 952). 2016[103]

Discontinuation

For the BAR 4 mg RA PC analysis set up to Week 24, most patients in baricitinib (4 mg QD) and placebo completed study drug treatment during the placebo-controlled period (Table 92). Similar numbers of patients discontinued baricitinib due to adverse events or death to placebo.

Safety Topics of Interest

Analysis was also undertaken for certain safety topics of interest based on the known or anticipated characteristics of:

- Patients with moderately to severely active RA (e.g. major adverse cardiovascular events [MACE; Section 4.12.1.3], malignancies [Section 4.12.1.4].
- Immunomodulatory DMARD therapy in general (e.g. treatment-emergent infections [Section 4.12.1.5])

Treatment with DMARDs with related mechanisms of action including JAK inhibitors (e.g. lipid elevations, elevations in creatine phosphokinase [CPK] [Section 4.12.1.6])

In order to perform a detailed analysis of these safety topics of interest, evidence is drawn from analysis sets additional to the BAR 4 mg RA PC set, as well as from individual trials. In general, safety results have been drawn from the BAR 4 mg RA PC and BAR2 mg vs 4 mg RA analysis sets, as 4 mg and 2 mg QD are the doses anticipated to be licensed. Incidence rates for adverse events have been calculated from the ALL BAR RA or ALL BAR analysis sets, as these sets present the largest patient exposure to baricitinib.

A summary of the safety profile of baricitinib with regards to each of the safety topics of interest is presented in Sections 4.12.1.3 to 4.12.1.6.

4.12.1.3 Major Adverse Cardiovascular Events

Given that patients with RA are at an increased risk of experiencing cardiovascular events,[104] major adverse cardiovascular events (MACE) are an expected topic of interest for a novel treatment in this indication.

During the baricitinib clinical studies, an independent clinical evaluation committee adjudicated potential MACE (cardiovascular death, myocardial infarction, stroke) and other cardiovascular events.[105]

The proportions of patients with a least 1 positively adjudicated MACE or other cardiovascular event in the BAR 4 mg RA PC and BAR 4 mg vs 2 mg RA analysis sets were not statistically significantly different between baricitinib (4 mg QD) and placebo, or baricitinib (4 mg QD) and

baricitinib (2 mg QD), respectively.[106] The exposure-adjusted incidence rate (EAIR) for MACE did not increase over time.[106] In addition, in RA-BEGIN (JADZ), which was not included in either analysis set, baricitinib (4 mg QD) alone or in combination with MTX did not increase the overall risk for MACE through 52 weeks of treatment compared to MTX monotherapy.[106] In RA-BEAM (JADV), baricitinib (4 mg) was not associated with a higher incidence of MACE through 52 weeks of treatment compared to adalimumab.[106]

The overall incidence rate of positively adjudicated MACE was 0.46 per 100 person-years for the All BAR RA analysis set.[107] The crude incidence rate of MACE was approximately constant throughout the course of baricitinib treatment, and event rates do not appear to exceed background rates for cardiovascular outcomes in RA.[107]

4.12.1.4 Malignancies

Given the physiologic role of immune surveillance in the prevention of malignancy,[108] in addition to the increased risk of cancer in RA patients, particularly non-melanoma skin cancer,[109,110] malignancy was deemed to be a safety topic of interest.

[redacted] cases of malignancy (excluding NMSC) in patients treated with baricitinib were observed in the ALL BAR RA analysis set (N=3,463).[111] Of these cases, [redacted] patients were receiving concomitant MTX.[112] In [redacted] patients, a causal role for baricitinib was unlikely as the malignancy occurred in the first 60 days of treatment in [redacted] patients, [redacted] patients had prior malignancies, and [redacted] had significant risk factors for malignancy.[112]

[redacted] cases of NMSC in patients treated with baricitinib were reported in the ALL BAR RA analysis set (N=3,363).[112] Of these, [redacted] patients were receiving MTX.[13] The influence of MTX and other medications cannot be quantified, but studies have shown that it may influence the rate of malignancies.[110]

The incidence rate for all malignancies (excluding NMSC malignancies) was [redacted] per 100pt/years for baricitinib and does not appear to exceed background rates in RA.[113] Furthermore, the incidence rates for NMSC and non-NMSC remained stable over time.[113] There were no statistically significant differences in the proportions of patients reporting all malignancies, non-NMSC, and NMSC between baricitinib (4 mg QD) and placebo, or baricitinib (4 mg QD) and baricitinib (2 mg QD) in the BAR 4 mg RA PC and BAR2 mg vs 4 mg RA analysis sets, respectively, albeit the number of malignancies observed during the placebo-controlled period of the studies is very small.[113] These data do not support the recognition of malignancy, malignancy excluding NMSC, or NMSC alone, as an identified or potential risk for baricitinib.

4.12.1.5 Treatment-Emergent Infections, Including Non-Serious and Serious Infections

An increase in overall (though not serious) infections, a large majority being mild or moderate in severity, was seen with baricitinib compared to placebo during the clinical trial programme.[99] Infections of the upper respiratory tract (URTIs) accounted for much of this imbalance [redacted] in the 16-week placebo-controlled portion of the program (BAR 4 mg RA PC analysis set).[114] Rates of URTI for baricitinib (4 mg QD) did not exceed those observed for active comparators.[99] The SmPC states that the risks and benefits of treatment with baricitinib should be carefully considered prior to initiating therapy in patients with active, chronic or recurrent infections. If an infection develops, the patient should be monitored carefully and therapy should be temporarily interrupted if the patient is not responding to standard therapy. Treatment should not be resumed until the infection resolves.

Non-serious herpes simplex [redacted][115] and herpes zoster infections (see below) were also reported statistically significantly more frequently with baricitinib compared to placebo.[2,13]

Infections requiring antibiotic treatment were more common with baricitinib (4 mg QD) as were temporary interruptions (due to upper respiratory tract infections and herpes simplex infections) and permanent discontinuations of study drug (attributed to the protocol requirement to discontinue with herpes zoster infections).[13]

There were no differences in the proportion of patients reporting serious infections with baricitinib (4 mg QD) compared to placebo through 24 weeks of treatment in the BAR 4 mg RA PC analysis set ([redacted] per 100pt/years for baricitinib vs. [redacted] for placebo).[116] Through 52 weeks of treatment in RA-BEAM (JADV), the proportion of patients experiencing a TEAE or serious infection was similar between baricitinib (4 mg QD) and adalimumab.[116] There was also no increased risk of serious infection with baricitinib (4 mg QD) alone or in combination with MTX compared to MTX monotherapy through 52 weeks of treatment in RA-BEGIN (JADZ).[116]

The EAIR for serious infections in the All BAR RA population was [redacted] and there was no evidence for an increase in the risk of serious infection with prolonged administration of baricitinib.[116] The EAIR for serious infections for baricitinib-treated patients does not appear to exceed background rates in patients with RA.[116]

Herpes Zoster

A statistically significantly larger proportion of patients reported the TEAE of herpes zoster with baricitinib (4 mg QD) compared to placebo in the BAR 4 mg RA PC analysis set.[117]

In the updated ALL BARI RA analysis set (as of 01 January 2016), [redacted] events of herpes zoster on treatment were reported. The majority of these (95%) cases were mild to moderate in severity and complicated cases were uncommon. No differences in the incidences of herpes simplex and herpes zoster infections were observed between patients treated with baricitinib 4 mg and baricitinib 2 mg in the extended baricitinib 4 mg versus baricitinib 2 mg RA analysis. After 18 months of treatment, the incidence of herpes simplex and herpes zoster infections (Figure 91) did not increase over time with prolonged treatment with baricitinib.[2,99]

The majority of cases of herpes zoster were mild or moderate in severity. Of the [redacted] reported cases in the ALL BAR RA analysis set (N=3,464), complicated or disseminated events (nerve palsy or dissemination beyond the primary or adjacent dermatomes) were only reported in [redacted] cases. There were no reported cases of internal organ involvement.[118]

Tuberculosis

Given that other immunomodulatory treatments for RA, e.g. TNF inhibitors, are associated with an increased risk of development of tuberculosis (TB), or reactivation of TB,[119] TB is a safety topic of interest.

Patients were excluded from participation in the Phase III clinical studies for baricitinib if they had evidence of active TB, as documented by a positive purified protein derivative (PPD) test, medical history, clinical symptoms or abnormal chest x-ray at screening.[120] Patients with evidence of latent TB (as documented by a positive PPD, no clinical symptoms consistent with active TB, and a normal chest x-ray at screening) could participate if they completed at least 4

weeks of appropriate treatment prior to randomisation and agreed to complete the remainder of treatment while in the study.[120]

During the controlled period of the RA studies, two events of clinically overt tuberculosis (TB) infection were reported, one among patients treated with baricitinib 4 mg and one among those treated with adalimumab. Six cases of TB (three unconfirmed by microbiology) were reported in the uncontrolled period (when all patients received baricitinib).[2,99]

No difference in the incidence of TB was detected between patients treated with baricitinib 4 mg and baricitinib 2 mg in the extended baricitinib 4 mg versus baricitinib 2 mg RA analysis. The overall exposure adjusted incidence rate (EAIR) of tuberculosis in RA patients treated with baricitinib 4 mg once daily was [redacted] [121] The CHMP noted in their assessment of baricitinib that these rates were lower than compared with adalimumab in RA-BEAM [redacted] and the expected background risk in the total population (0.64 events per 100 PY).[121]

All incidences of TB occurred in countries where prevalence was high, which is reflected in the rates of screen failure due to TB and the proportions of patients with latent TB at randomisation for these countries. The associated incidence rates were in keeping with expected background rates in these countries for patients with RA and do not point to TB as an identified risk for baricitinib.[121]

4.12.1.6 Laboratory Parameters

Elevations in Creatine Phosphokinase

In controlled studies, for up to 16 weeks, increases in CPK values were common. Significant increases (> 5 x ULN) occurred in [redacted] of patients treated with Olumiant and [redacted] of patients treated with placebo. A dose relationship was observed with CPK elevations ≥ 5 x ULN of normal reported in [redacted] and [redacted] of patients at 16 weeks in the 4 mg, 2 mg and placebo groups, respectively. Most cases were transient and did not require treatment discontinuation. In clinical trials, there were no confirmed cases of rhabdomyolysis. Elevations of CPK were observed at 4 weeks and remained stable at a higher value than baseline thereafter including in the long-term extension study.[2,121]

Elevations in Lipid Levels

Baricitinib treatment was associated with dose-dependent increases in lipid parameters including total cholesterol, triglycerides, LDL cholesterol, and HDL cholesterol. There was no change in the LDL/HDL ratio. Elevations were observed at 12 weeks and remained stable thereafter at a higher value than baseline including in the long-term extension study. In controlled studies, for up to 16 weeks, the following rates were observed for baricitinib vs placebo:[2]

- Increased total cholesterol ≥ 5.17 mmol/L: [redacted] vs.[redacted] respectively
- Increased LDL cholesterol ≥ 3.36 mmol/L: [redacted] vs. [redacted] respectively
- Increased HDL cholesterol ≥ 1.55 mmol/L: [redacted] vs [redacted] respectively
- Increased triglycerides ≥ 5.65 mmol/L: [redacted] vs. [redacted] respectively

Treatment with baricitinib 4 mg was associated with dose-dependant increases in lipid parameters including total cholesterol, triglycerides, LDL-C, and HDL-C (as assessed in the baricitinib 4 mg RA placebo-controlled analysis set).

Starting statin treatment in patients who experienced an increase in cholesterol while on baricitinib was effective in returning total cholesterol, LDL-C and triglycerides to baseline levels, while HDL-C remained increased.[13,122] Figure 44 presents the number of patients treated with statin therapy in the BAR 4-mg RA PC and BAR 2-mg/4-mg RA PC analysis sets, and the corresponding change in lipid analytes upon commencing and stopping statin use. In this regard it is of note that in the baricitinib clinical program, there were few cardiovascular events. As discussed in Section 4.12.1.3, data suggest that baricitinib treatment of patients with RA is not associated with an increased risk of MACE, including stroke, myocardial infarction and cardiovascular death. Taken together, the changes in lipids and the available data on MACE do not indicate a clear safety signal.[13,122]

Figure 44. Change in lipid analytes from baseline to initiation of statin therapy and end of statin use up to Week 24 for the BAR 4-mg RA PC and the BAR 2-mg/4-mg RA PC analysis set

[redacted]

Source: Eli Lilly and Company. Data on File (Clinical Safety Summary. Figure 2.7.4.18. Page 204). 2016[123]
Abbreviations: HDL = high density lipoprotein, LDL = low density lipoprotein, SD = standard deviation

Increases in LDL and triglycerides will be considered adverse reactions, however, as they may result in prescription of lipid-lowering therapies.[13,124] The SmPC for baricitinib will advise prescribers to assess lipid parameters at approximately 12 weeks following initiation of treatment and to manage patients according to international clinical guidelines for the management of hyperlipidaemia.[2]

Neutropenia

In controlled studies, for up to 16 weeks, decreases in neutrophil counts below 1 x 10^9 cells/L occurred in [redacted] of patients treated with baricitinib compared to [redacted] of patients treated with placebo. There was no clear relationship between decreases in neutrophil counts and the occurrence of serious infections. However, in clinical studies, treatment was interrupted in response to ANC < 1 x 10^9 cells/L. The pattern and incidence of decreases in neutrophil counts remained stable at a lower value than baseline over time including in the long-term extension study.[2]

4.12.2 Serious adverse events in RA-BEAM: baricitinib versus adalimumab

A summary of safety data from the RA-BEAM trial is presented in Table 95. Results from RA-BEAM demonstrated that the safety profile of baricitinib (4 mg QD) is comparable to that of adalimumab.[99] In RA-BEAM, the AE profiles between adalimumab and baricitinib (both on background MTX) in MTX-IR patients, were similar across clinically significant categories of risk.[13,99] Incidence of deaths, MACE and malignancies occurred infrequently and there were no clinically significant differences observed between groups.[2,13] Likewise, serious infections occurred at similar rates in placebo, adalimumab, and baricitinib groups after 24 weeks [redacted] and in adalimumab and baricitinib groups over 52 weeks of treatment. Increases in low-density lipoprotein cholesterol (LDL-C) occurred in both the adalimumab and baricitinib groups, with larger increases seen with baricitinib. Large elevations in ALT (≥3, ≥5 and ≥10xULN) were similar for both groups across 52 weeks of treatment, as were rates of upper

respiratory tract infections (URTI) and nausea.[99] There were also no significant differences in herpes zoster rates between adalimumab and baricitinib across 52 weeks.[99] However, through 52 weeks, treatment with baricitinib (4 mg QD) was associated with a statistically significant higher incidence of SAEs compared with adalimumab. It is worth noting that the proportion of patients reporting ≥1 SAE was higher among patients treated with placebo than adalimumab at Week 24.[99]

Table 95. Overview of adverse events in RA-BEAM (Week 0–24)

	TEAE[a], n [EAIR]		Serious AE[a] n [EAIR]	Permanent discont. due to AE/death[a] N [EAIR]	Temporary interruption due to AE[a] n [EAIR]	Death[a] n [EAIR]
	Overall	Severe				
Placebo (N=488)	295 [149.2]	19 [9.6]	22 [11.3]	17 [8.6]	45 [22.8]	0
Baricitinib (4 mg QD) (N=487)	347 [161.4]	21 [9.8]	23 [10.7]	25 [11.6]	48 [22.3]	2 (0.9)
Adalimumab (N=330)	224 [157.8]	6 [4.2]	6 [4.2]	7 [4.9]	29 [20.4]	0

Footnotes: [a] For deaths, the observation period includes the study follow-up period where available. Otherwise, data are during the treatment period. Patients with multiple occurrences of the same event are counted under the highest severity.
Source: Eli Lilly Data on File (Clinical Overview. Rheumatoid Arthritis. EMA Submission. Table 2.5.5.2. Page 64) 2016[99,100]
Abbreviations: AE = adverse event, TEAE = treatment-emergent adverse events, EAIR = exposure-adjusted incidence rate

4.12.3 Conclusions on the safety of baricitinib

As detailed in Table 92, treatment with baricitinib was generally well-tolerated, with similar incidences of overall and severe treatment-emergent adverse events, serious adverse events, temporary interruptions and permanent discontinuation of medication and death recorded up to Week 24 between baricitinib (4 mg QD) and placebo.

Despite a higher risk of cardiovascular disease, infection, and malignancy in the RA population, treatment with baricitinib did not result in increased risk of malignancy (including lymphoma), serious or opportunistic infections, or MACE.[13,99]

Immunomodulatory therapies may be associated with an increased risk of treatment-emergent infections, including serious infections. Baricitinib was observed to be associated with an increased risk of herpes zoster infection, however the majority observed cases were mild to moderate in severity, and less than 5% of cases were complicated. The incidence of serious infection, including TB was low across all baricitinib trials.[13,99]

Treatment with baricitinib also resulted in changes to haematology and clinical chemistry analytes. These included mean changes of greater magnitude for some analytes than seen with the active comparators, which are therefore likely to be related to the pharmacology of JAK inhibition (such as increase in lipids, CPK, and creatinine), as well as changes of similar magnitude to those seen with the active comparator adalimumab (such as small decrease in neutrophils and small increases in ALT).[13,99]

In general, the safety of baricitinib appeared comparable to adalimumab through 52 weeks of treatment based on the assessment of clinically important measures of safety and relevant measures of tolerability. An increase in AEs leading to permanent discontinuation of study drug

was observed in baricitinib 4 mg compared to adalimumab. The imbalance appeared to be driven by minor differences across a number of SOCs, with no single type of event appearing convincingly different. This increased rate in the baricitinib group was also observed at Week 24 at which time the rate of AEs leading to permanent discontinuation was lower in frequency for adalimumab than placebo. This observation of a lower rate of AE's leading to permanent discontinuation of treatment is inconsistent with prior published findings for adalimumab where AEs leading to permanent discontinuation have been higher in frequency for adalimumab than placebo. A slightly higher number of patients in baricitinib 4 mg compared to adalimumab experienced a temporary interruption of study due to an AE. The AEs were mild or moderate in severity and the Preferred Term accounting for the largest difference between groups was bronchitis. A higher percentage of patients experienced an increase in LDL to 'borderline high or higher' (≥3.36 mmol/L; ≥130 mg/dL). There were no other clinically relevant differences between groups for other parameters.[13,99]

The overall safety and tolerability profile of baricitinib 4 mg is also considered favourable as it did not differ in a clinically meaningful way compared to placebo, MTX, or baricitinib 2 mg with respect to designated critical measures of safety (i.e. death, AEs leading to permanent discontinuation of study drug, malignancies, serious infections, MACE, or large increases in ALT.

4.13 *Interpretation of clinical effectiveness and safety evidence*

4.13.1 Findings from clinical evidence

As summarised in Section 4.7, compared to placebo and to approved oral cDMARD (MTX) and injectable bDMARD (adalimumab) therapy—which together represent the established standards of care in moderate RA and severe RA—baricitinib 4 mg QD demonstrated clear, consistent and clinically meaningful improvements across all relevant domains of efficacy including signs and symptoms, low disease activity and remission, physical function, and patient reported outcomes. Improvements were evident from the earliest weeks of treatment and were maintained during prolonged administration. In addition, and importantly, baricitinib 4 mg QD significantly inhibited radiographic progression of structural joint damage compared to placebo in 3 completed confirmatory studies, and demonstrated numerically similar inhibition compared to adalimumab in the one head-to-head study versus this therapy. Efficacy has been demonstrated consistently for patient populations across the RA treatment continuum including patients refractory to prior cDMARDs and bDMARDs. In addition to a favourable efficacy profile, baricitinib brings further benefits to patients, including ease of administration in the form of an oral tablet, which was confirmed by a patient survey (Section 2) and lack of production of anti-drug antibodies. Given the current standard of care and unmet need in RA, the efficacy profile and additional benefits to patients affirms that it represents an important treatment alternative for RA patients.

In addition, the indirect treatment comparison (Section 4.10) demonstrated both doses of baricitinib to be statistically significantly superior to cDMARDs in both the cDMARD-IR and TNFi-IR populations, with the majority of comparisons to bDMARDs non-statistically significantly different with respect to EULAR moderate and good response, and ACR 20%, 50% and 70% response at Weeks 12 and 24.

As summarised in Section 4.12, baricitinib 4 mg QD was generally well-tolerated and a small proportion of patients discontinued from the baricitinib studies because of AEs. Safety topics of interest included adverse events commonly associated with RA, such as major adverse cardiovascular events and malignancies.These adverse events were observed to occur at frequencies no greater than in the background RA population in patients treated with baricitinib. Baricitinib 4 mg QD was observed to be associated with an increased risk of herpes zoster and other non-serious infections compared to placebo, however in an integrated analysis set of all baricitinib Phase I–III RA trials, only 5% of zoster cases were deemed to be complicated. Baricitinib was also associated with increased levels of creatine phosphokinase and blood lipids. However, these changes were not judged to be associated with AEs with Muscle Symptom query or MACE, respectively. Furthermore, as reflected in the SmPC, these changes in lipid levels are amenable to monitoring and management. In the context of the relevant comparator DMARDs for this appraisal, these data indicate an acceptable safety and tolerability profile for baricitinib 4 mg once daily and do not raise any significant concerns. In general, the safety of baricitinib appeared comparable to adalimumab through 52 weeks of treatment based on the assessment of clinically important measures of safety and relevant measures of tolerability. An increase in AEs leading to permanent discontinuation of study drug was observed in baricitinib 4 mg compared to adalimumab. The imbalance appeared to be driven by minor differences across a number of SOCs, with no single type of event appearing convincingly different. This increased rate in the baricitinib group was also observed at Week 24 at which time the rate of AEs leading to permanent discontinuation was lower in frequency for adalimumab than placebo

Across RA patient populations, therefore, substantial evidence of consistent, clinically meaningful benefit was demonstrated for baricitinib 4 mg once daily compared to placebo and to cDMARD and adalimumab across key domains of efficacy, without significant safety concerns or a concomitant increase in key risks. Therefore, baricitinib represents an important new efficacious oral treatment option for patients with moderately to severely active RA.

4.13.2 Strengths and limitations of the clinical evidence in this submission

Evidence from three high quality randomised controlled trials has been presented, as demonstrated by the results of the quality assessment in Section 4.7, which collectively cover patients corresponding to both the cDMARD-IR and anti-TNF-IR populations specified in the scope for this appraisal. A notable strength of the efficacy findings for baricitinib was the consistency of benefit observed for the 4 mg dose, across measures (including signs, symptoms, low disease activity, remission, HRQOL outcomes, and structural joint damage), over time, across studies, across methods of analysis, and across (placebo and active) comparators. Having individual, robustly designed phase III trials for each treatment position for baricitinib requested in the scope, and for each of these trials to demonstrate clear, consistent and clinically-relevant efficacy shows the significant strength of the baricitinib clinical evidence base.

A further strength of the clinical evidence base is the inclusion in the relevant phase III trial of a bDMARD active control in the form of adalimumab, in addition to the inclusion of cDMARDs as comparators in two trials. Each of the RCTs included well established and valid primary and secondary outcomes covering the signs and symptoms of RA, clinical decision-making criteria on low disease activity and remission, as well as physical function, and patient reported symptoms and HRQOL. The evidence base therefore provides robust evidence across objectively observed criteria as well as physician and patient-focussed measures that directly relate to clinical practice in the NHS. Notably, the RA-BEAM study considered relevant outcomes of ACR20 response and change from baseline in DAS28-hsCRP versus an active control of adalimumab as pre-specified secondary endpoints. The clinical development programme incorporated numerous features designed to answer clinically important questions across the spectrum of the RA treatment continuum, and to maximise the generalisability of the data without compromising the studies' ability to test hypotheses with integrity.

A number of weaknesses in the clinical evidence base may nonetheless be identified. For placebo-controlled trials the duration of placebo control is necessarily limited to that judged appropriate under current ethical standards. It is only possible, therefore, to demonstrate efficacy against placebo over a relatively short duration in the context of a chronic condition. This limitation, however, applies to all trials across relevant comparators and does not create any expectation that efficacy would not be confirmed in longer term studies. Furthermore, evidence for longer-term efficacy of baricitinib 4 mg is supported by results from the RA-BEYOND study presented in Section 4.7.4. The impact of rescue of placebo patients on interpretation of study results is also somewhat mitigated by the fact that the timepoint for rescue was Week 16 or Week 24 across all of the baricitinib phase III studies, and therefore after the pre-specified primary endpoint at Week 12. This trial design may complicate the evaluation of safety and efficacy by dose, as do protocol-designed dose switches such as the step-down from 4 mg to 2 mg, however, such trial designs are necessary to evaluate long-term efficacy and safety.

As with all newly investigated products, there is limited availability of long term data on the safety and efficacy of baricitinib in comparison to that available for long-established comparators; nonetheless the data available has not provided any signals indicating waning efficacy or significant changes in safety. Indeed, clinical efficacy has been sustained through to Week 52 across a range of relevant outcomes in Section 4.7. In line with comparator trials and international treatment guidelines,[38,39] the baricitinib trials recruited patients with moderately to severely active RA. In NHS England, previous NICE appraisals have resulted in approval of bDMARDs for restricted use in severe RA only, rather than across moderate to severe RA. Although the baricitinib trials provide robust evidence across its full license of moderate to severe RA, it may be noted that there was not a separate RCT that only recruited the moderate RA population, however, this is consistent with trial for comparators to baricitnib in the RA space. Therefore, it is necessarily subgroup analyses that have demonstrated efficacy specifically in the moderate population (see Section 4.8) and such analyses inherently have reduced patient numbers compared to the whole trial, which may be viewed as providing somewhat weaker evidence in this subgroup when compared to the evidence presented for the full licensed indication.

Overall, therefore, the weaknesses in the clinical evidence base for baricitinib are largely in line with those observed in the evidence for other comparators, being due to the constraints of appropriate trial design or the status of baricitinib as a newly investigated compound, and do not detract from the considerable strengths of the evidence presented.

End of life criteria are not relevant to baricitinib and/or rheumatoid arthritis.

4.14 *Ongoing studies*

No ongoing studies of baricitinib which have not been discussed in this submission are expected to be published in the next 12 months.

5 Cost-effectiveness

Summary of the Cost-Effectiveness Evaluation

- The developed model closely followed that developed by the Assessment Group in TA375, with necessary adaptations or additions in order to incorporate the modelling of baricitinib therapy and additional patient populations
- The model categorised patients based on EULAR response at a time point of 24 weeks (6 months), consistent with previous models in the disease area. Response rates were informed by the NMAs for the cDMARD-IR and TNF-IR populations presented in Section 4.10
- Initial changes in HAQ score dependent upon EULAR response status were derived from the BSRBR database for biologic therapies and baricitinib; the long-term HAQ trajectory was assumed to remain flat for patients receiving these interventions and to progress as per the non-linear latent class analysis from TA375 for those receiving cDMARDs
- Health-related quality of life was measured as EQ-5D utility index scores based on mapping from HAQ score, as per TA375. Various mapping algorithms were explored in scenario analyses
- Costs and resource use were implemented from a UK perspective, taking into account available discounts for comparator therapies where these were publically available. Baricitinib was considered with a confidential simple discount patient access scheme
- Base case cost-effectiveness results found that:
    - In the severe, cDMARD-IR population baricitinib dominated all comparators, with the exception of certolizumab pegol which had an ICER of £18,400 per QALY versus baricitinib
    - In the severe, anti-TNF-IR (rituximab ineligible) population, baricitinib was found to represent a cost-effective intervention at a £30,000 per QALY threshold versus a number of the comparator therapies currently used in the NHS
    - In the base case analysis for the anti-TNF-IR rituximab eligible populations the baricitinib sequence was dominated by the rituximab sequence.
    - In the moderate population, the basecase ICER for the baricitinib sequence vs the cDMARD sequence was around £37,400 per QALY
- Scenario analyses tested a wide range of assumptions employed in the base case analysis; the majority of scenario analyses demonstrated similar conclusions as the base case analyses. A linear HAQ progression assumption for the moderate population gave an ICER estimate of around £21,000 per QALY
- In summary, the economic evaluation presents a robust evaluation, closely aligned to that of TA375 previously reviewed by NICE, and finds baricitinib to represent a cost-effective treatment option versus a number of currently approved biologic therapies in both the severe, cDMARD-IR and severe, anti-TNF-IR (rituximab ineligible) populations

5.1 *Published cost-effectiveness studies*

Search strategy

A literature search was conducted in order to identify economic evaluations of bDMARDs or tofacitinib for the treatment of active RA. This search comprised an original search, which was completed on 5th November 2014, and an update search that searched for literature published from the beginning of November 2014 until October 2016. The update search followed the same methodology as the original search.

Literature was searched in electronic databases in accordance with NICE STA guidelines.[125] The following electronic databases were searched:

- Embase
- MEDLINE®
- MEDLINE® In-Process & Other Non-Indexed Citations
- EconLit
- The Cochrane Library:
    - National Health Service Economic Evaluations Database (NHS EED)

In addition, a pragmatic search for available economic models for biologics in RA submitted to the following HTA agencies was performed:

- National Institute for Health and Care Excellence (NICE)
- Scottish Medicines Consortium (SMC)
- Canadian Agency for Drugs and Technologies in Health (CADTH)
- The Pharmaceutical Benefits Advisory Committee (PBAC)
- The full search strategy for the literature review, including search terms, is presented in Appendix 20.

Study selection

Articles identified from the search were first screened based on the title and abstract for modelling studies and according to predefined eligibility criteria, as shown in Table 96. Studies were excluded if they focussed only on costs or if they were economic evaluations alongside clinical trials (EEACTs). Full-texts of all articles that met the eligibility criteria were then obtained and were subsequently screened for inclusion using the same eligibility criteria. All publications where there was uncertainty about inclusion were double reviewed, and any disagreement was resolved either through “reconciliation” (discussion between the two reviewers) or, through “arbitration” by a third independent reviewer, where the “majority view” determined inclusion or exclusion.

Table 96. Eligibility criteria for the economic literature review

	Inclusion	Exclusion
Population	Adults (≥18 years) with active RA	Non-human, juvenile population (aged 0-17 years), other forms of arthritis
Intervention	bDMARDs – abatacept, etanercept, infliximab, adalimumab, certolizumab pegol, golimumab, rituximab, or tocilizumab and targeted synthetic DMARDs – tofacitinib both as monotherapy and in combination with cDMARDs	
Comparator	• cDMARDs (e.g. methotrexate, sulfasalazine, leflunomide etc) • bDMARDs & targeted synthetic DMARDs (including the interventions being considered compared with each other) • Combination treatments of any of the above (including switching, adding, treat-to-target etc) • Supportive care (e.g. corticosteroids, NSAIDs, analgesics, or ongoing cDMARDs) • Placebo (including 'do nothing' option)	
Outcomes	• Resource use and direct and indirect cost parameters associated with RA medications, RA complications and adverse events • Reported utility (to include (dis)utility of RA medications and RA complications) • Measures of cost-effectiveness (cost per responder, cost per QALY, cost per day in remission, etc.)	
Study Design	• Cost-effectiveness analysis **OR** • Cost-utility analysis **OR** • Cost-minimisation **OR** • Cost-benefit analysis **OR** • Cost of illness study **OR** • Cost-consequence analysis **OR** • Systematic reviews of economic evalutions (EE)	• Editorials **OR** • Notes **OR** • Comments **OR** • Letters **OR**
Restrictions	**Language:** Titles and abstracts must be available in English **Year limitation:** *Original SLR:* None (search was completed on 5th November 2014) *Updated SLR:* November 2014 to October 2016	

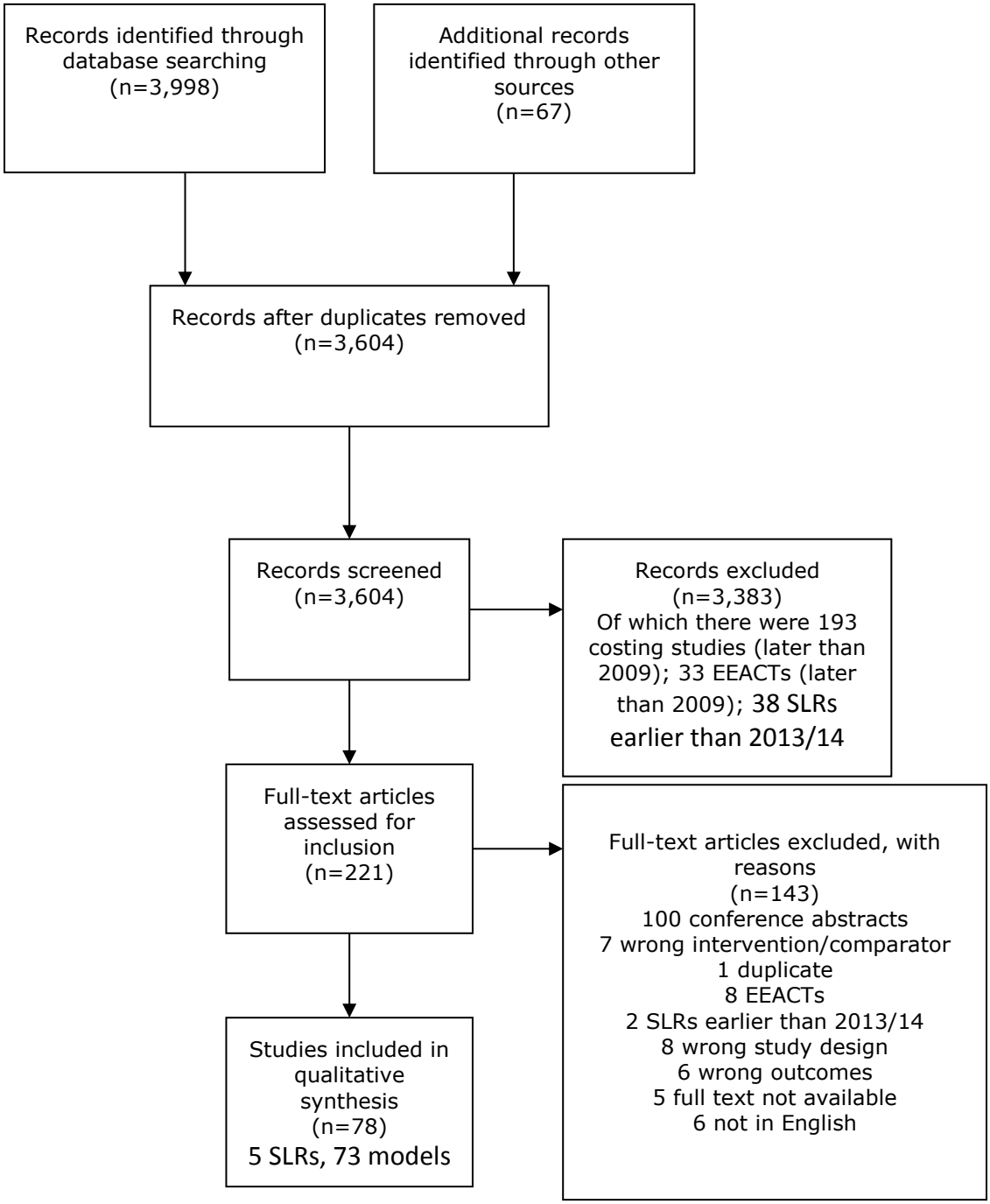
Abbreviations: RA = rheumatoid arthritis, DMARD = disease-modifying antirheumatic drug, bDMARD = biologic DMARD, cDMARD = conventional DMARD, NSAID = non-steroidal anti-inflammatory drug, QALY = quality-adjusted life year.

Description of identified studies

PRISMA diagrams for the original and update searches are presented in Figure 47 and Figure 48, respectively, below.

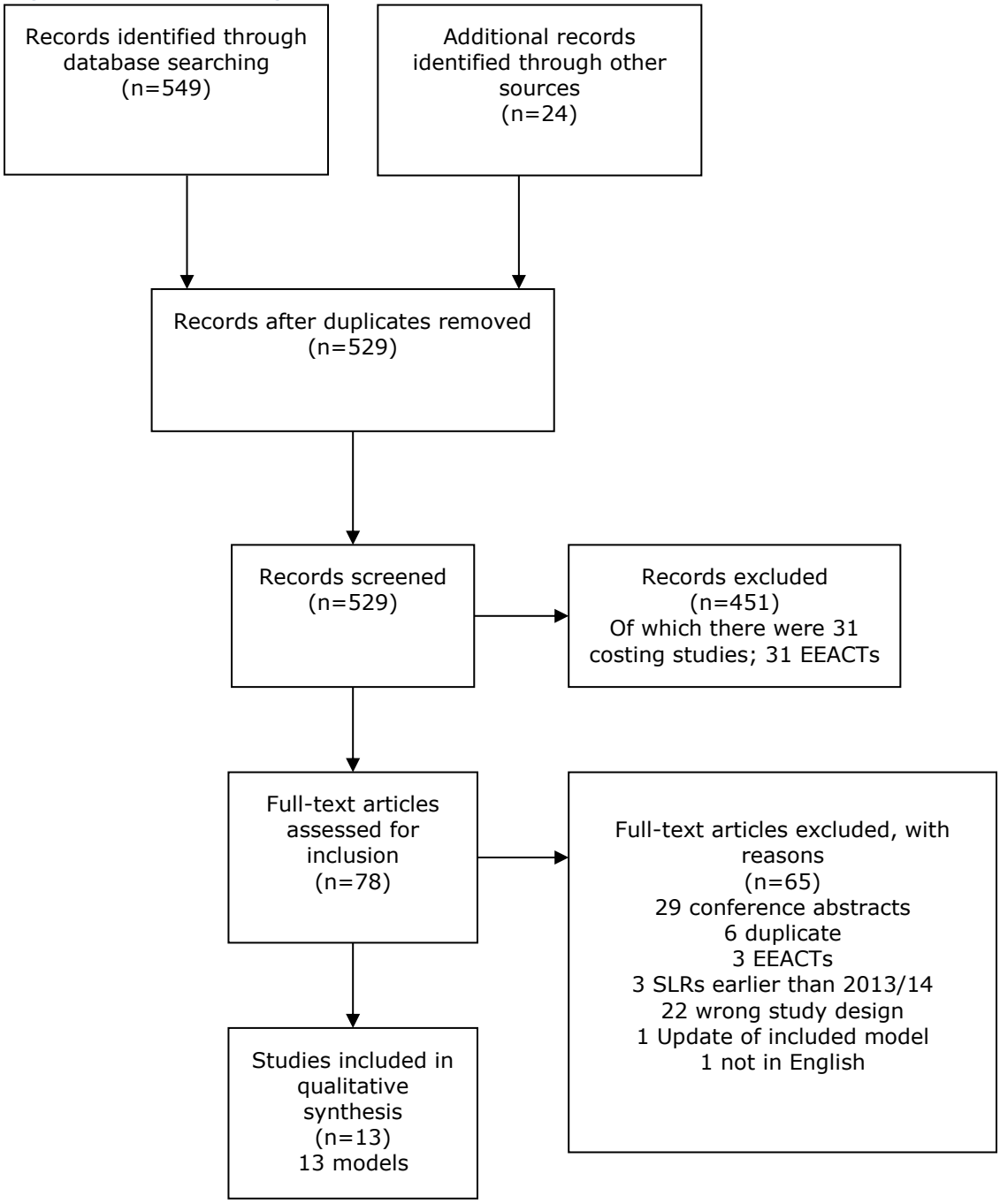
There were 4,133 studies identified initially from both searches after duplicates were removed. These records were screened and 3,834 of them were excluded based on title and abstract review either for not meeting the eligibility criteria or because they were not modelling studies. Studies were excluded if they focussed only on costs or if they were economic evaluations alongside clinical trials (EEACTs). This left 299 articles for full text review; 91 of them were found to meet the eligibility criteria and the following additional criteria: SLRs identified should be from 2013 or later and all other studies had to be models and not costing studies or EEACTs.

Figure 45. PRISMA Diagram for the economic literature review: original search



Abbreviations: PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses, EEACTs = Economic Evaluations Alongside Clinical Trials, SLR = systematic literature review.

Figure 46. PRISMA Diagram for the economic literature review: update search



Abbreviations: PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses, EEACTs = Economic Evaluations Alongside Clinical Trials, SLR = systematic literature review.

Published cost-effectiveness studies included in the review

Of the 91 studies that met the eligibility criteria, 5 were SLRs retrieved as part of the pragmatic search in the original review. These were included so as to capture any additional relevant models for which publications were not identified as part of our literature review. The remaining 86 studies were economic models. A summary of the 9 cost-effectiveness studies which relate to a UK context are presented in Table 97.

The other cost-effectiveness studies identified by the original and update searches are listed in Appendix 21 and Appendix 24 (update).

Of the 9 UK cost-effectiveness studies identified, one was the study conducted as part of NICE TA375. This technology appraisal was a large MTA involving multiple biologic therapies and the cost-effectiveness model informing the appraisal was developed by an independent Assessment Group. This model was accepted by NICE as a basis for decision-making in this indication as part of this appraisal. The model produced as part of TA375 is therefore highly relevant in terms of representing the latest preferences for modelling the use of bDMARDs in RA to inform NICE decision-making. The model produced for TA375 did not, however, include baricitinib. As such, it does not provide any estimates of the cost-effectiveness of this therapy.

The literature review described above had already been conducted by the time of development of this submission and unfortunately was not designed to identify economic evaluations of baricitinib in RA. As a result, a separate targeted literature review was performed to identify published economic evaluations of this therapy. For this review, the Ovid platform was used to search Embase, MEDLINE®, MEDLINE® In-Process & Other Non-Indexed Citations, EconLit and NHS EED (ie. the databases searched as part of the systematic literature review of economic evaluations described previously). The search terms used for this targeted literature review are provided in Appendix 22.

These searches returned 19 hits in total, which were screened against the eligibility criteria presented in Table 96. None of the 19 search results met these eligibility criteria and all 19 hits were therefore excluded. A list of the 19 excluded studies can be found in Appendix 23. Eli Lilly is not aware of any published economic evaluations of baricitinib, and this result therefore aligns with our expectations.

In summary, our searches identified no relevant economic evaluations providing estimates of the cost-effectiveness of baricitinib in RA. Therefore, development of a *de novo* model was required as part of this submission in order to provide these estimates. As noted above, the economic model produced by the AG as part of TA375 would be the most relevant for decision-making for moderate to severe RA in England and Wales and hence the *de novo* economic evaluation developed for this submission is based firmly on the AG's model. Any similarities and differences between the modelling approach for this submission and that of the AG in TA375 are discussed *in situ* in the description of the cost-effectiveness evaluation in the following sections. Although we acknowledge that systematic searches for economic evaluations of baricitinib were not performed, given the availability of details of the AG's model from TA375 and the high relevance of this model to NICE decision-making, we do not anticipate that this omission from the literature review will have had any impact on the design of our *de novo* model for baricitinib in RA.

Quality assessment of included cost-effectiveness studies

Critical appraisals of the 9 UK economic evaluations included in the literature review and considered relevant to this submission were conducted using the checklist adapted from Drummond *et al.* (1996),[126] as recommended by NICE.[125] The results of these critical appraisals are presented in Appendix 25.

Table 97. Summary list of published cost-effectiveness studies

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Diamantopoulos *et al.*[127]	2014	Non-Markov individual patient simulation • Utilities derived from EQ-5D mapped from HAQ scores based on Hernandez et al (2012) model[128] • Treatment sequence model • 6-month cycles • Efficacy in the model was determined by response to treatment, which was subsequently translated to a drop in HAQ score and change in VAS pain • Responders were categorised according to their level of ACR response	Moderate to severe RA (58 years)	Total QALYs Monotherapy: SoC = 8.0162 TCZ 1st line = 8.4987 TCZ 2nd line = 8.5194 Combination therapy: SoC = 8.8609 TCZ 1st line = 8.9050 TCZ 2nd line = 8.8983	Total costs Monotherapy: SoC = £139,008.09 TCZ 1st line = £142,525.23 TCZ 2nd line = £144,744.15 Combination therapy: SoC = £150,665.03 TCZ 1st line = £147,640.97 TCZ 2nd line = £150,127.31 Year of costing: 2011/2012	ICER Monotherapy: SoC TCZ 1st line = £7,289.63 / QALY TCZ 2nd line = £11,400.26 / QALY Combination therapy: SoC = Dominated by TCZ 1st line TCZ 1st line = Dominant TCZ 2nd line = Dominated by TCZ 1st line
NICE MTA TA375 (Stevenson et al)[67]	2013	Non-Markov individual patient simulation • Utilities derived from a two-step process for estimating EQ-5D values from HAQ values: the first step simulated the expected pain score associated with HAQ; the second step estimated EQ-5D based on both HAQ value and pain score • The assessment of treatment	Patients who are either MTX-experienced or naive with moderate or severe RA (NR)	NR	NR Year of costing: NR	PSA Median ICERS= bDMARDs vs MTX alone strategy (population who can receive MTX) Population 2 (severe MTX experienced) = £31,405 to £56,700 / QALY Population 3 (moderate MTX experienced) = £31,900 to £61,900 / QALY

		response was based upon EULAR response at six months • Time cycle was not employed • The model allows only legitimate HAQ scores (the 25 points defined in the 0 to 3 range) with time to a change in HAQ score being a competing risk • Costs and QALYs were assigned according to the HAQ score • Lifetime patient horizon • Evidence of the relationship between EULAR and ACR were sought using individual patient level data				bDMARDs vs MTX alone strategy (population who are treated with monotherapy) Population 2 = £35,500 to £76,100 / QALY Population 3 = £35,400 to £76,400 / QALY
NICE TA280[129]	2013	Non-Markov individual patient simulation • Utilities derived from HAQ score intervals mapped to utility values • Treatment sequence model • Health states were not used in the model. However, different costs and utilities were assigned to patients according to HAQ score at various time points in the model. Therefore, HAQ score intervals for disease-related costs and utilities were used as proxy health states to estimate results	Moderate to severe active RA (51.5 years)	QALYs: Abatacept = 6.06 cDMARD = 4.78 Infliximab = 5.84	Abatacept = redacted cDMARD = £77,199 Infliximab = £101,275 Year of costing: 2010	Abatacept vs cDMARD - £21,450 / QALY Infliximab vs cDMARD - £22,713 / QALY
NICE TA234[130]	2010-	Non-Markov individual patient	Moderate to	QALYs:	Abatacept =	Abatacept vs cDMARD -

	2011	simulation • Utilities derived from HAQ score intervals mapped to utility values • Treatment sequence model • Health states were not used in the model. However, different costs and utilities were assigned to patients according to HAQ score at various time points in the model. Therefore, HAQ score intervals for disease-related costs and utilities were used as proxy health states to estimate results	severe active RA (51.5 years)	Abatacept = 6.16 cDMARD = 4.88 Infliximab = 5.96	£114,548 cDMARD = £76,276 Infliximab = £109,419 Year of costing: 2010	£29,916 / QALY Abatacept vs Infliximab - £25,711 / QALY
NICE TA198[66]	2010	Markov individual patient simulation • Utilities derived from HAQ mapped to EQ-5D using a quadratic equation derived from work in a conference abstract • ACR response is used to reflect disease activity • HAQ is used to reflect disease progression	DMARD-IR (52.5 years); TNF-IR, severely active RA with insufficient response to cDMARDs or TNFi (53.7 years)	QALYs DMARD-IR Sequence with TCZ = 8.946 Sequence without TCZ = 7.775 TNF-IR Sequence with TCZ = 6.591 Sequence without TCZ = 5.381	Total Direct Medical Costs DMARD-IR Sequence with TCZ = £100,485 Sequence without TCZ = £77,231 TNF-IR Sequence with TCZ = £77,232 Sequence without TCZ = £50,592 Year of costing: 2008	ICER DMARD-IR £19,870 / QALY TNF-IR £22,003 / QALY
NICE TA186[131]	2010	Markov Model • Utilities derived from EQ-5D linked to HAQ using regression analysis	Moderate to severe RA who failed in cDMARDs (52.2 years)	QALYs Combination (with MTX) Certolizumab	Total costs Combination (with MTX)	ICER Versus combination (with MTX)

		• ACR response is the primary clinical method for monitoring disease activity • The use of HAQ-DI as a measure for disease progression is used because it is the most well-recorded measure of disease activity in the trials and in all recent cost-effectiveness analyses of RA		pegol = 2.903 Etanercept = 2.908 Adalimumab = 2.801 Rituximab = 2.77 Infliximab = 2.692 <u>Monotherapy</u> Certolizumab pegol = 2.736 Etanercept = 2.782 Adalimumab = 2.609	Certolizumab pegol = £96,417 Etanercept = £93,317 Adalimumab = £96,428 Rituximab = £92,936 Infliximab = £104,460 <u>Monotherapy</u> Certolizumab pegol = £91,820 Etanercept = £95,691 Adalimumab = £90,048 <u>Year of costing</u>: 2009	Etanercept = NA (£197,037 / QALY etanercept vs CTZ) Adalimumab = CTZ dominates Rituximab = £26,157 / QALY Infliximab = CTZ dominates <u>Monotherapy</u> Etanercept = NA (£82,695 / QALY; etanercept vs CTZ) Adalimumab=£13,982 / QALY
NICE TA195[68]	2010	<u>Non-Markov Individual patient simulation</u> • Utilities derived from HAQ mapped to utility values • The Birmingham Rheumatoid Arthritis Model (BRAM) has been further updated to allow for a non-linear relationship between HAQ and utility	Active RA (NR)	<u>Mean QALYs [95% CI]:</u> ADA = 2.89 [-2.25; 7.74] ETN = 2.81 [-2.29; 7.75] IFX = 2.81 [-2.44; 7.73] RTX = 3.10 [-1.91; 7.88] ABA = 3.28 [-1.67; 7.96] DMARDs = 2.14 [-3.47; 7.39]	<u>Mean cost, [95% CI]:</u> ADA = £74,500 [£68,400; £80,500] ETN = £74,800 [£68,700; £81,200] IFX = £72,800 [£65,900; £79,500] RTX = £69,100 [£62,400; £76,300] ABA = £92,800 [£86,000; £99,900] DMARDs = £48,800 [£43,100; £54,600]	<u>ICER:</u> ADA vs DMARDs = £34,300 ETN vs DMARDs = £38,800 IFX vs DMARDs = £36,200 RTX vs DMARDs = £21,200 ABA vs DMARDs = £38,600 ADA vs RTX = RTX dominates ETN vs RTX = RTX dominates IFX vs RTX = RTX dominates ABA vs RTX = £131,000 ADA vs ABA = £47,000

					Year of costing: 2008	ETN vs ABA = £38,400 IFX vs ABA = £42,100 ADA vs ETN = ADA dominates ADA vs IFX = £19,900 ETN vs IFX = £320,000
NICE TA225[64]	2011	Markov Model • Utilities derived from a published regression analysis comparing HAQ to EQ-5D • The model incorporated three health states, no response (sub ACR20), ACR20 and ACR50 • The model estimates a HAQ score for each health state • Adverse events are modelled indirectly	Moderate to severe RA who failed in cDMARDs: DMARD experienced (52 years), TNF inhibitor experienced (54 years)	QALYs DMARD experienced MTX = 4.569 ADA = 5.792 GOL = 5.827 IFX = 5.651 CTZ = 5.768 ETN = 6.133 TNF inhibitor experienced MTX = 3.129 GOL = 3.712 RTX = 3.523	Total costs DMARD experienced MTX = £35,869 ADA = £66,875 GOL = £67,747 IFX = £69,899 CTZ = £73,571 ETN = £74,208 TNF inhibitor experienced MTX = £33,673 GOL = £50,175 RTX = £50,206 Year of costing: 2008	ICER DMARD experienced Incremental ADA = £25,353 GOL = £24,914 IFX = dominated CTZ = £31,385 ETN = £1,745 Versus MTX ADA = £25,353 GOL = £25,346 IFX = £31,464 CTZ = £31,444 ETN = £24,514 TNF inhibitor experienced Incremental GOL = £28,286 RTX = dominated Versus MTX Golimumab = £28,286 Rituximab = £41,935

NICE TA130[132]	2007	Non-Markov Individual patient simulation • Utilities derived from HAQ mapped to utility values • Mortality risk is assumed to depend on current HAQ score, as well as age and gender • Individual variation in HAQ improvement on starting treatment • Time on treatment includes explicit consideration of early quitting; with early quitting owing to lack of effectiveness being correlated with poor HAQ improvement on starting treatment	Both early and late RA patients were examined (NR)	Single TNF inhibitor use vs base strategy of cDMARDs with no inhibitors (QALYs) – 3rd line (late RA data): ADA = 6.5016 ETN = 7.0374 ADA + MTX = 6.7638 ETN + MTX = 7.0347 IFX + MTX = 6.5284 cDMARDs = 6.5772	Single TNF inhibitor use vs base strategy of cDMARDs with no inhibitors – 3rd line (late RA data): ADA = £47,581 ETN = £59,467 ADA + MTX = £48,011 ETN + MTX = £59,671 IFX + MTX = £49,708 cDMARDs = £16,565 For further scenarios, see Tables 45–49 in the AG report. Year of costing: 2004	Single TNF inhibitor use (Cost/QALY gained) – 3rd line (late RA data): ADA vs cDMARDS: cDMARDs dominate ETN vs cDMARDs = £93,200 ADA + MTX vs cDMARDs = £169,000 ETN + MTX vs cDMARDs = £94,200 IFX + MTX vs cDMARDs = cDMARDs dominate ADA + MTX vs ADA = £1,640 ETN + MTX vs ETN = Inconclusive ETN vs ADA = £22,200 ETN + MTX vs ADA + MTX = £43,100 IFX + MTX vs ADA + MTX = ADA + MTX dominates ETN + MTX vs IFX + MTX = £19,700

Abbreviations: QALY = quality-adjusted life year, ICER = incremental cost-effectiveness ratio, HAQ = Health Assessment Questionnaire, VAS = Visual Analogue Scale, ACR = American College of Rheumatology, RA = rheumatoid arthritis, SoC = Standard of Care, TCZ = tocilizumab, MTA = Multiple Technology Assessment, EULAR = European League Against Rheumatism, MTX = methotrexate, PSA = probabilistic median sensitivity analysis, DMARD = disease-modifying antirheumatic drug, bDMARD = biologic DMARD, cDMARD = conventional DMARD, IR = inadequate response, TNF = tumour necrosis factor, BRAM = Birmingham Rheumatoid Arthritis Model, ADA = adalimumab, ETN = etanercept, IFX = infliximab, RTX = rituximab, ABA = abatacept, ACR20/50/70 = ACR 20%/50%/70% improvement.

5.2 *De novo analysis*

5.2.1 Patient population

As described above, a *de novo* economic model was developed for the cost-effectiveness evaluation.

The cost-effectiveness evaluation considers use of baricitinib in four patient populations:

1. Patients with **severely** active RA who have been previously treated with and failed on cDMARDs (cDMARD-IR) including methotrexate;
2. Patients with severely active RA who have been previously treated with and failed on TNFis (TNFi-IR), who are **ineligible for treatment with rituximab**).
3. Patients with **moderately** active RA who have been previously treated with and failed on cDMARDs (cDMARD-IR);
4. Patients with severely active RA who have been previously treated with and failed on TNFis (TNFi-IR) and who are **eligible for treatment with rituximab**);

There are therefore two cDMARD-IR populations (a moderate population and a severe population) and two TNFi-IR populations (both severe, one eligible for rituximab and one ineligible for rituximab) explored in the cost-effectiveness evaluation. Patient populations in the model are stratified by disease severity, based on the DAS28 cut-off points of 3.2–5.1 for moderate and >5.1 for severe RA patients. These populations are consistent with the anticipated marketing authorisation for baricitinib, encompass the populations outlined in the final scope issued by NICE for this appraisal and reflect differing treatment practices for different patient groups.[3]

Baseline characteristics were applied to each patient population based upon the respective baricitinib trials, as included in Section 4.5.2 and Appendix 8. Due to the lack of patient-level data for comparators not included in the baricitinib clinical trial programme, mean cohort characteristics obtained from the patient-level trial data in the baricitinib trials are used for comparator therapies. For cDMARD-IR patients, a weighted average of RA-BUILD and RA-BEAM were used, whilst for patients who were TNFi-IR, RA-BEACON was used. Summaries of the baseline characteristics for the moderate and severe subgroups of each of the simulated populations are presented in Table 98 and Table 99, respectively.

Table 98. Baseline patient characteristics for patients with moderate RA

Population of interest	Proportion of females (%)	Baseline age		Baseline HAQ		Source
		Mean	SD	Mean	SD	
Patients who have failed on cDMARDs	75.01	52.05	12.40	0.98	0.61	Weighted average of JADX[11] and JADV[9]
Patients who have failed on TNFis	81.60	57.13	10.96	1.31	0.51	JADW[12]

Footnotes: The above baseline characteristics show the values used for the comparisons of baricitinib 4 mg dose vs the respective active treatments per population of interest. When simulating comparisons for the baricitinib 2 mg dose, baseline values are changed accordingly to align with the active arm from the respective baricitinib trial, wherever applicable.

Abbreviations: HAQ = Health Assessment Questionnaire, DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, TNF = tumour necrosis factor.

Table 99. Baseline patient characteristics for patients with severe RA

Population of interest	Proportion of females (%)	Baseline age		Baseline HAQ		Source
		Mean	SD	Mean	SD	
Patients who have failed on cDMARDs	79.09	52.89	12.12	1.61	0.63	Weighted average of JADX[11] and JADV[9]
Patients who have failed on anti-TNFis	81.70	55.64	11.00	1.78	0.56	JADW[12]

Footnotes: The above baseline characteristics show the values used for the comparisons of baricitinib 4 mg dose vs the respective active treatments per population of interest. When simulating comparisons for the baricitinib 2 mg dose, baseline values are changed accordingly to align with the active arm from the respective baricitinib trial, wherever applicable.
Abbreviations: HAQ = Health Assessment Questionnaire, DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, TNF = tumour necrosis factor.

In addition, a scenario analysis was also performed in which baseline characteristics for the cDMARD-IR severe population were drawn from the British Society for Rheumatology Biologics Register (BSRBR) database, to explore an approach consistent with that taken in TA375. Such a scenario analysis was not explored for the cDMARD-IR moderate population as patients within the register received biologic therapy and therefore represent a severe population. The baseline characteristics from the BSRBR used in this scenario analysis are presented in Table 100.

Table 100. Baseline patient characteristics from the BSRBR

Population of interest	Proportion of females (%)	Baseline age (mean)	Source
cDMARD-IR, moderate	76.3%	56.096	NICE TA375[14]

5.2.2 Model structure

Model choice and rationale

The *de novo* cost-effectiveness model was constructed in Microsoft Excel (© Microsoft Corporation) with Visual Basic for Applications and is a discrete event simulation (DES) model. DES models use a continuous time approach, in which time jumps from event to event, decreasing simulation time compared to cycle-based approaches (such as Markov models), which record information at each cycle even if no event has occurred. Therefore, DES models may be used to describe complex conditions such as RA, where sequences of treatments need to be modelled. By simulating individual patients and considering their characteristics and disease histories, DES models afford more flexibility to be able to describe the diversity of a population.

The economic SLR (Section 5.1) determined that the majority of key models of relevance to the treatment of moderate to severe RA patients in the UK were developed using a DES approach, including the model developed by NICE for their recent multiple technology assessment regarding the use of biologics in DMARD-naïve and cDMARD-IR patients (TA375).[14] As such, DES was considered to be the most appropriate methodology for modelling cost-effectiveness of baricitinib in moderate-to-severely active RA.

Model structure and flow

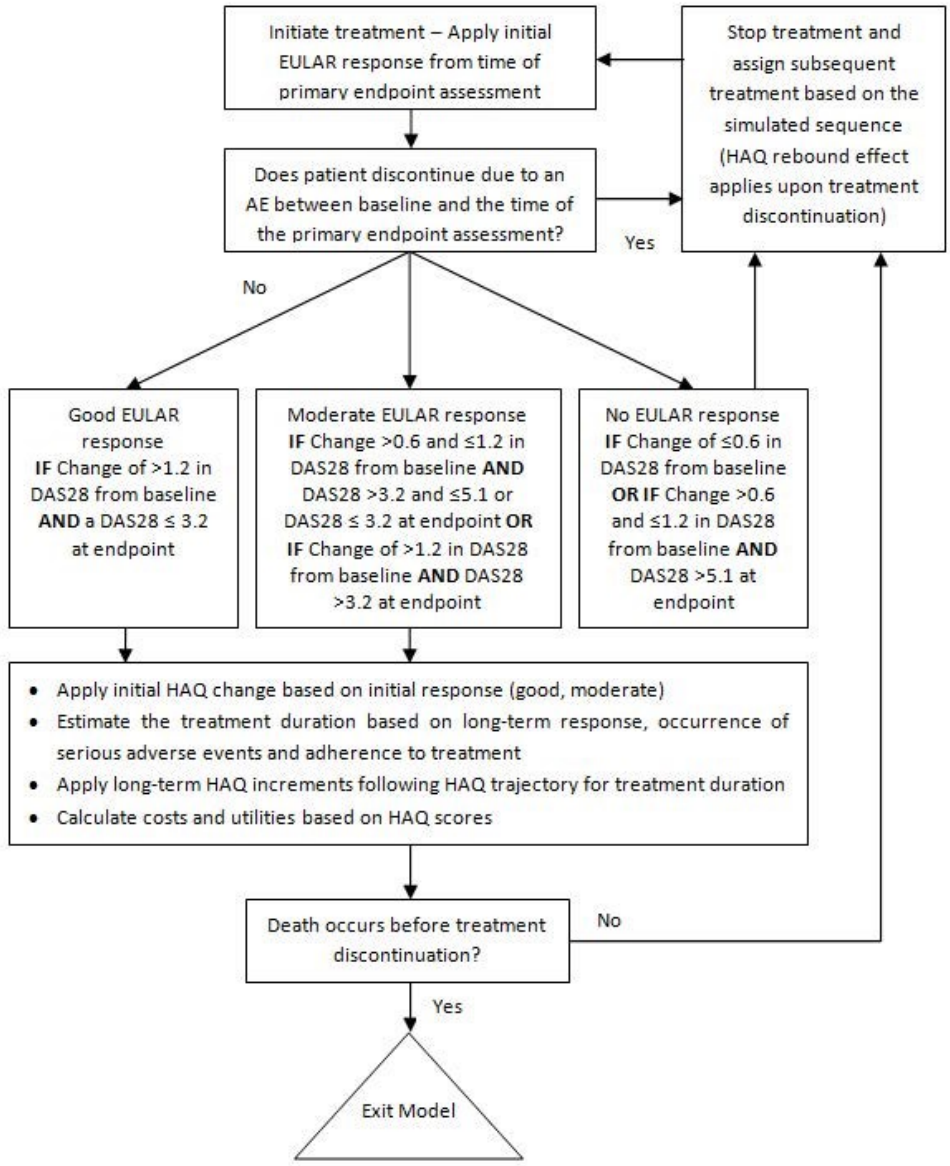
In order to be consistent with the AG approach in the recent bDMARD MTA (TA375), the baricitinib economic model was based directly upon the AG model. A schematic of the baricitinib model is presented in Figure 47, in which initial response to treatment is captured by the EULAR response criteria (good, moderate or non-responder), which are based upon the DAS28 score. A summary of the definition of the EULAR response criteria is presented in Appendix 26. As each patient enters the model, their initial clinical response is estimated from their EULAR response at 24 weeks, consistent with the AG model in TA375. If the patient does not respond to the intervention during the assessment period, the current treatment is terminated and the subsequent treatment sequence is initiated. EULAR response rates at 24 weeks are based upon the output from the 24 week base case NMA reported in Section 4.10.8. It should be noted that, as described in Section 4.10.5, EULAR response inputs for the NMA on the EULAR response outcome were derived from reported ACR scores for a number of interventions, due to a lack of reporting of EULAR data.

In the model, physical function is captured by the HAQ-DI score. HAQ scores, along with pain severity, are then used to determine patient health-related quality of life (HRQOL) within the model. Calculation of a patient's initial change in HAQ score over the duration of the initial assessment period is based on the patient's EULAR response category, in line with the approach adopted by the AG in TA375.[14] Following the initial change in HAQ score, a long-term HAQ trajectory is applied for the period beyond the primary endpoint assessment until either treatment discontinuation or death occurs. In line with the TA375 AG model, long-term HAQ progression in patients receiving biologics was assumed to remain flat whilst on treatment.[14] Conversely, HAQ progression for patients receiving cDMARDs was modelled based on a latent class approach, in an effort to replicate the approach taken by the AG in TA375. Within the model, each patient receiving cDMARDs was assigned to one of the four latent classes with a probability based on the proportion of patients modelled to belong to each class in NICE TA375. HAQ scores for patients receiving cDMARDs were modelled to increase as per the latent class trajectories until year 15. In line with TA375, zero HAQ progression was then assumed after year 15 for patients remaining on cDMARDs beyond this timepoint. Patients receiving baricitinib were modelled to have the same HAQ trajectory as patients receiving biologics, as discussed in Section 5.3.

Time on treatment is calculated by taking into account the rate of treatment discontinuation conditional upon the initial treatment response. Hospitalisation costs and mortality rates were derived from HAQ scores, as in the TA375 model.[14]

Patients experiencing death before treatment discontinuation exit the model and their accumulated costs and QALYs are saved. If treatment is discontinued before death occurs, the current treatment is terminated and the subsequent treatment in the pathway is initiated. A HAQ rebound effect is applied upon treatment discontinuation, assuming that the treatment effect is lost whenever a treatment is terminated.

Figure 47. Model structure



Abbreviations: EULAR = European League Against Rheumatism, HAQ = Health Assessment Questionnaire, DAS28 = Disease Activity Score 28.

Table 101. Features of the de novo analysis

Factor	Chosen values	Justification
Time horizon	45 years	Given the average age of patients in the model, 45 years is expected to represent a lifetime time horizon. This figure is consistent with the NICE reference case[125]
Were health effects measured in QALYs; if not, what was used?	Yes	Consistent with the NICE reference case[125]
Discount of 3.5% for utilities and costs	Yes	Consistent with the NICE reference case[125]
Perspective (NHS/PSS)	NHS	Consistent with the NICE reference case[125]

Abbreviations: NICE = National Institute for Health and Care Excellence, QALY = Quality-Adjusted Life Year, NHS = National Health Service, PSS = Personal Social Services.

5.2.3 Intervention technology and comparators

Baricitinib

In the economic evaluation, the posology of baricitinib has been considered as per the marketing authorisation, and is modelled at a dosage of 4 mg QD. The licence specifies that baricitinib may be administered as monotherapy or combination therapy with MTX. Within the economic evaluation, baricitinib combination therapy with MTX is considered only. The rationale behind this is the paucity of efficacy data in the baricitinib clinical trial programme for patients receiving baricitinib monotherapy, which would be insufficient to form a reliable estimate of efficacy in the modelled populations for baricitinib monotherapy. It may be noted that in the recent MTA regarding the use of biologics in DMARD-naïve and cDMARD-IR patients (TA375), the Committee agreed that the minority of (cDMARD-IR) patients with severely active rheumatoid arthritis who could not tolerate methotrexate should not be treated differently from other people with severe disease, as far as possible. The Committee concluded that biologic DMARDs should be recommended as a cost-effective use of NHS resources when used as monotherapy for severely active disease previously treated with DMARDs, where the marketing authorisation of the bDMARD allows for this recommendation to be made. The economic evaluation of baricitinib presented here assumes that a similar rationale will be applied to baricitinib monotherapy.

The licence is also anticipated to specify that patients aged ≥75 years should receive a 2 mg QD dose of baricitinib. The SmPC notes that there is limited clinical experience in this potentially frailer patient group. The SmPC also notes that patients with a history of chronic or recurrent infection should only be administered a dose of 2 mg QD, as well as patients taking Organic Anion Transporter 3 (OAT3) inhibitors with a strong inhibition potential. A dose of 2 mg once daily is also recommended in patients with renal impairment (creatinine clearance between 30 mL/min and 60 mL/min). Patients with renal impairment were randomised to the 2 mg dose of baricitinib in the trials in which this dose was a treatment option. These patients have not been treated as subgroups within the economic analysis; it is assumed that the outcomes modelled for the full population would be similar for these patient subgroups for whom the 2 mg dose is recommended.

In addition, the licence specifies that a dose of 2 mg once daily may also be considered for patients who have achieved sustained control of disease activity with 4 mg once daily and are eligible for dose tapering. The impact of this tapering for cost-effectiveness of baricitinib is explored through a scenario analysis. For this scenario analysis, patients assessed as EULAR good responders at Week 24 are modelled to continue therapy with baricitinib at the 2 mg dose. This scenario has no impact on costs, due to the flat pricing of the two doses. Howe\cver, the potential impact of a step down in dose on efficacy is modelled through the application of a HAQ decrement of 0.06 upon dose step-down. This HAQ decrement is derived from an analysis of the RA-BEYOND (JADY) study that found the difference in mean change from baseline in HAQ-DI score at week 24 in the step-down period between patients receiving baricitinib 4 mg and those receiving baricitinib 2 mg to be -0.06.

Comparators

In addition to baricitinib, the model enables inclusion of both cDMARDs and bDMARDs. The interventions included in the model are consistent with NICE recommendations for the RA treatment pathway and are in agreement with the NICE scope for baricitinib.[3] Comparator interventions in the two populations of interest are presented in Table 102.[3] All bDMARDs were modelled in combination with MTX. Biosimilars for etanercept and infliximab are approved for

marketing in the UK. They are thus included as comparator treatments in the analysis in place of their branded originator products, as the biosimilars are less expensive and hence this represents a conservative approach to evaluating cost-effectiveness of baricitinib. The modelled treatments were consistent with the treatment pathway presented in Section 3.3.

The source of efficacy data for comparator therapies is discussed in Section 5.3.

Table 102. Comparator treatments included in the model

Population of Interest	Treatment
cDMARD-IR moderate	Combination therapy with conventional DMARDs (cDMARDs) (including methotrexate and at least one other DMARD, such as sulfasalazine and leflunomide)
cDMARD-IR severe	Adalimumab, etanercept biosimilar, infliximab biosimilar, abatacept, golimumab, tocilizumab, certolizumab pegol (all in combination with MTX)
Anti-TNF-IR	*Rituximab-eligible patients*: Rituximab *Rituximab-ineligible patients:* tocilizumab, abatacept, certolizumab pegol, golimumab, etanercept biosimilar, infliximab biosimilar and adalimumab.

Abbreviations: cDMARD = conventional disease-modifying antirheumatic drug, IR = inadequate response, TNF = tumour necrosis factor.

Treatment sequences in the model

The vast majority of recent models in RA investigated comparisons of sequences of treatments. In line with both previous modelling approaches in RA and the latest NICE MTA (TA375),[14] the baricitinib economic model follows treatment sequences that reflect NICE guidance on the RA treatment pathway (see Section 3.3). The comparators to baricitinib in each patient population of interest are placed first in the treatment sequence. The treatment sequences used in the baricitinib model for the cDMARD-IR moderate, cDMARD-IR severe and bDMARD-IR severe populations are presented in Table 103, Table 104 and Table 105, respectively. It should be noted that in the cDMARD-IR severe population, it has been assumed that rituximab would be the second treatment in the sequence for patients treated with baricitinib, in line with the use of rituximab as the second-line treatment following comparator biologics. The principle taken in determining the sequences was that they were consistent between interventions in order to avoid spurious cost-effectiveness estimates driven by having different treatments in the sequence. Due to lack of data, there is no later-line adjustment of efficacy (i.e. the NMA estimates are propagated through the model regardless of the position of the treatment in the sequence).

Table 103. Treatment sequences for the cDMARD-IR moderate population

Sequence	First-line treatment	Second-line treatment	Third-line treatment	Fourth-line treatment/Rescue	Rescue
1	Baricitinib (4 mg or 2 mg QD)	Combination of cDMARDs	MTX	Pall	NA
2	Combination of cDMARDs	MTX	Pall	NA	NA

Abbreviations: DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, QD = once daily, Pall = palliative care, NA = not available

Clearly, patients with moderate disease could progress to severe disease so the comparison sequence is potentially artificial but predicated by current NICE guidance restricting treatment beyond cDMARDs to severe patients (i.e a DAS28 score >5.1). Therefore this sequence looks to assess cost-effectivness of baricitinib in the moderate population assuming that patients do not become eligible for biologic treatment over time.

Table 104. Treatment sequences for the cDMARD-IR severe population

Sequence	First-line treatment	Second-line treatment	Third-line treatment	Fourth-line treatment/Rescue	Rescue
1	Baricitinib (4 mg or 2 mg QD)+	RTX+	TCZ+	MTX	Pall
2	bDMARDs (excluding TCZ)+	RTX+	TCZ+	MTX	Pall
3	TCZ+	RTX+	ADA+	MTX	Pall

Footnote: + = plus MTX
Abbreviations: DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, QD = once daily. Pall = palliative care

Table 105. Treatment sequences for the bDMARD-IR severe population

Sequence	First-line treatment	Second-line treatment	Third-line treatment	Fourth-line treatment/Rescue	Rescue
Rituximab-eligible patients					
1	Baricitinib (4 mg or 2 mg QD)+	TCZ+	MTX	Pall	NA
2	RTX+	TCZ+	MTX	Pall	NA
Rituximab-ineligible patients					
1	Baricitinib (4 mg or 2 mg QD)+	TCZ+	MTX	Pall	NA
2	bDMARDs	TCZ+	MTX	Pall	NA
3	TCZ+	ADA+	MTX	Pall	NA

Footnote: + = plus MTX
Abbreviations: DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, QD = once daily, Pall = palliative care, NA = not available

Continuation rule

No continuation rule is specified in the SmPC for baricitinib. However, in its appraisals of biologic therapies for the treatment of RA, NICE has provided recommendations that include an assessment of response at 6 months to determine treatment continuation or discontinuation. For example, the NICE recommendation following TA375 states that treatment with biologic therapies should be continued “only if there is a moderate response measured using EULAR criteria at 6 months after starting therapy”. Although baricitinib is not a biologic treatment and a continuation rule is not mentioned in the SmPC, given that baricitinib would be used as an alternative to biologic therapies in the treatment pathway, it is anticipated that NICE would apply the same continuation rule for baricitinib as for comparator biologic therapies. Therefore, the base case cost-effectiveness analysis models the assessment of EULAR response at 24 weeks, with patients who exhibit no response being modelled to withdraw from baricitinib therapy and move on to the next treatment in the sequence, which is likely a reasonable assumption of how baricitinib may be used in clinical practice- i.e. if adequate patient response is not achieved, then therapy would be discontinued and an alternative treatment initiated. The timepoint of assessment is explored in a scenario analysis that considers assessment of EULAR response at 12 weeks using the 12 week NMA scenario analysis described in 4.10.9.

5.3 *Clinical parameters and variables*

Clinical response

Clinical response in the model is based on the EULAR response criteria. The use of EULAR response criteria as a measurement of treatment response reflects UK clinical practice; it is recommended in NICE guidance as well as by the British Health Professionals in Rheumatology and the British Society for Rheumatology.[37,133]

EULAR response rates at 6 months for baricitinib and all other active treatments were sourced from the NMA, presented in Section 4.10. The NMA provided EULAR response rates of relevant therapies for both the cDMARD-IR and TNFi-IR populations, for use with the respective modelled populations. NMA results were not stratified by disease severity and therefore results from NMAs were assumed relevant for the analyses in both the moderate and severe populations. For a number of comparators (adalimumab, certolizumab pegol, etanercept and infliximab), data were not available to allow for their inclusion in the NMA for the TNFi-IR population. In order to model these comparators in the economic analysis for the TNFi-IR population, these comparators were modelled as having the same EULAR response in the anti-TNF-IR population as in the cDMARD-IR population in the absence of alternative data. Mean EULAR response treatment effects from the NMA used for the cDMARD-IR population and TNFi-IR population in the base case are provided in Table 106 Moderate only response has been derived from the probabilities of good only and no response from the NMA.

In addition, given the availability of head-to-head data comparing baricitinib and adalimumab from the RA-BEAM RCT, a scenario analysis was performed for the comparison of baricitinib to adalimumab in the cDMARD-IR severe population for which EULAR response rates were sourced directly from this head-to-head comparison. Further details of this scenario analysis are provided in Section 5.8.3.

Table 106. Mean treatment effects for the cDMARD-IR population

Active Treatment	Good	Moderate	No response
BAR4+	[redacted]	[redacted]	[redacted]
BAR2+	[redacted]	[redacted]	[redacted]
ABA sc+	[redacted]	[redacted]	[redacted]
ADA+	[redacted]	[redacted]	[redacted]
CTZ+	[redacted]	[redacted]	[redacted]
ETN+	[redacted]	[redacted]	[redacted]
GOL+	[redacted]	[redacted]	[redacted]
IFX+	[redacted]	[redacted]	[redacted]
RTX+	[redacted]	[redacted]	[redacted]
TCZ+	[redacted]	[redacted]	[redacted]
MTX	[redacted]	[redacted]	[redacted]
COMB	[redacted]	[redacted]	[redacted]
Pall	[redacted]	[redacted]	[redacted]

Footnote: + = plus MTX.
Abbreviations: MTX = methotrexate, ABA sc = abatacept sc, ADA = adalimumab, COMB = combination of cDMARDs, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, Pall = Palliative care, TCZ = tocilizumab, TOFA = tofacitinib, RTX = rituximab.

Table 107. Mean treatment effects for the TNFi-IR population

Active Treatment	Good	Moderate	No response
BAR4+	[redacted]	[redacted]	[redacted]
BAR2+	[redacted]	[redacted]	[redacted]
ABA iv+	[redacted]	[redacted]	[redacted]
ABA sc+	[redacted]	[redacted]	[redacted]
ADA+ (*)	[redacted]	[redacted]	[redacted]
CTZ+ (*)	[redacted]	[redacted]	[redacted]
ETN+ (*)	[redacted]	[redacted]	[redacted]
GOL+	[redacted]	[redacted]	[redacted]
IFX+ (*)	[redacted]	[redacted]	[redacted]
RTX+	[redacted]	[redacted]	[redacted]
TCZ+	[redacted]	[redacted]	[redacted]
MTX	[redacted]	[redacted]	[redacted]
Pall	[redacted]	[redacted]	[redacted]

Footnotes: + = plus MTX, (*) = intervention not part of the anti-TNF-IR NMA and input therefore assumed to be the same as for the intervention in the cDMARD-IR population
Abbreviations: MTX = methotrexate, ABA iv = abatacept iv, ABA sc = abatacept sc, Pall = Palliative care, TCZ = tocilizumab, RTX = rituximab.

Initial change in HAQ score

As introduced in Section 4.3.6.2, physical function – and as a result HRQOL – are captured in the model by the HAQ-DI, a disease-specific patient reported outcome (PRO) measurement tool. HAQ-DI has shown good correlation with the generic EQ-5D HRQOL measurement tool.[134]

In the model, a patient's initial change in HAQ score as a result of a new treatment is calculated for the initial assessment period based on EULAR response category observed at the 6 month assessment time point.[14] A mean reduction in HAQ score of 0.317 (SE 0.048) for moderate responders and 0.673 (SE 0.112) for good responders was applied in the model. The modelling of an initial change in HAQ score dependent upon EULAR response, and the above figures used for initial HAQ change, were derived from analyses of patients treated with bDMARDs from the BSRBR database. This is consistent with the AG model in TA375.[14] The BSRBR database evaluated patients treated with biologic therapies but did not include patients treated with baricitinib; therefore use of this data requires an implicit assumption that initial HAQ changes with baricitinib would be expected to be similar to those observed with biologic therapies. As discussed in Section 4.7.1, in RA-BEAM baricitinib 4 mg QD demonstrated a statistically significant improvement in HAQ-DI score at Weeks 12 (p=0.005) versus adalimumab. Given these results, an assumption that the BSRBR data for initial HAQ scores apply to baricitinib would appear to be reasonable, if not conservative. This assumption is explored in a scenario analysis in which data from the baricitinib clinical trial programme is used to determine the initial change in HAQ score (see Section 5.8.3).

HAQ score trajectory

Subsequent to the initial change in HAQ score upon starting a new treatment, a long-term HAQ trajectory is then applied for the period beyond the assessment time point to the point at which either discontinuation of treatment or death occurs.

For HAQ progression whilst receiving cDMARDs or palliative care, previous economic models have commonly used an assumption whereby treatment benefit wanes (and HAQ score increases) linearly.[135] This assumption, however, results in patients' HAQ scores deteriorating to the worst possible state, which is seldom observed in clinical practice. To account for this fact, the approach taken by the AG in TA375[14] was used for this model. In the AG's approach, HAQ progression was stratified for different classes of patients using a latent class approach. The different classes were derived from a growth mixture model based on a study by Norton *et al.*[136] A modified analysis based on this study was performed including additional covariates to estimate the relevant class membership per patient according to individual characteristics.

An effort was made to replicate this stratified latent class approach for this baricitinib cost-effectiveness model, as the modelling of HAQ progression on cDMARDs in this manner is aligned to observations from the Early Rheumatoid Arthritis Study (ERAS), Early Rheumatoid Arthritis Network (ERAN), The Norfolk Arthritis Register (NOAR) and the US National Data Bank for Rheumatic Diseases (NDB) datasets.[14] For this model, data from Norton et al[136] were digitised and datasets were generated for all four patient classes based on the observed estimates. Excel was used to fit polynomial trendlines, which were found to be the best-fitting trendlines, and then parametric equations were generated from these trendlines to produce a curve for each latent class (Appendix 27). Subsequently, the area under the curve (AUC) for each of the four curves was calculated using a time-step of 0.01 years, starting from Year 2 onwards. In TA375, the ERAS cohort included the period in which patients experience a reduction in HAQ due to initiation of treatment. Since the cost-effectiveness model explicitly models this initial change in HAQ score, only values after that point were used; this was estimated to be from Year 2 onwards. After calculating the AUC for each of the four curves up to Year 15, the change in HAQ score was estimated for each of the latent classes at different time points. The generated tables of HAQ changes for fractions of time were then used to predict HAQ progression over the time for patients remaining on cDMARDs. In line with TA375, zero HAQ progression was assumed after Year 15 for patients remaining on cDMARDs for a longer period.[14]

Finally, to explore the impact of applying a simpler method of modelling HAQ progression on cDMARDs on cost-effectiveness results, a scenario analysis was conducted in which a linear rate of HAQ progression was applied for patients receiving cDMARDs.

For patients receiving bDMARDs, long-term HAQ progression after the assessment time point was assumed to remain flat while on treatment, and the same assumption is made for patients on baricitinib. This assumption is derived from three-year follow-up data collected as part of the BSRBR database, in which HAQ measurements were recorded every 6 months for up to 3 years. These data suggested no HAQ score progression during treatment with bDMARDs. This approach to the modelling of a flat HAQ trajectory whilst on treatment with bDMARDs is identical to the approach taken by the AG in their model in TA375.[14] The same assumption for baricitinib is supported by data from the extension study RA-BEYOND (JADY), demonstrating that patients' HAQ-DI scores are maintained on a long-term basis (Section 4.7.4). Nevertheless, a scenario analysis is additionally conducted to assess the impact of assuming a non-flat HAQ trajectory for patients receiving baricitinib. In this scenario analysis, a linear progression in HAQ is assumed for patients receiving baricitinib. In the absence of any empirical data to support such a rate, this scenario assumes that the linear rate of progression in HAQ with baricitinib is half the linear rate that is applied for the cDMARDs in the scenario analysis assuming linear HAQ progression for cDMARDs (see above). We consider that this represents a higly pessimistic assumption for baricitinib based on the clinical data on HAQ progression from RA-BEYOND, and it should also

be noted that the preference for modelling HAQ progression on cDMARDs is a latent class approach (see above) – assumptions of the same linear progression in HAQ for all patients therefore represent an over-simplification. This scenario analysis is therefore very much included as an exploratory analysis.

Long-term treatment discontinuation

The baricitinib cost-effectiveness model follows the approach from TA375 with regards to long-term discontinuation on treatment, in which the duration of treatment on the first biologic for adult RA patients was estimated using the BSRBR database (which records the dates on which therapies are initiated and ended). An advantage of using this approach is that the duration of treatment was calculated using separate analyses for good and moderate responders, resulting in separate discontinuation curves for the good and moderate EULAR responses. Due to potential confounding, treatment-specific discontinuation data from BSRBR was deemed inappropriate by the Assessment Group in TA375.[14] Thus, the same rate was applied for bDMARDS and cDMARDs.

The same approach is taken in the cost-effectiveness model for baricitinib presented here. The plots for continuation on therapy used in TA375 were digitised and a range of parametric models fitted to the digitization. Model fit statistics can be found in Appendix 28. The Weibull distribution was chosen as it provided good fit to the data and the most conservative estimates for long-term extrapolation compared to other models.

In order to explore an alternative approach to modelling discontinuation of baricitinib, based on data for baricitinib specifically, a scenario analysis was conducted in which a fixed annual discontinuation rate was applied for the baricitinib arm in the model. The discontinuation rate was derived from the relevant baricitinib trials for each respective population in which this scenario analysis was conducted.

The model assumes that initial HAQ improvement is lost whenever a treatment is terminated. Therefore, treatment effects on initial HAQ score following treatment discontinuation are lost, thus rebounding immediately to the level prior to initiation of the terminated therapy.

Mortality

Hazard ratios reported in TA375[14] were applied to the life tables from the UK Office for National Statistics (ONS).[137] The hazard ratios for mortality based on baseline HAQ category are presented in Table 108. This approach assumes that only baseline HAQ score, and not changes in HAQ, affect mortality and is consistent with the approach used by the AG in TA375.[138]

Table 108. Hazard ratios for mortality associated with HAQ category

Initial HAQ Category	Hazard Ratio (95% Confidence Interval)
0.000	1 (1 – 1) referent
0.125 – 0.375	1.4 (1.1 – 1.8)
0.500 – 0.875	1.5 (1.2 – 1.9)
1.000 – 1.375	1.8 (1.4 – 2.2)
1.500 – 1.875	2.7 (2.2 – 3.5)
2.000 – 2.375	4.0 (3.1 – 5.2)
2.500 – 3.000	5.5 (3.9 – 7.7)

Abbreviation: HAQ = Health Assessment Questionnaire.

5.4 *Measurement and valuation of health effects*

5.4.1 Health-related quality-of-life data from clinical trials

The EQ-5D-5L questionnaire was used to collect HRQOL data in the three baricitinib phase III trials presented in Section 4.7. Patients were scheduled to complete the EQ-5D-5L questionnaire during treatment visits at Week 1 (baseline) and then every 4 weeks from Week 4 onwards up to Week 52 for RA-BEAM[9] and up to Week 24 for RA-BUILD and RA-BEACON.[11,12]

The EQ-5D descriptive system comprises the following five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, which are assessed across 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems.[139] Using appropriate country-specific value weighting algorithms, a respondent's self-described health state can be converted into a utility representing the societal desirability of his/her own health. Patient-level EQ-5D-5L responses from the baricitinib phase III trials were converted to utility index-based scores using the UK-specific scoring algorithm as reported in Szende *et al.* (2007). A summary of the EQ-5D-5L results is included in Section 4.7, by individual baricitinib trial.

5.4.2 Mapping

Although EQ-5D data were available for baricitinib from the clinical trial programme, the established approach to modelling HRQOL in RA is via mapping of HAQ scores on to EQ-5D, as was done for TA375. This approach fits with the DES model framework in which HAQ progression is simulated over continuous time and in which there are not defined "health states" to which specific EQ-5D utility values can be directly attributed. Therefore, for the cost-effectiveness evaluation of baricitinib HRQOL was modelled using the standard approach of mapping to EQ-5D from HAQ.

Mapping algorithm based on RA-BEAM

Initially, development of a mapping algorithm to convert HAQ-DI scores to EQ-5D values using data from the RA-BEAM (JADV) baricitinib trial was explored, given that patient-level data were available for both baricitinib and a biologic comparator (adalimumab) from this trial. A summary of the exploration of this algorithm is provided below. Full details are provided in Appendix 29.

The individual patient level data contained HAQ-DI scores and EQ-5D-5L index values based on UK preference weights for each patient at baseline, Week 12 and Week 24. Based on this trial data for HAQ and EQ-5D scores, fixed and random effects regression models were explored using HAQ and HAQ^2 as explanatory variables for a dependent variable of EQ-5D score. The results of a Hausman test determined that the fixed effects model should be used.

This fixed effects model correctly predicted a negative relationship of HAQ-DI to EQ-5D and explained 53% of the variation in the EQ-5D estimates. However, predictive statistics demonstrated that there was large variation in the model's predicted EQ-5D estimates, leading to lower mean estimates. In addition, it was noted that the ceiling effect associated with EQ-5D, where most patients achieve an EQ-5D score of 1, was not captured. Therefore, it was deemed likely that the model was underestimating EQ-5D scores. Furthermore, use of this algorithm would not have been consistent with the approach taken in TA375. Therefore, the decision was taken to adopt the mapping approach taken in TA375 (see below) for the base case.[14] Use of the mapping algorithm based on RA-BEAM as described above was explored in a scenario analysis.

Mapping algorithm as per TA375 – base case

Estimates of EQ-5D scores from HAQ data were calculated by the Assessment Group in TA375[14] based upon the four-class mixture model of Hernandez et al,[134] as this study has been found to perform better than linear models for estimating EQ-5D. Both this mapping algorithm and an algorithm based on the three-class model reported previously in Hernandez et al 2012[128] were examined for use in the baricitinib cost-effectiveness model. Given that the detailed code of the two statistical models was not publicly available, an effort was made to replicate the two models based on the published coefficients and their underlying equations. The three-class model by Hernandez et al[128] was used in the model as the predicted values for selected combinations of covariates were reported in the published article, thus allowing for direct validation of these results with those generated from the replicated mapping algorithm.

The model uses data from the ERAS database, and describes three separate latent classes that demonstrate different relationships between the EQ-5D questionnaire, function, and pain. The Hernandez et al 2012 paper highlights how the inclusion of pain as a separate covariate within the models allows for better estimates of EQ-5D, and the AG in TA375 agrees with the importance with including pain as an explanatory variable for estimation of EQ-5D. The use of the Hernandez et al 2012 model for this cost-effectiveness analysis therefore aligns with the preferred approach of inclusion of pain as well as HAQ score in the estimation of EQ-5D.

In the three-class model of Hernandez et al 2012, Class 1 has a mean HAQ value of 2.26, indicating substantial disability, and a high pain score mean of 72.1, reflecting the element of the data distribution at the bottom of the EQ-5D questionnaire scale. Class 2 has the least disability as measured by HAQ and the least amount of pain, whilst class 3 has moderate pain (mean 33.3) and functional disability (mean HAQ value of 1.24). A summary of the parameter estimates for the three-class model, derived from Hernandez et al 2012,[128] is presented in Table 109. These include parameters relating to HAQ score, VAS pain score, age and gender.

Table 109. Parameter estimates for the three-class model for estimating EQ-5D from HAQ scores

Class	Parameter	Estimate (SE)
Within subject		
Latent class 1	HAQ	-0.062 (0.015)
	HAQ^2	- (-)
	VAS pain/100	-0.295 (0.030)
	σ_u^2	0.015 (0.002)
Latent class 2	HAQ	-0.245 (0.044)
	HAQ^2	0.068 (0.019)
	VAS pain/100	-0.105 (0.134)
	σ_u^2	0.006 (0.001)
Latent class 3	HAQ	-0.16 (0.013)
	HAQ^2	0.025 (0.005)
	VAS pain/100	-0.056 (0.018)
	σ_u^2	0.003 (0.000)
Between subject		
Latent class 1	Intercept	0.343 (0.037)

Latent class 2	Intercept	0.990 (0.025)
Latent class 3	Intercept	0.806 (0.011)
All classes	$\left(\frac{Age - 54.32}{10}\right)$	0.007 (0.002)
	$\left(\frac{Age - 54.32}{10}\right)^2$	0.004 (0.001)
	Male	-0.012 (0.006)
	σ_u^2	0.002 (0.000)
Within-subject categorical latent variables		
Latent class 1	Intercept	-5.201 (0.423)
	HAQ	2.868 (0.178)
	VAS pain/100	5.179 (0.433)
Latent class 2	Intercept	2.203 (0.312)
	HAQ	0.485 (0.214)
	VAS pain/100	-11.366 (4.227)
AIC	-2051.11	N/A
BIC	-1911.05	N/A
ME (SD)	-0.0003 (0.192)	N/A
MAE (SD) [% improvement]	0.1438 (0.128)	[4%]
RMSE [% improvement]	0.1923	[1%]

Abbreviations: AIC = Akaike information criteria; ALDVMM = adjusted limited dependent variable mixture model; BIC = Bayesian information criteria; EQ-5D = EuroQol five-dimensional; HAQ = Health Assessment Questionnaire; MAE = mean absolute error; ME = mean error; RE = random effects; RMSE = root mean square error; SE = standard error; VAS = visual analogue scale. **Source:** Hernandez et al 2012[128]

Mapping algorithm as per Malottki *et al.*

As described above, the mapping algorithm from Hernandez *et al.* was used in the base case in order to align most closely with the approach taken in TA375. A mapping algorithm derived from RA-BEAM trial data was considered in a scenario analysis. As a final exploration of the impact of using an alternative mapping algorithm from HAQ to EQ-5D, a scenario analysis was conducted using the quadratic mapping mechanism used by Malottki *et al.* in the Birmingham Rheumatoid Arthritis Model (BRAM) for a HTA of biologics for the treatment of RA following failure of TNFi inhibitor.[135]

5.4.3 Health-related quality-of-life studies

Search strategy

A systematic literature review was conducted to simultaneously identify health state utility values (HSUVs) and cost and resource use (CRU) data for patients with moderate-to-severe RA. The systematic review was performed in accordance with the methodological principles of conduct for systematic reviews as detailed in the University of York CRD's "Guidance for Undertaking Reviews in Health Care".[140] Details of the HSUV part of the review are presented here in Section 5.4.3. Details of the CRU part of the review are presented in Section 5.5.1.

The following electronic databases were searched on the 1st September 2016:

- MEDLINE, MEDLINE In-Process, MEDLINE Daily and MEDLINE Epub Ahead of Print (1946 to 1st September 2016)
- Embase (1974 to 1st September 2016)
- The Cochrane Library, including the following:
    - Health Technology Assessment (HTA) Database (Issue 3 of 4, July 2016)
    - NHS Economic Evaluation Database (NHS-EED) (Issue 2 of 4, April 2015)
- EconLit (1886 to September 2016)

Searches of MEDLINE, MEDLINE In-Process, MEDLINE Daily, MEDLINE Epub Ahead of Print and Embase were run simultaneously via the Ovid SP platform. The Cochrane Library databases were searched via the Wiley Online platform, and EconLit was searched using the EBSCO platform.

In addition to the electronic database searches, manual searches of abstracts presented in the last two years at the following major rheumatology and pharmacoeconomics congresses were conducted on 12th October 2016:

- European League Against Rheumatism (EULAR) – Annual European Congress of Rheumatology
- American College of Rheumatology (ACR) Annual Meeting
- The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) – Annual European Congress and Annual International Meeting

The NICE website was searched for previous relevant HTA submissions on 17th October 2016 and the following online databases were also searched on 17th October 2016 to ensure that no relevant publications had been missed in the other searches:

- The Cost-effectiveness Analysis (CEA) Registry, managed by Tufts Medical Center (available at healtheconomics.tuftsmedicalcenter.org/cear4/SearchingtheCEARegistry/SearchtheCEARegistry.aspx)
- The University of Sheffield Health Utilities Database (ScHARRHUD; available at www.scharrhud.org/)
- The EQ-5D Publications Database (available at www.euroqol.org/eq-5d-publications/search.html)

Finally, the bibliographies of systematic reviews and meta-analyses identified during the course of the review were hand-searched for references to other relevant studies for inclusion.

Full search strategies for the systematic review are provided in Appendix 30.

Study Selection

To be included in the utilities part of the systematic review, articles had to meet the pre-defined eligibility criteria detailed in Table 110.

The citations found through the searches were first assessed against the eligibility criteria by two independent reviewers based on abstract and title. Where the applicability of the inclusion criteria

was unclear, the article was included at this stage in order to ensure that all potentially relevant studies were captured. Full-text copies of publications potentially meeting the eligibility criteria were then obtained and reviewed in more detail by the two independent reviewers. At both the title/abstract and full-text review stages, any disagreements between the reviewers were resolved by discussion until a consensus was met, with a third reviewer making the final decision if necessary. For studies meeting the eligibility criteria after the second (full-text) screening stage, data were extracted by a single reviewer into a pre-specified data extraction grid and verified by a second individual.

Table 110. Eligibility criteria for the HSUV systematic review

Domain	Inclusion Criteria	Exclusion Criteria	Rationale
Population	Patients with moderate-to-severe RA, as defined by DAS28 criteria (>3.2DAS28≥5.1)	• Articles that do not include patients with moderate-to-severe RA defined by DAS28 • Articles reporting mixed populations were not included unless results were presented separately for those with moderate-to-severe RA	Only studies on patients with moderate-to-severe RA are relevant for the purposes of this submission
Intervention	Any or none	None	Both non-treatment specific and treatment specific utility values are relevant for the purposes of this submission
Comparator	Any or none	None	

Outcomes	Original health state utility data, obtained from EQ-5D or mapped from HAQ to EQ-5D*	• HSUV data not reported
		• HSUV data derived using other instruments eg. SF-6D, HUI3, TTO or SG, VAS
		• No useful HSUV data reported. For example:
		  ○ Article presents only previously published data
		  ○ Study is methodological only
			HSUV studies were required to align as closely as possible with the NICE reference case.

Study design	Primary research studies (eg. discrete choice experiments, observational studies, cross-sectional studies, randomised controlled trials [RCTs] and non-RCTs)	Publications without original data, comments, letters, editorials and non-systematic or narrative reviews, case studies, case reports or case series	The study designs specified as eligible for inclusion were those considered most likely to report relevant data for this submission
	Systematic reviews and HTAs were included at the title/abstract screening stage and used for identification of any additional primary studies not identified through the database searches, but were excluded during the full-text review if not presenting a novel analysis		
Language	English	Any other language	The review team did not have the

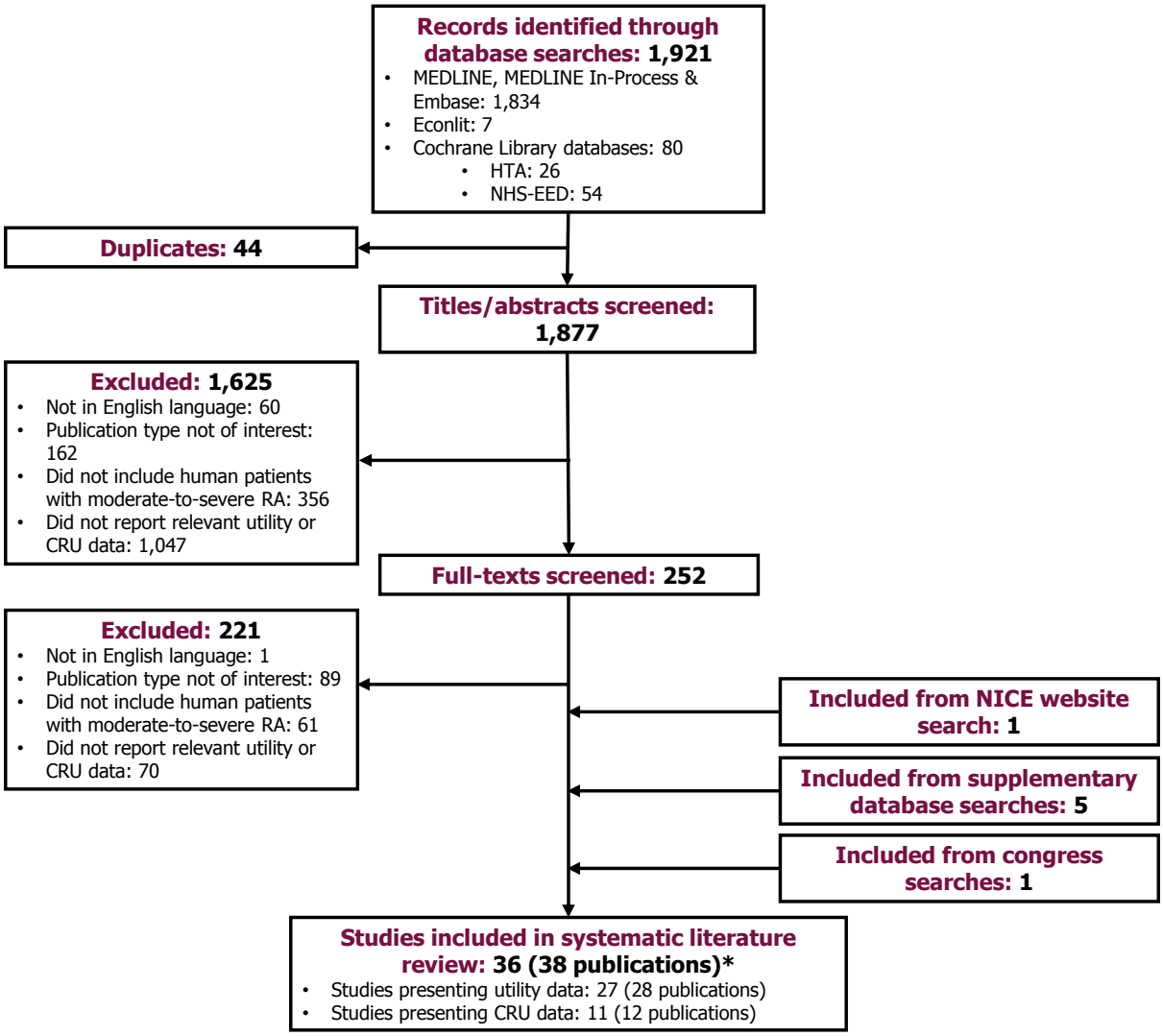
			linguistic capability to review non-English language articles; however, studies were not limited to those conducted in specific geographical locations

Footnotes: *In line with the specifications of the NICE reference case, which state that the EQ-5D is the preferred measure of HRQOL in adults, the SLR was designed to identify EQ-5D data as a priority. However, if a paucity of EQ-5D data was found, it was planned that the Systematic Reviewers could revisit utility studies excluded during the course of the review for the reason that they had used HRQOL instruments other than EQ-5D. Ultimately this action was unnecessary due to the abundance of relevant EQ-5D data found.
Abbreviations: EQ-5D = EuroQoL-5 Dimension, HAQ = health assessment questionnaire, HRQOL = Health-Related Quality of Life, HSUV = health state utility value, HTA = health technology assessment, HUI = Health Utilities Index, NICE = National Institute for Health and Care Excellence, RA = rheumatoid arthritis, SF-6D = Short Form-6 Dimension, SG = standard gamble, TTO = time trade-off, VAS = visual analog scale.

Results

The PRISMA flow diagram for the SLR is presented in Figure 48. A total of 38 publications were identified reporting on 36 separate studies. Of these, 28 publications reporting on 27 studies presented HSUV data.[141-168] Of the 28 publications reporting HSUV data, 2 of these also reported CRU data.[154,164]

Figure 48: PRISMA Flow Diagram of HSUV and CRU Studies



*Of the 28 publications presenting utility data, 2 publications (referring to 2 different studies) also presented CRU data.

Summaries of the studies reporting HSUV data are provided in Appendix 31 and a list of the 221 publications excluded at the full-text review stage is presented in Appendix 32. As previously noted, the established approach to modelling utility in DES models in RA is through mapping via HAQ score. This approach was also considered most appropriate for the cost-effectiveness evaluation of baricitinib. As such, literature-derived health state EQ-5D utility values were not considered to provide useful information for the health economic model, hence the presentation of the results of this search in the appendices.

5.4.4 Adverse reactions

Serious adverse events (SAEs) were not considered in the base case analyses due to the challenges faced conducting an NMA on SAEs, as these are either differentially defined or inconsistently reported in trials. This meant that reliable estimates of the relative rates of SAEs could not be obtained. Furthermore, in previous economic analyses in RA, including TA375, adverse events have not been assumed to be key model drivers and have sometimes been excluded on the assumption that there is no difference in the safety profiles of bDMARDs. In NICE MTA TA375, only serious infections were assumed to generate a significant cost and disutility and hence only these adverse events were included in the analysis.

The head-to-head comparison of baricitinib to adalimumab, presented in Section 4.12, demonstrated that baricitinib and adalimumab have comparable safety profiles. Furthermore, the absolute rate of SAEs observed in the baricitinib clinical trial programme was low (see Section 4.12). The exclusion of SAEs from the base case analysis was therefore considered appropriate.

In the baricitinib economic model, the impact of SAEs on cost-effectiveness was assessed in a sensitivity analysis of the cDMARD-IR population, since comparative data were available from the RA-BEAM trial. SAE rates observed for adalimumab were assumed to apply at a class level (i.e. to all biologics) whereas the rates observed for the placebo arm were used for the cDMARDs and the palliative care in a conservative assumption. To this end, rates of SAEs per patient-year for use in the model were derived from the RA-BEAM trial data. This gave rise to rates of 0.152, 0.049 and 0.095 for baricitinib 4 mg QD, all biologics and cDMARDs/palliative care, respectively. These rates, alongside a random number generated from a uniform distribution, were then applied to the inverse exponential distribution to generate the time to a serious adverse event, assuming a constant hazard rate.

5.4.5 Health-related quality of life data used in cost-effectiveness analysis

It is preferential to use HRQOL data mapped from patients treated with baricitinib, and relevant comparators, over HRQOL data from databases of RA patients, which may not take into account other treatment-specific factors affecting HRQOL. Use of a mapping algorithm based on patient-level data from the RA-BEAM (JADV) clinical trial was therefore explored; however it was deemed that the algorithm developed underestimated patient EQ-5D scores and use of this mapping algorithm based on trial data would also be inconsistent with the approach taken in TA375 (see Section 5.4.2). Therefore, the HAQ-DI to EQ-5D mapping approach taken by the AG in TA375 was adopted, as described in Section 5.4.2.[14] Mapping was based on a three class model reported by Hernandez *et al.* 2012, the parameters for which are presented in Section 5.4.2.

5.5 *Cost and healthcare resource use identification, measurement and valuation*

Resource use included in the model was as follows:

- Drug acquisition costs
- Drug administration costs
- Drug monitoring costs
- Hospitalisation cost (dependent upon HAQ score)

This aligns with the resource use inputs considered in the AG's model in TA375. The only exception to this is the cost of SAEs, which was included in the base case of the AG's model but is included as a scenario analysis here as discussed in Section 5.8.3.

5.5.1 Resource identification, measurement and valuation studies

Search strategy and study selection

As previously described, a single systematic literature review was conducted to simultaneously identify HSUVs and cost and resource use data for patients with moderate-to-severe RA. Details of the search strategy and screening methodology are provided in Section 5.4.3 and Appendix 30.

It should be noted that the cost and resource use inputs used in the de novo economic analysis presented in this submission were derived from previously published models developed in support of NICE HTAs, and which have previously been reviewed and accepted by NICE. Therefore, when designing the systematic review, a pragmatic date limit was applied to limit the cost and resource use data identified to that collected since 2011, with an aim of identifying only cost and resource use data published since the previous models were developed.

To be included in the cost and resource use part of the systematic review, articles had to meet the pre-defined eligibility criteria detailed in Table 111.

Table 111. Eligibility criteria for the CRU systematic review

Domain	Inclusion Criteria	Exclusion Criteria	Rationale
Population	Patients with moderate-to-severe RA	• Articles that do not include patients with moderate-to-severe RA • Articles reporting mixed populations were not included unless results were presented separately for those with moderate-to-severe RA	Only studies on patients with moderate-to-severe RA are relevant for the purposes of this submission
Intervention	Any or none	None	Both non-treatment specific and treatment specific costs and resource use data are relevant for the purposes of this
Comparator	Any or none	None	

			submission
Outcomes	Original costs and resource use data published in or since 2011 which were relevant to the UK NHS and PSS, and to a cost-effectiveness model for baricitinib. Specifically: • Hospitalisation • Inpatient visits • Joint replacement • Medication use and administration • Monitoring • Management of serious infections • Productivity losses	Articles not reporting original, relevant cost or resource use data	Only costs and resource use relevant to the cost-effectiveness model for baricitinib, the UK NHS and PSS were required for the purposes of this submission
Study design	Primary research publications (eg. observational studies, cross-sectional studies, randomised controlled trials [RCTs] and non-RCTs)	Publications without original data, comments, letters, editorials and non-systematic or narrative reviews, case studies, case reports or case series	The study designs specified as eligible for inclusion were those considered most likely to report relevant data for this submission
	Systematic reviews and HTAs were included at the title/abstract screening stage and used for identification of any additional primary studies not identified through the database searches, but were excluded during the full-text review if not presenting a novel analysis		
Language	English	Any other language	The review team did not have the linguistic capability to review non-English language articles; however, studies were not limited to those conducted in specific geographical locations

Abbreviations: EQ-5D = EuroQoL-5 Dimension, HAQ = Health Assessment Questionnaire, HSUV = health state utility value, HTA = health technology assessment, HUI = Health Utilities Index, NHS = National Health Service, NICE = National Institute for Health and Care Excellence, PSS = Personal and Social Services, RA = rheumatoid arthritis, SF-6D = Short Form-6 Dimension, SG = standard gamble, TTO = time trade-off, UK = United Kingdom, VAS = visual analog scale.

Results

A total of 12 publications reporting on 11 studies presented CRU data. The PRISMA flow diagram is presented in Figure 48 in Section 5.4.3.

Summaries of the studies reporting CRU data are provided in Appendix 33 and a list of the 221 publications excluded at the full-text review stage is presented in Appendix 32.

5.5.2 Intervention and comparators' costs and resource use

The perspective adopted for the analysis was that of the UK NHS and personal social services. All costs were reported in British Pound Sterling with a 2016 price base. Wherever costs were not available for the most recent year but only from previous years, values were inflated to 2016 prices using the Personal Social Services Research Unit (PSSRU) hospital & community health services (HCHS) inflation rate index.[169]

Drug treatment costs

Treatment costs accounted for in the model consist of drug acquisition costs, administration costs and monitoring costs. For some drugs, additional start-up costs apply in the first year of treatment associated with loading dose requirements. Annual and start-up costs for each drug considered in the model are presented in Table 113 below.

For weight-based treatments, which include tocilizumab (intravenous formulation), abatacept (intravenous formulation), infliximab and golimumab, the average baseline weight of patients from all treatment arms of the baricitinib clinical trial(s) corresponding to the appropriate population was used, as presented in Table 112.

For tocilizumab, a subcutaneous formulation is available in addition to the intravenous formulation; however, only the intravenous formulation was modelled as the subcutaneous formulation is associated with a confidential PAS and it is therefore not possible to model the cost of this comparator accurately. For abatacept, both subcutaneous and intravenous formulations are also available. For simplicity, only the subcutaneous formulation was considered in the economic analysis.

Finally, as detailed in Section 2.3, a Simplified discount PAS scheme has been submitted to PASLU for baricitinib.

Table 112. Average baseline weight of patients from baricitinib phase III trials

Population	Trials	Baseline Weight (kg)
cDMARD-IR, moderate	RA-BEAM and RA-BUILD	70.54
cDMARD-IR, severe	RA-BEAM and RA-BUILD, weighted average for baricitinib 4 mg	72.15
bDMARD-IR, moderate	RA-BEACON	82.48
bDMARD-IR, severe	RA-BEACON	81.80

Table 113. Drug acquisition costs

	Unit per vial/tab	Vials / tablets per pack	Cost per pack	Cost per vial / tab	Vials /tab year 1	Vials/tabs per subsequent year	Additional doses in year 1	Annual cost	Add. cost in Year 1	Source	Comment
BAR (list price)	4 mg	28	£805.56	£28.77	365	365	0	£10,501.05	£0.00		List price, not taking into account PAS
BAR (list price)	2 mg	28	£805.56	£28.77	365	365	0	£10,501.05	£0.00		List price, not taking into account PAS. Included in scenario analysis where baricitinib dose tapered to 2 mg.
BAR (PAS price)	4 mg	28	[redacted]	[redacted]	365	365	0	[redacted]	£0.00		Takes into account confidential PAS.
BAR (PAS price)	2 mg	28	[redacted]	[redacted]	365	365	0	[redacted]	£0.00		Takes into account confidential PAS. Included in scenario analysis where baricitinib dose tapered to 2 mg.
ABA sc	125 mg	1	£302.40	£302.40	52	52	0	£15,724.80	£0.00	BNF (May 2016)	Assumed weight 60-100 kg to align with baricitinib baseline patient characteristics. No loading dose assumed. It should be noted that abatacept is provided to the NHS with a PAS. As the discount associated

											with this PAS is confidential this PAS cannot be incorporated into the cost-effectiveness model.
ADA	40 mg	2	£704.28	£352.14	26	26	0	£9,155.64	£0.00	BNF (May 2016)	-
COMB	NA			£0.00		NA	0	£816.47	£0.00	Table 162, NICE TA375[14]	Cost of hydroxycholoroquine + MTX + prednisolone + SSZ. No start-up cost assumed based on Table 162, NICE TA375[14]. Annual cost of £816 based on inflation of value of £786.94 reported in TA375.[14]
CTZ	200 mg	2	£715.00	£357.50	29	26	3	£9,295.00	-£2,502.00	BNF (May 2016)	-Certolizumab pegol is associated with a PAS that provides the first 12 weeks of certolizumab pegol (currently 10 pre-loaded syringes of 200 mg each) free of charge. This PAS is incorporated into the year 1 acquisition costs for certolizumab pegol in this analysis.
ETN	50 mg	4	£656.00	£164.00	52	52	0	£8,528.00	£0.00	BNF	-

										(May 2016)	
GOL	50 mg	1	£762.97	£762.97	12	12	0	£9,155.64	£0.00	BNF (May 2016)	Adult (body-weight up to 100 kg) to align with baricitinib baseline patient characteristics. Additionally, the 100 mg dose of golimumab that may be considered for patients weighting >100 kg is, in any case, offered at the same price as the 50 mg dose used in patients weighing <100 kg.
IFX	100 mg	1	£377.66	£377.66	24	21	3	£7,930.86	£1,132.98	Cost: BNF (May 2016), dose: TA195[69]	Six doses per year; one additional dose in first year; three vials per dose
MTX	10 mg	100	36.78	£0.37	104	104	0	£38.25	£0.00	MIMS (May 2016)	Assumed max dose of 20 mg per week based on Table 162, TA375.[14] No start-up cost
PaII	NA			£0.00		NA		£747.02	£0.00	Table 162, NICE TA375[14]	An approximation of monthly 'post biologic' cDMARD therapy (leflunomide, gold, cyclosporine etc.). Annual cost of £747 based on

											inflation of value of £720 reported for TA375.[14]
RTX	500 mg	1	873.15	£873.15	5.3	5.3	0.0	£4,656.80	£0.00	Cost: BNF (May 2016), dose: Table 162, NICE TA375[14]	2 g every 9 months following NICE TA375[14], 1 g given in two weeks' intervals according to BNF (May 2016)
TCZ iv	80 mg	1	102.4	£102.40	104	104	0	£10,649.60	£0.00	BNF (May 2016)	8 mg/kg (70 kg assumed weight) once every 4 weeks, drug wastage not included

Footnotes: Annual costs are rounded to the nearest whole Great British Pound
Abbreviations: BAR = baricitinib; ABA iv = abatacept iv; ABA sc = abatacept sc; ADA = adalimumab; COMB = combination cDMARDs; CTZ = certolizumab pegol; ETN = etanercept; GOL = golimumab; IFX = infliximab; MTX = methotrexate; NA = Not applicable; Pall = Palliative care; RTX = rituximab; TCZ iv = tocilizumab.

Administration costs associated with each treatment were based on data from TA375[14] and inflated to current price levels using the Hospital and Community Health Services (HCHS) Index.[169] Thus, a cost of £159.78 was applied per intravenous injection and a cost of £2.71 per subcutaneous injection, dependent on the method of administration for each of the treatments modelled.

Monitoring costs included full blood count (FBC), erythrocyte sedimentation rate (ESR), biochemical profile (BCP), chest x-ray and urine analysis. In line with TA375,[14] the same monitoring costs are applied to both cDMARDs and bDMARDs, and are also assumed to apply to baricitinib. As discussed in Section 2.4, similarly to the bDMARDs baricitinib is anticipated to be associated with tests for tuberculosis and hepatitis prior to initiation of therapy and with monitoring requirements that would be captured by routine patient management in the NHS. No further additional monitoring requirements specific to baricitinib are specified in the SmPC. Cost data were sourced from TA375[14] and inflated to current price levels using the HCHS Index, as shown in Table 114.[169] This amounts to a monitoring cost of £176.38 before each treatment initiation, £1,763.79 for the first six months of treatment monitoring, and a monthly monitoring cost of £139.03 for the steady condition state. Monitoring costs were applied for all treatments equally.

Table 114. Inflated monitoring costs

Monitoring component	FBC (£2)	ESR (£3)	BCP (£3)	CXR (£33)	Urinalysis (£0.09)	Hospital outpatient attendance (£128)	Total cost	2016 inflated cost
Pre-treatment monitoring	1	1	1	1	0	1	£170	£176.38
First 6 months of treatment monitoring	10	0	10	0	0	10	£1,700	£1,763.79
Monthly monitoring cost	1	0	1	0	0	1	£134	£139.03

Abbreviations: FBC = full blood count, ESR = erythrocyte sedimentation rate, BCP = biochemical profile, CXR = chest x-ray.
Source: TA375 Assessment Group Report[138]

5.5.3 Health-state unit costs and resource use

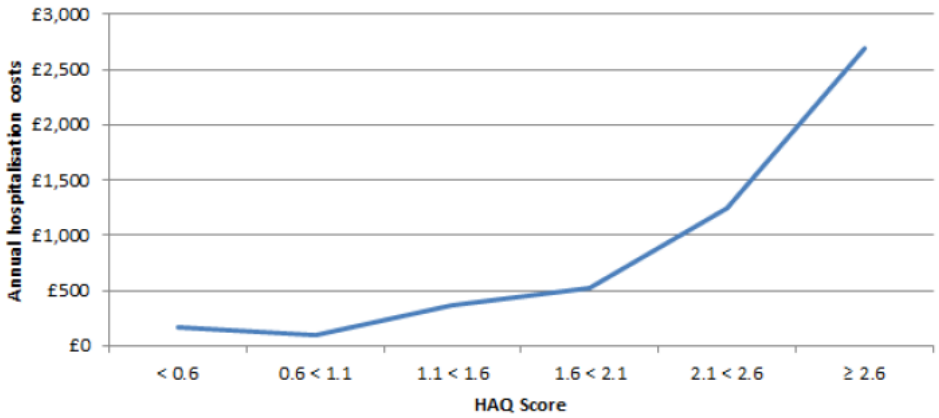
Annual hospitalisation costs per HAQ

For hospitalisation costs, the model uses the same approach as TA375[14] in applying hospitalisation costs per HAQ score.

In order to do this, the AG in TA375[14] used data on inpatient visits and joint replacements from the NOAR database to estimate disease-related hospitalisation costs per HAQ score band.[170] These hospitalisation activities were then multiplied by NHS reference costs to derive an assumed relationship between hospitalisation cost and HAQ score, as presented below in Figure 49.

For the baricitinib cost-effectiveness model, the plot of the assumed relationship between annual hospitalisation costs and HAQ score from NICE TA375[14] (Figure 49) was digitised using Plot Digitizer v2.6.6 and a polynomial model fitted. Extrapolated values were used to supplement any gaps from the dataset derived from the digitisation of the plot. This process provided continuous outcomes in terms of hospitalisation cost per HAQ score; however, since the HAQ score is based on discrete scores, only hospitalisation cost values associated with discrete HAQ scores were used in the model.

Figure 49. Relationship between annual hospitalisation costs and HAQ



Abbreviation: HAQ = Health Assessment Questionnaire
Source: NICE TA375 AG Report.[138]

Adverse reaction unit costs and resource use

As described previously, SAEs were not included in the base case analysis. However, a scenario analysis was conducted in which SAEs were included for the cDMARD-IR population. Rates of SAEs were taken from RA-BEAM, as previously described in Section 5.4.4. The costs associated with SAEs were £1,789 per episode, based on an average of the costs for celullitus and herpes zoster as the preferred term SAEs observed in more than 0.3% of any population in any of the study arms of RA-BEAM.

5.5.4 Miscellaneous unit costs and resource use

No additional costs or resource use items were included in the model that which have not already been listed above.

5.6 *Summary of base-case de novo analysis inputs and assumptions*

5.6.1 Summary of base-case de novo analysis inputs

A list of inputs used in the base case analysis is presented in Table 115.

Table 115: Summary of variables applied in the economic model

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution	Reference to section in submission
Model Settings			
Time horizon	Lifetime (45 years)	N/A – not varied in the PSA	Section 5.2.2
Discount rate (costs and outcomes)	3.5%	N/A – not varied in the PSA	
Mean age at baseline, years			
Moderate, cDMARD-IR population	52.05	SD: 12.40	Section 5.2.1
Severe, cDMARD-IR population	52.89	SD: 12.12	
Severe, anti-TNF-IR population (RTX-eligible and ineligible)	55.64	SD: 11.00	
Mean HAQ score at baseline			
Moderate, cDMARD-IR population	0.98	SD: 0.61	Section 5.2.1
Severe, cDMARD-IR population	1.61	SD: 0.63	
Severe, anti-TNF-IR population (RTX-eligible and ineligible)	1.78	SD: 0.56	
Percentage female			
Moderate, cDMARD-IR population	75.01	N/A – not varied in the PSA	Section 5.2.1
Severe, cDMARD-IR population	79.09	N/A – not varied in the PSA	
Severe, anti-TNF-IR population (RTX-eligible and ineligible)	81.70	N/A – not varied in the PSA	
Mean weight at baseline, kg			
Moderate, cDMARD-IR population	70.54	N/A – not varied in the PSA	Section 5.2.2
Severe, cDMARD-IR population	72.15	N/A – not varied in the PSA	
Severe, anti-TNF-IR population (RTX-	81.80	N/A – not varied in the PSA	

eligible and ineligible)			
Clinical Inputs			
EULAR good response (cDMARD-IR population)			
NMA CODA	NMA CODA	NMA CODA	NMA CODA
COMB	0.19		
Pall	0		
MTX	0.19		
ABTSMTX	0.37		
ADAMTX	0.38		
CTZMTX	0.48		
ETNMTX	0.39		
GOLMTX	0.38		
IFXMTX	0.35		
RTXMTX	0.43		
TCZMTX	0.52		
EULAR good response (anti-TNF-IR population)			
BAR4MTX	0.19	NMA CODA	Section 5.3
Pall	0		
ABTSMTX	0.21		
RTXMTX	0.31		
TCZMTX	0.26		
MTX	0.08		
GOLMTX	0.15		
ADAMTX	0.38		
ETNMTX	0.39		
IFXMTX	0.35		
CTZMTX	0.48		
EULAR moderate response (cDMARD-IR population)			
BAR4MTX	0.32	NMA CODA	Section 5.3
COMB	0.31		
Pall	0		
MTX	0.31		
ABTSMTX	0.33		
ADAMTX	0.33		
CTZMTX	0.31		
ETNMTX	0.33		
GOLMTX	0.33		
IFXMTX	0.33		
RTXMTX	0.32		
TCZMTX	0.3		
EULAR moderate response (anti-TNF-IR population)			

BAR4MTX	0.32	NMA CODA	Section 5.3
Pall	0		
ABTSMTX	0.34		
RTXMTX	0.36		
TCZMTX	0.36		
MTX	0.23		
GOLMTX	0.31		
ADAMTX	0.33		
ETNMTX	0.33		
IFXMTX	0.33		
CTZMTX	0.31		
HAQ Improvement (cDMARD-IR and anti-TNF-IR populations)			
Good	0.672	SE: 0.112	Section 5.3
Moderate	0.317	SE: 0.048	
Cost and Resource Inputs			
Annual drug cost, £			
BAR4MTX	[redacted]	NA	Section 5.5.2
Pall	747.02		
ABTSMTX	15724.80		
RTXMTX	4656.80		
TCZMTX	10649.60		
MTX	38.25		
GOLMTX	9155.64		
ADAMTX	9155.64		
ETNMTX	8528		
IFXMTX	7930.86		
CTZMTX	9295		
Additional costs of administration and monitoring in Year 1, £			
BAR4MTX	1106	SE: 56.4285	Section 5.5.2
COMB	1106	SE: 56.4285	
Pall	1106	SE: 56.4285	
ABTSMTX	1106	SE: 56.4285	
ADAMTX	1106	SE: 64.58046	
CTZMTX	1106	SE: 56.4285	
ETNMTX	1106	SE: 56.4285	
GOLMTX	1106	SE: 56.4285	
IFXMTX	1265.776991	SE: 56.4285	
RTXMTX	1106	SE: 56.4285	
TCZMTX	1106	SE: 64.58046	
MTX	1106	SE: 56.4285	
Annual administration and monitoring costs, £			

BAR4MTX	1668	SE: 85.11916	Section 5.5.2
COMB	1668	SE: 85.11916	
Pall	1668	SE: 85.11916	
ABTSMTX	1704	SE: 86.91524	
ADAMTX	1739	SE: 88.71132	
CTZMTX	1739	SE: 88.71132	
ETNMTX	1809	SE: 92.30347	
GOLMTX	1701	SE: 86.77708	
IFXMTX	2787	SE: 142.1829	
RTXMTX	2094	SE: 106.8577	
TCZMTX	3745	SE: 191.0946	
MTX	1668	SE: 85.11916	
Cost of hospitalisation per HAQ score, £			
0	257.18	SE: 13.12144	Section 5.5.3
0.125	189.89	SE: 9.688484	
0.25	163.07	SE: 8.320033	
0.375	146.26	SE: 7.462315	
0.5	127.86	SE: 6.523224	
0.625	109.98	SE: 5.611257	
0.75	94.23	SE: 4.807485	
0.875	159.29	SE: 8.12688	
1	226.47	SE: 11.55477	
1.125	295.79	SE: 15.09116	
1.25	363.50	SE: 18.54587	
1.375	402.51	SE: 20.53619	
1.5	442.58	SE: 22.58075	
1.625	480.52	SE: 24.51651	
1.75	520.07	SE: 26.53395	
1.875	702.62	SE: 35.84778	
2	885.17	SE: 45.1616	
2.125	1063.46	SE: 54.25813	
2.25	1247.07	SE: 63.6262	
2.375	1607.72	SE: 82.0266	
2.5	1961.45	SE: 100.0741	
2.625	2328.48	SE: 118.8	
2.75	2688.06	SE: 137.1461	
2.875	3381.78	SE: 172.5399	
3	4065.17	SE: 207.4068	
Linear HAQ Trajectory			
Time to increase HAQ by 0.125			
Comb	2.7	SE: 0.137755	Section 5.3

Pall	2	SE: 0.102041	
MTX	2.7	SE: 0.137755	
Mortality per HAQ score			
0	1.0	SE: 0.00000	Section 5.3
0.125	1.4	SE: 0.17857	
0.25	1.4	SE: 0.17857	
0.375	1.4	SE: 0.17857	
0.5	1.5	SE: 0.17857	
0.625	1.5	SE: 0.17857	
0.75	1.5	SE: 0.17857	
0.875	1.5	SE: 0.17857	
1	1.8	SE: 0.20408	
1.125	1.8	SE: 0.20408	
1.25	1.8	SE: 0.20408	
1.375	1.8	SE: 0.20408	
1.5	2.7	SE: 0.33163	
1.625	2.7	SE: 0.33163	
1.75	2.7	SE: 0.33163	
1.875	2.7	SE: 0.33163	
2	4.0	SE: 0.53571	
2.125	4.0	SE: 0.53571	
2.25	4.0	SE: 0.53571	
2.375	4.0	SE: 0.53571	
2.5	5.5	SE: 0.96939	
2.625	5.5	SE: 0.96939	
2.75	5.5	SE: 0.96939	
2.875	5.5	SE: 0.96939	
3	5.5	SE: 0.96939	
Long-term Treatment Discontinuation			
Weibull parameter lambda			
Good	0.00069	SE: 0.00004	Section 5.3
Moderate	0.00151	SE: 0.00008	
Weibull shape parameter gamma			
Good	0.86206	SE: 0.04398	Section 5.3
Moderate	0.82591	SE: 0.04214	
Single-curve scale and shape parameters			
Scale (lambda)	0.00069	SE: 0.00004	Section 5.3
Shape (gamma)	0.86206	SE: 0.04398	
Coefficients for the Mapping HAQ-DI to EQ-5D (Hernandez *et al*. 2012)			
Within subject			
Latent class 1			

HAQ	-0.062	SE: 0.015	Section 5.4.2
HAQ^2	N/A	N/A	
VASpain/100	-0.295	SE: 0.030	
σ_u^2	0.015	SE: 0.002	
Latent class 2			
HAQ	-0.245	SE: 0.044	Section 5.4.2
HAQ^2	0.068	SE: 0.019	
VASpain/100	-0.105	SE 0.134	
σ_u^2	0.006	SE: 0.001	
Latent class 3			
HAQ	-0.16	SE: 0.013	Section 5.4.2
HAQ^2	0.025	SE: 0.005	
VASpain/100	-0.056	SE: 0.018	
σ_u^2	0.003	SE: 0.000	
Between-subject			
Latent class 1 Intercept	0.343	SE: 0.037	Section 5.4.2
Latent class 2 Intercept	0.990	SE: 0.025	
Latent class 3 Intercept	0.806	SE: 0.011	
All classes			
$\left(\frac{Age - 54.32}{10}\right)$	0.007	SE: 0.002	Section 5.4.2
$\left(\frac{Age - 54.32}{10}\right)^2$	0.004	SE: 0.001	
Male	-0.012	SE: 0.006	
σ_u^2	0.002	SE: 0.000	
Within-subject categorical latent variables			
Latent class 1			
Intercept	-5.201	SE: 0.423	Section 5.4.2
HAQ	2.868	SE: 0.178	
VASpain/100	5.179	SE: 0.433	
Latent class 2			
Intercept	2.203	SE: 0.312	Section 5.4.2
HAQ	0.485	SE: 0.214	
VASpain/100	-11.366	SE: 4.227	
Pain to HAQ score			
0	11.834	SE: 0.60	Section 5.4.2
0.125	18.316	SE: 0.93	
0.25	19.380	SE: 0.99	
0.375	22.571	SE: 1.15	
0.5	24.951	SE: 1.27	
0.625	27.644	SE: 1.41	
0.75	30.463	SE: 1.55	

0.875	32.398	SE: 1.65	
1	35.197	SE: 1.80	
1.125	37.550	SE: 1.92	
1.25	41.380	SE: 2.11	
1.375	44.065	SE: 2.25	
1.5	46.831	SE: 2.39	
1.625	50.069	SE: 2.55	
1.75	53.287	SE: 2.72	
1.875	55.401	SE: 2.83	
2	57.409	SE: 2.93	
2.125	58.925	SE: 3.01	
2.25	61.817	SE: 3.15	
2.375	63.938	SE: 3.26	
2.5	67.747	SE: 3.46	
2.625	69.329	SE: 3.54	
2.75	67.727	SE: 3.46	
2.875	61.371	SE: 3.13	
3	58.020	SE: 2.96	

Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib 4 mg (QD), cDMARD = conventional disease-modifying antirheumatic drug, COMB = combination cDMARDs, CTZ = certolizumab pegol, EQ-5D = EuroQoL five-dimensional; ETN = etanercept, GOL = golimumab, HAQ-DI = health assessment questionnaire disability index; IFX = infliximab, IR = inadequate response, MTX = methotrexate, N/A = not available, Pall = palliative care, RTX = rituximab, SE = standard error; TCZ = tocilizumab, TNF = tumour necrosis factor, VAS = visual analogue scale.

5.6.2 Assumptions

A list of assumptions made in the baricitinib cost-effectiveness model is presented in Table 116.

Table 116. List of model assumptions and their justifications

Model assumption	Justification
Long-term HAQ trajectory was assumed to remain flat whilst on treatment for patients receiving biologics or baricitinib.	For biologics this is consistent with the AG model in TA375.[14] For baricitinib this assumption is justified based on long-term HAQ-DI data from patients entering the long-term extension study, RA-BEYOND (JADV), see Section 4.7.4). A scenario analysis is also presented in which an assumption of linear HAQ progression is applied in the baricitinib arm.
HRQOL was assumed to be a function of HAQ score and pain severity.	Consistent with the AG model in TA375.[14]
Hospitalisation costs were assumed to be dependent upon HAQ score.	Consistent with the AG model in TA375.[14]
Mortality was assumed to be dependent upon baseline HAQ, but not upon change in HAQ.	Consistent with the AG model in TA375.[14]
Upon discontinuation of treatment, HAQ score was assumed to rebound to its level prior to initiation of therapy.	Consistent with the AG model in TA375.[14]
It was assumed that long-term HAQ progression	Consistent with the AG model in TA375.[14]

on cDMARDs could be modelled with a latent class approach, with no further HAQ change assumed after 15 years.	
Dosing of infliximab was modelled as per NICE TA195 and therefore assumed to comprise 6 doses of infliximab per patient per year, consisting of 3 vials per dose.	Simplifying assumption in line with TA195.[68]
Initial HAQ changes for moderate and good EULAR responders based on analysis of the BSRBR database were assumed to apply to baricitinib	Initial HAQ change was 0.672 (SE: 0.112) for EULAR good responders and 0.317 (SE: 0.048) for EULAR moderate responders. In line with TA375.[14] A scenario anlysis using barictinibib trial data was carried out.
Long-term treatment discontinuation was modelled to be treatment independent, based upon fitting a parametric curve (Weibull curve) to data from the BSRBR database, and was assumed to apply to baricitinib.	This assumption is justified as discontinuation rates from RA-BEAM were very similar between baricitinib and adalimumab (6.8% and 5.6%, respectively). Similarly, for the mITT population, 87.7% and 86.7% of patients completed the study up to Week 52 for baricitinib and adalimumab, respectively. In addition, a scenario analysis is conducted to explore the impact of alternatively assuming a fixed annual discontinuation rate for baricitinib, derived from baricitinib trial data.

5.7 *Base-case results*

5.7.1 Base-case incremental cost-effectiveness analysis results

The base case cost-effectiveness results for the following populations are presented below

- Severe, cDMARD-IR;
- Severe, anti-TNF-IR (rituximab-ineligible);
- Moderate, cDMARD-IR;
- Severe, anti-TNF-IR (rituximab eligible).

Severe, cDMARD-IR population

The results of the base case analysis for the severe, cDMARD-IR population are presented in Table 117. In the pairwise analysis, baricitinib (4 mg QD + MTX) dominated all comparators with the exception of certolizumab pegol + MTX, which was associated with an ICER of £18,400 per QALY versus baricitinib (4 mg QD).

Table 117. Base case cost-effectiveness results for the severe, cDMARD-IR population

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Referent	Referent
ETN-bMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	£18,400	£18,400
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
IFX-bMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN-b etanercept biosimilar, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX-b = infliximab biosimilar, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

Severe, anti-TNF-IR population (rituximab-ineligible)

The results of the base case analysis for the severe, anti-TNF-IR (rituximab-ineligible) population are presented in Table 118. In the incremental analysis, certolizumab pegol was found to be the cost-effective option at a cost-effectiveness threshold of £30,000 per QALY, with an ICER of £16,201 per QALY in the incremental analysis. In the pairwise analyses, baricitinib (4 mg QD + MTX) was found to dominate golimumab + MTX; the ICERs for the remaining comparisons varied, with some ICERs greater and some lower than £30,000 per QALY versus baricitinib.

Table 118. Base-case cost-effectiveness results for the severe, anti-TNF-IR (rituximab-ineligible) population

Technology sequence	Total costs (£)	Total QALYs	Total LYG	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (QALYs)	ICER (£) incremental (QALYs)
BAR4MTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	Referent	Referent
GOLMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	Dominated	Dominated
ETN-bMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£19,874	Ext Dominated
CTZMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£16,201	£16,201
ADAMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£27,008	Dominated
IFX-bMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£34,942	Dominated
TCZMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£36,757	Dominated
ABTSMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£484,782	Dominated

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN-b etanercept biosimilar, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX-b = infliximab biosimilar, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

Moderate, cDMARD-IR population

The results of the base case analysis for the moderate, cDMARD-IR population are presented in Table 119. The ICER for baricitinib (4 mg QD) versus combination cDMARD therapy was found to be £37,420 per QALY.

Table 119. Base-case cost-effectiveness results for the moderate, cDMARD-IR population

Technology sequence	Total costs (£)	Total QALYs	Total LYG	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (QALYs)
COMB+MTX+Pall	[redacted]	[redacted]	16.04	[redacted]	[redacted]	-
BAR4MTX+COMB+MTX+Pall	[redacted]	[redacted]	16.03	[redacted]	[redacted]	£37,420

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: BAR4 = baricitinib 4 mg (QD), COMB = combination cDMARDs, ICER = incremental cost-effectiveness ratio, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, LYG = life years gained

Severe, anti-TNF-IR population (rituximab-eligible)

The results of the base case analysis for the severe, anti-TNF-IR (rituximab-eligible) population are presented in Table 120. Baricitinib (4 mg QD + MTX) was dominated by rituximab + MTX.

Table 120. Base-case cost-effectiveness results for the severe, anti-TNF-IR (rituximab-eligible) population

Technologies	Total costs (£)	Total QALYs	Total LYG	Incremental costs (£)	Incremental QALYs	ICER (£) versus baseline (per QALY)
RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	5.35	[redacted]	[redacted]	-
BAR4MTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	5.25	[redacted]	[redacted]	Dominated

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: BAR4 = baricitinib (4 mg QD), ICER = incremental cost-effectiveness analysis, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

5.7.2 Clinical outcomes from the model

As the model was a DES where costs and outcomes were determined by HAQ-DI score it was not feasible to perfrom a comparison with clinical outcomes. This was further complicated by the sequence approach taken. It should be noted that the EULAR reponse estimates from the NMA were similar to those from the baricitinib clinical studies, therefore, as this is the key pararmeter determining treatment response, it can be assumed that the model outcomes would be similar to clinical outcomes.

5.7.3 Disaggregated results of the base case incremental cost-effectiveness analysis

The disaggregated base case cost results by cost type are presented below for each of the populations of interest (Table 121 to Table 124). Due to the fact that health states were not included in the model, a summary of costs and QALY gains by health state is not presented. Drug costs were seen to be the largest category of cost for all interventions in all populations with the exception of the sequence of a combination of cDMARDs followed by methotrexate followed by palliative care, for which monitoring costs were the largest contributor to the total costs.

Table 121. Summary of predicted resource use by category of cost in the severe, cDMARD-IR population

Treatment Sequence	Drugs	Monitoring	Hospitalisation	Total
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
ETN-bMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
IFX-bMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN-b = etanercept biosimilar, GOL = golimumab, IFX-b = infliximab biosimilar, MTX = methotrexate, Pall = palliative care, RTX = rituximab, TCZ = tocilizumab

Table 122. Summary of predicted resource use by category of cost in the severe, anti-TNF-IR (rituximab-ineligible) population

Treatment Sequence	Drugs	Monitoring	Hospitalisation	Total
BAR4MTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
GOLMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
ETN-bMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
CTZMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
ADAMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
IFX-bMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
TCZMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
ABTSMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN-b = etanercept biosimilar, GOL = golimumab, IFX-b = infliximab biosimilar, MTX = methotrexate, Pall = palliative care, RTX = rituximab, TCZ = tocilizumab

Table 123. Summary of predicted resource use by category of cost in the moderate, cDMARD-IR population

Treatment Sequence	Drugs	Monitoring	Hospitalisation	Total
BAR4MTX+COMB+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
COMB+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BAR4 = baricitinib (4 mg QD), COMB = combination cDMARDs, MTX = methotrexate, Pall = palliative care, RTX = rituximab, TCZ = tocilizumab

Table 124. Summary of predicted resource use by category of cost in the severe, anti-TNF-IR population (rituximab-eligible) population

Treatment Sequence	Drugs	Monitoring	Hospitalisation	Total
BAR4MTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]
RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BAR4 = baricitinib (4 mg QD), MTX = methotrexate, Pall = palliative care, RTX = rituximab, TCZ = tocilizumab

5.8 *Sensitivity analyses*

5.8.1 Probabilistic sensitivity analysis

The results of the probabilistic sensitivity analyses for the severe, cDMARD-IR population and the anti-TNF-IR (rituximab-ineligible) population are presented in Table 125 and Table 126, respectively. The cost-effectiveness acceptability curves for these populations are presented in Figure 50 and Figure 51, respectively.

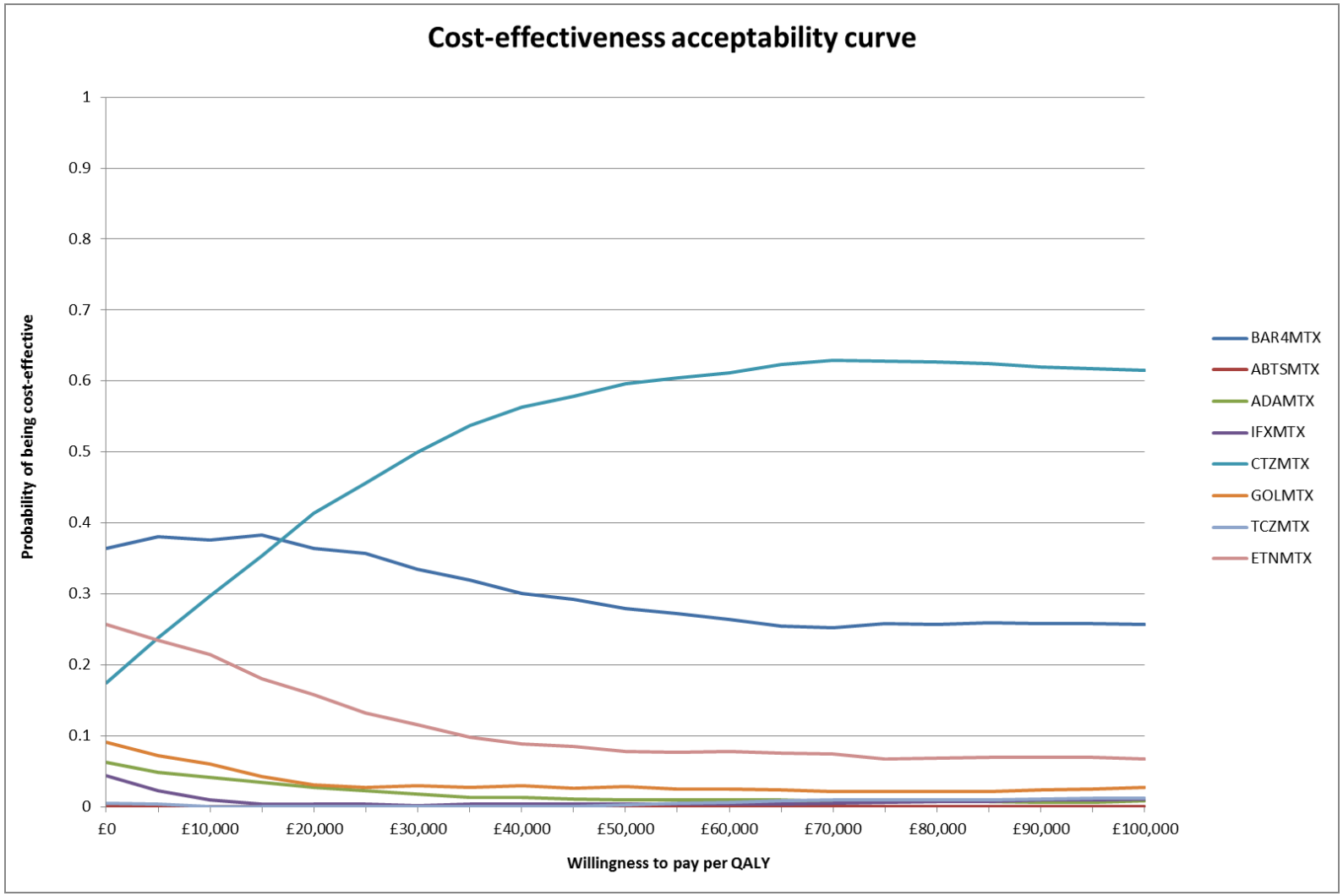
In the severe, cDMARD-IR population, cost-effectiveness results for the probabilistic analysis were seen to be closely aligned to deterministic results in terms of ICERs versus baricitinib in the pairwise analyses. In the severe, anti-TNF-IR (rituximab-ineligible) population, the probabilistic analysis resulted in a different ordering of therapies in the incremental analysis to that of the base case analysis. Baricitinib was seen to be associated with a pairwise ICER of £20,612 versus the reference product (etanercept). In the incremental analysis, baricitinib was extendedly dominated. Again, it should be noted that this analysis is subject to significant uncertainty as data was not available for all comparators in the scope and it is likely that the treatment effect of comparators for which there was no data will have been overestimated as the cDMARD-IR efficacy estimates were used (e.g adalimumab etc.)

Table 125. Results of the probabilistic sensitivity analysis for the severe, cDMARD-IR population

Treatment Sequence	Total Discounted Costs	Total Discounted QALYs	Total Discounted LYG	Incremental Costs	Incremental QALYS	ICER vs. Baseline	Incremental ICER
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	Referent	Referent
ETN-bMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	Dominated	Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	£18,414	£18,414
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.71	[redacted]	[redacted]	Dominated	Dominated
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	Dominated	Dominated
IFX-bMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.71	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	Dominated	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	Dominated	Dominated

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN-b = etanercept biosimilar, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX-b = infliximab biosimilar, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

Figure 50. Cost-effectiveness accepatibility curve for the severe, cDMARD-IR population



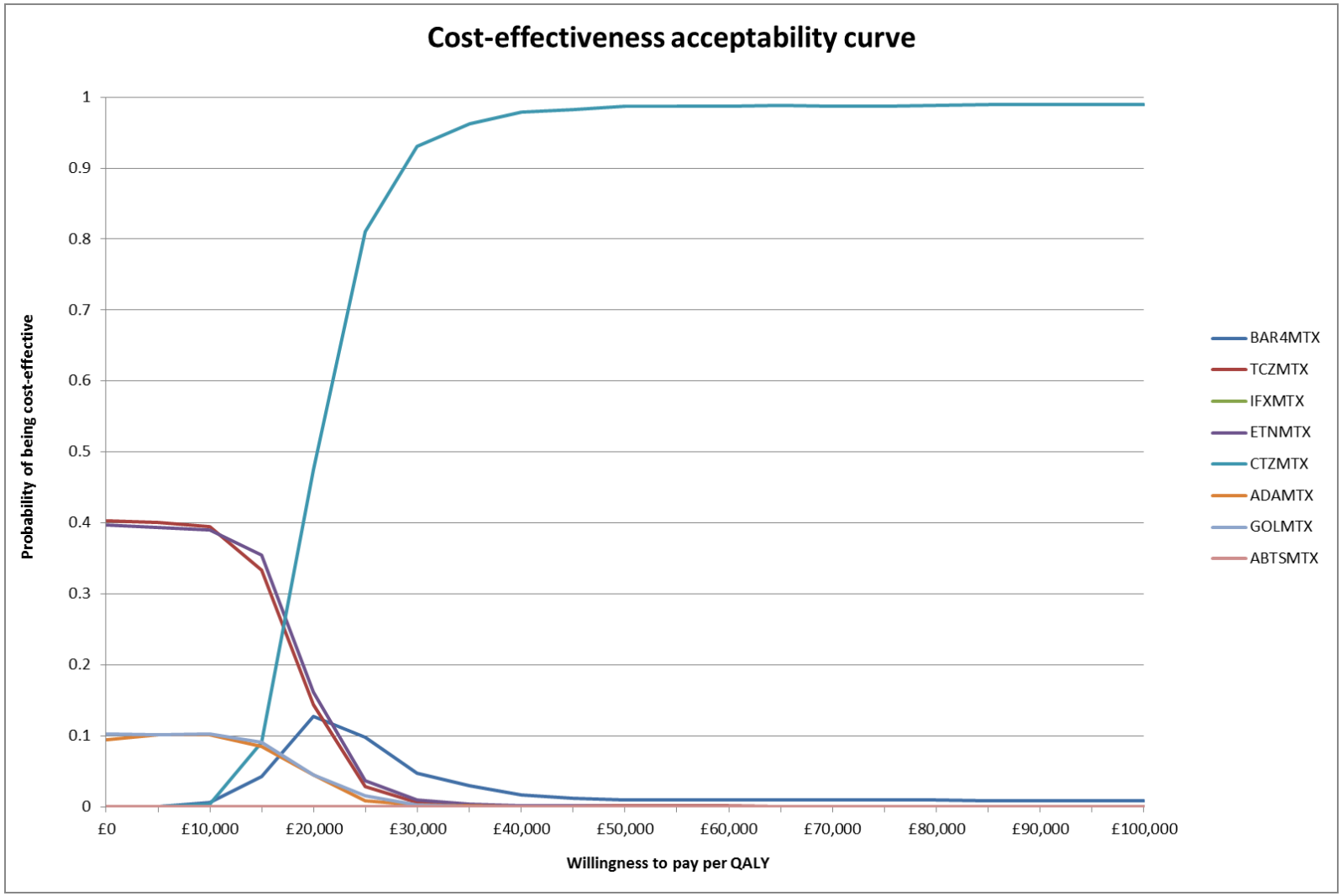
Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, IFX = infliximab, MTX = methotrexate, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

Table 126. Results of the probabilistic sensitivity analysis for the anti-TNF-IR (rituximab-ineligible) population

Treatment Sequence	Total Discounted Costs	Total Discounted QALYs	Total Discounted LYG	Incremental Costs	Incremental QALYS	ICER vs. Baseline	Incremental ICER
ETN-bMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.53	[redacted]	[redacted]	Referent	Referent
GOLMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.53	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.53	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	13.52	[redacted]	[redacted]	Dominated	Dominated
IFX-bMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.52	[redacted]	[redacted]	£45,998	Ext Dominated
BAR4MTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.52	[redacted]	[redacted]	£20,612	Ext Dominated
CTZMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.52	[redacted]	[redacted]	£18,738	£18,738
ABTSMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.52	[redacted]	[redacted]	£62,763	Dominated

Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN-b = etanercept biosimilar, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX-b = infliximab biosimilar, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

Figure 51. Cost-effectiveness acceptability curve for the anti-TNF-IR (rituximab-ineligible) population



Footnote: Presented costs and QALYs are discounted at 3.5%. The convention [interventionMTX] represents the intervention in combination with MTX.
Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR4 = baricitinib (4 mg QD), CTZ = certolizumab pegol, ETN = etanercept biosimilar, GOL = golimumab, IFX = infliximab biosimilar, MTX = methotrexate, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab

5.8.2 Deterministic sensitivity analysis

It was considered that deterministic sensitivity analysis would provide limited insight relative to the extensive computational burden required to produce a tornado diagram for each pairwise comparison in each population. A number of the pairwise analyses resulted in a conclusion of dominance of baricitinib under the base case settings, which presents difficulties in presentation and interpretation for tornado diagrams. As such, deterministic sensitivity analyses are not presented in the submission. However, a number of scenario analyses have been performed to evaluate the underlying assumptions in the model, and are presented in Section 5.8.3.

5.8.3 Scenario analysis

The economic evaluation explored a number of scenario analyses in which key model assumptions or parameters were altered. The scenario analyses considered in the economic modelling are presented in Table 127.

Table 127. Summary of scenario analyses

#	Scenario analysis	Description of scenario analysis
Severe, cDMARD-IR population		
1	Alternative discount rates applied to the costs and benefits	Costs: 6% Benefits: 1.5%
2	Linear HAQ progression on cDMARDs/palliative care	An assumption is made that patients receiving cDMARDs/palliative care follow a linear HAQ progression instead of the latent class approach described in Section 5.3. The mean rates of HAQ increase modelled in the BRAM by Malottki *et al*.[135] (0.045/year on cDMARDs and 0.06/year on palliative care) are used. These rates are modelled as mean times to increase in HAQ score (by 0.125) of 2.7 years and 2 years, respectively.
3	Malottki mapping equation	The HAQ-DI to EQ-5D quadratic mapping mechanism used by Malottki *et al*. [135] in the BRAM was adopted in this scenario.
4	RA-BEAM mapping equation	Please see Section 5.4.2.
5	Inclusion of costs for SAEs	Please see Sections 5.4.4 and 5.6.1.
6	4 mg-2 mg step-down	This scenario explores the cost-effectiveness of baricitinib in patients who respond well to the 4 mg QD dose and are tapered down to the 2 mg QD dose. Patients assessed as good EULAR responders at Week 24 are assumed to continue therapy with baricitinib (2 mg QD). An initial HAQ decrement of [redacted] is applied to these patients, based on data from RA-BEYOND in patients randomised to the 2 mg QD dose, please see Section 4.7.4.
7	Week 12 cost-effectiveness results	EULAR response data from the NMA sensitivity analysis for the Week 12 time point (cDMARD-IR population) are used to inform the model, please see Section 4.10.8.2.
8	Non-flat linear HAQ progression for baricitinib	To explore the uncertainty around the assumption that the HAQ progression for baricitinib is equivalent to bDMARDs, i.e. flat, a mean time for a HAQ increase of 0.125 of 5 years is assumed. This is approximately half the rate of progression of cDMARDs modelled in scenario 2.
9	Drug wastage	This scenario adds an additional 4 weeks of cost to the 1st year of treatment with baricitinib to allow consideration of the possibility of drug wastage with baricitinib

10	BSRBR baseline patient characteristics	Patient baseline characteristics from the BSRBR database are used, please see Section 5.2.1.
11	Time horizon: 15 years	An alternative time horizon of 15 years is evaluated.
12	Initial change in HAQ score based on baricitinib clinical trial programme data	The value for the initial change in HAQ score for patients assigned to baricitinib is based upon data from the baricitinib clinical trial programme instead of the BSRBR database, see Section 5.3.
13	Baricitinib discontinuation rate based on baricitinib clinical trial programme data	An alternative rate of treatment discontinuation is modelled for patients receiving baricitinib, based on discontinuation rates from RA-BEAM.
RTX-ineligible, severe, anti-TNF-IR population		
1	Linear HAQ progression on cDMARDs/palliative care	Please see scenario 2 for the severe, cDMARD-IR population.
2	Inclusion of SAE costs	This scenario uses the same assumptions as the scenario for the cDMARD-IR population in the absence of alternative data
3	Initial change in HAQ score based on baricitinib clinical trial programme data	Please see scenario 12 for the severe, cDMARD-IR population.
4	Baricitinib discontinuation rate based on baricitinib clinical trial programme data	Please see scenario 13 for the severe, cDMARD-IR population.
5	Week 12 cost-effectiveness results	Please see scenario 7 for the severe, cDMARD-IR population.
Moderate, cDMARD-IR population		
1	Linear HAQ progression for cDMARDs/pall	Please see scenario 2 for the severe, cDMARD-IR population.
2	Inclusion of SAE costs	Please see scenario 5 for the severe, cDMARD-IR population.
3	Non-flat linear HAQ progression for baricitinib	Please see scenario 8 for the severe, cDMARD-IR population.
4	Initial change in HAQ score based on baricitinib clinical trial programme data	Please see scenario 12 for the severe, cDMARD-IR population.
RTX-eligible, severe, anti-TNF-IR population		
1	Linear HAQ progression on cDMARDs/palliative care	Please see scenario 2 for the severe, cDMARD-IR population.
2	Inclusion of SAE costs	This scenario uses the same assumptions as the scenario for the cDMARD-IR population in the absence of alternative data
3	Initial change in HAQ score based on baricitinib clinical trial programme data	Please see scenario 12 for the severe, cDMARD-IR population.
4	Baricitinib discontinuation rate based on baricitinib clinical trial programme data	Please see scenario 13 for the severe, cDMARD-IR population.
5	Week 12 cost-effectiveness results	Please see scenario 7 for the severe, cDMARD-IR population.
MTX-IR population		
1	Baricitinib versus adalimumab head-to-head comparison	This scenario explores a head-to-head comparison between baricitinib (4 mg QD) and adalimumab in the MTX-IR population. Data from the RA-BEAM trial were used to inform the baseline characteristics inputted into the model and EULAR response rates (data from the mITT population were used). In addition, this scenario uses the HAQ-DI to EQ-5D mapping algorithm based on RA-BEAM.

Abbreviations: BRAM = Birmingham Rheumatoid Arthritis Model, cDMARD = conventional disease-modifying antirheumatic drug, EQ-5D = EuroQoL-Five Dimensions, HAQ = health assessment questionnaire, IR = inadequate response, mITT = modified intention to treat, QD = once daily, SAE = serious adverse event, TNF = tumour necrosis factor.

Results of the scenario analyses are presented below for the respective populations.

Severe, cDMARD-IR population

The results of the scenario analyses for the severe, cDMARD-IR population are presented in Table 128, Table 129 and Table 130.

Table 128. Scenario analysis results for the severe, cDMARD-IR population (1/2)

	ETN-b+MTX			ADA+MTX			TCZ+MTX			IFX-b+MTX		
Scenario	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER
Base case	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
1	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
2	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
3	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
4	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
5	[redacted]	[redacted]	£10,275	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
6	[redacted]	[redacted]	£13,800	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£196,500	[redacted]	[redacted]	Dominated
7	[redacted]	[redacted]	£8,278	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
8	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£2,323	[redacted]	[redacted]	£14,313	[redacted]	[redacted]	£4,739
9	[redacted]	[redacted]	£14,700	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
10	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
11	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
12	[redacted]	[redacted]	£10,447	[redacted]	[redacted]	£12,979	[redacted]	[redacted]	£17,868	[redacted]	[redacted]	£14,558
13	[redacted]	[redacted]	£4,086	[redacted]	[redacted]	£5,725	[redacted]	[redacted]	£9,322	[redacted]	[redacted]	£6,503

Footnote: Results are presented for the pairwise analysis versus baricitinib (4 mg QD).
Abbreviations: ADA = adalimumab, ETN-b = etanercept biosimilar, Inc. = Incremental, QALY = Quality adjusted life years, ICER = incremental cost=effectiveness ratio, infliximab-b = infliximab biosimilar, MTX = methotrexate

Table 129. Scenario analysis results for the severe, cDMARD-IR population (2/2)

	GOL+MTX			CTZ+MTX			ABTS+MTX		
	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER
Base case	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£18,400	[redacted]	[redacted]	Dominated
1	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£9,440	[redacted]	[redacted]	Dominated
2	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£13,739	[redacted]	[redacted]	Dominated
3	[redacted]	[redacted]	£15,214	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	Dominated
4	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£19,448	[redacted]	[redacted]	Dominated
5	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£10,275	[redacted]	[redacted]	Dominated
6	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£9,925	[redacted]	[redacted]	Dominated
7	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£14,519	[redacted]	[redacted]	Dominated
8	[redacted]	[redacted]	£1,927	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£71,714
9	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£2,700	[redacted]	[redacted]	Dominated
10	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£16,950	[redacted]	[redacted]	Dominated
11	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£15,851	[redacted]	[redacted]	Dominated
12	[redacted]	[redacted]	£12,803	[redacted]	[redacted]	£10,212	[redacted]	[redacted]	£43,443
13	[redacted]	[redacted]	£5,606	[redacted]	[redacted]	£4,384	[redacted]	[redacted]	£26,036

Footnote: Results are presented for the pairwise analysis versus baricitinib (4 mg QD).
Abbreviations: ABTS = subcutaneous abatacept, CTZ = certolizumab pegol, GOL = golimumab, Inc. = Incremental, QALY = Quality adjusted life years, ICER = incremental cost=effectiveness ratio, MTX = methotrexate

cDMARD-IR, RA-BEAM head to head data

Table 130: Scenario analysis for the cDMARD-IR population, RA-BEAM head to head data

Treatment Sequence	Total Discounted Costs	Total Discounted QALYs	Total Discounted LYG	Incremental Costs	Incremental QALYS	ICER vs. Baseline	Incremental ICER
BAR4MTX+Pall	[redacted]	[redacted]	14.67	[redacted]	[redacted]	Referent	Referent
ADAMTX+Pall	[redacted]	[redacted]	14.68	[redacted]	[redacted]	Dominated	Dominated

Anti-TNF-IR, RTX-ineligible population

The results of the scenario analyses for the anti-TNF-IR, RTX-ineligible population are presented in Table 131.

Table 131. Scenario analysis results for the anti-TNF-IR, RTX-ineligible population (1/2)

	GOL+MTX			ETN-b+MTX			CTZ+MTX			ADA+MTX		
Scenario	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER
Base case	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£19,874	[redacted]	[redacted]	£16,201	[redacted]	[redacted]	£27,008
1	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£14,261	[redacted]	[redacted]	£11,418	[redacted]	[redacted]	£19,864
2	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£18,100	[redacted]	[redacted]	£14,849	[redacted]	[redacted]	£25,092
3	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£18,085	[redacted]	[redacted]	£15,204	[redacted]	[redacted]	£24,658
4	[redacted]	[redacted]	£64,285	[redacted]	[redacted]	£91,602	[redacted]	[redacted]	£110,454	[redacted]	[redacted]	£86,564
5	[redacted]	[redacted]	Dominated	[redacted]	[redacted]	£21,490	[redacted]	[redacted]	£16,508	[redacted]	[redacted]	£31,500

Footnote: Results are presented for the pairwise analysis versus baricitinib (4 mg QD).
Abbreviations: ADA = adalimumab, CTZ = certolizumab pegol, ETN-b = etanercept biosimilar, GOL = golimumab, Inc. = Incremental, QALY = Quality adjusted life years, ICER = incremental cost=effectiveness ratio, MTX = methotrexate

The results of the scenario analyses for the anti-TNF-IR, RTX-ineligible population are presented in Table 136.

Table 132. Scenario analysis results for the anti-TNF-IR, RTX-ineligible population (2/2)

	IFX-b+MTX			TCZ+MTX			ABTS+MTX		
	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER	Inc. costs	Inc. QALYs	ICER
Base case	[redacted]	[redacted]	£34,942	[redacted]	[redacted]	£36,757	[redacted]	[redacted]	£484,782
1	[redacted]	[redacted]	£27,322	[redacted]	[redacted]	£28,833	[redacted]	[redacted]	£439,967
2	[redacted]	[redacted]	£32,505	[redacted]	[redacted]	£34,240	[redacted]	[redacted]	£357,743
3	[redacted]	[redacted]	£32,250	[redacted]	[redacted]	£34,928	[redacted]	[redacted]	£440,667
4	[redacted]	[redacted]	£81,408	[redacted]	[redacted]	£83,659	[redacted]	[redacted]	£88,070
5	[redacted]	[redacted]	£44,863	[redacted]	[redacted]	£44,667	[redacted]	[redacted]	Dominated

Moderate, cDMARD-IR population

The results of the scenario analyses for the moderate, cDMARD-IR population are presented in Table 133.

Table 133. Scenario analysis results for the moderate, cDMARD-IR population

	COMB+MTX+Pall		
Scenario	**Inc. costs**	**Inc. QALYs**	**ICER**
Base case	[redacted]	[redacted]	£37,420
1	[redacted]	[redacted]	£20,965
2	[redacted]	[redacted]	£37,018
3	[redacted]	[redacted]	£30,280
4	[redacted]	[redacted]	£32,303

Abbreviations: DMARD = Disease-modifying antirheumatic drug, cDMARD = conventional DMARD, IR = inadequate response, MTX = methotrexate, COMB = combination therapy, Pall = palliative care, QALY = quality-adjusted life year, ICER = incremental cost-effectiveness ratio.

Anti-TNF-IR, RTX-eligible population

The results of the scenario analyses for the anti-TNF-IR, RTX-eligible population are presented in Table 134.

Table 134. Scenario analysis results for the Anti-TNF-IR, RTX-eligible population

	RTXMTX+TCZMTX+MTX+Pall		
Scenario	**Inc. costs**	**Inc. QALYs**	**ICER**
Base case	[redacted]	[redacted]	Dominated
1	[redacted]	[redacted]	Dominated
2	[redacted]	[redacted]	Dominated
3	[redacted]	[redacted]	£374,150
4	[redacted]	[redacted]	£58,747
5	[redacted]	[redacted]	Dominated

Abbreviations: IR = inadequate response, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, ICER = incremental cost-effectiveness ratio, RTX = rituximab, TCZ = tocilizumab, TNF = tumour necrosis factor

5.8.4 Summary of sensitivity analyses results

The probabilistic sensitivity analysis, taking into account combined uncertainty in model parameters, produced similar results to those of the deterministic analysis for the cDMARD-IR population. In the anti-TNF-IR (rituximab-ineligible) population, probabilistic analysis resulted in a different ordering of the therapies in the incremental analysis and hence resulted in different cost-effectiveness conclusions. However, the differences in costs and QALYs remained small and, as previously noted, the lack of data for several comparators in the scope (adalimumab, infliximab, etanercept) means that there is considerable uncertainty in this analysis. Furthermore, certolizumab data was only available for 12 weeks rather than the 24 week base case assumption As per the base case deterministic results, pairwise comparisons from the probabilistic sensitivity analysis found baricitinib to be a cost-effective therapy at a cost-effectiveness threshold of £30,000 per QALY versus at least one of the biologic therapies currently approved by NICE and used in the NHS for the treatment of RA patients.

An extensive range of scenario analyses was performed in order to test key model assumptions. In both the severe, cDMARD-IR and anti-TNF-IR (rituximab-ineligible) populations, the majority of results of pairwise analyses from scenario analyses were aligned to the results of the base case analysis; baricitinib (4 mg QD) was found to be a dominant or cost-effective treatment option versus a number of the currently approved biologic therapies across scenario analyses in these two populations. Only an extreme assumption around progression of disease on baricitinib treatment appeared to have a significant impact on results.

5.9 *Subgroup analysis*

No specific subgroup analyses in addition to the results presented for each of the four populations described in Section 5.2.1 are included in the submission.

5.10 *Validation*

5.10.1 Validation of de novo cost-effectiveness analysis

The model was validated by four experts, who were asked to judge the appropriateness of the conceptual model. The experts included two rheumatologists and two health economists with experience in health economic modelling within the therapeutic area of RA. Both clinicians, who are from two different European countries, are Professors for Rheumatology with more than 20 years of professional experience each and have acted as Principal Investigators of landmark studies in RA. An advisory panel with all four experts was organised prior to model development to gain an understanding on the appropriateness of the conceptual model framework. In addition, an SLR of published economic evaluations in RA, and accompanying targeted literature review of published economic evaluations of baricitinib in RA, were conducted in order to establish the most appropriate conceptual framework for the analysis. The results of these literature reviews are reported in Section 5.1. Finally, the de novo cost-effectiveness analysis follows closely the economic analysis developed by the AG and considered by the NICE Committee in TA375. Where appropriate, the assumptions applied in the model aimed to be consistent with those applied in TA375, as detailed in the preceding sections.

The final computerized model was subjected to several rounds of rigorous review and quality control (QC). These processes included review of data inputs against sources, review of calculations for accuracy, review of code in Visual Basic for Applications (VBA) for accuracy and speed, review of the implementation and translation of the conceptual model and its underlying assumptions and ‘extreme value’ testing of the model. A QC checklist was also completed, assessing components of the model such as accuracy, consistency, functionality, clarity, input sheets calculations and sensitivity analyses. Detected outages were recorded, subsequently rectified, and the model then re-reviewed to ensure correct implementation of error rectification. This was conducted by three experienced modellers: (a) the modeller responsible for model development, (b) a modeller who was involved in the conceptual model framework development, and (c) a modeller not involved in model development. Furthermore, numerous individual patients were tracked through the model to determine whether the model logic is correct. This was done by running the model code and monitoring the flow of sampled, individual patients through each line of the VBA code and throughout their disease progression (i.e. either to the end of follow-up or to death). In addition, the model prints the calculated costs and outcomes of a sample of individual modelled patients.

The outputs of the model are costs and QALYs, whereas the data sources used for the model in terms of efficacy are trials or evidence synthesis of RCTs which reported outcomes such as EULAR or ACR response but not QALYs. This means that the feasibility of dependent validation of model outcomes against the data sources on which it is based is limited. As an approximate validation, the direction of results from the modelled comparisons was assessed against the data sources (i.e. JADV trial, NMA results for each modelled population) and found to be consistent. In addition, baseline demographics of modelled patients (i.e. age, gender distribution, HAQ score) were compared with the relevant baricitinib clinical trials and found to be comparable.

5.11 *Interpretation and conclusions of economic evidence*

The cost-effectiveness analysis presented here provides a robust economic evaluation in the multiple RA patient populations requested in the final scope of this appraisal. Comparators and treatment sequences have been matched to expected clinical practice in each population and, where possible, data have been drawn from the specific populations under consideration for each intervention. As such, the economic evaluation provides relevant cost-effectiveness results for each population that might be considered for treatment with baricitinib in clinical practice in the NHS.

In addition to the consideration of relevant populations, the analysis can be considered generalisable to the NHS with regards to its inputs. In particular, as far as possible the inputs used in the recent model produced for TA375 and considered relevant by NICE for decision-making have been used. For example, drug costs have been sourced from UK formularies and monitoring costs applied have been inflated from those used in TA375.

As well as the generalisability to the UK population, a major strength of the economic evaluation is the use of robust relative efficacy estimates informed by NMA. In the cDMARD-IR, severe population in particular, there is a considerable amount of published clinical data on baricitinib and relevant comparator therapies, allowing a comprehensive network to be formed to provide evidence of the relative efficacy of baricitinib and comparator therapies. In the TNF-IR populations, data availability for comparator therapies is more limited; a weakness of the economic evaluation in the severe, TNF-IR populations is therefore the lack of availability of data on some comparators in this population and the resultant requirements for strong assumptions over comparator efficacy.

Another strength of the analysis is the range of scenario analyses performed to test key assumptions underlying the base case analysis. In the base case analysis, baricitinib trial data is used to inform patient baseline characteristics, but a scenario analysis explores the use of the BSRBR database to further align with NICE TA375 and the likely characteristics of UK patients who would receive biologic therapy. Scenario analyses also provide a robust analysis of the impact of using alternative utility algorithms and different assumptions around HAQ progression both with cDMARDs and palliative care, and also baricitinib therapy. These have represented key points of discussion in previous models in this disease area.

An inevitable weakness of the analysis is the limited long-term data on HAQ progression with baricitinib in order to validate the appropriateness of the HAQ trajectory (assumed flat) applied for this intervention in the base case analysis. The RA-BEYOND study provides longer-term evidence for the maintenance of HAQ-DI scores with baricitinib as well as similar outcomes with respect to radiographic progression compared to adalimumab in RA-BEAM. Nevertheless, longer-term data collection on such outcomes in the future would further support the validity of this assumption.

As has been highlighted previously, there is a limitation in data availability for all comparators in the scope when considering the TNF-IR population who are not eligible for rituximab. Despite this uncertainty, cost and QALY outcomes were similar across interventions.

In the moderate population, cost-effectivness was demonstrated when linear progression on cDMARDs was assumed. This is analgous to the ‘rapid-progressor’ population discussed in TA375 and the associated appeal. Work summarised in section 2.4 indicates that this population is likely to be relatively small (perhaps less than 20% of the overall moderate population) and that a number of factors could be considered in identifying these patients such as response to cDMARD treatment, ACPA status and ultrasound. Clinical judgment in the indentification of these patients and extended treatment options available could be considered.

In summary, the economic evaluation presented aims to replicate the analysis performed for TA375 as far possible, with adjustments made where appropriate to incorporate additional populations relevant to the decision problem and to clinical practice in the UK and to incorporate the extensive body of baricitinib clinical trial data in order to model this intervention. Doing so has demonstrated that baricitinib is likely a cost-effective option in a number of the populations outlined in the scope for this appraisal.

6 Assessment of factors relevant to the NHS and other parties

Summary of Budget Impact Analysis

- An analysis was conducted to explore the budget impact of the introduction of baricitinib in two patient populations between the years 2017–2021: severe, cDMARD-IR patients and anti-TNF-IR patients.
- The analysis was based on predicted market share estimates and took into account drug acquisition costs (including drug wastage), drug administration costs and monitoring costs.
- In the severe, cDMARD-IR population, the introduction of baricitinib is anticipated to be associated with cost savings, ranging from [redacted] in Year 1 to [redacted] in Year 5.
- In the anti-TNF-IR population, the introduction of baricitinib is also anticipated to be associated with cost savings, ranging from [redacted] in Year 1 to [redacted] in Year 5.
- These results demonstrate the significant cost savings that could be achieved by the NHS through the introduction of baricitinib, presenting further evidence of the value of baricitinib in addition to the evidence of clinical and cost-effectiveness.

6.1 *Projected prevalence and incidence of populations of interest*

Two patient populations were considered in this budget impact analysis:

1. Patients with severely active RA who have failed cDMARDs in England;
2. Patients with moderately to severely active RA who have failed anti-TNF treatment in England.

In order to calculate the prevalence and incidence of the two populations of interest, data for the total adult (≥18 years) population of England were first obtained from the Office of National Statistics (ONS).[171] A population growth factor of 0.80% was applied to estimate the national population in future years based on the average annual growth in England as reported by the ONS.[172]

A prevalence rate for RA of 0.86% was taken from a study by Symmonds *et al.* 2002,[46] a UK-specific RA prevalence study, and an incidence of 0.47 cases per 1,000 person-years was identified from the National Audit Office.[173]

The projected prevalent and incident population of RA in adults in England between the years 2017–2021 is presented in Table 135.

Table 135. Projected prevalent and incident population of RA in England 2017–2021

	Base	2017	2018	2019	2020	2021
Projected population in England aged ≥18 years	-	43,865,113	44,216,034	44,569,762	44,926,320	45,285,731
Projected prevalence of RA in adults in England	0.86%	377,240	397,892	418,709	439,693	460,844
Projected incidence of RA in adults in England (cases per 1,000 person-years)	0.47 cases per 1000 person-years	20,652	20,817	20,984	21,152	21,321

Abbreviation: RA = rheumatoid arthritis

The following sections illustrate how the number of patients in each specific population of interest was calculated.

6.1.1 Projected prevalence and incidence in the severe, cDMARD-IR population

The prevalence and incidence of patients with severely active RA was identified to be 13%.[174] It was assumed that 100% of these patients would be diagnosed and that all diagnosed patients would be eligible for, and receive treatment. Finally, it was assumed that treatment with cDMARDs would fail for 100% of patients with severely active RA. The prevalence and incidence calculations for the severe, cDMARD-IR population are presented in Table 136 and Table 137.

Table 136. Projected prevalent population of severe, cDMARD-IR patients in England 2017–2021

	Base	2017	2018	2019	2020	2021
Prevalence of patients with severely active RA	-	49,041	51,726	54,432	57,160	59,910
Prevalence of patients diagnosed with severely active RA	100%	49,041	51,726	54,432	57,160	59,910
Prevalence of patients eligible for treatment with cDMARDs	100%	49,041	51,726	54,432	57,160	59,910
Prevalence of patients treated with cDMARDs	100%	49,041	51,726	54,432	57,160	59,910
Prevalence of patients for whom cDMARDs have failed	100%	49,041	51,726	54,432	57,160	59,910

Abbreviations: RA = rheumatoid arthritis, cDMARDs = conventional disease-modifying antirheumatic drugs, IR = inadequate response

Table 137. Projected incident population of severe, cDMARD-IR patients in England 2017–2021

	Base	2017	2018	2019	2020	2021
Incidence of patients with severely active RA	-	2,685	2,706	2,728	2,750	2,772
Incidence of patients diagnosed with severely active RA	100%	2,685	2,706	2,728	2,750	2,772
Incidence of eligible patients for treatment with cDMARDs	100%	2,685	2,706	2,728	2,750	2,772
Incidence of patients treated with cDMARDs	100%	2,685	2,706	2,728	2,750	2,772
Incidence of patients for whom cDMARDs have failed	100%	2,685	2,706	2,728	2,750	2,772

Abbreviations: cDMARDs = conventional disease-modifying antirheumatic drugs, IR = inadequate response, RA = rheumatoid arthritis,

Estimated number of severe, cDMARD-IR patients treated with bDMARDs

The interventions currently available in UK clinical practice for the treatment of patients with severely active RA for whom treatment with cDMARDs has failed comprise a set of bDMARDs (see Section 3.3). It is these patients who currently receive bDMARD therapy following failure of cDMARDs that would be considered for baricitinib as an alternative to current bDMARDs.

In the NICE costing template for tocilizumab in TA198,[175] it was identified that 77% of severe, cDMARD-IR patients receive treatment with bDMARDs. The projected prevalent and incident

populations for the population of severe, cDMARD-IR patients treated with bDMARDs are provided in Table 138 and Table 139, respectively.

Table 138. Projected population of severe, cDMARD-IR patients treated with bDMARDs

	Base	2017	2018	2019	2020	2021
Proportion of patients treated with bDMARDs	77%	37,724	39,789	41,871	43,969	46,084

Abbreviation: bDMARDs= biologic disease-modifying antirheumatic drugs

Table 139. Projected incident population of severe, cDMARD-IR patients treated with bDMARDs

	Base	2017	2018	2019	2020	2021
Proportion of incident patients treated with bDMARDs	77%	2,065	2,082	2,098	2,115	2,132

Abbreviation: cDMARD = conventional disease-modifying antirheumatic drug, IR = inadequate response, bDMARDs= biologic disease-modifying antirheumatic drugs

6.1.2 Projected prevalence and incidence in the moderate to severe anti-TNF-IR population

The prevalence and incidence of patients with moderately to severely active RA was identified to be 62%.[174] It was assumed that 100% of these patients would be diagnosed and that all diagnosed patients would be eligible for, and receive, treatment with cDMARDs. It was further assumed that cDMARD therapy would fail for 100% of these patients. Based on the NICE costing template for tocilizumab in TA198, 16% of these patients would subsequently be treated with bDMARDs.[175] Also based on the NICE costing statement for tocilizumab (TA198), it was estimated that anti-TNF treatment would fail for 30% of patients.[175] The projected prevalent and incident populations for this population of interest are presented Table 140 and Table 141, respectively.

Table 140. Projected prevalent population of severe, anti-TNF-IR population treated with subsequent bDMARDs in England 2017–2021

	Base	2017	2018	2019	2020	2021
Prevalence of patients with moderately to severely active RA	62%	233,889	246,693	259,600	272,609	285,723
Prevalence of patients diagnosed with moderately to severely active RA	100%	233,889	246,693	259,600	272,609	285,723
Proportion of patients eligible for treatment with cDMARDs	100%	233,889	246,693	259,600	272,609	285,723
Proportion of patients treated with cDMARDs	100%	233,889	246,693	259,600	272,609	285,723
Proportion of patients for whom cDMARDs fail	100%	233,889	246,693	259,600	272,609	285,723

Proportion of patients treated with bDMARDs	16%	37,724	39,789	41,871	43,969	46,084
Proportion of patients for whom anti-TNFs fail	30%	11,317	11,937	12,561	13,191	13,825

Abbreviations: TNF = tumor necrosis factor, IR = inadequate response, bDMARD = biologic disease-modifying antirheumatic drugs, cDMARDs = conventional disease-modifying antirheumatic drugs, RA = rheumatoid arthritis

Table 141. Projected incident population of severe, anti-TNF-IR population treated with subsequent bDMARDs in England 2017–2021

	Base	2017	2018	2019	2020	2021
Incidence of patients with moderately to severely active RA	62%	12,804	12,907	13,010	13,114	13,219
Incidence of patients diagnosed with moderately to severely active RA	100%	12,804	12,907	13,010	13,114	13,219
Proportion of incident patients eligible for treatment with cDMARDs	100%	12,804	12,907	13,010	13,114	13,219
Porportion of incident patients treated with cDMARDs	100%	12,804	12,907	13,010	13,114	13,219
Proportion of incident patients for whom cDMARDs fail	100%	12,804	12,907	13,010	13,114	13,219
Proportion of incident patients treated with bDMARDs	16%	2,065	2,082	2,098	2,115	2,132
Proportion of incident patients for whom anti-TNFs fail	30%	620	625	630	635	640

Abbreviations: TNF = tumor necrosis factor, IR = inadequate response, bDMARD = biologic disease-modifying antirheumatic drugs, cDMARDs = conventional disease-modifying antirheumatic drugs, RA = rheumatoid arthritis

Estimated number of moderate to severe, anti-TNF-IR patients treated with further bDMARDs

The interventions currently available in UK clinical practice for the treatment of patients with moderately to severely active RA who have failed anti-TNF treatment comprise a set of multiple bDMARDs (see Section 3.3). It is these patients who currently receive a second bDMARD after their initial bDMARD therapy fails that would be considered for treatment with baricitinib as an alternative to current bDMARDs.In the population calculations, it was assumed that 100% of moderate to severe, anti-TNF-IR patients are treated with further bDMARDs. The projected prevalent and incident populations for the population of moderate to severe anti-TNF-IR patients treated with bDMARDs are presented in Table 142 and Table 143, respectively.

Table 142. Projected prevalent population of moderate to severe, anti-TNF-IR patients treated with bDMARDs

	Base	2017	2018	2019	2020	2021
Patients treated with bDMARDs	100%	11,317	11,937	12,561	13,191	13,825

Abbreviation: bDMARDs= biologic disease-modifying antirheumatic drugs, TNF = tumor necrosis factor, IR = inadequate response

Table 143. Projected incident population of moderate to severe, anti-TNF-IR patients treated with bDMARDs

	Base	2017	2018	2019	2020	2021
Incident patients treated with bDMARDs	100%	620	625	630	635	640

Abbreviation: bDMARDs= biologic disease-modifying antirheumatic drugs, TNF = tumor necrosis factor, IR = inadequate response

6.1.3 Predicted uptake of baricitinib

Two sets of market share projections are used in the model for each population of interest: one describing an NHS without baricitinib and the other describing an NHS with baricitinib. In the NHS with baricitinib, baricitinib displaces other treatment regimens by an amount equivalent to its uptake. Default market shares were calculated using sales data obtained by Lilly UK for drugs used in RA (Table 144 and Table 145).[174]

Table 144. Severe, cDMARD-IR population: NHS without baricitinib

	2017	2018	2019	2020	2021
ABTS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CTZ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN-b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX-b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TCZ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviation: cDMARDs= conventional disease-modifying antirheumatic drugs, IR = inadequate response, NHS = National Health Service

Table 145. Severe, cDMARD-IR population: NHS with baricitinib

	2017	2018	2019	2020	2021
BAR2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAR4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ABTS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CTZ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN-b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX-b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TCZ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviation: cDMARDs= conventional disease-modifying antirheumatic drugs, IR = inadequate response, NHS = National Health Service

Table 146 and Table 147 show the market share projections used in the budget impact analyses for the severe, TNF-IR population.

Table 146. Moderate to severe, anti-TNF-IR population: NHS without baricitinib

	2017	2018	2019	2020	2021
ABTS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CTZ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN-b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX-b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TCZ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
RTX	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviation: TNF = tumour necrosis factor, IR = inadequate response, NHS = National Health Service

Table 147. Moderate to severe, anti-TNF-IR population: NHS with baricitinib

	2017	2018	2019	2020	2021
BAR2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAR4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ABTS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ADA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CTZ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ETN-b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
GOL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
IFX-b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TCZ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
RTX	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviation: TNF = tumour necrosis factor, IR = inadequate response, NHS = National Health Service

6.1.4 Costs included in the budget impact analysis

The model includes costs for drug acquisition, drug administration and patient monitoring, further details of which are provided in the sections below.

Drug acquisition costs

Drug acquisition costs, other than baricitinib, are from 2016 and were derived from the British National Formulary (BNF).[176] The acquisition costs of all interventions included in the analysis for either population are presented in Table 148. Table 149 provides information on weight-based

dosing, which influences drug acquisition costs and drug wastage. The price of baricitinib in the model was the price when applying the confidential Patient Access Scheme discount.

Table 148. Drug acquisition costs for interventions included in the budget impact analysis

Drug	Unit per vial / tab	Unit of measure	Mg per pack	Vials / tablets per pack	Cost per pack	Cost per vial / tab	Source	Comment
ABTS	125	mg	125	1	£302.40	£302.40	BNF (Nov 2016)	125 mg weekly starting, assumed weight 60–100 kg to align with BAR baseline patient characteristics (no IV loading dose assumed)
ADA	40	mg	80	2	£704.28	£352.14	BNF (Nov 2016)	Initially 80 mg, then 40 mg every 2 weeks, to be started 1 week after initial dose
BAR2	2	mg	56	28	[redacted]	[redacted]	Eli Lilly	UK Olumiant Launch Price (PAS)
BAR4	4	mg	112	28	[redacted]	[redacted]	Eli Lilly	UK Olumiant Launch Price (PAS)
CTZ	200	mg	400	2	£715.00	£357.50	BNF (May 2016)	400 mg every 2 weeks for 3 doses then 200 mg every 2 weeks (SC); first three months of treatment at no charge
ETN	50	mg	200	4	£715.00	£178.75	BNF (Nov 2016)	50 mg weekly by SC (priced according to Enbrel 50 mg PFS)
ETN-b	50	mg	200	4	£656.00	£164.00	BNF (Nov 2016)	Benepali 50 mg /mL pre-filled syringe (x4)
GOL	50	mg	50	1	£762.97	£762.97	BNF (Nov 2016)	Initially 200 mg, then 100 mg after 2 weeks then 50mg every 4 weeks, assumed body weight upto 80 kg; price for Simponi 50 mg PFS used
IFX	100	mg	100	1	£419.62	£419.62	BNF (May 2016), dose: TA195 (2009)	3 mg/kg, then 3 mg/kg after 2 weeks then 3 mg/kg after 4 weeks then 3 mg/kg every 8 weeks; assumed body weight >67 kg and <100 kg
IFX-b	100	mg	100	1	£377.66	£377.66	BNF (Nov 2016)	Average cost of Inflectra 100 mg and Remsima 100 mg vials
RTX	500	mg	500	1	£873.15	£873.15	BNF (May 2016), dose: Table 162, NICE MTA report	2000 mg every 9 months following NICE MTA, 1 g given in two weeks intervals according to BNF (Nov 2016)
TCZ	80	mg	80	1	£102.40	£102.40	BNF (Nov 2016)	8 mg/kg (70 kg assumed weight) once every 4 weeks, drug wastage not included (IV)

Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR2 = baricitinib (2 mg QD), BAR4 = baricitinib 4 mg (QD), CTZ = certolizumab pegol, ETN =etanercept, ETN-b = etanercept biosimilar, GOL = golimumab, IFX = infliximab, IFX-b = infliximab biosimilar, PAS = patient access scheme, RTX = rituximab, TCZ = tocilizumab

Weight-based dosing and wastage

The three trials in the baricitinib clinical trials programme, RA-BEAM (JADV), RA-BUILD (JADX) and RA-BEACON (JADW), were used to calculate the average weight of patients for the cDMARD-IR and TNF-IR populations, as presented in Table 149. The average weights were used to calculate drug doses and wastage for weight-based treatments in the model (i.e. infliximab/infliximab biosimilar and tocilizumab).

Table 149. Patient weight data used for weight-based dosing

Severity	Population	Patient weight (kg), BAR (2 mg QD)	Patient weight (kg), BAR (4 mg QD)	Average weight	Source/comment	Population
Severe	Patients who have failed on cDMARDs	74.94	72.70	73.82	JADV & JADX, weighted average for BAR (4 mg QD)	Patients who have failed on cDMARDs
Total	Patients who have failed on anti-TNFs	83.01	80.67	81.84	JADW	Patients who have failed on anti-TNFs

Abbreviations: BAR = baricitinib, cDMARD = conventional disease-modifying antirheumatic drugs, QD = once daily, TNF = tumour necrosis factor

In order to model drug wastage, it was assumed that any excess medication left in a vial after administration of the recommended weight-based dose would be discarded to clinical waste. Of all comparator treatments included in the model, drug wastage only applies to infliximab/infliximab biosimilar and tocilizumab. Table 150 presents the number of units (vials) used in year 1 (accounting for loading doses) and in subsequent years.

Table 150. Wastage for infliximab, infliximab biosimilar and tocilizumab (units per year)

Treatment	Loading dose (units in Year 1)	Maintenance dose (units per year)
IFX/ IFX-b	24.00	21.00
TCZ	117.00	117.00

Abbreviations: IFX = infliximab, IFX-b = infliximab biosimilar, TCZ = tocilizumab

Administration costs

The budget impact analysis also considers administration costs, which are split into two categories:

1. Cost per intravenous (IV) injection: £159.78; this cost was sourced from TA375,[14] and is based on the assumption that an IV administration requires 60 minutes.

2. Cost per subcutaneous (SC) injection: £2.71; this cost was sourced from TA375, [14] and it was assumed that 10% of patients require administration support by a district nurse.

Oral drugs are assumed to be self-administered and incur no charge.

Administration costs for each intervention are presented in Table 151.

Table 151. Administration costs

Drug	Route of administration	Annual cost	Add. cost in Year 1	Annual IV/SC administrations	Additional administrations in Year 1
ABTS	SC	£35.23	£0.00	13	0
ADA	SC	£70.46	£0.00	26	0
BAR2	Oral	£0.00	£0.00	0	0
BAR4	Oral	£0.00	£0.00	0	0
CTZ	SC	£70.46	£0.00	26	0
ETN	SC	£140.92	£0.00	52	0
ETN-b	SC	£140.92	£0.00	52	0
GOL	SC	£32.52	£0.00	12	0
IFX	IV	£1,118.46	£159.78	7	1
IFX-b	IV	£1,118.46	£159.78	7	1
RTX	IV	£426.08	£0.00	2.7	0
TCZ	IV	£2,077.14	£0.00	13	0

Abbreviations: ABTS = subcutaneous abatacept, ADA = adalimumab, BAR2 = baricitinb (2 mg QD), BAR4=baricitnib 4 mg (QD), CTZ = certolizumab pegol, ETN =etanercept, ETN-b = etanercept biosimilar, GOL = golimumab, IFX = infliximab, IFX-b = infliximab biosimilar, IV = intravenous, RTX = rituximab, SC = subcutaneous, TCZ = tocilizumab.

Monitoring costs

The BIM calculates the monitoring costs for patients at three different time points:

1. Before treatment initiation: £176.38
2. In the first 6 months of treatment: £1,763.79
3. At monthly intervals: £139.03

These costs are used to assign annual costs and additional costs in Year 1. Costs were derived from the NICE MTA report Table 163 (inflated to 2016).[14] As monitoring requirements were assumed to be the same for all bDMARDs, for all treatments included in the BIM the annual monitoring cost was £1,668.34 and the additional cost in year was £1,106.00.

6.1.5 Additional cost savings

There is a cost saving associated with the oral administration of baricitinib in comparison to bDMARDs, which are subcutaneously or intravenously administered and have associated administration costs. This has been incorporated into the above unit costs. No other additional resource savings are anticipated.

6.1.6 Results of the budget impact analysis

Severe, cDMARD-IR population

The annual budget impact anticipated in the severe, cDMARD-IR population following the introduction of baricitinib with a PAS applied is presented in Table 152. Baricitinib (with PAS) is anticipated to result in a negative budget impact each year, i.e. annual cost saving rising from [redacted] in Year 1 to [redacted] in Year 5.

Table 152. Budget impact results for the severe, cDMARD-IR population 2017–2021 – with PAS

	2017	2018	2019	2020	2021
NHS without baricitinib					
ABTS	£21,175,221	£22,279,997	£23,393,611	£24,516,134	£25,647,637
ADA	£134,404,108	£132,235,468	£129,207,676	£125,310,198	£120,532,386
CTZ	£31,190,708	£32,831,700	£34,485,821	£36,153,175	£37,833,868
ETN	£143,305,763	£136,753,220	£128,861,244	£119,613,522	£108,993,568
ETN-b	£17,662,612	£22,933,696	£28,646,334	£34,805,795	£41,417,403
GOL	£24,295,209	£30,129,371	£36,429,644	£43,201,578	£50,450,787
IFX	£3,487,990	£8,577,834	£14,156,735	£20,230,558	£26,805,228
IFX-b	£17,866,187	£23,329,341	£29,250,491	£35,635,102	£42,488,699
TCZ	£68,660,655	£78,291,165	£88,554,807	£99,459,064	£111,011,498
Total	£462,048,453	£487,361,794	£512,986,363	£538,925,125	£565,181,072
NHS with baricitinib					
BAR2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAR4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ABTS	£20,963,469	£21,611,597	£22,223,930	£22,064,520	£21,800,491
ADA	£133,060,067	£137,162,571	£141,038,255	£140,016,944	£138,332,696
CTZ	£30,878,800	£31,846,749	£32,761,530	£32,537,858	£32,158,787
ETN	£141,872,706	£146,247,527	£150,380,486	£149,292,043	£147,496,697
ETN-b	£17,485,986	£18,024,900	£18,534,013	£18,399,620	£18,178,127
GOL	£24,052,257	£24,793,816	£25,494,376	£25,309,745	£25,005,282
IFX	£3,453,110	£3,558,687	£3,658,407	£3,631,157	£3,586,789
IFX-b	£17,687,525	£18,228,072	£18,738,629	£18,598,844	£18,371,403
TCZ	£67,974,049	£70,073,570	£72,057,094	£71,538,500	£70,680,877
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Net budget impact (with PAS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Anti-TNF-IR population

The annual budget impact anticipated in the moderate to severe, anti-TNF-IR population following the introduction of baricitinib with a PAS applied is presented in Table 153. Baricitinib (with PAS) is anticipated to result in a negative budget impact each year, i.e. annual cost saving rising from [redacted] in Year 1 to [redacted] in Year 5.

Table 153. Budget impact results for the anti-TNF-IR population 2017–2021 – with PAS

	2017	2018	2019	2020	2021
NHS without baricitinib					
ABTS	£5,661,279.75	£5,956,645.91	£6,254,375.00	£6,554,485.92	£6,856,997.73
ADA	£35,933,473.44	£35,054,339.61	£33,915,631.75	£32,514,175.09	£30,846,760.66
CTZ	£8,338,959.84	£8,777,685.82	£9,219,921.60	£9,665,695.27	£10,115,035.13
ETN	£38,313,366.32	£36,103,904.71	£33,490,663.51	£30,468,733.49	£27,033,152.73
ETN-b	£4,722,169.58	£6,273,468.49	£7,956,993.17	£9,774,322.47	£11,727,052.02
GOL	£6,495,420.86	£8,204,358.44	£10,052,798.45	£12,042,404.09	£14,174,856.26
IFX	£932,527.98	£980,854.79	£1,029,568.21	£1,078,671.34	£1,128,167.30
IFX-b	£4,776,596.19	£6,385,323.99	£8,131,208.64	£10,015,888.08	£12,041,017.71
TCZ	£18,356,699.56	£21,128,975.18	£24,090,255.41	£27,242,777.69	£30,588,803.14
RTX	£8,844,116.01	£10,158,468.02	£11,561,604.91	£13,054,577.39	£14,638,447.31
Total	£132,374,609.54	£139,024,024.96	£145,703,020.65	£152,411,730.84	£159,150,290.00
NHS with baricitinib					

BAR2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BAR4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ABTS	£5,604,666.96	£5,777,946.54	£5,941,656.25	£5,899,037.33	£5,828,448.07
ADA	£35,574,138.71	£36,670,959.35	£37,707,138.88	£37,434,087.40	£36,983,796.99
CTZ	£8,255,570.25	£8,514,355.24	£8,758,925.52	£8,699,125.75	£8,597,779.86
ETN	£37,930,232.65	£39,099,858.60	£40,204,821.71	£39,913,822.09	£39,433,829.06
ETN-b	£4,674,947.89	£4,819,028.69	£4,955,142.29	£4,919,211.57	£4,859,994.71
GOL	£6,430,466.65	£6,628,725.49	£6,816,022.87	£6,766,661.16	£6,685,261.81
IFX	£923,202.70	£951,429.15	£978,089.80	£970,804.21	£958,942.20
IFX-b	£4,728,830.23	£4,873,347.69	£5,009,847.00	£4,972,474.95	£4,911,667.85
TCZ	£18,173,132.56	£18,734,447.86	£19,264,751.04	£19,126,102.87	£18,896,813.96
RTX	£8,755,674.85	£9,024,416.83	£9,278,274.03	£9,210,051.57	£9,098,324.62
Total	£132,259,387.27	£138,908,860.44	£145,589,050.62	£151,899,494.62	£158,203,734.89
Net budget impact (with PAS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

6.1.7 Limitations of budget impact analysis

The presented budget impact analysis is limited by the difficulty in obtaining population estimates and market share data for the patient populations corresponding to the exact positions in the treatment pathway where baricitinib may be used in clinical practice. As such, the populations included in the budget impact analysis (severe cDMARD-IR and moderate to severe TNF-IR) do not fully correspond to the three separate populations for which cost-effectiveness estimates are presented in Section 4 (moderate, cDMARD-IR; severe cDMARD-IR and severe anti-TNF-IR).

Additionally, some therapies included in the budget impact analysis are associated with a PAS discount that is confidential, meaning that their true cost to the NHS cannot be accurately represented in the BIM.

6.1.8 Summary of results

The budget impact of using baricitinib with the PAS resulted in a substantial estimated cost saving in Years 1–5 in both the severe, cDMARD-IR population and moderate to severe, anti-TNF-IR populations. In the former population, cost savings ranged from [redacted] in Year 1 to [redacted] in Year 5, and in the latter population [redacted] in Year 1 to [redacted] in Year 5. These results demonstrate the significant cost savings that could be achieved by the NHS through the introduction of baricitinib, presenting further evidence of the value of baricitinib in addition to the evidence of clinical and cost-effectiveness.

7 References

1. Fridman JS, Scherle PA, Collins R, Burn TC, Li Y, Li J, et al. Selective inhibition of JAK1 and JAK2 is efficacious in rodent models of arthritis: preclinical characterization of INCB028050. J Immunol. 2010;184(9):5298-307.
2. Eli Lilly and Company. Olumiant Summary of Product Characteristics. 2016.
3. National Institute for Health and Care Excellence. Baricitinib for treating moderate to severe rheumatoid arthritis: draft scope. 2016.
4. Moreland L. Unmet needs in rheumatoid arthritis. Arthritis Res Ther. 2005;7 Suppl 3:S2-8.
5. Alzabin S, Abraham SM, Taher TE, Palfreeman A, Hull D, McNamee K, et al. Incomplete response of inflammatory arthritis to TNFalpha blockade is associated with the Th17 pathway. Ann Rheum Dis. 2012;71(10):1741-8.
6. Curtis JR, Singh JA. Use of biologics in rheumatoid arthritis: current and emerging paradigms of care. Clin Ther. 2011;33(6):679-707.
7. Mocci G, Marzo M, Papa A, Armuzzi A, Guidi L. Dermatological adverse reactions during anti-TNF treatments: focus on inflammatory bowel disease. J Crohns Colitis. 2013;7(10):769-79.
8. Dougados M, van der Heijde D, Chen Y-C, 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. Annals of the Rheumatic Diseases. 2016.
9. Eli Lilly and Company. Data on File (JADV Clinical Study Report). 2016.
10. Genovese MC, Kremer J, Zamani O, Ludivico C, Krogulec M, Xie L, et al. Baricitinib in Patients with Refractory Rheumatoid Arthritis. N Engl J Med. 2016;374(13):1243-52.
11. Eli Lilly and Company. Data on File (JADX Clinical Study Report). 2016.
12. Eli Lilly and Company. Data on File (JADW Clinical Study Report). 2016.
13. Eli Lilly and Company. Data on File (Summary of Clinical Safety). 2016.
14. 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.
15. Eli Lilly and Company. JADY Abbreviated Clinical Study Report. 2016.
16. Felson DT, LaValley MP. The ACR20 and defining a threshold for response in rheumatic diseases: too much of a good thing. Arthritis Research and Therapy. 2014;16(101).
17. Leonard WJ, O'Shea JJ. Jaks and STATs: biological implications. Annu Rev Immunol. 1998;16:293-322.
18. Mohr A, Chatain N, Domoszlai T, Rinis N, Sommerauer M, Vogt M, et al. Dynamics and non-canonical aspects of JAK/STAT signalling. Eur J Cell Biol. 2012;91(6-7):524-32.
19. Kisseleva T, Bhattacharya S, Braunstein J, Schindler CW. Signaling through the JAK/STAT pathway, recent advances and future challenges. Gene. 2002;285(1-2):1-24.
20. Shi JG, Chen X, Lee F, Emm T, Scherle PA, Lo Y, et al. The pharmacokinetics, pharmacodynamics, and safety of baricitinib, an oral JAK 1/2 inhibitor, in healthy volunteers. J Clin Pharmacol. 2014;54(12):1354-61.

21. Shuai K, Liu B. Regulation of JAK-STAT signalling in the immune system. Nat Rev Immunol. 2003;3(11):900-11.
22. O'Shea JJ, Plenge R. JAK and STAT signaling molecules in immunoregulation and immune-mediated disease. Immunity. 2012;36(4):542-50.
23. Ariza-Ariza R, Navarro-Sarabia F, Hernandez-Cruz B, Rodriguez-Arboleya L, Navarro-Compan V, Toyos J. Dose escalation of the anti-TNF-alpha agents in patients with rheumatoid arthritis. A systematic review. Rheumatology (Oxford). 2007;46(3):529-32.
24. Ali T, Kaitha S, Mahmood S, Ftesi A, Stone J, Bronze MS. Clinical use of anti-TNF therapy and increased risk of infections. Drug, Healthcare and Patient Safety. 2013;5:79-99.
25. Chiu YM, Lang HC, Lin HY, Yang MT, Fang CH, Yang YW, et al. Risk of tuberculosis, serious infection and lymphoma with disease-modifying biologic drugs in rheumatoid arthritis patients in Taiwan. Int J Rheum Dis. 2014;17 Suppl 3:9-19.
26. Chatzidionysiou K, van Vollenhoven RF. When to initiate and discontinue biologic treatments for rheumatoid arthritis? J Intern Med. 2011;269(6):614-25.
27. Montag K, Gingold M, Boers A, Littlejohn G. Disease-modifying anti-rheumatic drug usage, prescribing patterns and disease activity in rheumatoid arthritis patients in community-based practice. Intern Med J. 2011;41(6):450-5.
28. Chilton F, Collett RA. Treatment choices, preferences and decision-making by patients with rheumatoid arthritis. Musculoskeletal Care. 2008;6(1):1-14.
29. Louder AM, Singh A, Saverno K, Cappelleri JC, Aten AJ, Koenig AS, et al. Patient Preferences Regarding Rheumatoid Arthritis Therapies: A Conjoint Analysis. American Health & Drug Benefits. 2016;9(2):84-93.
30. Bolge SC, Goren A, Tandon N. Reasons for discontinuation of subcutaneous biologic therapy in the treatment of rheumatoid arthritis: a patient perspective. Patient preference and adherence. 2015;9:121-31.
31. Darmawan J, Rasker JJ, Nuralim H. Reduced burden of disease and improved outcome of patients with rheumatoid factor positive rheumatoid arthritis compared with dropouts. A 10 year observational study. J Rheumatol Suppl. 2003;67:50-3.
32. Centres for Disease Control and Prevention. Rheumatoid Arthritis (RA) 2015. Available from: http://www.cdc.gov/arthritis/basics/rheumatoid.htm.
33. Isik A KS, Ozturk A, Mermi O Anxiety and depression in patients with rheumatoid arthritis. Clinical Rheumatology. 2007;26(6):872-8.
34. Jorge R BC, Jones A, Natour J. Body image in patients with rheumatoid arthritis. Modern Rheumatology. 2010;20(5):491-5.
35. Listing J, Pattloch D, Kekow J, Kaufmann J, Burmester G-R, Zink A, et al. OP0047 Successful control of disease activity and treatment with biologics increase the life expectancy in rheumatoid arthritis patients. Annals of the Rheumatic Diseases. 2013;71(Suppl 3):69.
36. National Institute for Health and Care Excellence. Rheumatoid arthritis in adults: management 2009. Available from: https://www.nice.org.uk/guidance/cg79.
37. National Institute for Health and Care Excellence. NICE Pathway: Rheumatoid arthritis overview 2016. Available from: http://pathways.nice.org.uk/pathways/rheumatoid-arthritis.

38. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis. 2014;73(3):492-509.
39. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. Arthritis Care Res (Hoboken). 2016;68(1):1-25.
40. National Institute for Health and Care Excellence. Baricitinib for treating moderate to severe rheumatoid arthritis: final scope. 2016.
41. Rubbert-Roth A, Finckh A. Treatment options in patients with rheumatoid arthritis failing initial TNF inhibitor therapy: a critical review. Arthritis Res Ther. 2009;11 Suppl 1:S1.
42. Kahlenberg JM, Fox DA. Advances in the medical treatment of rheumatoid arthritis. Hand Clin. 2011;27(1):11-20.
43. National Rheumatoid Arthritis Society. Know your disease activity score: a patient guide to the DAS28 2015. Available from: http://www.nras.org.uk/data/files/Publications/DAS%20patient%20guide.pdf.
44. Colmegna I, Ohata BR, Menard HA. Current understanding of rheumatoid arthritis therapy. Clin Pharmacol Ther. 2012;91(4):607-20.
45. Borigini M. What is the Difference Between Mild, Moderate and Severe Rheumatoid Arthritis? 2010 [cited 2016 09/12]. Available from: http://www.healthcentral.com/rheumatoid-arthritis/c/53/112480/severe/.
46. 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 (Oxford). 2002;41(7):793-800.
47. National Audit Office. Services for people with rheumatoid arthritis 2009. Available from: https://www.nao.org.uk/report/services-for-people-with-rheumatoid-arthritis/.
48. National Rheumatoid Arthritis Society. The economic burden of rheumatoid arthritis. 2010.
49. Schneider M, Krüger K. Rheumatoid Arthritis—Early Diagnosis and Disease Management. Dtsch Arztebl Int. 2013;110(27-28):477-84.
50. Khurana R BS. Clinical aspects of rheumatoid arthritis. Pathophysiology. 2005;12(3):153-65.
51. National Rheumatoid Arthritis Society. Employment and rheumatoid arthritis. A national picture. 2007.
52. Altman D. Rheumatoid Arthritis (RA): Merck; 2015. Available from: http://www.merckmanuals.com/professional/musculoskeletal-and-connective-tissue-disorders/joint-disorders/rheumatoid-arthritis-ra.
53. Ryan S. Psychological effects of living with rheumatoid arthritis. Nursing Standard. 2014;29(13):52-9.
54. National Rheumatoid Arthritis Society. NRAS Survey 2013: Relationships, emotions, sexuality 2013. Available from: http://www.nras.org.uk/emotions-relationships-and-sexuality-booklet.
55. Peters MJ SD, McCarey D, Dijkmans BA, Nicola P, Kvien TK, et al. EULAR evidence-based recommendations for cardiovascular risk management in patients with rheumatoid arthritis and other forms of inflammatory arthritis. Ann Rheum Dis. 2010;69(325-331).
56. National Rheumatoid Arthritis Society. How is lifespan affected by RA 2016. Available from: http://www.nras.org.uk/how-is-lifespan-affected-by-ra.

57. Cutolo M, Kitas GD, van Riel PL. Burden of disease in treated rheumatoid arthritis patients: going beyond the joint. Semin Arthritis Rheum. 2014;43(4):479-88.
58. Li X GM, Anis AH. The indirect costs of arthritis resulting from unemployment, reduced performance, and occupational changes while at work. Med Care. 2006;44(4):304-10.
59. Bansback N ZW, Walsh D, et al. Factors associated with absenteeism, presenteeism and activity impairment in patients in the first years of RA. . Rheumatology. 2012;51(2):375-84.
60. Riemsma RP, Taal E, Rasker JJ, Klein G, Bruyn GAW, Wouters JMGW, et al. The burden of care for informal caregivers of patients with rheumatoid arthritis. Psychology & Health. 1999;14(5):773-94.
61. Werner B. F. Brouwer, N. Job A. van Exel, Bernard van de Berg, Huib J. Dinant, Marc A. Koopmanschap, Bos GAMvd. Burden of caregiving: Evidence of objective burden, subjective burden, and quality of life impacts on informal caregivers of patients with rheumatoid arthritis. Arthritis Care and Research. 2004;51(4):570-7.
62. American College of Rheumatology. Homepage [06/09/2016]. Available from: http://www.rheumatology.org/.
63. European League Against Rheumatism. EULAR Mission Statement [06/09/2016]. Available from: http://www.eular.org/eular_mission.cfm.
64. National Institute for Health and Care Excellence. Golimumab for the treatment of rheumatoid arthritis after the failure of previous disease-modifying anti-rheumatic drugs (TA225) 2011. Available from: https://www.nice.org.uk/guidance/ta225.
65. National Institute for Health and Care Excellence. Certolizumab pegol for treating rheumatoid arthritis after inadequate response to a TNF-alpha inhibitor (TA415) 2016. Available from: https://www.nice.org.uk/guidance/ta415.
66. National Institute for Health and Care Excellence. Tocilizumab for the treatment of rheumatoid arthritis (TA198) 2010. Available from: https://www.nice.org.uk/guidance/ta198.
67. National Institute for Health and Care Excellence. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying anti-rheumatic drugs and after the failure of conventional disease-modifying anti-rhuematic drugs only: systematic review and economic evaluation. 2013.
68. National Institute for Health and Care Excellence. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor (TA195) 2010. Available from: https://www.nice.org.uk/guidance/ta195.
69. National Institute for Health and Care Excellence. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor 2010. Available from: https://www.nice.org.uk/guidance/ta195.
70. National Institute for Health and Care Excellence. Tocilizumab for the treatment of rheumatoid arthritis 2012. Available from: https://www.nice.org.uk/guidance/ta247.
71. Rendas-Baum R, Wallenstein GV, Koncz T, Kosinski M, Yang M, Bradley J, et al. Evaluating the efficacy of sequential biologic therapies for rheumatoid

arthritis patients with an inadequate response to tumor necrosis factor-α inhibitors. Arthritis Research & Therapy. 2011;13(1):R25-R.
72. Olson AL, Swigris JJ, Sprunger DB, Fischer A, Fernandez-Perez ER, Solomon J, et al. Rheumatoid Arthritis–Interstitial Lung Disease–associated Mortality. American Journal of Respiratory and Critical Care Medicine. 2011;183(3):372-8.
73. European League Against Rheumatism. EULAR Recommendations for the management of rheumatoid arthrits - 2016 update. 2016.
74. Kiely P, Walsh D, Williams R, Young A, Early Rheumatoid Arthritis N. Outcome in rheumatoid arthritis patients with continued conventional therapy for moderate disease activity--the early RA network (ERAN). Rheumatology (Oxford). 2011;50(5):926-31.
75. Nikiphorou E CL, Norton S et al. Different levels of moderate disease in rheumatoid arthritis (ra) are associated with varying risk for joint destruction and failure. Time to update das cut-offs for biologic DMARD use? Ann Rheum Dis. 2015;74:139.
76. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264-9, w64.
77. Tacconelli E. Systematic reviews: CRD's guidance for undertaking reviews in health care. The Lancet Infectious Diseases.10(4):226.
78. Eli Lilly and Company. Data on File (JADZ Clinical Study Report). 2016.
79. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010;62(9):2569-81.
80. Mercer LK, Galloway JB, Lunt M, Davies R, Low ALS, Dixon WG, et al. Risk of lymphoma in patients exposed to antitumour necrosis factor therapy: results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis. Annals of the Rheumatic Diseases. 2016.
81. Eli Lilly and Company. Data on File (JADV Table 11.3. Page 237). 2016.
82. Eli Lilly and Company. Data on File (JADV CSR Table JADV.14.63). 2016.
83. Eli Lilly and Company. Data on File (JADX CSR Table JADX.14.61). 2016.
84. Eli Lilly and Company. Data on File (JADW CSR Table JADW.14.46.). 2016.
85. The University of Sheffield. School of Health and Related Research. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying anti-rheumatic drugs and after the failure of conventional disease modifying anti-rhuematic drugs only: systematic review and economic evaluation [7th December 2016]. Available from: https://www.nice.org.uk/guidance/ta375/history.
86. National Institute for Health and Care Excellence (NICE). Single technology appraisal: User guide for company evidence submission template [7th December 2016]. Available from: https://www.nice.org.uk/article/pmg24/resources/non-guidancesingle-technology-appraisal-user-guide-for-company-evidence-submissiontemplate-pdf.
87. Dias SW, N.J., Sutton, A.J., Ades, A.E.;. NICE DSU Technical Support Document 1: Introduction to Evidence Synthesis for Decision Making 2011a.

Available from: http://www.nicedsu.org.uk/TSD1%20Introduction.final.08.05.12.pdf.
88. Dias SW, N.J., Sutton, A.J., Ades, A.E.;. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framewokr for Pairwise and Network Meta-Analysis of Randomised Controlled Trials 2011b. Available from: http://www.nicedsu.org.uk/TSD2%20General%20meta%20analysis%20corrected%202Sep2016v2.pdf.
89. Dias SW, N.J., Sutton, A.J., Ades, A.E.;. NICE DSU Technical Support Document 3: Heterogeneity: Subgroups, Meta-Regression, Bias and Bias-Adjustment. 2011c.
90. Carlin BPL, T.A.;. Bayesian methods for data analysis. 3rd ed. Hall BRC, editor2008.
91. Spiegelhalter DJB, N.G.; Carlin, B.P.; van der Linde, A.; . Bayesian measures of model complexity and fit 2002 [cited 64 4]. 583-639]. Available from: http://www.biostat.umn.edu/~brad/DICpaper_SBCVdL02.pdf.
92. Gelman A, Rubin, DB., . Inference from iterative simulation using multiple sequences 1992 [cited 7 4]. 457-511]. Available from: http://www2.stat.duke.edu/~scs/Courses/Stat376/Papers/ConvergeDiagnostics/GelmanRubinStatSci1992.pdf.
93. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Figure 2.7.4.3. Page 46). 2016.
94. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.1.2.4 Page 49). 2016.
95. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.2. Page 30). 2016.
96. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.3. Page 31). 2016.
97. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.34. Page 197). 2016.
98. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline - Clinical safety data management: definitions and standards for expedited reporting E2A 1994. Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf.
99. Eli Lilly and Company. Clinical Overview Baricitinib Rheumatoid Arthritis. 2016.
100. Eli Lilly and Company. Data on File (Clinical Overview Baricitinib Rheumatoid Arthritis. Table 2.5.5.2. Page 64). 2016.
101. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.2.1.1.1. Page 59). 2016.
102. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.54. Page 903). 2016.
103. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.65. Page 952). 2016.
104. Voskuyl AE. The heart and cardiovascular manifestations in rheumatoid arthritis. Rheumatology (Oxford). 2006;45 Suppl 4:iv4-7.
105. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.4.3. Page 214). 2016.

106. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.4.6 Page 220). 2016.
107. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.4.6 Page 221). 2016.
108. Swann JB, Smyth MJ. Immune surveillance of tumors. J Clin Invest. 2007;117(5):1137-46.
109. Askling J. Malignancy and rheumatoid arthritis. Curr Rheumatol Rep. 2007;9(5):421-6.
110. Scott FI, Mamtani R, Brensinger CM, Haynes K, Chiesa-Fuxench ZC, Zhang J, et al. Risk of Nonmelanoma Skin Cancer Associated With the Use of Immunosuppressant and Biologic Agents in Patients With a History of Autoimmune Disease and Nonmelanoma Skin Cancer. JAMA Dermatol. 2016;152(2):164-72.
111. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.2.4.1 Page 180). 2016.
112. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.2.4.1 Page 187). 2016.
113. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.2.4.1 Page 188-189). 2016.
114. Eli Lilly and Company. Data on File (Clinical Overview Baricitinib Rheumatoid Arthritis. Table 2.5.5.2. Page 69). 2016.
115. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.111. Page 1383). 2016.
116. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.1.9. Page 173). 2016.
117. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.1.6. Page 163). 2016.
118. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.1.6. Page 164). 2016.
119. Miller EA, Ernst JD. Anti-TNF immunotherapy and tuberculosis reactivation: another mechanism revealed. J Clin Invest. 2009;119(5):1079-82.
120. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Section 2.7.4.5.9.1.5.Page 159). 2016.
121. Committee for Medicinal Products for Human Use (CHMP). CHMP Assessment Report. Olumiant. . 2016.
122. Eli Lilly and Company Limited. Data on File (Summary of Clinical Safety Section 2.7.4.5.9.3.4. Page 211). 2016.
123. Eli Lilly and Company. Data on File (Clinical Safety Summary. Figure 2.7.4.18. Page 204). 2016.
124. Eli Lilly and Company. Data on File (Summary of Clinical Safety Section 2.7.4.5.9.3.4. Page 213). 2016.
125. National Institute for Health and Care Excellence. PMG9: Guide to the methods of technology appraisal 2013 [21/10/2016]. Available from: https://www.nice.org.uk/article/pmg9/chapter/Foreword.
126. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ. 1996;313.
127. Diamantopoulos A, Finckh A, Huizinga T, Sungher DK, Sawyer L, Neto D, et al. Tocilizumab in the treatment of rheumatoid arthritis: a cost-effectiveness analysis in the UK. Pharmacoeconomics. 2014;32(8):775-87.

128. Hernández Alava M, Wailoo AJ, Ara R. Tails from the Peak District: Adjusted Limited Dependent Variable Mixture Models of EQ-5D Questionnaire Health State Utility Values. Value in Health. 2012;15(3):550-61.
129. National Institute for Health and Care Excellence. Abatacept for treating rheumatoid arthritis after the failure of conventional disease-modifying anti-rheumatic drugs (rapid review of technology appraisal guidance 234) (TA280) 2013. Available from: https://www.nice.org.uk/guidance/ta234.
130. National Institute for Health and Care Excellence. Abatacept for the treatment of rheumatoid arthritis after the failure of conventional disease-modifying anti-rheumatic drugs (TA234) 2011. Available from: https://www.nice.org.uk/guidance/ta234.
131. National Institute for Health and Care Excellence. Certolizumab pegol for the treatment of rheumatoid arthritis (TA186) 2010. Available from: https://www.nice.org.uk/guidance/ta186.
132. National Institute for Health and Care Excellence. Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis. 2007. Available from: https://www.nice.org.uk/guidance/ta130.
133. Chakravarty K, McDonald H, Pullar T, Taggart A, Chalmers R, Oliver S, et al. BSR/BHPR guideline for disease-modifying anti-rheumatic drug (DMARD) therapy in consultation with the British Association of Dermatologists. Rheumatology (Oxford). 2008;47(6):924-5.
134. Hernandez Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. Rheumatology (Oxford). 2013;52(5):944-50.
135. Malottki K, Barton P, Tsourapas A, Uthman AO, Liu Z, Routh K, et al. 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 Technol Assess. 2011;15(14):1-278.
136. 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 (Oxford). 2013;52(11):2016-24.
137. Office for National Statistics. National Life Tables, United Kingdom: 2011-2013 2014. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2014-09-25.
138. National Institute for Health and Care Excellence. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, abatacept and tocilizumab - review: assessment report 2016. Available from: https://www.nice.org.uk/guidance/TA375/documents/rheumatoid-arthritis-adalimumab-etanercept-infliximab-certolizumab-pegol-golimumab-abatacept-and-tocilizumab-review-assessment-group-report2.
139. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
140. Centre for Reviews and Dissemination (CRD). Systematic Reviews: CRD's guidance for undertaking reviews in health care 2009 [9th November 2016]. Available from: www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf.
141. Alemao E, Joo S, Kawabata H, Al MJ, Allison PD, Rutten-Van Molken MPMH, et al. Effects of Achieving Target Measures in Rheumatoid Arthritis on

Functional Status, Quality of Life, and Resource Utilization: Analysis of Clinical Practice Data. Arthritis Care and Research. 2016;68(3):308-17.
142. Buitinga L, Braakman-Jansen LMA, Taal E, Kievit W, Visser H, Van Riel PLCM, et al. Comparative responsiveness of the EuroQol-5D and Short Form 6D to improvement in patients with rheumatoid arthritis treated with tumor necrosis factor blockers: Results of the Dutch Rheumatoid Arthritis Monitoring registry. Arthritis Care and Research. 2012;64(6):826-32.
143. Chang HC, Chen LC, Tsai WC. Assessing health-related quality of life in taiwanese rheumatoid arthritis patients by applying time-trade-off and EQ-5D methods. Value in Health. 2010;13 (3):A130.
144. Cho SK, Kim D, Jun JB, Bae SC, Sung YK. Factors influencing quality of life (QOL) for Korean patients with rheumatoid arthritis (RA). Rheumatology International. 2013;33(1):93-102.
145. Dasgupta B, Combe B, Louw I, Wollenhaupt J, Zerbini CAF, Beaulieu A, et al. Patient and physician expectations of add-on treatment with golimumab for rheumatoid arthritis: Relationships between expectations and clinical and quality of life outcomes. Arthritis Care and Research. 2014;66(12):1799-807.
146. Han C, Bingham CO, Westhovens R, Weinblatt ME, Kim L, Baker D, et al. A comparison of EQ5D index from the UK, US, and Japan preference weights model, and mapping algorithm from clinical outcomes in patients with rheumatoid arthritis: Results of from simponi aria study [Conference Abstract]. 2015 [updated June; cited 74]. 225-6]. Available from: http://ard.bmj.com/content/74/Suppl_2/225.3.abstract?sid=d3d8d12e-20e2-49ad-9af2-ee56cfb07ab3
http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=emed13&AN=72151665.
147. Han C, Bingham ICO, Westhovens R, Weinblatt M, Kim L, Baker D, et al. A comparison of EQ5D index from the UK, US, and Japan preference weights model, and mapping algorithm from clinical outcomes in patients with rheumatoid arthritis: Results from golimumab intravenous study. Arthritis and Rheumatology Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP. 2015;67(no pagination).
148. Harrison MJ, Davies LM, Bansback NJ, McCoy MJ, Verstappen SMM, Watson K, et al. The comparative responsiveness of the EQ-5D and SF-6D to change in patients with inflammatory arthritis. Quality of Life Research. 2009;18(9):1195-205.
149. Horslev-Petersen K, Hetland ML, Junker P, Podenphant J, Ellingsen T, Ahlquist P, et al. Adalimumab added to a treat-to-target strategy with methotrexate and intra-articular triamcinolone in early rheumatoid arthritis increased remission rates, function and quality of life. The OPERA Study: an investigator-initiated, randomised, double-blind, parallel-group, placebo-controlled trial. Ann Rheum Dis. 2014;73(4):654-61.
150. Hoshi D, Tanaka E, Igarashi A, Inoue E, Kobayashi A, Sugimoto N, et al. Profiles of EQ-5D utility scores in the daily practice of Japanese patients with rheumatoid arthritis; Analysis of the IORRA database. Modern Rheumatology. 2016;26(1):40-5.
151. Hoshi D, Tanaka E, Inoue E, Shidara K, Shimizu Y, Kobayashi A, et al. Minimally important difference in the European quality of life-five dimensions

in patients with rheumatoid arthritis. Arthritis and Rheumatology. 2014;66:S183-S4.
152. Jobanputra P, Maggs F, Deeming A, Carruthers D, Rankin E, Jordan AC, et al. A randomised efficacy and discontinuation study of etanercept versus adalimumab (RED SEA) for rheumatoid arthritis: a pragmatic, unblinded, non-inferiority study of first TNF inhibitor use: outcomes over 2 years. BMJ Open. 2012;2(6).
153. Kievit W, Fransen J, Adang EM, den Broeder AA, Bernelot Moens HJ, Visser H, et al. Long-term effectiveness and safety of TNF-blocking agents in daily clinical practice: results from the Dutch Rheumatoid Arthritis Monitoring register. Rheumatology (Oxford). 2011;50(1):196-203.
154. Kirkham B, Chan E, Vincent A, Elliott A. REAL-Life Cost of UK Healthcare Resource for Patients with Rheumatoid Arthritis, Comparing High and Low/Remission Disease States [abstract]. Arthritis and Rheumatology. 2016;68(Suppl 10).
155. Kobelt G, Eberhardt K, Geborek P. TNF inhibitors in the treatment of rheumatoid arthritis in clinical practice: Costs and outcomes in a follow up study of patients with Ra treated with etanercept or infliximab in southern Sweden. Annals of the Rheumatic Diseases. 2004;63(1):4-10.
156. Lie E, Lillegraven S, Fagerli KM, Uhlig T, Kvien TK. TNF inhibitor treatment in rheumatoid arthritis (RA) patients with moderate versus high disease activity at baseline: A comparison of utility gains, response and remission rates. Arthritis and Rheumatism. 2012;64:S220.
157. Linde L, Sorensen J, Ostergaard M, Horslev-Petersen K, Hetland ML. Does clinical remission lead to normalization of EQ-5D in patients with rheumatoid arthritis and is selection of remission criteria important? Journal of Rheumatology. 2010;37(2):285-90.
158. Nakajima A, Inoue E, Shimizu Y, Kobayashi A, Shidara K, Sugimoto N, et al. Presence of comorbidity affects both treatment strategies and outcomes in disease activity, physical function, and quality of life in patients with rheumatoid arthritis. Clinical Rheumatology. 2015;34(3):441-9.
159. NICE. Golimumab for the Treatment of Rheumatoid Arthritis after Failure of Previous Disease-Modifying Antirheumatic Drugs [TA225]. Health Technology Assessment Database. 2010
160. Panasyuk E, Nasonov E, Nasonova VA. Tocilizumab improved of the quality of the life patients with severe/moderate RA (LORNET Study). Annals of the Rheumatic Diseases. 2015;74:722-3.
161. Pavelka K, Szekanecz Z, Damjanov N, Majdan M, Nasonov E, Mazurov V, et al. Induction of response with etanercept-methotrexate therapy in patients with moderately active rheumatoid arthritis in Central and Eastern Europe in the PRESERVE study. Clinical Rheumatology. 2013;32(9):1275-81.
162. Pentek M. Acceptability of Less Than Perfect Health States in Rheumatoid Arthritis: The Patients' Perspective. European Journal of Health Economics. 2014;15:73-82.
163. Pentek M, Gulacsi L, Rojkovich B, Brodszky V, van Exel J, Brouwer WB. Subjective health expectations at biological therapy initiation: a survey of rheumatoid arthritis patients and rheumatologists. The European journal of health economics : HEPAC : health economics in prevention and care. 2014;15:S83-S92.

164. Porter D, van Melckebeke J, Dale J, Messow CM, McConnachie A, Walker A, et al. Tumour necrosis factor inhibition versus rituximab for patients with rheumatoid arthritis who require biological treatment (ORBIT): an open-label, randomised controlled, non-inferiority, trial. Lancet (London, England). 2016;388(10041):239-47.
165. Salaffi F, Carotti M, Ciapetti A, Gasparini S, Grassi W. A comparison of utility measurement using EQ-5D and SF-6D preference-based generic instruments in patients with rheumatoid arthritis. Clinical and experimental rheumatology. 2011;29(4):661-71.
166. Tang CH, Hsu PN, Hsu JY, Fang CH. Health-state utilities in measuring health-related quality of life among patients with rheumatoid arthritis in Taiwan. Value in Health. 2013;16 (7):A568.
167. Teng YKO, Verburg RJ, Sont JK, Van Den Hout WB, Breedveld FC, Van Laar JM. Long-term followup of health status in patients with severe rheumatoid arthritis after high-dose chemotherapy followed by autologous hematopoietic stem cell transplantation. Arthritis and Rheumatism. 2005;52(8):2272-6.
168. Traki L, Rostom S, Tahiri L, Bahiri R, Harzy T, Abouqal R, et al. Responsiveness of the EuroQol EQ-5D and Hospital Anxiety and Depression Scale (HADS) in rheumatoid arthritis patients receiving tocilizumab. Clin Rheumatol. 2014;33(8):1055-60.
169. Personal Social Services Research Unit. Unit Costs of Health and Social Care 2015 2015. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2015/index.php.
170. Symmons DP, Silman AJ. The Norfolk Arthritis Register (NOAR) 2003 [updated Sep-Oct; cited 21 5 Suppl 31]. 2004/02/19:[S94-9]. Available from: http://www.clinexprheumatol.org/article.asp?a=2197.
171. Office for National Statistics. Overview of the UK population: February 2016. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/february2016#how-are-the-characteristics-of-the-uk-population-changing. Accessed: 10/01/2016.
172. Office for National Statistics. Overview of the UK Population. Available at: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/mid-2014/sty---overview-of-the-uk-population.html Accessed: 10/01/2017.
173. National Audit Office. Services for people with RA, July 2009.
174. Kantar Health. Kantar Health, 2016. Market Understanding & Market Sizing in RA.
175. National Institute for Health and Care Excellence. Tocilizumab for the treatment of rheumatoid arthritis: review of NICE technology appraisal guidance 198. Costing statement. [10/01/2017]. Available from: https://www.nice.org.uk/guidance/ta247/resources.
176. British National Formulary 2016. Available at: https://www.bnf.org/. Accessed 2016.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis [ID979]

Dear Eli Lilly,

The Evidence Review Group, the University of Sheffield's School of Health and Related Research (ScHARR), and the technical team at NICE have looked at the submission received on 6 February 2017 from Lilly. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **14 March 2017**. Your response and any supporting documents should be uploaded to NICE Docs/Appraisals.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Helen Powell, Technical Lead (Helen.Powell@nice.org.uk). Any procedural questions should be addressed to Stephanie Yates, Project Manager (Stephanie.Yates@nice.org.uk).

Yours sincerely

Frances Sutcliffe
Associate Director – Appraisals
Centre for Health Technology Evaluation

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Encl. checklist for confidential information

Section A: Clarification on effectiveness data

Systematic review and baricitinib RCTs

A1. How many reviewers conducted the data extraction and quality assessment?

A2. Please confirm that there are no further data regarding the effectiveness of baricitinib used in monotherapy other than that contained in RA-BEGIN. Please clarify why the company did not try to estimate the effectiveness of baricitinib used in monotherapy assuming that the relative efficacy observed in RA-BEGIN between baricitinib + MTX and baricitinib monotherapy was generalisable to patients who were cDMARD experienced and anti-TNF experienced.

A3. Please clarify the doses of background MTX used in the treatment arms in the included baricitinib RCTs.

Related to the network meta-analysis

A4. **Priority question**: Please clarify why the following studies were excluded from the network meta-analysis (NMA): ACQUIRE (NCT00559585); SURPRISE (NCT01120366); NCT01001832; CERTAIN (NCT00674362); SAMURAI (NCT00144508); SWEFOT (NCT00764725) and SWITCH (NCT01295151).

A5. **Priority question**: Please clarify why the treatment effect of MTX was assumed to be the same as cDMARD but SSZ and HCQ were assumed to be different from cDMARD in the NMA for the cDMARD-IR population. Please also clarify why the treatment effect of MTX was assumed to be different from cDMARD in the NMA for anti-TNF-IR population.

A6. **Priority question**: Please provide the ACR and EULAR response data at week 24 in the cDMARD-IR and anti-TNF-IR populations. These data need to be in a format ready to use in the NMA.

A7. **Priority question**: Please clarify if any procedure has been carried out to ensure that the information in the baseline model does not propagate to the relative treatment effect model when these two models were run simultaneously. If not, please provide the results generated by running the baseline model and relative treatment effect model separately.

A8. **Priority question**: Please clarify the process of converting ACR data to EULAR data. Was ACR data converted to EULAR data and then synthesised with reported EULAR data? If so, were any checks performed to assess the appropriateness of the

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

transformation, that is where a paper reported both EULAR and ACR data. If the data was not converted prior to synthesis, please clarify the method used. It is noted that in TA375 a different approach was used in that where ACR data were transformed this was done following a synthesis of purely ACR data. As there was no mixing of EULAR and ACR data there were two distinct sets of NMA results, and two corresponding results.

A9. **Priority question**: Please provide all the relative treatment effects NMA results on the probit scale, using forest plots if possible.

A10. **Priority question**: Please provide the total residual deviance for each of the NMA analyses, and comments on the goodness-of-fit.

A11. **Priority question**: Please provide the results of the random effects (RE) model for the TNF-IR population using a weakly informative or informative prior for the between-study standard deviation.

A12. **Priority question**: Please provide estimates and 95% credible intervals for between-study standard deviation when the RE model was used.

A13. Please provide the network diagrams for EULAR outcomes at week 24 in the cDMARD-IR and anti-TNF-IR populations.

A14. Section 4.10.7, page 189. Please provide the definition of δi, bk. It has been interpreted as the pooled effect of the experimental treatment versus the cDMARD arm of the included studies. Is this correct? Please clarify how the pooled effect was modelled when the control arm of the study was not cDMARD.

A15. Section 4.10.7, page 190. Please clarify how the cut-offs zij were modelled.

A16. Section 4.10.9, page 215. Please clarify whether the node-splitting results presented in Table 91 were for the ACR or EULAR NMA. Please provide all the node-splitting results when there were closed loops for the cDMARD-IR population for both ACR and EULAR outcome measures.

A17. Please provide the NMA code for the node-splitting analysis.

A18. Section 4.10.8, page 192 onwards. Where results are presented that use an RE model, were these results from a posterior distribution rather than a predictive distribution?

A19. Section 4.10.9, page 214. Please provide details on how Higgins' I^2 was calculated for each study. Please clarify what attempt was made to explain the heterogeneity between studies and what the predictive distribution is for the effect in a new study.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Section B: Clarification on cost-effectiveness data

B1. **Priority question**: Please clarify why estimated time of death is recalculated after every event. This produces different life year gains estimated from the model, which is not in line with the assumptions used in TA375.

B2. **Priority question**: Section 5.3, page 254. Please clarify why the predicted HAQ changes for patients remaining on cDMARDs are rounded to the nearest 0.125. This method will cause markedly different results if the predicted HAQ change between events was consistently 0.0620 compared with when it was 0.0630. Please amend the model if appropriate.

B3. **Priority question**: Section 5.8.3, page 301. Please clarify the reasons why in Table 133 the assumption of a non-flat HAQ progression for baricitinib reduces the ICER compared with combination cDMARDs.

B4. Section 5.2.2, page 247. Please clarify why it was assumed that a person could belong to only one latent class rather than the method used in TA375 where the probabilities of being in each latent class were used to form a weighted average for progression.

B5. Section 5.2.3, page 251. Please clarify why adalimumab + MTX was selected to follow rituximab + MTX in the tocilizumab + MTX sequence (Table 104)? Please clarify how the results change if a bDMARD with more favourable midpoint EULAR responses was selected.

B6. Please clarify why SAEs were not included in scenario analyses for the non-linear HAQ progression but were included for the linear HAQ progression.

B7. In relation to the HAQ improvement based on EULAR response, please clarify why the samples for the alpha and beta parameters are not correlated.

B8. Please clarify why average weights were used in the baseline model rather than distributions of these weights? This can lead to errors: see Hatswell AJ, Porter J, Lee D, Hertel N & Latimer NR (2016) The Cost of Costing Treatments Incorrectly: Errors in the Application of Drug Prices in Economic Evaluation Due to Failing to Account for the Distribution of Patient Weight. Value in Health, 19: 1055-8.

B9. Please clarify why subcutaneous tocilizumab was not modelled. Many comparators, which have been modelled, have confidential patient access schemes so this does not appear to be a valid reason.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B10. Please clarify what correlations have been incorporated into the model for patient characteristics, and for dependent parameter values (such as the shape and scale of a Weibull distribution).

B11. Section 5.6.1, page 276. Please clarify why the standard error for the administration and monitoring costs of ADA and TCZ in year 1 are different to the other interventions in Table 115. Please clarify why the monitoring costs of RTX, which is given less frequently, are the same as other bDMARDs.

B12. Section 5.6.2, table 116, page 281. Please clarify why the dosing of infliximab is in line with TA195 rather than TA375.

B13. Please clarify how many patients were run through the model to generate probabilistic results. Please comment on whether this number of patients were sufficient to generate stable results.

B14. Please clarify how many PSA configurations were run and whether this was sufficient to produce robust results.

B15. Please clarify why ABT IV was not considered within the submission. Please answer from the perspectives of i) ABT IV as a relevant comparator and ii) to allow further information to be gained regarding the comparison of subcutaneous ABT and cDMARDs through linking to ABT IV using the ACQUIRE study.

B16. Please clarify why results were not presented for patients receiving 2 mg baricitinib from initiation.

B17. Please clarify why results were not presented for patients who may step down from 4 mg to 2 mg of baricitinib.

B18. Please clarify how adherence, which was stated to be expected to be higher for an oral treatment, was incorporated into the model. It appears that the time to discontinuation when using a baricitinib-only curve is less than for other bDMARDs.

B19. Please clarify why it was assumed that the results from the NMA for patients who did not respond adequately to cDMARDs were generalisable to those patients who did not respond adequately to an anti-TNF (where data were not available). For those interventions where data were available for both patient populations it is clear that the EULAR responses are worse where a patient has already received an anti-TNF. Please present results for the population who did not respond adequately to an anti-TNF without ADA, CTZ, ETN, and IFX.

B20. Please clarify why it is assumed that the HAQ change associated with EULAR response happens at time zero. In TA375, this occurred at 6 months.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B21. Related to question A2. If it is assumed that the relative efficacy observed in RA-BEGIN between baricitinib + MTX and baricitinib monotherapy was generalisable to patients who were cDMARD experienced and anti-TNF experienced, please provide results for baricitinib monotherapy.

Section C: Textual clarifications and additional points

C1. Please clarify the number of baricitinib studies included in the submission. Figure 5 (page 75) suggests that six studies were identified however Table 12 (page 77) suggests that there are only five relevant studies.

C2. Section 4.7.1, page 107. Please clarify whether there is significant difference in ACR20 scores between baricitinib and adalimumab at Week 52. Figure 13 and the accompanying explanatory text appear to be contradictory.

C3. Section 5.2.8, page 249 (and elsewhere). It is stated that a HAQ decrement of 0.06 is applied when there is a step down from 4 mg to 2 mg. Please confirm that this is actually an increase of 0.06 (i.e. a worsening in HAQ score).

C4. Section 5.3, page 256. It is stated that following treatment discontinuation the HAQ score would rebound immediately to the level prior to initiation of the terminated therapy. Please clarify that this is applicable for bDMARDs (and baricitinib in the base case) but not MTX or the scenario analysis where baricitinib is associated with a HAQ increase across time.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis [ID979]

Dear Eli Lilly,

The Evidence Review Group, the University of Sheffield's School of Health and Related Research (ScHARR), and the technical team at NICE have looked at the submission received on 6 February 2017 from Lilly. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **14 March 2017**. Your response and any supporting documents should be uploaded to NICE Docs/Appraisals.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Helen Powell, Technical Lead (Helen.Powell@nice.org.uk). Any procedural questions should be addressed to Stephanie Yates, Project Manager (Stephanie.Yates@nice.org.uk).

Yours sincerely

Frances Sutcliffe
Associate Director – Appraisals
Centre for Health Technology Evaluation

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Encl. checklist for confidential information

Section A: Clarification on effectiveness data

Systematic review and baricitinib RCTs

A1. How many reviewers conducted the data extraction and quality assessment?

Data was extracted by one reviewer and was verified against the original source by a second researcher. Quality assessment was conducted by one reviewer.

A2. Please confirm that there are no further data regarding the effectiveness of baricitinib used in monotherapy other than that contained in RA-BEGIN. Please clarify why the company did not try to estimate the effectiveness of baricitinib used in monotherapy assuming that the relative efficacy observed in RA-BEGIN between baricitinib + MTX and baricitinib monotherapy was generalisable to patients who were cDMARD experienced and anti-TNF experienced.

RA-BUILD enrolled 684 patients who had had an insufficient response, were intolerant to or contraindicated to cDMARDs including MTX, and who had not been previously treated with biologic DMARDs. In the patient cohort, there were 48 patients who were enrolled into the study without background cDMARDs and were therefore treated with monotherapy baricitinib. The CHMP assessment for baricitinib took note of the data from RA-BEGIN and this led to the following statement being included in section 5.1 of the SPC: 'The RA-BEGIN Study in MTX-naïve patients is supportive for the target population of patients with an inadequate response to, or intolerance to, other DMARDs…'

Whilst from a regulatory perspective, monotherapy is in the indication statement, the decision not to attempt to use RA-BEGIN data in any relative effectiveness estimates was made for the following reasons:

- In clinical practice, EULAR guidelines recommend methotrexate as first-line therapy, with other cDMARDs added after MTX failure. Therefore, the RA-BEGIN monotherapy population is not congruent with these guidelines as these patients were naïve to methotrexate.
- Unknown heterogeneity impact of attempting to **quantitatively** combine data from a naïve population to those with prior cDMARD and/or TNFi experience. It is highly likely that the exchangeability assumptions underpinning standard indirect comparison approaches would be compromised

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

- TA375 recommends adalimumab, etanercept, certolizumab or tocilizumab as monotherapy where methotrexate is contra-indicated or not tolerated. The Committee discussion in the guidance states that the basis of this recommendation is that the minority of people with severe rheumatoid arthritis who cannot tolerate methotrexate should not be treated differently from other people with severe disease. Therefore, the conclusion of the Committee was that biological DMARDs could be recommended as monotherapy within their marketing authorisation. It would be logical to assume that similar rationale should apply to baricitinib should it gain a recommendation for severe cDMARD-IR and TNF-IR patients

A3. Please clarify the doses of background MTX used in the treatment arms in the included baricitinib RCTs.

JADV – RA-BEAM

Use of concomitant oral MTX for at least 12 weeks with treatment at a stable dose of 7.5-25mg/week (or the equivalent injectable dose) for at least 8 weeks prior to entry into the study was required for study participation. Patients remained on MTX throughout Part A (double-blind, placebo- and active-controlled period from Week 0 through Week 24) AND Part B (double-blind, active-controlled period from the end of Week 24 through Week 52.) of the study; however, the MTX dose could be adjusted for safety reasons.

At baseline; the mean weekly dose of MTX across all patients (N=1,305) was 14.8mg/week.

JADW – RA-BEACON

Treatment with concomitant DMARDs during the study was required as outlined below:

- Patients who took concomitant oral MTX must have been taking MTX for at least 12 weeks and been on a stable dose for at least 8 weeks prior to entry into the study. Patients remained on the same dose throughout the study unless the MTX dose was adjusted for safety reasons

At baseline; the mean weekly dose of MTX across all patients (N=527) was 16.3mg/week.

JADX – RA-BUILD

Treatment with concomitant DMARDs during the study was permitted as outlined below:

- Use of concomitant DMARD for at least 12 weeks with a continuous, non-changing dose for at least 8 weeks prior to study entry. For patients receiving MTX, a stable, unchanging dose of 7.5 to 25 mg/week (or the equivalent injectable dose) was administered for at least 8 weeks prior to entry into the study. Patients remained on the

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

same dose of the permitted DMARD or MTX throughout the study unless the dose was adjusted for safety reasons.

At baseline; the mean weekly dose of MTX across all patients (N=684) was 16.2mg/week

Related to the network meta-analysis

A4. **Priority question**: Please clarify why the following studies were excluded from the network meta-analysis (NMA): ACQUIRE (NCT00559585); SURPRISE (NCT01120366); NCT01001832; CERTAIN (NCT00674362); SAMURAI (NCT00144508); SWEFOT (NCT00764725) and SWITCH (NCT01295151).

- ACQUIRE: This study compared S/C vs IV abatacept. The search strategy specified that studies were to include two different comparators of interest to be included in the shortlist
- SURPRISE: This was an open-label study that compared tocilizumab combination therapy to tocilizumab monotherapy and was therefore excluded
- NCT0100132: This study mirrored the design of ACQUIRE in a Japanese only population and was therefore excluded on the same basis as ACQUIRE
- CERTAIN: This study was in patients with low to moderate disease activity rather than moderate to severe patients which was the population of interest for baricitinib
- SAMURAI: This was an open-label study comparing tocilizumab monotherapy to DMARDs. Only 52 week data was available so this study could not be included for the 12/24 week time-points in the NMA
- SWEFOT: This was an open-label study in patient with early rheumatoid arthritis (less than a year since diagnosis) and was therefore excluded. Additionally, the infliximab arm, allowed an increase in dose frequency (to every 6 weeks) or a switch to etanercept and it does not appear that reported results take this into account
- SWITCH: The only publication related to this study that was found in the search was that of the study protocol- no results appear to be available

A5. **Priority question**: Please clarify why the treatment effect of MTX was assumed to be the same as cDMARD but SSZ and HCQ were assumed to be different from cDMARD in the NMA for the cDMARD-IR population. Please also clarify why the treatment effect of MTX was assumed to be different from cDMARD in the NMA for anti-TNF-IR population.

For the studies included in the cDMARD-IR network, the majority of studies compared to methotrexate. Where the comparators were cDMARDs, the main analysis included specific

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

named treatments where this information was available and included unspecified cDMARDs in the methotrexate node (with the assumption that the majority of patients would be on methotrexate being felt reasonable given that it is the most widely used cDMARD). In order to investigate the impact of this assumption, a sensitivity analysis was conducted that combined all treatments into a single cDMARD node and this was found to have no notable difference in results.
For the TNF-IR population, the network of evidence was much sparser and in order to ensure that there was a connected network, the central node in the essentially 'star-shaped' network was required to be cDMARDs.

A6. **Priority question**: Please provide the ACR and EULAR response data at week 24 in the cDMARD-IR and anti-TNF-IR populations. These data need to be in a format ready to use in the NMA.

The data for ACR for the cDMARD-IR and TNF-IR populations is in appendix 14 of the company submission. EULAR data is provided in an addendum to this clarification response (please also refer to the response for question A8).

A7. **Priority question**: Please clarify if any procedure has been carried out to ensure that the information in the baseline model does not propagate to the relative treatment effect model when these two models were run simultaneously. If not, please provide the results generated by running the baseline model and relative treatment effect model separately.

Simultaneous models were used in the main analysis for the following reasons:

- Data for the baseline and treatment effects came from the same sources
- Some networks had zero cells and using simultaneous models increased stability
- In the case of the TNF-IR network, the evidence base was sparse

Pre-planned sensitivity analyses with separate models were considered in the broader NMA plan- select results for the cDMARD-IR population are presented in tables 1 and 2 with the EULAR rates for separate models for fixed-effects and random effects.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 1: EULAR Response rates with separate fixed effect models

Intervention	EULAR Response	Median	2.5 percentile	97.5 percentile
Cdmard	Moderate+ Good	XXX	XXX	XXX
Cdmard	Good	XXX	XXX	XXX
Bari 4mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Bari 4mg + Cdmard	Good	XXX	XXX	XXX
Bari 2mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Bari 2mg + Cdmard	Good	XXX	XXX	XXX
Tcz 8mg	Moderate+ Good	XXX	XXX	XXX
Tcz 8mg	Good	XXX	XXX	XXX
Tcz 8mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Tcz 8mg + Cdmard	Good	XXX	XXX	XXX
Ada 40mg	Moderate+ Good	XXX	XXX	XXX
Ada 40mg	Good	XXX	XXX	XXX
Aba 10mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Aba 10mg + Cdmard	Good	XXX	XXX	XXX
Aba Subcut + Cdmard	Moderate+ Good	XXX	XXX	XXX
Aba Subcut + Cdmard	Good	XXX	XXX	XXX
Ada 40mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Ada 40mg + Cdmard	Good	XXX	XXX	XXX
Ifx 3mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Ifx 3mg + Cdmard	Good	XXX	XXX	XXX

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Intervention	EULAR Response	Median	2.5 percentile	97.5 percentile
Tcz Subcut + Cdmard	Moderate+ Good	XXX	XXX	XXX
Tcz Subcut + Cdmard	Good	XXX	XXX	XXX
Placebo	Moderate+ Good	XXX	XXX	XXX
Placebo	Good	XXX	XXX	XXX
ETN	Moderate+ Good	XXX	XXX	XXX
ETN	Good	XXX	XXX	XXX
ETN + SSZ	Moderate+ Good	XXX	XXX	XXX
ETN + SSZ	Good	XXX	XXX	XXX
SSZ	Moderate+ Good	XXX	XXX	XXX
SSZ	Good	XXX	XXX	XXX
Etn + Cdmard	Moderate+ Good	XXX	XXX	XXX
Etn + Cdmard	Good	XXX	XXX	XXX
Rtx 1000mg	Moderate+ Good	XXX	XXX	XXX
Rtx 1000mg	Good	XXX	XXX	XXX
Rtx 1000mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Rtx 1000mg + Cdmard	Good	XXX	XXX	XXX
Gol 50mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Gol 50mg + Cdmard	Good	XXX	XXX	XXX
Czp + Cdmard	Moderate+ Good	XXX	XXX	XXX
Czp + Cdmard	Good	XXX	XXX	XXX
Tofa 10mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Tofa 10mg + Cdmard	Good	XXX	XXX	XXX
Tofa 5mg + Cdmard	Moderate+ Good	XXX	XXX	XXX

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Intervention	EULAR Response	Median	2.5 percentile	97.5 percentile
Tofa 5mg + Cdmard	Good	XXX	XXX	XXX
Ssz + Hcq + Cdmard	Moderate+ Good	XXX	XXX	XXX
Ssz + Hcq + Cdmard	Good	XXX	XXX	XXX
Rtx 2000mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Rtx 2000mg + Cdmard	Good	XXX	XXX	XXX
TCZ SUBCUT	Moderate+ Good	XXX	XXX	XXX
TCZ SUBCUT	Good	XXX	XXX	XXX

Table 2: EULAR Response rates with separate random effects models

Intervention	EULAR Response	Median	2.5 percentile	97.5 percentile
Cdmard	Moderate+ Good	XXX	XXX	XXX
Cdmard	Good	XXX	XXX	XXX
Bari 4mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Bari 4mg + Cdmard	Good	XXX	XXX	XXX
Bari 2mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Bari 2mg + Cdmard	Good	XXX	XXX	XXX
Tcz 8mg	Moderate+ Good	XXX	XXX	XXX
Tcz 8mg	Good	XXX	XXX	XXX
Tcz 8mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Tcz 8mg + Cdmard	Good	XXX	XXX	XXX

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Intervention	EULAR Response	Median	2.5 percentile	97.5 percentile
Ada 40mg	Moderate+ Good	[redacted]	[redacted]	[redacted]
Ada 40mg	Good	[redacted]	[redacted]	[redacted]
Aba 10mg + Cdmard	Moderate+ Good	[redacted]	[redacted]	[redacted]
Aba 10mg + Cdmard	Good	[redacted]	[redacted]	[redacted]
Aba Subcut + Cdmard	Moderate+ Good	[redacted]	[redacted]	[redacted]
Aba Subcut + Cdmard	Good	[redacted]	[redacted]	[redacted]
Ada 40mg + Cdmard	Moderate+ Good	[redacted]	[redacted]	[redacted]
Ada 40mg + Cdmard	Good	[redacted]	[redacted]	[redacted]
Ifx 3mg + Cdmard	Moderate+ Good	[redacted]	[redacted]	[redacted]
Ifx 3mg + Cdmard	Good	[redacted]	[redacted]	[redacted]
Tcz Subcut + Cdmard	Moderate+ Good	[redacted]	[redacted]	[redacted]
Tcz Subcut + Cdmard	Good	[redacted]	[redacted]	[redacted]
Placebo	Moderate+ Good	[redacted]	[redacted]	[redacted]
Placebo	Good	[redacted]	[redacted]	[redacted]
ETN	Moderate+ Good	[redacted]	[redacted]	[redacted]
ETN	Good	[redacted]	[redacted]	[redacted]
ETN + SSZ	Moderate+ Good	[redacted]	[redacted]	[redacted]
ETN + SSZ	Good	[redacted]	[redacted]	[redacted]
SSZ	Moderate+ Good	[redacted]	[redacted]	[redacted]
SSZ	Good	[redacted]	[redacted]	[redacted]
Etn + Cdmard	Moderate+ Good	[redacted]	[redacted]	[redacted]
Etn + Cdmard	Good	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Intervention	EULAR Response	Median	2.5 percentile	97.5 percentile
Rtx 1000mg	Moderate+ Good	XXX	XXX	XXX
Rtx 1000mg	Good	XXX	XXX	XXX
Rtx 1000mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Rtx 1000mg + Cdmard	Good	XXX	XXX	XXX
Gol 50mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Gol 50mg + Cdmard	Good	XXX	XXX	XXX
Czp + Cdmard	Moderate+ Good	XXX	XXX	XXX
Czp + Cdmard	Good	XXX	XXX	XXX
Tofa 10mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Tofa 10mg + Cdmard	Good	XXX	XXX	XXX
Tofa 5mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Tofa 5mg + Cdmard	Good	XXX	XXX	XXX
Ssz + Hcq + Cdmard	Moderate+ Good	XXX	XXX	XXX
Ssz + Hcq + Cdmard	Good	XXX	XXX	XXX
Rtx 2000mg + Cdmard	Moderate+ Good	XXX	XXX	XXX
Rtx 2000mg + Cdmard	Good	XXX	XXX	XXX
TCZ SUBCUT	Moderate+ Good	XXX	XXX	XXX
TCZ SUBCUT	Good	XXX	XXX	XXX

A8. **Priority question**: Please clarify the process of converting ACR data to EULAR data. Was ACR data converted to EULAR data and then synthesised with reported EULAR

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

data? If so, were any checks performed to assess the appropriateness of the transformation, that is where a paper reported both EULAR and ACR data. If the data was not converted prior to synthesis, please clarify the method used. It is noted that in TA375 a different approach was used in that where ACR data were transformed this was done following a synthesis of purely ACR data. As there was no mixing of EULAR and ACR data there were two distinct sets of NMA results, and two corresponding results.

For studies that reported **ACR data only**, these data were transformed into EULAR data using the mapping algorithm derived from the Veteran Affairs Rheumatoid Arthritis (VARA) database, following the information provided in TA375[1]. The two-way table developed for TA375 is replicated below (Table X):

Table 3: EULAR and ACR Responses from VARA- from TA375

	Less than ACR20[a]	ACR20[a]	ACR50	ACR70	Total[a]
EULAR ESR, all patients					
EULAR None	755	4	2	0	759
EULAR Moderate	136	27	2	2	163
EULAR Good	57	26	10	2	83
EULAR ESR, severe active					
EULAR None	72	2	0	0	74
EULAR Moderate	33	19	0	0	52
EULAR Good	3	9	5	1	12

[a] The total number of patients are the sum of the first two columns (Less than ACR20, ACR20), as ACR50 and ACR70 are both part of ACR20.

As all studies in the NMA were in the moderate to severe population (including baricitinib studies), the data for all patients was used. The studies where ACR data was converted to EULAR were then included in the network with studies that did report EULAR data for the EULAR evidence synthesis network. To confirm, all included studies reported ACR data, therefore the ACR network used trial reported data in the evidence synthesis. As reported in the submission, there are two sets of results- ACR data only, and EULAR response where the evidence synthesis network includes ACR only reporting studies converted to EULAR response in the method described above. For clarity, the conversion was applied to 29 out of 45 studies in the cDMARD-IR network and 4 out of 6 studies in the TNF-IR network.

With respect to the appropriateness of the transformation used, the primary deciding factor was consistency with TA375- no alternative approaches were discussed in TA375 as far as is apparent and similarly there was limited discussion of the appropriateness of the

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

transformation. Given that patient-level data was available for both EULAR and ACR response was available from the baricitinib studies, some exploratory analyses were conducted. A mixture of approaches were undertaken (column probabilities only, row probabilities only, a combination of column and row probabilities, and on total probabilities). Column probabilities were found to be the better performing and all algorithms were tested using data from other studies in the NMA reporting both ACR and EULAR data. 'Misclassifications' in these validation tests ranged from XXX% with the baricitinib study derived transformations compared to XX% with the VARA transformation from TA375. Therefore, as noted, this supported the use of the VARA dataset in the submission for consistency reasons

A9. **Priority question**: Please provide all the relative treatment effects NMA results on the probit scale, using forest plots if possible.

Figs 1 and 2 show the relative treatment effects (to cDMARDs) on the probit scale for the cDMARD-IR and TNF-IR populations, in the form of forest plots.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

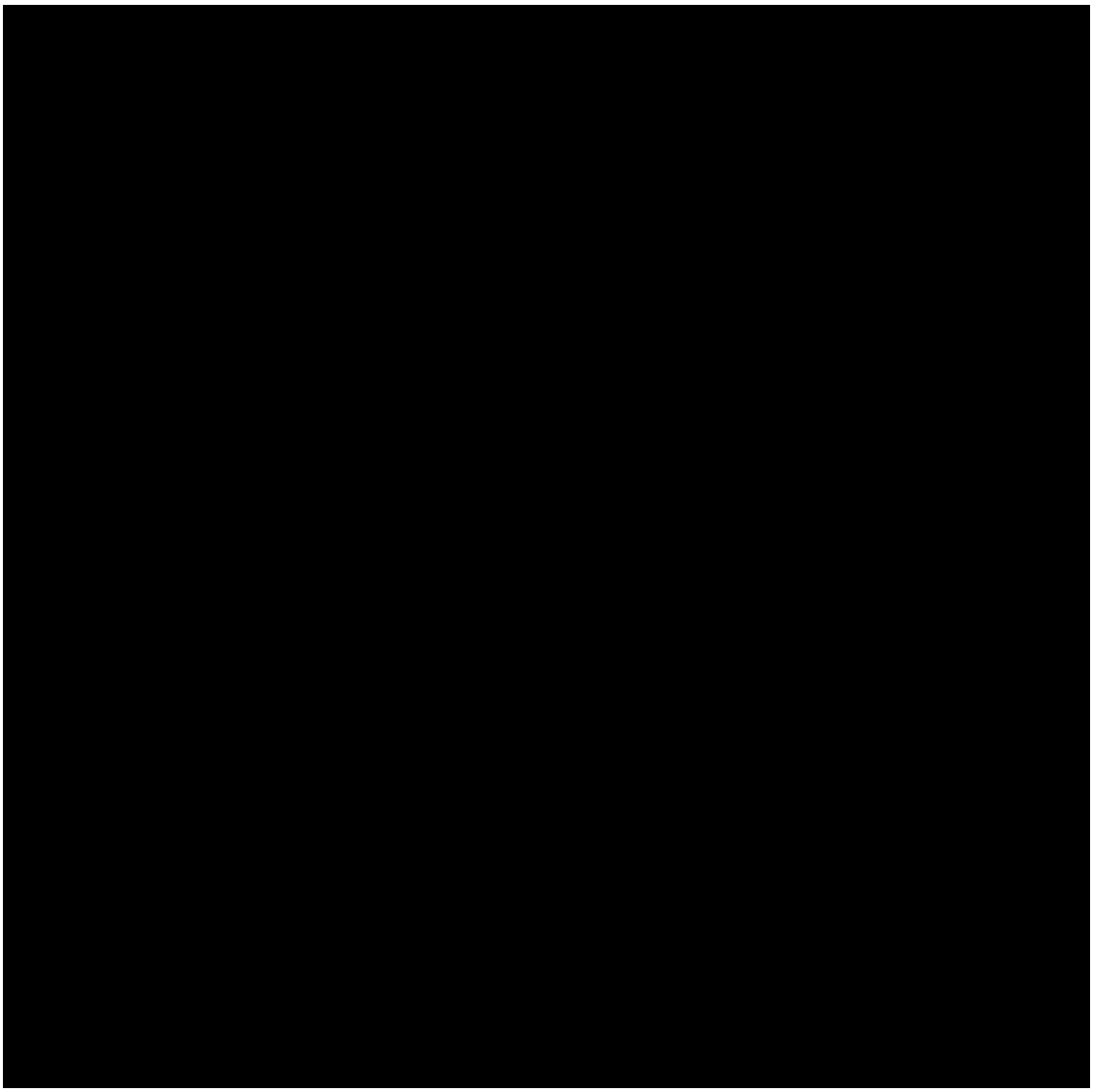


Fig 1- Treatment differences on the probit scale forest plot- cDMARD-IR (RE model)

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

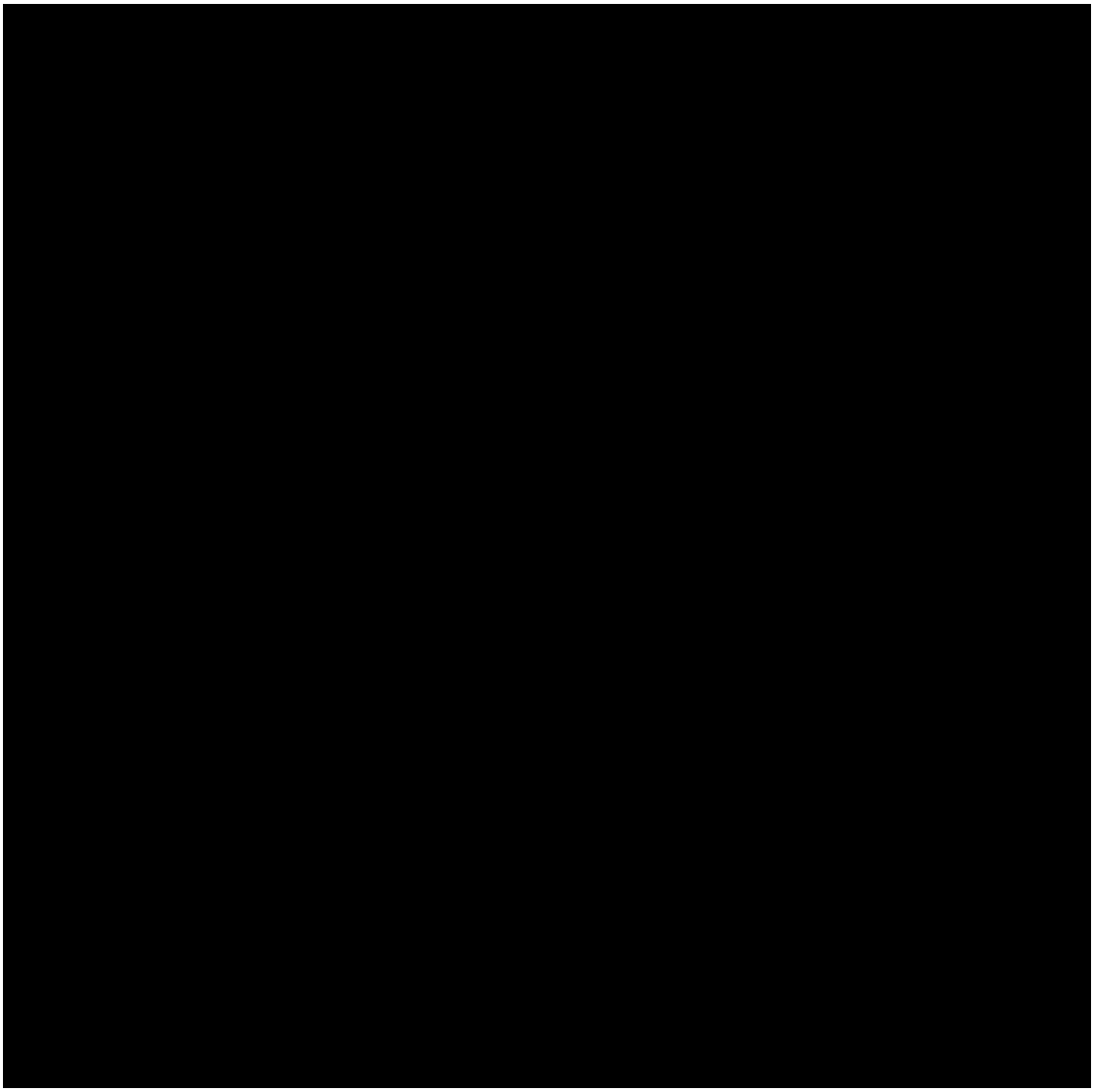


Fig 2- Treatment differences on probit scale forest plot, TNF-IR (FE Model)

A10. **Priority question**: Please provide the total residual deviance for each of the NMA analyses, and comments on the goodness-of-fit.

Table X and Y provide the DIC and residual deviance for the ACR and EULAR models for each population (cDMARD-IR and TNF-IR. As would be expected given the approach taken to transform ACR to EULAR data, the models only using ACR data as reported in the included studies have a better fit.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 4: DIC and Residual Deviance cDMARD-IR

Outcome	Model Type	DIC	Residual Deviance	median SD (95% CrI)
probit_24wks_ACR	RE	508.008	427.509	0.17 (0.10, 0.27)
probit_24wks_EULAR	RE	798.299	717.651	0.20 (0.13, 0.30)

Table 5: DIC and Residual Deviance TNF-IR

Outcome	Model Type	DIC	Residual Deviance	median SD (95% CrI)
probit_24wks_ACR	FE	70.694	56.639	NA
probit_24wks_EULAR	FE	105.501	92.333	NA

A11. **Priority question**: Please provide the results of the random effects (RE) model for the TNF-IR population using a weakly informative or informative prior for the between-study standard deviation.

As highlighted in the response to question AX, the small number of studies in the TNF-IR population led to convergence issues, even with informative priors. For information, EULAR ORs are provided in table X and Y for model with a uniform (0,2) prior and a log-normal prior (-2.34,1.62). Other uniform priors were considered by results were even poorer with respect to convergence.

Table 6- EULAR ORs TNF-IR RE Model, Uniform Prior (0.2)

Intervention	Cdmard	Bari 4mg + Cdmard	Bari 2mg + Cdmard	Aba 10mg + Cdmard	Tcz 162mg + Cdmard	Gol 50mg + Cdmard	Tcz 8mg + Mtx	Rtx 1000mg + Mtx
Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bari 4mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bari 2mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Aba 10mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tcz 162mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gol 50mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tcz 8mg + Mtx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Rtx 1000mg + Mtx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 7: EULAR ORs, TNF-IR RE model, log-normal prior (-2.34, 1.62)

Intervention	Cdmard	Bari 4mg + Cdmard	Bari 2mg + Cdmard	Aba 10mg + Cdmard	Tcz 162mg + Cdmard	Gol 50mg + Cdmard	Tcz 8mg + Mtx	Rtx 1000mg + Mtx
Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bari 4mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bari 2mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Aba 10mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tcz 162mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gol 50mg + Cdmard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tcz 8mg + Mtx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Rtx 1000mg + Mtx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

A12. **Priority question**: Please provide estimates and 95% credible intervals for between-study standard deviation when the RE model was used.

This is given in column 4 of table 4 provided in response to question A10.

A13. Please provide the network diagrams for EULAR outcomes at week 24 in the cDMARD-IR and anti-TNF-IR populations.

As clarified in the response to question A8, the approach taken to imputing data for EULAR response where only ACR response was reported means that in effect, the trials included in the evidence network are effectively the same for ACR and EULAR. Table 69 in the company submission states which studies reported ACR and EULAR response.

A14. Section 4.10.7, page 189. Please provide the definition of δi, bk. It has been interpreted as the pooled effect of the experimental treatment versus the cDMARD arm of the included studies. Is this correct? Please clarify how the pooled effect was modelled when the control arm of the study was not cDMARD.

The interpretation stated above is correct. The approach followed was that outlined in TSD2 (A general linear modelling framework for pair-wise and network meta-analysis of randomised controlled trials) from the NICE DSU[2]- for the analysis conducted using a probit link function, the TSD states '*In this setup, the pooled effect of taking the experimental treatment instead of the control is to change the probit score (or Z score) of the control arm, by δi,bk standard deviations'* Where the control arm was non-cDMARD, its' effect was subtracted out so no further adjustment was necessary

A15. Section 4.10.7, page 190. Please clarify how the cut-offs zij were modelled

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The zij cut-offs were modelled as increasing fixed effects that were uniformly distributed over (0,5). This approach was again concurrent with the NICE DSU TSD2. The NMA code, provided in appendix X of the company submission also shows how the cut-offs were modelled (for example, pg225, second to last code line group)

A16. Section 4.10.9, page 215. Please clarify whether the node-splitting results presented in Table 91 were for the ACR or EULAR NMA. Please provide all the node-splitting results when there were closed loops for the cDMARD-IR population for both ACR and EULAR outcome measures.

Table 91 shows the node-splitting results for the ACR NMA- that is using ACR data as reported in the included studies as previously clarified. The EULAR results in table 91 are those from the NMA that transformed ACR data to EULAR responses before the evidence synthesis took place.

Full node splitting results are provided in the addendum to this response.

A17. Please provide the NMA code for the node-splitting analysis.

The relevant code is provided below:

```
#Node Splitting Model in JAGS for a simultanous fixed treatment effects
model.
#pair is a vector describing the node split (ie, pair=c(1,8)).
#ns_data.txt contains an example dataset.
model{
  for(i in 1:ns){    # LOOP THROUGH STUDIES
    phi[i] <- mu[i]
    mu[i] ~ dnorm(m, tau_phi_prec)  #Random effects baseline.
    for(k in 1:na[i]){                # LOOP THROUGH ARMS
      r[i,k] ~ dbin(ph[i,k],n[i,k])       # Binomial likelihood
      logit(ph[i,k]) <- phi[i] + delta[i,k]

      #index[i,k]=1 if the study i has the split node (pair[1],pair[2])
      # and arm k is one of pair[1] or pair[2]).
      index[i,k] <- split[i]*(equals(trt[i,k], pair[1]) + equals(trt[i,k],
pair[2]))

      #Deviance contribution , imputed ==1 for imputed arm, 0 for observed
arm
      rhat[i,k] <- ph[i,k]*n[i,k] # expected value of the numerators
      dev[i,k] <- 2*(r[i,k]*(log(r[i,k])-log(rhat[i,k])) + (n[i,k]-
r[i,k])*(
      log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))*(1 - imputed[i,k])
```

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

```
    }
    # summed residual deviance contribution for this trial
    devstudy[i] <- sum(dev[i,1:na[i]])

    for(k in 1:na[i]){ # LOOP THROUGH ARMS, with direct evidence parameter
when index=1
      delta[i,k] <- (d[trt[i,k]] - d[trt[i,1]])*(1-index[i,k]) +
direct*index[i,k]
    }
  }

  Dbar <- sum(devstudy[]) #Total Residual Deviance
  d[1] <-0 # treatment effect is zero for reference treatment

  direct ~ dnorm(0,.0001) # vague prior for direct comparison parameter
  diffDI <- direct - d[pair[2]] + d[pair[1]] #difference between direct and
indirect evidence

  #Fixed treatment effects.
  for(k in 2:M){
    d[k] ~ dnorm(0,.0001)
  } # vague priors for treatment effects
  m ~ dnorm(0,.0001)          # vague prior for mean baseline response

  #Predictive baseline
  mu.new ~ dnorm(m, tau_phi_prec)

  #between-trial precision for baseline = (1/between-trial variance)
  tau_phi_prec <- pow(sd.m,-2)

  #vague prior for baseline response
  sd.m ~ dunif( 0 , 2 )

  #All pairwise differece
  for(c in 1:(M-1)) {
    for(k in (c+1):M){
      dif[c,k] <- d[k] - d[c]
      pdif[c,k]<- step(dif[c,k])
    }
  }

  #Pairwise OR and LORs for all possible pair-wise comparisons.
  for(c in 1:(M-1)) {
    for(k in (c+1):M){
      OR[c,k] <- max(min(exp(d[k]-d[c]), .999/.001),.001/.999)
      LOR[c,k]<- d[k]-d[c]
    }
  }

  #Provide estimates of treatment effects T[k] on the probability scale
  for(k in 1:M) {
    logit(T[k])<-mu.new+d[k]
```

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A18. Section 4.10.8, page 192 onwards. Where results are presented that use an RE model, were these results from a posterior distribution rather than a predictive distribution?

All results are presented for a posterior distribution.

A19. Section 4.10.9, page 214. Please provide details on how Higgins' I^2 was calculated for each study. Please clarify what attempt was made to explain the heterogeneity between studies and what the predictive distribution is for the effect in a new study.

The I^2 values on pg. 214 were calculated from a meta-analysis of studies which contained the same treatment arms, for example the TEMPO, SATORI and ARMADA studies. The pair-wise meta-analyses that indicated potentially heterogeneity were used to inform which studies to remove from the network in the sensitivity analyses presented in section 4.10.8.2 of the company submission.

A predictive distribution for the effect in a new study was not considered as only the results for posterior distributions were presented.

In the broader NMA, several pre-specified meta-regressions were undertaken but found to only have a small explanatory effect. The variables considered in the ACR/EULAR cDMARD-IR networks were: year of publication, low vs. normal MTX dose, early vs. established disease and duration of disease.

Section B: Clarification on cost-effectiveness data

B1. **Priority question**: Please clarify why estimated time of death is recalculated after every event. This produces different life year gains estimated from the model, which is not in line with the assumptions used in TA375.

Consistent with the approach adopted in TA375, it was assumed that there was no treatment related mortality effect for any of the interventions or simulated events considered. It could therefore be argued that the model may generate less uncertainty in results if time to death was simply calculated at baseline for each patient (and was therefore independent of any event and thus identical for all arms). However, the model was built to incorporate the possibility that events or treatment may impact mortality should such data become available, and therefore time to death was recalculated following each event. Further, given that a sufficient number of patients are run through the model in order to achieve convergence, this

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

is deemed not to have an impact on results as the simulated difference between arms should be minimal and the result of random variation, as the below convergence graph shows.

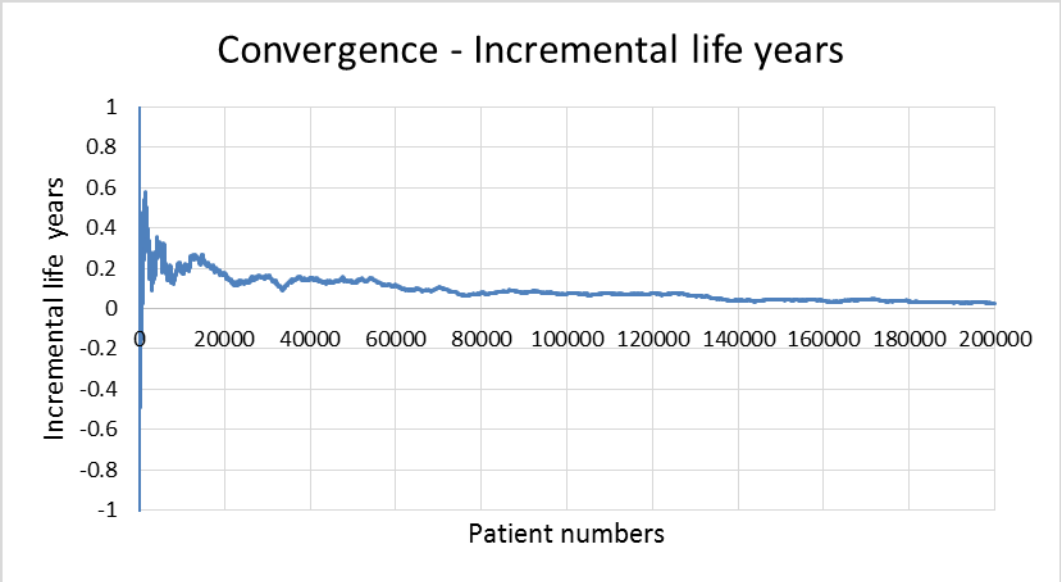


Fig 3- Incremental life years convergence graph

B2. **Priority question**: Section 5.3, page 254. Please clarify why the predicted HAQ changes for patients remaining on cDMARDs are rounded to the nearest 0.125. This method will cause markedly different results if the predicted HAQ change between events was consistently 0.0620 compared with when it was 0.0630. Please amend the model if appropriate.

As noted in TA375, HAQ-DI is a discrete scale with step values of 0.125, resulting in 25 possible values on the HAQ scale (Stevenson et al). When estimating change in HAQ for cDMARDs or palliative care the increase in HAQ was rounded to the nearest 0.125 value to be consistent with the step values of the HAQ discrete scale. As the discrete event simulation (DES) models individual patients, and since HAQ is a discrete scale, patients can only be attributed to a discrete point in the scale, which is why the rounding is applied.

B3. **Priority question**: Section 5.8.3, page 301. Please clarify the reasons why in Table 133 the assumption of a non-flat HAQ progression for baricitinib reduces the ICER compared with combination cDMARDs.

Due to the approach taken in programming the model, it was not readily feasible to apply different on-treatment HAQ trajectories between baricitinib and biologic interventions. The

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

alternative approach taken to explore non-flat HAQ progression for baricitinib was to apply linear progression on cDMARDs and then apply an assumption of slower but still linear HAQ progression for baricitinib. Therefore, the appropriate comparison to make in this case is with scenario 1 in table 133, which is the ICER when there is linear progression with cDMARDs and no progression for baricitinib. The ICER for scenario 3 (£30,280 per QALY) is higher than scenario 1 (£20,965 per QALY) which would be as expected given the alternative assumption of HAQ progression for baricitinib. It should be noted again that this is likely an extreme scenario with the progression rate assumed to be around 50% of that with cDMARDs.

B4. Section 5.2.2, page 247. Please clarify why it was assumed that a person could belong to only one latent class rather than the method used in TA375 where the probabilities of being in each latent class were used to form a weighted average for progression.

It was assumed that consistent with the nature of a DES, the model should track each patient individually and thus assigned each patient to a specific latent class rather than modelling an average patient following a weighted average for progression. By running the model with a sufficiently large number of patients, overall results are anticipated to reflect the distribution of probabilities of class membership and thus be roughly equivalent to a weighted average approach for progression.

B5. Section 5.2.3, page 251. Please clarify why adalimumab + MTX was selected to follow rituximab + MTX in the tocilizumab + MTX sequence (Table 104)? Please clarify how the results change if a bDMARD with more favourable midpoint EULAR responses was selected.

As is apparent, a pragmatic approach was taken in determining the treatment sequences. In the tocilziumab sequence referred to in the question, it was assumed to be plausible that a TNF-inhibitor would be an option if a TNF-inhibitor had not been used in the treatment pathway until the failure of the second-line treatment. As adalimumab has the largest market share amongst TNF-inhibitors, it was selected at this point in this particular treatment sequence. Table X replicates the base-case cDMARD-IR analysis, with certolizumab substituted for adalimumab in the tocilizumab sequence. This sequence results in greater QALY benefit vs. the baricitinib sequence compared to the base base, but the difference in costs results in a pair-wise ICER of around £84000 per QALY.

Table 8- Incremental cost-effectiveness analysis using CTZ in the TCZ sequence

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Referent	Referent
ETNMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	£18,400	£18,400
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
IFXMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+CTZMTX+MTX+Pall	[redacted]	[redacted]	14.72	[redacted]	[redacted]	£84,106	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated

B6. Please clarify why SAEs were not included in scenario analyses for the non-linear HAQ progression but were included for the linear HAQ progression.

Scenario analyses provided amended a single assumption in each scenario- i.e. SAE were included whilst all other assumptions remained as per the base case and similarly, linear HAQ progression was included as an alternative assumption in the relevant scenario analysis. This approach was taken in order to investigate the impact of each alternative assumption upon the cost-effectiveness estimates. The results of each scenario are presented again below for the cDMARD-IR population as well as those for a combined scenario that includes both linear HAQ progression and SAE

Table 9-Incremental Cost-effectiveness including cost of SAEs

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
ETNMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.72	[redacted]	[redacted]	Referent	Referent
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	£7,691	£7,691
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.71	[redacted]	[redacted]	£8,810	£9,438
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.72	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.72	[redacted]	[redacted]	Dominated	Dominated
IFXMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.72	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.71	[redacted]	[redacted]	£553,034	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated

Table 10 Incremental Cost-effectiveness with linear progression with cDMARDs

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.72	[redacted]	[redacted]	Referent	Referent
ETNMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.71	[redacted]	[redacted]	Dominated	Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.72	[redacted]	[redacted]	£13,739	£13,739
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.72	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.72	[redacted]	[redacted]	Dominated	Dominated
IFXMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.71	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.71	[redacted]	[redacted]	Dominated	Dominated

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 11- Incremental Cost-effectiveness with linear progression on cDMARDs and SAE costs

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
ETNMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	Referent	Referent
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.69	[redacted]	[redacted]	£7,499	£7.499
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	£6,190	£5,730
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	Dominated	Dominated
IFXMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.70	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.71	[redacted]	[redacted]	Dominated	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.71	[redacted]	[redacted]	Dominated	Dominated

B7. In relation to the HAQ improvement based on EULAR response, please clarify why the samples for the alpha and beta parameters are not correlated.

For HAQ improvement based on EULAR response, the model uses data provided in TA375 (HTA 20[35] April 2016, p.249)[1]. TA375 does not provide a variance-covariance matrix, therefore there was a lack of data that precluded accounting for correlation of samples for alpha and beta parameters. Model functionality was added to prevent overlapping of realised values in the VBA code. While inclusion of such functionality would be ideal, it is not expected to impact overall results.

B8. Please clarify why average weights were used in the baseline model rather than distributions of these weights? This can lead to errors: see Hatswell AJ, Porter J, Lee D, Hertel N & Latimer NR (2016) The Cost of Costing Treatments Incorrectly: Errors in the Application of Drug Prices in Economic Evaluation Due to Failing to Account for the Distribution of Patient Weight. Value in Health, 19: 1055-8.

In the analyses presented, weight-based dosing would have only impacted on infliximab and tocilizumab. As a vial wastage approach was used, the cost for infliximab at a dose of 3mg/kg covered patients weighing between 66.7-100kg, there for accounting for a fairly broad weight range. With respect to tocilzumab, an increase in weight by 10kg led to the requirement of an additional 80mg vial at a dose of 8mg/kg- therefore the vial size available implies that there will be limited impact from taking into account weight distribution.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

As noted in the reference above, a number of factors influence the impact of this assumption such as vial wastage and the nature of the weight distribution applicable- the paper concludes that using a mean weight approach often underestimates the cost, therefore in this case, an underestimate of costs for infliximab and tocilizumab would be negative toward baricitinib, therefore this may be of limited relevance, as baricitinib is not dosed by weight.

B9. Please clarify why subcutaneous tocilizumab was not modelled. Many comparators, which have been modelled, have confidential patient access schemes so this does not appear to be a valid reason.

To confirm, two comparators have confidential patient access schemes (abatacept and tocilizumab). Patient access schemes for the other comparators have been included as appropriate. A pragmatic decision was made to model only IV tocilizumab which was influenced by the fact that only one study was found in the literature review and therefore included in the evidence synthesis for sub-cutaneous tocilizumab and this study was combination with cDMARDs- (including patients not just on background methotrexate)- which was a potential source of uncertainty. Secondly, this study informing the S/C efficacy estimate from the NMA resulted in a lower point estimate for a EULAR good response than the IV version- therefore including the IV was thought to be conservative. Finally, the difference in annual costs based on list price between the IV and S/C versions of tocilizumab is relatively small (approximately £1000). Although, when factoring in the administration cost of an IV drug, the IV version does become more expensive, the cost difference remains of a similar magnitude. The model also assumes no difference in monitoring costs between IV and SC administered drugs. For these reasons, and given that there were already a relatively large number of sequences in the model, including only IV tocilizumab only was felt to be a reasonable choice, with it likely to be representative of the costs and outcomes associated with the S/C version.

B10. Please clarify what correlations have been incorporated into the model for patient characteristics, and for dependent parameter values (such as the shape and scale of a Weibull distribution).

Model functionality has subsequently been added to account for correlation between included patient baseline characteristics (i.e. gender, age, weight and HAQ), using a multivariate normal distribution. It is noted that the omission of correlations for dependent parameter values (i.e. shape and scale of the Weibull distribution for long-term discontinuation) is a limitation but this was not possible due to the lack of correlation co-efficient reported in TA375 without access to patient level data. The cost-effectiveness results with this modelling scenario are shown in table 12. The order of treatments in the incremental analysis remain as per the base case with only small differences between sequences with respect to QALYs.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 12- Incremental Cost-effectiveness with patient characteristic correlations incorporated

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.67	[redacted]	[redacted]	Referent	Referent
ETNMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.60	[redacted]	[redacted]	Dominated	Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.61	[redacted]	[redacted]	£26,987	£26,987
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.59	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.59	[redacted]	[redacted]	Dominated	Dominated
IFXMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.58	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.61	[redacted]	[redacted]	Dominated	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.59	[redacted]	[redacted]	Dominated	Dominated

B11. Section 5.6.1, page 276. Please clarify why the standard error for the administration and monitoring costs of ADA and TCZ in year 1 are different to the other interventions in Table 115. Please clarify why the monitoring costs of RTX, which is given less frequently, are the same as other bDMARDs.

There has been a transcription error taking this information from the model into the company submission. In the model, the same SE applies to all interventions apart from IV abatacept and infliximab. The cells in question are I46-I59 on the data input sheet. The model assumed equal monitoring cost for methotrexate and all biologic DMARDs (as well as baricitinib). In the absence of any clear data supporting the proposition that rituximab should have lower monitoring costs, the choice was made to assume equivalence. This is supported by the expectation that it is unlikely that monitoring will be impacted by dosing frequency, particularly as rituximab is only licensed in combination with methotrexate whose monitoring requirements will likely be consistent regardless of the combination components.

B12. Section 5.6.2, table 116, page 281. Please clarify why the dosing of infliximab is in line with TA195 rather than TA375.

As per TA375 the dosing regimen used in the model is as listed in the British National Formulary (BNF) and the Summary of Product Characteristics (SPC) for Remicade 100mg powder: 3mg/kg at weeks 0, two and six, followed by eight weekly administrations. As

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

specified above, the model accounts for wastage and so only whole vials are included. The model therefore allows for eight administrations in year one and seven in subsequent years which, when rounded up, is in line with TA375.

B13. Please clarify how many patients were run through the model to generate probabilistic results. Please comment on whether this number of patients were sufficient to generate stable results.

In the PSA results presented in the submission, each PSA loop simulated 500 patients per treatment arm (i.e. a total of 1000 patients across two treatment arms). The analysis used 1000 sampling loops for the parameters included in the probabilistic model. With respect to stability of results, whilst assessment of convergence suggested that analyses should be conducted with a total of approximately 55,000 patients (Fig 4), a degree of difficulty was presented by the computational burden of the model so necessary limits needed to be applied to the number of patients and probabilistic runs. It should be noted that the base case results in the cDMARD-IR population were similar between the deterministic and probabilistic models, that the PSA setting compared favourably with TA375 (i.e. 1,000 patients in the severe subpopulation and 100 iterations) and that O'Hagan et al[3] postulated that the most efficient method of generating expected values for cost-effectiveness would be to generate only one patient per PSA iteration.

Convergence diagnostics were run for total costs and QALYs, and shown below in Figs 4 and 5.

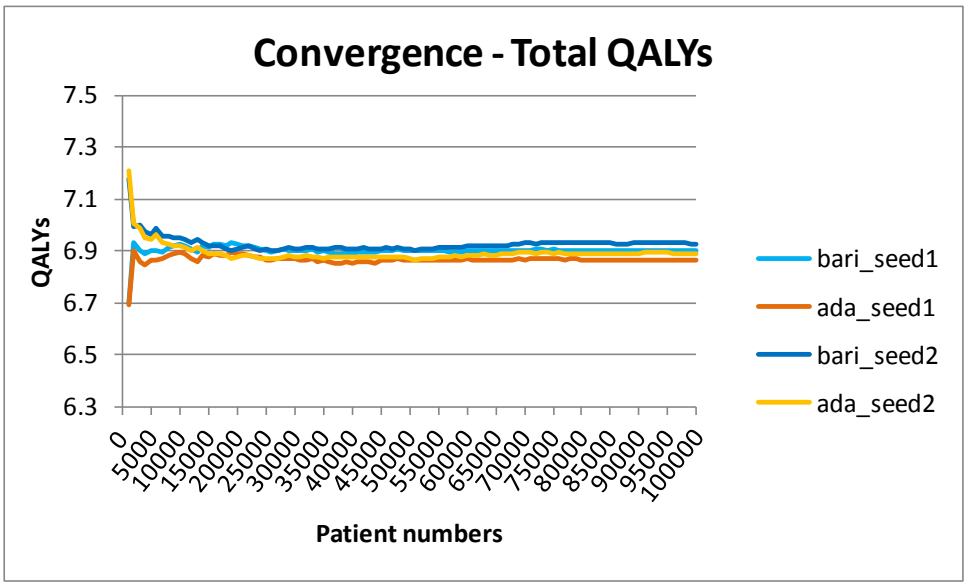


Fig 4- Convergence graph- Total QALYs

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

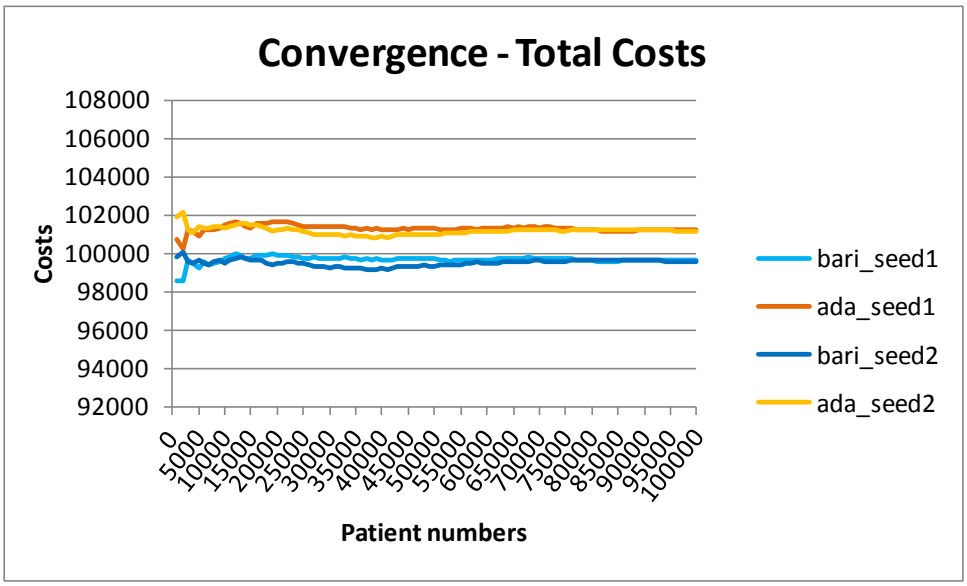


Fig 5- Convergence graph- Total Costs

B14. Please clarify how many PSA configurations were run and whether this was sufficient to produce robust results.

Please see the response to B13 which we believe also addresses this question.

B15. Please clarify why ABT IV was not considered within the submission. Please answer from the perspectives of i) ABT IV as a relevant comparator and ii) to allow further information to be gained regarding the comparison of subcutaneous ABT and cDMARDs through linking to ABT IV using the ACQUIRE study.

Intravenous abatacept (ABT IV) was included in the evidence synthesis network (studies AIM and ATTEST were included). Again, a pragmatic decision was made to attempt to limit the number of sequences included in the submission where it was possible that inclusion of the same intervention with different administration routes was unlikely to be informative. Given similar list prices between the S/C and I/V forms once administration costs had been taken into account, only the S/C version was included in the base-case sequences. The table below includes the IV sequence and shows only small differences between the estimates in costs and outcomes between the two different abatacept administration routes (whilst noting that a confidential patient access scheme applies to abatacept). Please see the response to clarification question A4 with respect to the ACQUIRE study.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 13- Incremental cost-effectiveness including ABT IV sequence

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Referent	Referent
ETNMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	£18,400	£18,400
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
IFXMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
ABTIMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated

B16. Please clarify why results were not presented for patients receiving 2 mg baricitinib from initiation.

Section 4.2 of the SPC states that 'The recommended dose of baricitinib is 4mg once daily' A dose of 2mg daily is suggested for patients above the age of 75 or those with a history of chronic or recurrent infections. Additionally, patients with renal impairment causing creatinine clearance to be between 30 and 60ml/min are recommended to take 2mg daily. Therefore, the majority of patients commencing treatment would start on the 4mg dose, hence why only the 4mg dose from initiation was considered.

B17. Please clarify why results were not presented for patients who may step down from 4 mg to 2 mg of baricitinib.

This scenario was presented and is described in detail as scenario 6 in table 127 of the company submission. The results are reproduced in more detail in table X below. Please also note the response to clarification question C3 below.

Table 14- Incremental Cost-effectiveness for baricitinib step-down scenario

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Referent	Referent
ETNMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	£13,347	Ext Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	£9,656	£9,656
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
IFXMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	£180,573	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.73	[redacted]	[redacted]	Dominated	Dominated

B18. Please clarify how adherence, which was stated to be expected to be higher for an oral treatment, was incorporated into the model. It appears that the time to discontinuation when using a baricitinib-only curve is less than for other bDMARDs.

It should be clarified that adherence was discussed in a very much hypothetical perspective as there is no compelling evidence that an oral vs. injectable route of administration results in better adherence for the oral dosage form. And at this stage, real world data on adherence will of course be limited so it is not possible to make an assessment on the long-term adherence profile of baricitinib.

The scenario which explores an alternate assumption on the discontinuation rate uses the observed data for the mITT population in the JADV study, for which the 52-week discontinuation rate was 12.5% Applying scale and shape parameters to approximate this rate results in patients having more time on treatment with baricitinib compared to biologic treatments which leads higher costs and benefits for baricitinib in this scenario and the incremental ICER of £35,393.

B19. Please clarify why it was assumed that the results from the NMA for patients who did not respond adequately to cDMARDs were generalisable to those patients who did not respond adequately to an anti-TNF (where data were not available). For those interventions where data were available for both patient populations it is clear that the EULAR responses are worse where a patient has already received an anti-TNF. Please present results for the population who did not respond adequately to an anti-TNF without ADA, CTZ, ETN, and IFX.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

This was a pragmatic decision, although a number of alternative approaches could have been taken, (e.g. arbitrary downgrade of efficacy, assumption of similar efficacy to a treatment with data available etc.) but as each would be subject to uncertainty, cDMARD-IR data was used for these interventions as this was likely to represent the best case scenario for these treatments, given, as noted in the question, responses in patients with prior TNF-inihibitor treatment appear to trend lower with treatments where both cDMARD-IR and TNF-IR data is available. The requested analysis is presented in table X.

Table 15- Incremental Cost-effectiveness, TNF-IR excluding ADA, CTX, ETN and IFX

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
BAR4MTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	-	-
GOLMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+ABTSMTX+MTX+Pall	[redacted]	[redacted]	13.50	[redacted]	[redacted]	£430,301	£430,301
ABTSMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	13.49	[redacted]	[redacted]	£484,782	Dominated

B20. Please clarify why it is assumed that the HAQ change associated with EULAR response happens at time zero. In TA375, this occurred at 6 months.

The model applies the HAQ change associated with EULAR response at baseline rather than at 6 months, contrary to TA375. The clinical data from the JADV trial, in Table 15 below, demonstrates that the largest clinical improvement is gained within 12 weeks (Month 3), rather than Week 24 (Month 6) for both baricitinib and adalimumab. NICE guidance states that assessment of response takes place at 24 weeks for biologics but it could be argued that applying the benefit of HAQ at that point underestimates the benefits that would already have been accrued. In addition to the data below, there is evidence that clinical response to bDMARDs in RA is often rapid, with patients potentially experiencing improvements in symptoms within a few weeks of treatment initiation[4,5,6,7], perhaps even as early as 48 hours[8] after commencement. Therefore, the model applied HAQ change at time zero based on an assessment of response at week 24 as it is unlikely that a patient would experience not treatment benefit at all until the assessment point at week 24.

Additionally, as this approach is being applied across all treatments arms and sequences, the impact on incremental costs and QALYs is likely to be minimal.

Table 16- HAQ Change over time, JADV ITT population

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Mean (SD) HAQ-DI, JADV ITT population	PBO + MTX N=488	BAR 4mg + MTX N=487	ADA + MTX N=330	Total N=1305
Baseline	1.547 (0.671)	1.566 (0.678)	1.587 (0.702)	1.565 (0.681)
Week 12	1.216 (0.723)	0.914 (0.690)	1.028 (0.688)	1.056 (0.713)
Week 24	1.206 (0.735)	0.839 (0.698)	0.959 (0.726)	1.007 (0.736)
Week 52	--	0.852 (0.710)	0.948 (0.756)	0.891 (0.730)

B21. Related to question A2. If it is assumed that the relative efficacy observed in RA-BEGIN between baricitinib + MTX and baricitinib monotherapy was generalisable to patients who were cDMARD experienced and anti-TNF experienced, please provide results for baricitinib monotherapy.

Please see the response to question A2 for a rationale as to why cost-effectiveness results have not been presented for monotherapy.

Section C: Textual clarifications and additional points

C1. Please clarify the number of baricitinib studies included in the submission. Figure 5 (page 75) suggests that six studies were identified however Table 12 (page 77) suggests that there are only five relevant studies.

Please note the baricitinib submission references five key PIII studies:

- RA-BEAM (JADV)
- RA-BUILD (JADX)
- RA-BEACON (JADW)
- RA-BEGIN (JADZ) (unlicensed population)
- RA-BEYOND (JADY)

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The PRISMA diagram in Figure 7 of the submission includes some additional secondary references which can be seen below.

Table 17 – List of included baricitinib studies

Trial Acronym (or Author, Year)	Primary Reference	Secondary Reference(s)
Fleischmann (2015)	Fleischmann R, Takeuchi T, Schlichting DE, Macias WL, Rooney T, Gurbuz S, et al. Baricitinib, methotrexate, or baricitinib plus methotrexate in patients with early rheumatoid arthritis who had received limited or no treatment with disease-modifying anti-rheumatic drugs (DMARDs): phase 3 trial results. Arthritis Rheumatol. 2015;67(suppl 10). Abstract Number 1045.	None
RA-BEACON	Genovese MC, Kremer J, Zamani O, Ludivico C, Krogulec M, Xie L, et al. Baricitinib in patients with refractory rheumatoid arthritis. N Engl J Med. 2016 Mar 31;374(13):1243-52.	Genovese MC, Kremer J, Zamani O, Ludicico C, Krogulec M, Xie L, et al. Baricitinib, an oral janus kinase (JAK)1/JAK2 inhibitor, in patients with active rheumatoid arthritis (RA) and an inadequate response to TNF inhibitors: results of the phase 3 RA-BEACON study. Ann Rheum Dis. 2015;74(suppl 2):75.
RA-BUILD	Dougados M, van der Heijde D, Chen YC, Greenwald M, Drescher E, Liu J, et al. Baricitinib, an oral janus kinase (JAK)1/JAK2 inhibitor, in patients with active rheumatoid arthritis (RA) and an inadequate response to cDMARD therapy: results of the phase 3 RA-BUILD study. Ann Rheum Dis. 2015;74(suppl 2):79.	Emery P, Gaich C, DeLozier A, de Bono S, Liu J, Chang C, et al. Patient-Reported Outcomes from a Phase 3 Study of Baricitinib in Patients with Rheumatoid Arthritis with Inadequate Response to Conventional Synthetic Disease-Modifying Antirheumatic Drugs. Arthritis Rheumatol. 2015e;67((suppl 10)).
Taylor (2015)	Taylor PC, Keystone EC, van der Heijde D, Tanaka Y, Ishii T, Emoto K, et al. Baricitinib versus placebo or adalimumab in patients with active rheumatoid arthritis (RA) and an inadequate response to background methotrexate therapy: results of a phase 3 study. Arthritis Rheumatol. 2015;67(suppl 10).	None

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

C2. Section 4.7.1, page 107. Please clarify whether there is significant difference in ACR20 scores between baricitinib and adalimumab at Week 52. Figure 13 and the accompanying explanatory text appear to be contradictory.

To clarify, there was a statistically significant difference in favour of baricitinib in ACR20 scores at Week 52. The accompanying text only refers to ACR20 scores at Week 12 as this was the primary outcome time point for the study.

C3. Section 5.2.8, page 249 (and elsewhere). It is stated that a HAQ decrement of 0.06 is applied when there is a step down from 4 mg to 2 mg. Please confirm that this is actually an increase of 0.06 (i.e. a worsening in HAQ score).

To confirm, the 0.06 decrement mentioned in the submission refers to an increase in HAQ score of 0.06 (i.e. worsening in HAQ score).

C4. Section 5.3, page 256. It is stated that following treatment discontinuation the HAQ score would rebound immediately to the level prior to initiation of the terminated therapy. Please clarify that this is applicable for bDMARDs (and baricitinib in the base case) but not MTX or the scenario analysis where baricitinib is associated with a HAQ increase across time.

In the base case, treatment discontinuation after any bDMARD or baricitinib will cause the HAQ score to rebound to the level prior to initiation of the terminated therapy. In the version of the model submitted to NICE, this was also the case with MTX. Results with an amended version of the model are in table 18- please note that these results also include the correlations for which model sensitivity was discussed in the response to clarification question B10.
Again, the order of treatments in the incremental analysis remains as per the original base case and the difference in QALYs between treatments small. To confirm, in this scenario discontinuation of MTX (or palliative care) does not cause the HAQ to rebound to the level at initiation of that treatment, rather it continues to progress.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 18- Incremental Cost-effectiveness, cDMARD-IR population with no HAQ rebound on MTX

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.67	[redacted]	[redacted]	Referent	Referent
ETNMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.60	[redacted]	[redacted]	Dominated	Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.61	[redacted]	[redacted]	£40,181	£40,181
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.59	[redacted]	[redacted]	Dominated	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.59	[redacted]	[redacted]	Dominated	Dominated
IFXMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.58	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.61	[redacted]	[redacted]	Dominated	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.59	[redacted]	[redacted]	Dominated	Dominated

Table 19 shows the results for the amended model version for the scenario where baricitinib is associated with HAQ increase across time with the same rule applied to baricitinib as for MTX (and palliative care) i.e. the HAQ score does not rebound to the pre-treatment level, rather it continues to progress. In this scenario, the HAQ score for patients treated with other bDMARDs does rebound, as described for the base case.

This scenario is not included as a model option and was conducted by manually altering the time to increase HAQ by 0.125. The method is only applicable when a linear HAQ trajectory is selected.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 19- Incremental Cost-effectiveness, TNF-IR

Technology sequence	Total costs (£)	Total QALYs	LYG	Inc costs (£)	Inc QALYs	ICER (£) v. baseline (per QALY)	Inc Analysis
BAR4MTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.64	[redacted]	[redacted]	Referent	Referent
ETNMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.66	[redacted]	[redacted]	£10,103	Ex Dominated
CTZMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.67	[redacted]	[redacted]	£9,147	£9,147
GOLMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.66	[redacted]	[redacted]	£52,484	Dominated
ADAMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.66	[redacted]	[redacted]	£55,837	Dominated
IFXMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.65	[redacted]	[redacted]	Dominated	Dominated
TCZMTX+RTXMTX+ADAMTX+MTX+Pall	[redacted]	[redacted]	14.65	[redacted]	[redacted]	£119,669	Dominated
ABTSMTX+RTXMTX+TCZMTX+MTX+Pall	[redacted]	[redacted]	14.66	[redacted]	[redacted]	£866,407	Dominated

References

1. Stevenson M, Archer R, Tosh J, Simpson E, Everson-Hock E, Stevens J, Hernandez-Alava M, Paisley S, Dickinson K, Scott D, Young A & Wailoo A.. Adalimumab, etanercept, infliximab,certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying antirheumatic drugs and after the failure of conventional disease-modifying antirheumatic drugs only: systematic review and economic evaluation. Health Technology Assessment 20(35), May 2016. Available at: https://njl-admin.nihr.ac.uk/document/download/2003416

2. Dias, S., Welton, N.J., Sutton, A.J. & Ades, A.E. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011; last updated September 2016; available from http://www.nicedsu.org.uk

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

3. O'Hagan A, Stevenson M, Madan J. Monte Carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using ANOVA. Health Econ 2007;16:1009–23

4. Cohen SB, Emery P, Greenwald MW, 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 Rheum.

5. Moreland LW, Schiff MH, Baumgartner SW, et al. Etanercept therapy in rheumatoid arthritis. A randomized, controlled trial. Ann Intern Med. 1999;130:478–486

6. Genovese MC, Becker JC, Schiff M, et al. Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. N Engl J Med. 2005;353:1114–1123.

7. Weinblatt ME, Keystone EC, Furst DE, 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 Rheum. 2003;48:35–45.

8. Shergy WJ, Isern RA, Cooley DA, et al. Open label study to assess infliximab safety and timing of onset of clinical benefit among patients with rheumatoid arthritis. J Rheumatol. 2002;29:667–677.

.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Appendix 1- EULAR NMA Data Inputs

cDMARD-IR Network:

EULAR response data: population: cDMARD-IR

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
Abe (2006)	cDMARD	14	47	17	19.1	63.8	Y		Y
Abe (2006)	IFX 3mg + cDMARD	14	49	34.7	28.6	36.7	Y		Y
ACCOMPANY	ABA SUBCUT	12	-	-	-	-	N	Missing data	-
ACCOMPANY	ABA SUBCUT	24	-	-	-	-	N	Missing data	-
ACCOMPANY	ABA SUBCUT + cDMARD	12	-	-	-	-	N	Missing data	-
ACCOMPANY	ABA SUBCUT + cDMARD	24	-	-	-	-	N	Missing data	-
ACT-RAY	TCZ 8mg	12	276	33	31.5	35.5	Y		Y
ACT-RAY	TCZ 8mg	24	276	51.4	34.8	13.8	Y		-
ACT-RAY	TCZ 8mg + cDMARD	12	277	36.8	31	32.1	Y		Y
ACT-RAY	TCZ 8mg + cDMARD	24	277	61.7	27.8	10.5	Y		-
ADACTA	ADA 40mg	12	-	-	-	-	N	Missing data	-
ADACTA	ADA 40mg	24	162	19.8	35.2	45.1	Y		-
ADACTA	TCZ 8mg	12	-	-	-	-	N	Missing data	-
ADACTA	TCZ 8mg	24	163	51.5	26.4	22.1	Y		-
AIM	ABA 10mg + cDMARD	12	424	35.6	28.1	36.3	Y		Y

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
AIM	ABA 10mg + cDMARD	24	424	38.2	30.9	30.9	Y		Y
AIM	cDMARD	12	214	20.1	27.1	52.8	Y		Y
AIM	cDMARD	24	214	23.8	25.2	50.9	Y		Y
AMBITION	cDMARD	12	284	27.1	26.4	46.5	Y		Y
AMBITION	cDMARD	24	284	14.8	50	35.2	Y		-
AMBITION	TCZ 8mg	12	286	36	26.6	37.4	Y		Y
AMBITION	TCZ 8mg	24	286	37.8	44.4	17.8	Y		-
AMPLE	ABA SUBCUT + cDMARD	12	318	33.6	30.5	35.8	Y		Y
AMPLE	ABA SUBCUT + cDMARD	24	319	37.3	30.4	32.3	Y		Y
AMPLE	ADA 40mg + cDMARD	12	328	36	27.7	36.3	Y		Y
AMPLE	ADA 40mg + cDMARD	24	329	37.1	30.7	32.2	Y		Y
APPEAL	cDMARD	12	103	32	28.2	39.8	Y		Y
APPEAL	ETN + cDMARD	12	197	43.7	25.9	30.5	Y		Y
ARMADA	ADA 40mg + cDMARD	12	67	35.8	32.8	31.3	Y		Y
ARMADA	ADA 40mg + cDMARD	24	67	41.8	25.4	32.8	Y		Y
ARMADA	cDMARD	12	62	16.1	21	62.9	Y		Y
ARMADA	cDMARD	24	62	12.9	17.7	69.4	Y		Y
ATTEST	ABA 10mg + cDMARD	28	156	37.8	30.1	32.1	Y		Y
ATTEST	cDMARD	28	109	24.8	25.7	49.5	Y		Y
ATTEST	IFX 3mg + cDMARD	28	165	33.3	30.9	35.8	Y		Y
ATTRACT	cDMARD	14	-	-	-	-	N	Missing data	-
ATTRACT	cDMARD	26	-	-	-	-	N	Missing data	-
ATTRACT	IFX 3mg	14	-	-	-	-	N	Missing data	-
ATTRACT	IFX 3mg	26	-	-	-	-	N	Missing data	-
BEAM	ADA 40mg + cDMARD	12	330	21.2	55.8	23	Y		-

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
BEAM	ADA 40mg + cDMARD	24	330	33.6	41.8	24.5	Y		-
BEAM	BARI 4mg + cDMARD	12	487	23.8	58.9	17.2	Y		-
BEAM	BARI 4mg + cDMARD	24	487	31.6	50.3	18.1	Y		-
BEAM	cDMARD	12	488	6.8	43.9	49.4	Y		-
BEAM	cDMARD	24	488	9.2	39.1	51.6	Y		-
BREVACTA	cDMARD	24	173	20.8	26	53.2	Y		Y
BREVACTA	TCZ SUBCUT + cDMARD	24	348	38.2	27.9	33.9	Y		Y
BUILD	BARI 2mg + cDMARD	12	229	21.4	50.7	27.9	Y		-
BUILD	BARI 2mg + cDMARD	24	229	27.9	43.2	28.8	Y		-
BUILD	BARI 4mg + cDMARD	12	227	21.1	55.1	23.8	Y		-
BUILD	BARI 4mg + cDMARD	24	227	30.8	44.1	25.1	Y		-
BUILD	cDMARD	12	228	6.6	43.9	49.6	Y		-
BUILD	cDMARD	24	228	9.6	42.1	48.2	Y		-
CAMEO	ETN	24	-	-	-	-	N	Missing data	-
CAMEO	ETN + cDMARD	24	-	-	-	-	N	Missing data	-
CHANGE	ADA 40mg	12	91	23.1	29.7	47.3	Y		Y
CHANGE	ADA 40mg	24	91	26.4	25.3	48.4	Y		Y
CHANGE	Placebo	12	87	11.5	18.4	70.1	Y		Y
CHANGE	Placebo	24	87	12.6	17.2	70.1	Y		Y
CNTO 148	cDMARD	12	35	22.9	22.9	54.3	Y		Y
CNTO 148	cDMARD	20	35	25.7	28.6	45.7	N	Request	Y
CNTO 148	GOL 50mg + cDMARD	12	35	34.3	25.7	40	Y		Y
CNTO 148	GOL 50mg + cDMARD	20	35	42.9	28.6	28.6	N	Request	Y
Combe (2006)	ETN	12	104	36.5	28.8	34.6	N	Request	Y
Combe (2006)	ETN	24	103	42.7	30.1	27.2	Y		Y
Combe (2006)	ETN + SSZ	12	101	37.6	25.7	36.6	N	Request	Y

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
Combe (2006)	ETN + SSZ	24	101	43.6	28.7	27.7	Y		Y
Combe (2006)	SSZ	12	50	20	26	54	N	Request	Y
Combe (2006)	SSZ	24	50	20	20	60	Y		Y
De Filippis (2006)	ETN + cDMARD	14	15	20	40	40	Y		Y
De Filippis (2006)	ETN + cDMARD	22	15	33.3	26.7	40	Y		Y
De Filippis (2006)	IFX 3mg + cDMARD	14	14	42.9	28.6	28.6	Y		Y
De Filippis (2006)	IFX 3mg + cDMARD	22	16	37.5	25	37.5	Y		Y
Edwards (2004)	cDMARD	24	40	22.5	25	52.5	Y		Y
Edwards (2004)	RTX 1000mg	24	40	35	32.5	32.5	Y		Y
Edwards (2004)	RTX 1000mg + cDMARD	24	40	40	32.5	27.5	Y		Y
GO-FORTH	cDMARD	14	88	18.2	21.6	60.2	Y		Y
GO-FORTH	cDMARD	24	87	20.7	23	56.3	Y		Y
GO-FORTH	GOL 50mg + cDMARD	14	86	40.7	31.4	27.9	Y		Y
GO-FORTH	GOL 50mg + cDMARD	24	87	37.9	34.5	27.6	Y		Y
GO-FORWARD	cDMARD	14	133	19.5	24.8	55.6	Y		Y
GO-FORWARD	cDMARD	24	133	18.8	21.1	60.2	Y		Y
GO-FORWARD	GOL 50mg + cDMARD	14	88	34.1	25	40.9	Y		Y
GO-FORWARD	GOL 50mg + cDMARD	24	90	34.4	28.9	36.7	Y		Y
Hobbs (2015)	cDMARD	12	104	20.2	20.2	59.6	N	Request	Y
Hobbs (2015)	ETN + cDMARD	12	105	25.7	22.9	51.4	N	Request	Y
J-RAPID	cDMARD	12	77	18.2	22.1	59.7	Y		Y
J-RAPID	cDMARD	24	77	20.8	15.6	63.6	Y		Y
J-RAPID	CZP + cDMARD	12	82	41.5	34.1	24.4	Y		Y
J-RAPID	CZP + cDMARD	24	82	43.9	29.3	26.8	Y		Y
JESMR	ETN	12	69	42	50.7	7.2	N	Request	-
JESMR	ETN	24	69	33.3	37.7	29	Y		-
JESMR	ETN + cDMARD	12	73	21.9	50.7	27.4	N	Request	-

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
JESMR	ETN + cDMARD	24	73	52.1	43.8	4.1	Y		-
Kang (2013)	cDMARD	24	40	22.5	15	62.5	Y		Y
Kang (2013)	CZP + cDMARD	24	81	39.5	27.2	33.3	Y		Y
Keystone (2004)	ADA 40mg + cDMARD	12	207	33.8	28	38.2	Y		Y
Keystone (2004)	ADA 40mg + cDMARD	24	206	36.4	29.6	34	Y		Y
Keystone (2004)	cDMARD	12	200	16.5	20.5	63	Y		Y
Keystone (2004)	cDMARD	24	200	18.5	23	58.5	Y		Y
Kim (2007)	ADA 40mg + cDMARD	12	65	35.4	24.6	40	Y		Y
Kim (2007)	ADA 40mg + cDMARD	24	65	36.9	27.7	35.4	Y		Y
Kim (2007)	cDMARD	12	63	17.5	20.6	61.9	Y		Y
Kim (2007)	cDMARD	24	62	21	25.8	53.2	Y		Y
Kremer (2002)	cDMARD	24	133	14.3	20.3	65.4	N	Request	Y
Kremer (2002)	LEF 10mg + cDMARD	24	130	28.5	24.6	46.9	N	Request	Y
Lan (2004)	cDMARD	12	29	20.7	24.1	55.2	Y		Y
Lan (2004)	ETN + cDMARD	12	29	55.2	27.6	17.2	Y		Y
Li (2013)	cDMARD	24	132	13.6	18.2	68.2	Y		Y
Li (2013)	GOL 50mg + cDMARD	24	132	25.8	25	49.2	Y		Y
Lim (2012)	cDMARD	24	40	2.5	32.5	65	N	Request	-
Lim (2012)	TCZ 8mg + cDMARD	24	40	52.5	32.5	15	N	Request	-
LITHE	cDMARD	12	393	16.8	22.4	60.8	Y		Y
LITHE	cDMARD	24	393	18.3	21.6	60.1	Y		Y
LITHE	TCZ 8mg + cDMARD	12	398	29.1	27.9	43	Y		Y
LITHE	TCZ 8mg + cDMARD	24	398	33.7	26.6	39.7	Y		Y
Machado (2014)	cDMARD	12	142	31	28.2	40.8	Y		Y
Machado (2014)	cDMARD	24	142	12	52.8	35.2	Y		-
Machado (2014)	ETN + cDMARD	12	279	47	29	24	Y		Y
Machado (2014)	ETN + cDMARD	24	279	47	44.8	8.2	Y		-

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
Moreland (1999)	ETN	13	79	38	25.3	36.7	Y		Y
Moreland (1999)	ETN	26	78	37.2	24.4	38.5	Y		Y
Moreland (1999)	Placebo	13	80	15	21.2	63.7	Y		Y
Moreland (1999)	Placebo	26	80	11.2	17.5	71.2	Y		Y
Nishimoto (2004)	Placebo	12	53	0	18.9	81.1	Y		-
Nishimoto (2004)	TCZ 8mg	12	55	18.2	72.7	9.1	Y		-
OPTION	cDMARD	12	204	17.6	23.5	58.8	Y		Y
OPTION	cDMARD	24	204	2.9	31.9	65.2	Y		-
OPTION	TCZ 8mg + cDMARD	12	205	35.1	29.3	35.6	Y		Y
OPTION	TCZ 8mg + cDMARD	24	205	38	41.5	20.5	Y		-
ORAL Scan	cDMARD	13	161	17.4	23	59.6	Y		Y
ORAL Scan	cDMARD	26	80	17.5	21.2	61.3	Y		Y
ORAL Scan	TOFA 10mg + cDMARD	13	317	36.9	31.2	31.9	Y		Y
ORAL Scan	TOFA 10mg + cDMARD	26	316	37.3	26.9	35.8	Y		Y
ORAL Scan	TOFA 5mg + cDMARD	13	321	32.4	28	39.6	Y		Y
ORAL Scan	TOFA 5mg + cDMARD	26	320	31.6	25.3	43.1	Y		Y
ORAL Standard	ADA 40mg + cDMARD	12	204	31.4	29.4	39.2	Y		Y
ORAL Standard	ADA 40mg + cDMARD	24	204	30.4	22.5	47.1	Y		Y
ORAL Standard	cDMARD	12	108	16.7	23.1	60.2	Y		Y
ORAL Standard	cDMARD	24	108	19.4	21.3	59.3	Y		Y
ORAL Standard	TOFA 10mg + cDMARD	12	201	31.8	31.8	36.3	Y		Y
ORAL Standard	TOFA 10mg + cDMARD	24	202	32.7	25.7	41.6	Y		Y
ORAL Standard	TOFA 5mg + cDMARD	12	204	36.3	27	36.8	Y		Y
ORAL Standard	TOFA 5mg + cDMARD	24	204	32.8	24	43.1	Y		Y
RA-SCORE	cDMARD	24	63	19	22.2	58.7	Y		-
RA-SCORE	RTX 1000mg	24	60	35	41.7	23.3	Y		-
RACAT	ETN + cDMARD	24	163	33.7	25.8	40.5	Y		Y

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
RACAT	SSZ + HCQ + cDMARD	24	159	32.7	26.4	40.9	Y		Y
RAPID 1	cDMARD	12	199	14.1	19.1	66.8	Y		Y
RAPID 1	cDMARD	24	199	12.6	17.6	69.8	Y		Y
RAPID 1	CZP + cDMARD	12	394	35.5	30.7	33.8	Y		Y
RAPID 1	CZP + cDMARD	24	393	33.8	29.3	36.9	Y		Y
RAPID 2	cDMARD	12	127	11.8	17.3	70.9	Y		Y
RAPID 2	cDMARD	24	126	9.5	16.7	73.8	Y		Y
RAPID 2	CZP + cDMARD	12	246	34.1	33.3	32.5	Y		Y
RAPID 2	CZP + cDMARD	24	246	32.9	28.5	38.6	Y		Y
REALISTIC	cDMARD	12	212	17	21.2	61.8	Y		Y
REALISTIC	CZP + cDMARD	12	851	30.8	28.2	41	Y		Y
RED SEA	ADA 40mg + cDMARD	12	63	-	-	-	N	Missing data	-
RED SEA	ADA 40mg + cDMARD	24	63	-	-	-	N	Missing data	-
RED SEA	ETN + cDMARD	12	62	-	-	-	N	Missing data	-
RED SEA	ETN + cDMARD	24	62	-	-	-	N	Missing data	-
RELIEF	LEF 20mg + SSZ	24	56	14.3	25	60.7	N	Request	Y
RELIEF	SSZ	24	49	16.3	20.4	63.3	N	Request	Y
ROSE	cDMARD	12	206	18.9	21.8	59.2	N	Request	Y
ROSE	cDMARD	24	205	18	20	62	N	Request	Y
ROSE	TCZ 8mg + cDMARD	12	408	28.4	27.5	44.1	N	Request	Y
ROSE	TCZ 8mg + cDMARD	24	409	28.1	24.2	47.7	N	Request	Y
SATORI	cDMARD	12	64	15.6	25	59.4	Y		Y
SATORI	cDMARD	24	64	3.1	35.9	60.9	Y		-
SATORI	TCZ 8mg + cDMARD	12	62	38.7	35.5	25.8	Y		Y
SATORI	TCZ 8mg + cDMARD	24	61	65.6	31.1	3.3	Y		-
SERENE	cDMARD	24	172	4.7	29.1	66.3	Y		-
SERENE	RTX 1000mg + cDMARD	24	167	17.4	49.1	33.5	Y		-

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
SERENE	RTX 2000mg + cDMARD	24	170	11.8	51.2	37.1	Y		-
STAR	ADA 40mg + cDMARD	12	318	29.9	27	43.1	Y		Y
STAR	ADA 40mg + cDMARD	24	318	30.5	28	41.5	Y		Y
STAR	cDMARD	12	318	18.2	23.9	57.9	Y		Y
STAR	cDMARD	24	318	20.8	24.8	54.4	Y		Y
START	cDMARD	22	340	16.8	22.4	60.9	Y		Y
START	IFX 3mg + cDMARD	22	342	33.6	28.1	38.3	Y		Y
SUMMACTA	TCZ 8mg	12	537	38.2	31.5	30.4	Y		Y
SUMMACTA	TCZ 8mg	24	538	41.4	32	26.6	Y		Y
SUMMACTA	TCZ SUBCUT	12	558	38.7	30.6	30.6	Y		Y
SUMMACTA	TCZ SUBCUT	24	558	40.5	29.4	30.1	Y		Y
SURPRISE	TCZ 8mg	24	115	40.9	30.4	28.7	N	Request	Y
SURPRISE	TCZ 8mg + cDMARD	24	118	43.2	31.4	25.4	N	Request	Y
TEMPO	cDMARD	12	228	35.1	30.3	34.6	Y		Y
TEMPO	cDMARD	24	228	41.7	30.3	28.1	Y		Y
TEMPO	ETN	12	223	38.1	30	31.8	Y		Y
TEMPO	ETN	24	223	40.4	30.5	29.1	Y		Y
TEMPO	ETN + cDMARD	12	231	41.1	35.1	23.8	Y		Y
TEMPO	ETN + cDMARD	24	231	45.9	33.8	20.3	Y		Y
TOWARD	cDMARD	12	415	15.2	21.7	63.1	Y		Y
TOWARD	cDMARD	24	413	16.7	21.3	62	Y		Y
TOWARD	TCZ 8mg + cDMARD	12	805	31.6	27.8	40.6	Y		Y
TOWARD	TCZ 8mg + cDMARD	24	803	34.9	29.4	35.7	Y		Y
van de Putte (2004)	ADA 40mg	12	113	8.8	58.4	32.7	Y		-
van de Putte (2004)	ADA 40mg	26	113	8.8	55.8	35.4	Y		-
van de Putte (2004)	Placebo	12	110	0	24.5	75.5	Y		-
van de Putte (2004)	Placebo	26	110	3.6	26.4	70	Y		-

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
Weinblatt (1999)	cDMARD	12	30	16.7	30	53.3	Y		Y
Weinblatt (1999)	cDMARD	24	30	16.7	23.3	60	Y		Y
Weinblatt (1999)	ETN + cDMARD	12	60	40	26.7	33.3	Y		Y
Weinblatt (1999)	ETN + cDMARD	24	59	40.7	30.5	28.8	Y		Y

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

TNF-IR Network:

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
ATTAIN	ABA 10mg + cDMARD	12	-	-	-	-	N	Missing data	-
ATTAIN	ABA 10mg + cDMARD	24	256	27.3	29.7	43	Y		Y
ATTAIN	cDMARD	12	-	-	-	-	N	Missing data	-
ATTAIN	cDMARD	24	134	13.4	21.6	64.9	Y		Y
BEACON	BARI 2mg + cDMARD	12	174	12.6	49.4	37.9	Y		-
BEACON	BARI 2mg + cDMARD	24	174	11.5	39.1	49.4	Y		-
BEACON	BARI 4mg + cDMARD	12	177	11.9	55.4	32.8	Y		-
BEACON	BARI 4mg + cDMARD	24	177	16.9	38.4	44.6	Y		-
BEACON	cDMARD	12	176	4	37.5	58.5	Y		-
BEACON	cDMARD	24	176	7.4	27.3	65.3	Y		-
BREVACTA	cDMARD	24	47	17	14.9	68.1	Y		Y
BREVACTA	TCZ 162mg + cDMARD	24	89	29.2	25.8	44.9	Y		Y
GO-AFTER	cDMARD	14	154	13.6	19.5	66.9	Y		Y
GO-AFTER	cDMARD	24	154	12.3	20.1	67.5	Y		Y
GO-AFTER	GOL 50mg + cDMARD	14	153	20.9	25.5	53.6	Y		Y
GO-AFTER	GOL 50mg + cDMARD	24	153	20.9	24.2	54.9	Y		Y
Manders (2015)	ABA 10mg + cDMARD	13	-	-	-	-	N	Missing data	-
Manders (2015)	ABA 10mg + cDMARD	26	43	23.3	23.3	53.5	N	Request	-
Manders (2015)	RTX 1000mg + MTX	13	-	-	-	-	N	Missing data	-
Manders (2015)	RTX 1000mg + MTX	26	46	19.6	39.1	41.3	N	Request	-
Manders (2015)	TNF	13	-	-	-	-	N	Missing data	-
Manders (2015)	TNF	26	50	34	22	44	N	Request	-

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study	Treatment arm (as labelled in NMA)	Time-point (week)	Numbers analyzed	% with EULAR no response	% with EULAR moderate response	% with EULAR good response	Data used in NMA ?	Exclusion reason	EULAR converted from ACR?
ORAL Step	cDMARD	12	131	16.8	20.6	62.6	Y		Y
ORAL Step	TOFA 10mg + MTX	12	134	29.9	24.6	45.5	Y		Y
ORAL Step	TOFA 10mg + MTX	24	-	-	-	-	N	Missing data	-
ORAL Step	TOFA 5mg + MTX	12	133	26.3	24.1	49.6	Y		Y
ORAL Step	TOFA 5mg + MTX	24	-	-	-	-	N	Missing data	-
RADIATE	cDMARD	12	160	11.9	19.4	68.8	Y		Y
RADIATE	cDMARD	24	160	10.6	16.9	72.5	Y		Y
RADIATE	TCZ 8mg + MTX	12	175	26.9	24	49.1	Y		Y
RADIATE	TCZ 8mg + MTX	24	176	30.1	26.1	43.8	Y		Y
REFLEX	cDMARD	12	201	13.9	20.9	65.2	Y		Y
REFLEX	cDMARD	24	201	2	19.9	78.1	Y		-
REFLEX	RTX 1000mg + MTX	12	298	28.9	25.5	45.6	Y		Y
REFLEX	RTX 1000mg + MTX	24	298	15.1	50	34.9	Y		-

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Appendix 2- ACR and EULAR Node-splitting

Table A2.1 ACR20 Treatment Effects at 24 Weeks with and without Node-Splitting Random Relative Treatment Effect Model with Simultaneous Baseline Treatment Effect

Comparison	sigma.mean	sigma. median	MTC.mean	MTC.sd	DIR.mean	DIR.sd	INDIR .mean	INDIR.sd	diffDI. mean	diffDI.sd	P
1,4	0.388	0.381	1.19	0.235	0.722	0.856	1.08	0.433	-0.36	0.962	0.706
1,5	0.753	0.74	1.48	0.184	-0.449	0.454	0.647	1.04	-1.1	1.14	0.329
1,9	0.712	0.699	1.17	0.161	-0.235	0.349	1.69	1.75	-1.93	1.85	0.291
1,10	0.503	0.494	1.13	0.252	0.174	0.475	1.12	0.932	-0.942	1.05	0.366
1,15	0.258	0.253	0.732	0.264	1.78	0.577	0.477	0.436	1.3	0.788	0.0948
1,17	0.52	0.51	1.5	0.218	0.881	0.443	2.03	0.669	-1.15	0.671	0.0827
2,9	**0.346**	**0.339**	**-0.161**	**0.271**	**0.757**	**0.451**	**0.405**	**0.42**	**0.352**	**0.367**	**0.325**
4,5	0.31	0.304	0.289	0.257	1.45	0.816	0.919	0.384	0.529	0.897	0.55
4,6	0.351	0.344	-0.761	0.309	1.09	0.844	-1.8	0.783	2.89	1.15	0.0122
6,13	0.498	0.488	-1.61	0.308	0.191	0.626	-2.66	1.04	2.85	1.22	0.0198
7,10	0.345	0.338	-0.106	0.316	0.108	0.485	0.38	0.511	-0.272	0.512	0.577
9,22	0.338	0.33	0.189	0.285	-0.178	0.372	0.367	0.466	-0.545	0.528	0.288
9,23	0.343	0.337	-0.0594	0.283	-0.078	0.358	-0.218	0.433	0.139	0.451	0.75
10,17	0.341	0.335	0.374	0.323	0.973	0.93	0.285	0.379	0.689	1	0.488
13,15	0.525	0.514	1.91	0.362	-0.978	0.876	1.02	0.964	-2	1.31	0.123
15,17	0.421	0.412	0.767	0.275	0.843	0.532	2	0.673	-1.16	0.531	0.0331
18,19	0.35	0.343	0.354	0.434	0.732	0.588	0.597	0.628	0.136	0.649	0.838

posterior means (mean) and standard deviations (sd) of the Log Odds Ratios calculated using all the evidence (MTC) and when direct (DIR) and indirect (INDIR) evidence on each node is split and their difference with a 2-sided probability, P, measuring agreement between direct and indirect evidence for each split node

Treatment arm coding: 1 for Cdmard, **2 for Bari 4mg + Cdmard**, 4 for Tcz 8mg, 5 for Tcz 8mg + Cdmard, 6 for Ada 40mg, 7 for Aba 10mg + Cdmard**, 9 for Ada 40mg + Cdmard**, 10 for Ifx 3mg + Cdmard, 13 for Placebo, 15 for ETN, 17 for Etn + Cdmard, 18 for Rtx 1000mg, 19 for Rtx 1000mg + Cdmard, 22 for Tofa 10mg + Cdmard, 23 for Tofa 5mg + Cdmard.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table A2.2 ACR50 Treatment Effects at 24 Weeks with and without Node-Splitting Random Relative Treatment Effect Model with Simultaneous Baseline Treatment Effect

Comparison	sigma.mean	Sigma .median	MTC.mean	MTC.sd	DIR.mean	DIR.sd	INDIR. mean	INDIR.sd	diffDI. mean	diffDI.sd	P
1,4	0.276	0.271	1.16	0.22	0.948	0.831	1.26	0.348	-0.31	0.902	0.73
1,5	0.783	0.769	1.63	0.177	-0.53	0.46	0.694	1.08	-1.22	1.17	0.286
1,9	0.806	0.791	1.4	0.16	-0.262	0.369	1.76	2.03	-2.02	2.12	0.331
1,10	0.482	0.472	1.23	0.248	0.225	0.477	1.48	0.976	-1.25	1.09	0.245
1,14	0.223	0.22	0.877	0.25	1.11	0.531	0.893	0.407	0.214	0.72	0.759
1,16	0.532	0.521	1.48	0.2	0.703	0.461	1.18	0.672	-0.477	0.655	0.459
2,9	**0.298**	**0.292**	**0.0249**	**0.255**	**0.591**	**0.424**	**0.606**	**0.397**	**-0.0145**	**0.347**	**0.959**
4,5	0.224	0.222	0.466	0.243	1.22	0.829	1.25	0.324	-0.0352	0.89	0.971
4,6	0.343	0.336	-0.76	0.298	1.36	0.821	-1.39	0.835	2.75	1.18	0.0188
6,12	0.394	0.385	-1.54	0.352	0.107	0.632	-1.98	0.917	2.09	1.12	0.0583
7,10	0.296	0.29	-0.00433	0.303	0.122	0.461	0.445	0.484	-0.323	0.496	0.499
9,21	0.315	0.308	0.44	0.279	-0.00653	0.393	0.651	0.494	-0.657	0.549	0.226
9,22	0.309	0.303	0.171	0.277	0.134	0.374	-0.0587	0.437	0.193	0.433	0.642
10,16	0.297	0.29	0.254	0.309	0.688	0.929	0.135	0.355	0.553	0.995	0.572
12,14	0.446	0.436	2.02	0.391	-0.421	0.897	1.65	0.878	-2.07	1.26	0.0936
14,16	0.355	0.347	0.607	0.248	0.926	0.5	2.09	0.635	-1.16	0.478	0.019
17,18	0.3	0.293	0.32	0.43	0.729	0.619	0.319	0.65	0.411	0.7	0.556

posterior means (mean) and standard deviations (sd) of the Log Odds Ratios calculated using all the evidence (MTC) and when direct (DIR) and indirect (INDIR) evidence on each node is split and their difference with a 2-sided probability, P, measuring agreement between direct and indirect evidence for each split node

Treatment arm coding: 1 for Cdmard, **2 for Bari 4mg + Cdmard**, 4 for Tcz 8mg, 5 for Tcz 8mg + Cdmard, 6 for Ada 40mg, 7 for Aba 10mg + Cdmard, **9 for Ada 40mg + Cdmard**, 10 for Ifx 3mg + Cdmard, 12 for Placebo, 14 for ETN, 16 for Etn + Cdmard, 17 for Rtx 1000mg, 18 for Rtx 1000mg + Cdmard, 21 for Tofa 10mg + Cdmard, 22 for Tofa 5mg + Cdmard.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table A2.3 ACR70 Treatment Effects at 24 Weeks with and without Node-Splitting Random Relative Treatment Effect Model with Simultaneous Baseline Treatment Effect

Comparison	sigma.mean	sigma. median	MTC.mean	MTC.sd	DIR.mean	DIR.sd	INDIR. mean	INDIR.sd	diffDI. mean	diffDI.sd	P
1,4	0.245	0.242	1.55	0.213	1.16	1.02	1.98	0.376	-0.821	1.08	0.439
1,5	0.884	0.868	1.83	0.19	-0.602	0.565	0.953	1.24	-1.55	1.37	0.243
1,9	0.784	0.768	1.45	0.17	-0.0995	0.437	1.43	2.25	-1.53	2.35	0.513
1,10	0.433	0.424	1.26	0.258	0.683	0.522	1.75	1.59	-1.06	1.68	0.533
1,14	0.232	0.231	0.813	0.261	0.448	0.634	0.656	0.496	-0.208	0.857	0.806
1,16	0.529	0.518	1.53	0.207	0.745	0.563	1.02	0.815	-0.271	0.734	0.704
2,9	**0.266**	**0.26**	**-0.206**	**0.24**	**0.391**	**0.459**	**0.492**	**0.443**	**-0.101**	**0.386**	**0.778**
4,5	0.103	0.0887	0.279	0.228	1.62	0.974	1.2	0.316	0.413	1.02	0.673
4,6	0.3	0.295	-0.703	0.296	1.82	0.95	-0.621	1.11	2.44	1.47	0.0955
6,12	0.351	0.343	-2.44	0.6	0.278	0.743	-2.23	1.03	2.51	1.28	0.0433
7,10	0.25	0.247	0.074	0.296	0.206	0.523	0.112	0.532	0.0941	0.565	0.87
9,21	0.365	0.357	0.963	0.299	0.498	0.559	1.25	0.684	-0.751	0.72	0.286
9,22	0.32	0.314	0.596	0.302	0.537	0.521	0.565	0.585	-0.0283	0.508	0.953
10,16	0.232	0.229	0.275	0.324	-0.24	1.28	0.0709	0.359	-0.311	1.33	0.823
12,14	0.292	0.288	2.41	0.623	-0.0423	1.02	2.09	0.932	-2.13	1.39	0.118
14,16	0.387	0.379	0.721	0.242	1.29	0.613	2.06	0.783	-0.774	0.566	0.165
17,18	0.227	0.224	0.166	0.518	0.889	0.803	-0.346	0.836	1.23	0.911	0.168

posterior means (mean) and standard deviations (sd) of the Log Odds Ratios calculated using all the evidence (MTC) and when direct (DIR) and indirect (INDIR) evidence on each node is split and their difference with a 2-sided probability, P, measuring agreement between direct and indirect evidence for each split node

Treatment arm coding: 1 for Cdmard, 2 for Bari 4mg + Cdmard, 4 for Tcz 8mg, 5 for Tcz 8mg + Cdmard, 6 for Ada 40mg, 7 for Aba 10mg + Cdmard, 9 for Ada 40mg + Cdmard, 10 for Ifx 3mg + Cdmard, 12 for Placebo, 14 for ETN, 16 for Etn + Cdmard, 17 for Rtx 1000mg, 18 for Rtx 1000mg + Cdmard, 21 for Tofa 10mg + Cdmard, 22 for Tofa 5mg + Cdmard.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table A2.4 EULAR No Reponse Treatment Effects at 24 Weeks with and without Node-Splitting Random Relative Treatment Effect Model with Simultaneous Baseline Treatment Effect

Comparison	sigma.mean	sigma. median	MTC.mean	MTC.sd	DIR.mean	DIR.sd	INDIR .mean	INDIR.sd	diffDI. mean	diffDI.sd	P
1,4	0.46	0.45	-1.33	0.25	-0.713	0.695	-1.04	0.512	0.326	0.866	0.705
1,5	0.809	0.795	-1.6	0.215	0.435	0.356	-1.34	1.14	1.77	1.19	0.136
1,9	0.63	0.617	-0.863	0.174	0.235	0.287	-1.08	1.55	1.31	1.62	0.41
1,10	0.482	0.472	-0.815	0.277	-0.0409	0.415	-1.16	0.905	1.12	1	0.259
1,14	0.33	0.322	-0.572	0.264	-1.56	0.578	0.181	0.503	-1.74	0.831	0.0362
1,16	0.545	0.535	-1.34	0.228	-0.511	0.38	-1.83	0.696	1.32	0.709	0.0579
2,9	**0.409**	**0.399**	**0.415**	**0.31**	**-0.843**	**0.469**	**-0.0524**	**0.459**	**-0.791**	**0.418**	**0.0551**
4,5	0.406	0.397	-0.269	0.294	-1.95	0.624	-0.61	0.476	-1.34	0.784	0.0911
4,6	0.468	0.458	0.899	0.33	-1.18	0.68	1.19	0.909	-2.37	1.14	0.0367
6,12	0.538	0.527	1.17	0.313	-0.24	0.507	1.45	1.1	-1.69	1.21	0.159
7,10	0.411	0.401	0.0607	0.356	-0.219	0.476	-0.345	0.56	0.126	0.556	0.812
9,21	0.399	0.389	-0.0907	0.317	0.194	0.383	-0.133	0.505	0.327	0.569	0.549
9,22	0.4	0.391	0.1	0.318	0.161	0.363	0.267	0.473	-0.105	0.493	0.826
10,16	0.4	0.39	-0.521	0.347	-0.52	0.799	-0.505	0.421	-0.0147	0.901	0.984
12,14	0.48	0.471	-1.32	0.354	0.667	0.71	-1.06	0.903	1.73	1.15	0.133
14,16	0.434	0.424	-0.763	0.291	-0.95	0.437	-2.09	0.639	1.14	0.542	0.0404
17,18	0.411	0.402	-0.0599	0.472	-0.177	0.574	-0.231	0.672	0.0544	0.689	0.932

posterior means (mean) and standard deviations (sd) of the Log Odds Ratios calculated using all the evidence (MTC) and when direct (DIR) and indirect (INDIR) evidence on each node is split and their difference with a 2-sided probability, P, measuring agreement between direct and indirect evidence for each split node

Treatment arm coding: 1 for Cdmard, **2 for Bari 4mg + Cdmard**, 4 for Tcz 8mg, 5 for Tcz 8mg + Cdmard, 6 for Ada 40mg, 7 for Aba 10mg + Cdmard, **9 for Ada 40mg + Cdmard**, 10 for Ifx 3mg + Cdmard, 12 for Placebo, 14 for ETN, 16 for Etn + Cdmard, 17 for Rtx 1000mg, 18 for Rtx 1000mg + Cdmard, 21 for Tofa 10mg + Cdmard, 22 for Tofa 5mg + Cdmard.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table A2.5 EULAR Moderate Reponse Treatment Effects at 24 Weeks with and without Node-Splitting Random Relative Treatment Effect Model with Simultaneous Baseline Treatment Effect

Comparison	sigma.mean	sigma. median	MTC.mean	MTC.sd	DIR.mean	DIR.sd	INDIR. mean	INDIR.sd	diffDI. mean	diffDI.sd	P
1,4	0.0559	0.0473	0.19	0.125	0.908	0.392	0.541	0.201	0.367	0.441	0.397
1,5	0.116	0.114	0.24	0.0948	0.077	0.234	-0.415	0.3	0.492	0.381	0.193
1,9	0.0918	0.0858	0.142	0.0926	-0.114	0.157	0.179	0.643	-0.293	0.687	0.671
1,10	0.0969	0.0918	0.261	0.146	0.0672	0.233	0.158	0.732	-0.0905	0.772	0.932
1,14	0.0784	0.0709	0.032	0.158	0.0896	0.345	-0.32	0.298	0.41	0.492	0.396
1,16	0.0786	0.0712	0.119	0.128	0.374	0.228	0.585	0.368	-0.211	0.342	0.533
2,9	**0.11**	**0.107**	**-0.251**	**0.132**	**0.383**	**0.239**	**0.17**	**0.237**	**0.213**	**0.201**	**0.278**
4,5	0.0602	0.0514	0.0494	0.136	0.228	0.436	0.445	0.198	-0.217	0.481	0.643
4,6	0.0929	0.087	0.401	0.193	0.149	0.436	0.362	0.497	-0.213	0.67	0.751
6,12	0.16	0.158	-0.971	0.21	0.309	0.319	-0.908	0.573	1.22	0.66	0.0611
7,10	0.086	0.0794	0.00937	0.187	0.0456	0.278	0.0216	0.27	0.024	0.312	0.941
9,21	0.0849	0.0773	0.0943	0.177	-0.056	0.239	0.049	0.276	-0.105	0.324	0.741
9,22	0.0851	0.078	0.00748	0.176	-0.0938	0.228	-0.0365	0.254	-0.0573	0.269	0.837
10,16	0.0831	0.0766	-0.142	0.191	-0.103	0.587	-0.205	0.21	0.102	0.623	0.865
12,14	0.085	0.0783	0.412	0.248	-0.378	0.453	0.168	0.418	-0.546	0.62	0.375
14,16	0.0813	0.074	0.0868	0.162	0.356	0.27	0.19	0.345	0.165	0.266	0.531
17,18	0.0824	0.075	0.0921	0.315	0.00506	0.409	0.26	0.392	-0.255	0.471	0.591

posterior means (mean) and standard deviations (sd) of the Log Odds Ratios calculated using all the evidence (MTC) and when direct (DIR) and indirect (INDIR) evidence on each node is split and their difference with a 2-sided probability, P, measuring agreement between direct and indirect evidence for each split node

Treatment arm coding: 1 for Cdmard, **2 for Bari 4mg + Cdmard**, 4 for Tcz 8mg, 5 for Tcz 8mg + Cdmard, 6 for Ada 40mg, 7 for Aba 10mg + Cdmard, **9 for Ada 40mg + Cdmard**, 10 for Ifx 3mg + Cdmard, 12 for Placebo, 14 for ETN, 16 for Etn + Cdmard, 17 for Rtx 1000mg, 18 for Rtx 1000mg + Cdmard, 21 for Tofa 10mg + Cdmard, 22 for Tofa 5mg + Cdmard.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table A2.6 EULAR Good Reponse Treatment Effects at 24 Weeks with and without Node-Splitting Random Relative Treatment Effect Model with Simultaneous Baseline Treatment Effect

Comparison	sigma.mean	sigma. median	MTC.mean	MTC.sd	DIR.mean	DIR.sd	INDIR. mean	INDIR.sd	diffDI. mean	diffDI.sd	P
1,4	0.544	0.534	1.35	0.247	-0.235	0.761	1.26	0.57	-1.5	0.955	0.114
1,5	0.879	0.863	1.65	0.213	-0.728	0.4	1.71	1.21	-2.44	1.28	0.057
1,9	0.688	0.675	1.01	0.181	-0.0486	0.33	1.05	1.72	-1.1	1.81	0.536
1,10	0.513	0.503	0.803	0.287	0.179	0.44	0.929	0.951	-0.749	1.05	0.479
1,14	0.398	0.39	0.703	0.264	1.42	0.586	0.555	0.518	0.862	0.824	0.287
1,16	0.565	0.555	1.15	0.222	0.349	0.422	0.99	0.674	-0.641	0.67	0.331
2,9	**0.432**	**0.423**	**-0.217**	**0.316**	**0.393**	**0.485**	**-0.42**	**0.487**	**0.813**	**0.442**	**0.0675**
4,5	0.467	0.458	0.293	0.282	1.55	0.703	0.435	0.542	1.11	0.888	0.209
4,6	0.553	0.543	-1.37	0.346	1.4	0.724	-1.37	1.08	2.78	1.3	0.0317
6,12	0.499	0.489	-0.862	0.382	-0.441	0.566	-0.962	1.04	0.521	1.2	0.668
7,10	0.472	0.463	-0.0106	0.358	0.327	0.524	0.491	0.63	-0.164	0.624	0.783
9,21	0.467	0.457	0.00569	0.325	-0.0867	0.443	0.0744	0.589	-0.161	0.666	0.797
9,22	0.469	0.459	-0.121	0.325	-0.0304	0.424	-0.254	0.553	0.224	0.571	0.691
10,16	0.464	0.455	0.345	0.35	0.83	0.851	0.345	0.475	0.484	0.977	0.612
12,14	0.528	0.518	1.58	0.4	-0.297	0.791	1.47	0.997	-1.76	1.28	0.162
14,16	0.425	0.417	0.444	0.279	0.776	0.461	2.04	0.655	-1.27	0.533	0.0223
17,18	0.471	0.462	0.157	0.476	0.435	0.606	0.627	0.796	-0.192	0.818	0.814

posterior means (mean) and standard deviations (sd) of the Log Odds Ratios calculated using all the evidence (MTC) and when direct (DIR) and indirect (INDIR) evidence on each node is split and their difference with a 2-sided probability, P, measuring agreement between direct and indirect evidence for each split node

Treatment arm coding: 1 for Cdmard, **2 for Bari 4mg + Cdmard,** 4 for Tcz 8mg, 5 for Tcz 8mg + Cdmard, 6 for Ada 40mg, 7 for Aba 10mg + Cdmard, **9 for Ada 40mg + Cdmard**, 10 for Ifx 3mg + Cdmard, 12 for Placebo, 14 for ETN, 16 for Etn + Cdmard, 17 for Rtx 1000mg, 18 for Rtx 1000mg + Cdmard, 21 for Tofa 10mg + Cdmard, 22 for Tofa 5mg + Cdmard.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Baricitinib for treating moderate to severe rheumatoid arthritis [ID979]

Thank you for agreeing to give us a statement on your view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name:

Dr Christopher Holroyd

Name of your organisation

[redacted]

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? √
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)? √
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)?
- other? (please specify)

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry:

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

Rheumatoid arthritis is a chronic incurable inflammatory arthritis that causes joint pain, swelling and stiffness. If inadequately treated will result in irreversible joint damage, deformity and a significant impact on quality of life.

At present rheumatoid arthritis is treated in line with NICE guideline (CG79). Most rheumatology departments adopt a very similar approach. Patients are treated promptly at diagnosis with disease modifying anti-rheumatic drugs (DMARDs) usually including methotrexate, following a treat to target approach, with regular review and assessment of disease activity. DMARD medication is then titrated according to patient's response with an aim of achieving disease remission; i.e. their disease is so well controlled that they have minimal tender and swollen joints and are able to enjoy a high quality of life without disease progression.

Conventional DMARD therapy however is insufficient in a significant proportion of patients to achieve adequate disease control. These patients may be eligible for targeted biological therapies (such as anti-TNF agents) if they fulfil criteria defined by NICE (NICE TA375). At present, there are 7 biologics approved by NICE as a first-line therapy after conventional DMARD failure (infliximab, etanercept, adalimumab, certolizumab pegol, golimumab, tocilizumab and abatacept). If patients do not respond to an initial biologic, they are eligible to be switched to an alternative biological therapy (NICE TA195).

Although it is apparent that a considerable number of patients do not respond adequately to their first biologic, there are few tools available to predict response, or

to help the clinical decision about which biologic to be used first in a particular patient. It is well recognisde that patients may respond better to different classes of biologic; the reasons behind this are not clear. This is presently an area of research interest.

Baricitinib is a tyrosine kinase protein that inhibits JAK1 and JAK2. It has been shown in clinical trials to be effective in treating RA patients across a variety of scenarios (with methotrexate, without methotrexate, after DMARD failure, after anti-TNF failure) and has shown superiority to methotrexate and similar results to those observed with currently available biologics. It has also been compared against adalimumab in a direct head to head RCT, and achieved superiority in some but not all endpoints. Baricinib is novel, in the sense that there are no other drugs with this mechanism of action available in the UK.

The current alternatives to baricitinib are the other available biological therapies (infliximab, etanercept, adalimumab, certolizumab pegol, golimumab, abatacept, tocilizumab and rituximab). All of these are delivered either subcutaneously via a pre-filled syringe or pen, or via an intra-venous infusion. In contrast, baricitinib is an oral medication hence more suitable in cohorts of patients who are needle phobic.

As with other advanced therapies for RA, baricitinib should be prescribed via secondary care for patients with RA in line with NICE TA 375 and 195. As all rheumatology departments in the country will be familiar with the use of advanced therapies in RA, no specific additional professional input will be needed.

This technology is currently unavailable in the UK

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

Baricitinib has demonstrated similar efficacy to other NICE approved biological therapies for RA, and possible superiority (at some, but not all endpoints) to adalimumab. I do not foresee the technology being any more difficult to use than the current alternatives, and may even be preferred by patients and easier to use, as it is an oral medication rather than subcutaneous or intravenous (of which all the other biologics are).

As this is an advanced therapy for RA, I expect that the pre-defined "rules" for commencing a biologic in this scenario will be applicable here (as defined in NICE TA 375)

Baricitinib has undergone several studies of its efficacy and safety in patients with RA, across a broad range of scenarios. Clinical trial patients are often a far more homogenous group of patients than those encountered in a real life setting, however the patients included in the trials are broadly reflective of those seen in the UK, and comparable to those included in previous controlled trails for other NICE approved advanced therapies.

Important outcomes in RA are disease activity, disease remission, ACR (American College of Rheumatology) improvement scores, disability (HAQ) scores, joint

damage scores and safety utcomes. These measures have been included in the baricitinib trials.

With regards to adverse events, now new safety signals compared to the other commonly used biologics have been encountered, however it is crucial that real world data from registries (such as BSB-BR) is collected for patients who receive this therapy.

Equality and Diversity

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that this appraisal:

- Could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed;
- Could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology;
- Could lead to recommendations that have any adverse impact on people with a particular disability or disabilities

Please tell us what evidence should be obtained to enable the Committee to identify and consider such impacts

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

I do not believe that this appraisal has any specific issues with regards to equality and diversity.

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

I have no other evidence to report that is of significance to this appraisal.

Implementation issues

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

I expect that this technology would be treated in line with the other NICE approved biological therapies, which are commonly prescribed in the UK. As all rheumatology departments are very used to this line of therapy, no extra staff education or training would be needed, outside of the relevant drug information. No extra equipment or facilities would be required.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Baricitinib for treating moderate to severe rheumatoid arthritis [ID979]

Thank you for agreeing to give us a statement on your view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: Professor Peter C. Taylor

Name of your organisation: [redacted]

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? ✓
- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)? ✓
- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)?
- other? (please specify)

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry:

None

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

Rheumatoid Arthritis (RA) is managed in the NHS according to national (NICE CG79[1]) and international (EULAR2013[2] and EULAR2016 update[3]) guidelines as detailed in the relevant section below. Treatment of RA is aimed at controlling inflammatory disease activity as early and optimally as possible, initially with csDMARDs and bridging steroid as required. For patients failing to achieve the desirable target disease activity of remission, or where that is not possible, low disease activity (DAS28≤3.2), and where poor prognostic factors are present (such as persistent high disease activity, imaging evidence of joint erosions or seropositivity for rheumatoid factor or anti-CCP antibodies, especially at high titre) the EULAR 2013 recommendations advocate addition of a bDMARD. In the NHS, access to bDMARD is set at a higher threshold of disease activity (DAS28 of 5.1), in accordance with the relevant NICE Technology Appraisal (e.g. MTA375, TA195, TA 225, TA247).
Of note, the EULAR 2016 update recommendations suggest that in poor prognosis patients, after initial failure of csDMARDs, that a JAK inhibitor may be considered as an alternative to a bDMARD.
The recommendations of NICE MTA/TAs have resulted in there being little variation in overall access to bDMARDs, However, local and regional pathways have been developed between rheumatologists and commissioners and the choice of first bDMARD has often been influenced by the recently emerging availability of biosimilars.
Currently, in UK practice, patients who fail to respond adequately to their initial bDMARD, which will often be an anti-TNF therapy, should next receive rituximab in combination with methotrexate. But if either rituximab or methotrexate are contra-indicated, a second anti-TNF or alternative mechanism of action bDMARD may be prescribed.
The EULAR 2013 recommendations state that tofacitinib (the only Jak inhibitor available in 2013, although not approved in most of Europe at that time) may be considered after a bDMARD treatment has failed. The 2016 update of the EULAR recommendations recommend that a Jak inhibitor (such as the technology under appraisal) may be considered as an option after failure of a bDMARD.

Management of RA takes place largely within the secondary care setting using a multidisciplinary team approach and supervised by a consultant rheumatologist (although in some UK circumstances the consultant may be community-based).

The technology under appraisal is not currently available in the UK. A major advantage of this technology is that it is an oral therapy with convenient once daily dosing. It also has a fast kinetic of clinical response.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology?

RA is a highly heterogeneous in its presentation, disease course and impact on aspects of life that matter to the individual. Thus, it is not readily possible to describe a "typical" patient. However, poorer prognosis patients can be identified in clinical practice by the presence of various factors that include persistently high disease activity, persistently high systemic inflammation as assessed by acute phase response markers, the presence on imaging of joint damage/erosions early in the disease course, seropositive status, high levels of functional loss, and inadequate response to csDMARD treatment.
There is no evidence that I am aware of to indicate that baricitinib offers greater or less therapeutic benefit in any patient subgroup (by for example, disease severity or antibody status). There appears to be a greater risk of herpes zoster infection following baricitinib exposure in some Asian populations, notably Japanese and Korean, than in the case of white Caucasian subjects.

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)?

Patients receiving baricitinib will need to do so under the supervision of a consultant rheumatologist. This will normally be in a secondary care setting which has the support of clinical nurse specialists in rheumatology although some consultant rheumatologists provide community based care. It is not anticipated that any service requirements additional to those currently offered would be necessary to support this technology.

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur?

The technology is not currently available in the NHS.

Please tell us about any relevant *clinical guidelines* and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

RA is managed in the NHS according to national and international guidelines. The national guidelines include NICE CG79, originally posted in 2009 and last updated in December 2015[1], and the European League Against Rheumatism (EULAR) recommendations for the management of RA were last published in 2013[2] and updated in 2016 (manuscript in press). The draft revised EULAR guidelines were presented at the 2016 EULAR meeting[3]. These recommendations were based on systematic literature reviews and focused on indications for the use of, and suggestions for, differential and strategic employment of conventional synthetic

disease-modifying anti-rheumatic drugs (csDMARDs) and biologic disease-modifying anti-rheumatic drugs (bDMARDs) based on treatment targets, disease risk assessment, safety aspects and contraindications.

1. NICE Clinical Guideline 79. NICE. 2009. https://www.nice.org.uk/guidance/cg79
2. Smolen JS, Landewé R, et al. Ann Rheum Dis; doi: 10.1136/annrheumdis-2013-204573.
3. Smolen JS, Landewé R, et al. http://www.eular.org/myUploadData/files/EULAR%20RA%20Management%20recommendations%202016%20update%20June%202016-c_2.pdf

THE ADVANTAGES AND DISADVANTAGES OF THE TECHNOLOGY

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

Baricitinib is a low molecular weight synthetic DMARD which is orally available and delivered as a once daily tablet. Clinical trial data suggests that the efficacy of baricitinib taken with concomitant once weekly methotrexate is similar to or greater than that of subcutaneously administered adalimumab taken with concomitant once weekly methotrexate. In early phase RA it has been shown that baricitinib can also be used as an efficacious monotherapy[4]. This is advantageous as many people with RA struggle to tolerate methotrexate and poor adherence of this drug is well recognised. As an orally available drug, baricitinib is anticipated to be easier to use than parenterally (iv or sc) administered bDMARDs for many patients. The ease of use of a once daily oral tablet is also anticipated to carry a lower administrative burden than is required to train an individual to self-administer a subcutaneous injection or to deliver an intravenous injection with the associated facilities that the latter necessitates. It is not anticipated that any additional tests will be required over and above those routinely ordered for patients on csDMARDs.

4. Fleischmann R et al, Baricitinib, Methotrexate, or Combination in Patients with Rheumatoid Arthritis and No or Limited Prior Disease-Modifying Antirheumatic Drug Treatment. Arthritis & Rheumatology 2016 Oct 9. doi: 10.1002/art.39953. [Epub ahead of print]

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

There are as yet no established rules for starting or stopping the use of this technology. However, based on the efficacy data from clinical trials and EULAR recommendations[3], baricitinib could potentially be started in patients with active

disease and an inadequate response to csDMARDs after 6 months of treatment or in patients' refractory to a first or subsequent bDMARD. In an eligible patient population, it would be expected that baricitinib would be stopped in patients who do not exhibit a good clinical response within 3-6 months or if a clinical response is not maintained. In the absence of data regarding the safety of baricitinib in pregnancy, the drug would be discontinued prior to a planned conception and consideration might be given to temporary discontinuation prior to planned surgery.

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

Baricitinib has undergone an extensive phase III trial programme. Key studies in support of this STA include the RA-BEAM study[5] which compared baricitinib 4mg od with placebo and with adalimumab 40mg every other week in patients with active RA despite background methotrexate. In the RA-BUILD study[6], patients with active RA and an insufficient response (despite prior therapy) or intolerance to ≥1 csDMARDs were assigned 1:1:1 to placebo or baricitinib (2 or 4 mg) once daily. Most of these patients were receiving background methotrexate, either alone (49%) or in combination with another csDMARD (23%). Approximately, 16% were receiving a single non-MTX csDMARD. Thus the patient populations investigated in RA-BEAM and RA-BUILD are reflective of UK patient populations prior to bDMARD treatment. It is commonly the case in UK practice that patients will be treated with combination csDMARD treatment prior to consideration of a bDMARD.
In the RA-BEACON study[7], baricitinib at 2mg od or 4mg od was compared with placebo in patients with active RA who were refractory to bDMARDs including anti-TNFs. At study entry, patients had been taking one or more csDMARDs regularly for at least the preceding 12 weeks, with stable doses for at least the preceding 8 weeks. This population is broadly reflective of the bDMARD refractory patients found in current UK practice with respect to demographics, disease activity, patient reported outcome measures, other clinical indices and functional deficit at baseline.
For each of these three pivotal studies, the primary endpoint was ACR20 response at week 12. The categorical ACR20 response metric is widely used in RA clinical trials and will thus be familiar to UK rheumatologists. It is not a measure used in routine clinical practice, but it is the case that a clinical assessment at around 3 months after initiating a targeted therapy reflects clinical practice. Assessment of response in such a routine setting is usually be based on a composite score of disease activity of which one of the variants of DAS28 (for which there are different formulae based on use of either ESR or CRP as an acute phase response measure, and with or without inclusion of the patient global health assessment visual analogue scale). In all three pivotal phase III trials, DAS28CRP change at week 12 was included as a secondary endpoint, confirming clinical efficacy with a metric familiar to rheumatologists in routine clinical practice. Early clinical response has been shown to be predictive of long term response to baricitinib[8].
Other outcome measures important to physicians include those indicative of long term inhibition of structural damage to joints and preservation of function. The RA-

BEAM study in methotrexate refractory RA and RA-BUILD study in csDMARD refractory RA patients also demonstrated significant inhibition of structural damage to joints at 6 months. This information is of great importance to rheumatologists but formal assessment of radiographic structural damage inhibition is not routinely measured at 6 months of treatment intervention. Change in physical function, assessed by HAQ-DI, was also measured as a major secondary endpoint after 12 weeks of treatment in both the RA-BEAM and RA-BUILD studies and in patients' refractory to bDMARD in RA-BEACON.
And of importance to patients and their physicians are patient reported outcomes that assess the impact of various aspects of life that are important to people with RA, including pain and the length and severity of early morning joint stiffness. There is a considerable body of data that demonstrates statistically significant and clinically meaningful improvements in these parameters[9,10,11].

5. Taylor P C et al.; presented at ACR 2015. Abstract 2L. Baricitinib Versus Placebo or Adalimumab in Patients with Active Rheumatoid Arthritis (RA) and an Inadequate Response to background Methotrexate Therapy: Results of a Phase 3 Study. (full manuscript in press).
6. Dougados M et al., Baricitinib in patients with inadequate response or intolerance to conventional synthetic DMARDs: results from the RA-BUILD study. Ann Rheum Dis 2017 Jan;76(1):88-95.
7. Genovese M C et al.; Baricitinib in Patients with Refractory Rheumatoid Arthritis. N Engl J Med. 2016 Mar 31;374(13):1243-52.
8. Weinblatt M et al; Response to baricitinib at 4 weeks predicts response at 12 and 24 weeks in patients with rheumatoid arthritis: results from two phase 3 studies. presented at EULAR 2016. Ann Rheum Dis 75(Suppl 2):255.2-256
9. Keystone E et al. Patient-reported outcomes from a phase 3 study of baricitinib versus placebo or adalimumab in patients with active rheumatoid arthritis and an inadequate response to background methotrexate therapy. Presented at EULAR 2016. *Ann Rheum Dis* 2016;75:412-413 doi:10.1136/annrheumdis-2016-eular.1239
10. Taylor PC et al; presented at ACR2016. Abstract 1599. Speed of Onset of Effect on Patient-Reported Outcomes Assessed through Daily Electronic Patient Diaries in the Baricitinib Phase 3 RA Clinical Program.
11. Smolen JS et al. Patient-reported outcomes from a randomised phase III study of baricitinib in patients with rheumatoid arthritis and an inadequate response to biological agents (RA-BEACON). Ann Rheum Dis. 2016 Oct 31. pii: annrheumdis-2016-209821

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

The overall benefit:risk profile of baricitib and positive impact on overall quality of life to have emerged from clinical trials is favourable and broadly comparable to that observed with bDMARDs taken as a whole. One particular infectious adverse event of note observed in international clinical trials of baricitinib (and other Jak inhibitors) is herpes zoster although most cases were monodermatomal and non-serious. There is also an increased risk of herpes zoster with anti-TNF monoclonal antibodies. But in the case of baricitinib, the risk seems to be greatest in subpopulations of people of

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Asian descent (Japanese and Korean). Depending on findings to emerge from trials of vaccination against zoster, a possible outcome is that vaccination might be recommended for those patients to be given baricitinib who are considered to be at risk of zoster.

Equality and Diversity

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that this appraisal:

- **Could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed;**
- **Could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology;**
- **Could lead to recommendations that have any adverse impact on people with a particular disability or disabilities**

Please tell us what evidence should be obtained to enable the Committee to identify and consider such impacts

I am not aware of any impact this appraisal might have on people protected by the equality legislation who might fall within the patient population for which this treatment will be licensed.

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

Of the references cited in the text above, the following have been presented at international meetings and published in abstract form:

5. Taylor P C et al.; presented at ACR 2015. Abstract 2L. Baricitinib Versus Placebo or Adalimumab in Patients with Active Rheumatoid Arthritis (RA) and an Inadequate Response to background Methotrexate Therapy: Results of a Phase 3 Study. (full manuscript in press).

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

8. Weinblatt M et al; Response to baricitinib at 4 weeks predicts response at 12 and 24 weeks in patients with rheumatoid arthritis: results from two phase 3 studies. presented at EULAR 2016. Ann Rheum Dis 75(Suppl 2):255.2-256
9. Keystone E et al. Patient-reported outcomes from a phase 3 study of baricitinib versus placebo or adalimumab in patients with active rheumatoid arthritis and an inadequate response to background methotrexate therapy. Presented at EULAR 2016. *Ann Rheum Dis* 2016;**75**:412-413 doi:10.1136/annrheumdis-2016-eular.1239
10. Taylor PC et al; presented at ACR2016. Abstract 1599. Speed of Onset of Effect on Patient-Reported Outcomes Assessed through Daily Electronic Patient Diaries in the Baricitinib Phase 3 RA Clinical Program.

IMPLEMENTATION ISSUES

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within
3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

If this STA is approved, NICE guidance is not anticipated to impose any new burden on NHS resources such as the need for additional staff or training. Rather, it is anticipated that over time, it might favourably impact on the resources required to train people with RA to administer subcutaneous injections and administrative burdens associated with bDMARD prescribing and delivery to patients.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Patient/carer organisation submission (STA)

Baricitinib for treating moderate to severe rheumatoid arthritis [ID979]

Thank you for agreeing to give us your views on this treatment that is being appraised by NICE and how it could be used in the NHS. Patients, carers and patient organisations can provide a unique perspective on conditions and their treatment that is not typically available from other sources. We are interested in hearing about:

- the experience of having the condition or caring for someone with the condition
- the experience of receiving NHS care for the condition
- the experience of having specific treatments for the condition
- the outcomes of treatment that are important to patients or carers (which might differ from those measured in clinical studies, and including health-related quality of life)
- the acceptability of different treatments and how they are given
- expectations about the risks and benefits of the treatment.

To help you give your views, we have provided a questionnaire. You do not have to answer every question — the questions are there as prompts to guide you. The length of your response should not normally exceed 10 pages.

1. About you and your organisation

Your name: Ailsa Bosworth

Name of your organisation: National Rheumatoid Arthritis Society

Your position in the organisation: CEO

Brief description of the organisation: We are the national organisation representing people with RA and children and young people and their families living with JIA. We also support the health profiessionals who treat those with RA and JIA.

(For example: who funds the organisation? How many members does the organisation have?)

We have approx 5,500 members including health professional members. We have a wide range of income streams with the majority of our funding coming from grant-giving trusts and foundations, events, legacy income and a maximum cap which we impose of 15% of annual income comes from projects funded by pharmaceutical industry, although to date such funding has never reached as much as 15%.

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: None

2. Living with the condition

What is it like to live with the condition or what do carers experience when caring for someone with the condition?

Being diagnosed with an incurable, painful disease like RA can be extremely distressing as it is life-changing and as you can be diagnosed at any age post 16, it can have a major impact on your future life plans, dreams and aspirations, although being diagnosed today has significantly better potential outcomes than when I was diagnosed over 35 years ago when treatments and the way the disease was treated were quite different. RA impacts on every area of life and impacts both physical and emotional wellbeing. Health beliefs, how you come to diagnosis (how long it takes to be diagnosed), the network of support you have and how aggressive the disease is will all impact on how you come to terms with your diagnosis and cope day to day. It can be very

distressing for a partner of someone with RA to witness their loved-one in severe pain and suffering the debilitating effects of fatigue and so this disease does very much impact on the whole family. As ¾ of people are diagnosed when of working age, anxiety over job-loss due to their disease is a significant factor and whilst we are making steps towards seeing work as a health outcome, we are far from a situation where rheumatology teams pay enough attention to how worried patients may be about their job particularly at time of diagnosis when they may have already had quite a lot of time off work in the process of finding out what is wrong and may already be at risk of losing their job. For young people who are not yet in a permanent relationship, it can be very hard to come to terms with the fact that they have a long term condition and we know from our own research that RA can have a huge impact, making them feel less desirable, much less confident and worried that they will not find a partner. For older people diagnosed as they approach retirement for example, dreams of being able to travel and look after grand-children can suddenly seem unachievable. Diagnosed in mid-years with young children to care for can also be incredibly challenging. Imagine not being able to pick up your baby and change its nappy. For whilst much has been done in terms of new and innovative therapies coming into rheumatology and the way in which we now treat the disease, there remains a lot of pain and distress at all stages of this disease.

3. Current practice in treating the condition

Which treatment outcomes are important to patients or carers? (That is, what would patients or carers like treatment to achieve?) Which of these are most important? If possible, please explain why.

People simply want their life back. They want a reduction in pain, want to prevent permanent disability, want reduction in fatigue, and above all want to maintain independence and ability to work, if of working age, and carry out all the normal activities of daily living. Side effects of some drugs can be quite debilitating, however, by comparison to methotrexate for example, side effects from biologics are generally fewer in our experience. In my own experience and also listening to many thousands of people over the last 15.5 years

running NRAS, one of the most important things people want is to be able to maintain their independence. Pain and fatigue are the two most common symptoms and therefore the most major barriers to being able to live independently and without having to rely heavily on others for a myriad of things.

What is your organisation's experience of currently available NHS care and of specific treatments for the condition? How acceptable are these treatments and which are preferred and why?

One of the key issues associated with current care is the variability of access to best, evidence-based care and access to all the relevant members of a consultant-led multi-disciplinary team. This has been demonstrated in the past by the Kings Fund and National Audit Office reports into services for people with RA and most recently by the 3-year audit results from the HQIP audit into early RA. People do experience different levels of care and not all, by any means, have access to research studies for example. In the early stages of their disease, people don't know what good looks like or what they should be able to ask for or expect and they are also vulnerable at that time as a consequence. This is where we come in – our goal is to be there at the start of everyone's journey and whenever they need us along the way. We try to emphasise the importance of supported self-management early on as the more you know about the disease and the more you can do to help yourself in a positive way, the better your outcomes are likely to be. Unfortunately, whilst there is a lot of rhetoric about self-management for people with LTCs, we still live in a very medical management model where investment in patient education, support and self-management by commissioners is far too low. That's one of the reasons it is essential that health professionals sign-post patients to organisations who can help and support like NRAS. Access to treatment where there are specific eligibility criteria – ref the biologics and biosimilars – is better than pre-NICE, however, with the introduction of biosimilars, the market has changed and there is a lot of confusion at the moment with local procurement deals ensuring that what is available in one area, may not be the same as the next. Even with all the new treatments

available, the heterogeneity of this disease syndrome means that there remains unmet need. Even with cheaper drugs available and many people thinking that therefore more people will be able to get the treatment they need, this is not the case unless NICE change the eligibility criteria which currently apply.

4. What do patients or carers consider to be the advantages of the treatment being appraised?

Benefits of a treatment might include its effect on:

- the course and/or outcome of the condition
- physical symptoms
- pain
- level of disability
- mental health
- quality of life (such as lifestyle and work)
- other people (for example, family, friends and employers)
- ease of use (for example, tablets rather than injection)
- where the treatment has to be used (for example, at home rather than in hospital)
- any other issues not listed above

Please list the benefits that patients or carers expect to gain from using the treatment being appraised.

The key driver of RA is inflammation which can result quite quickly in bone erosion leading ultimately to joint destruction and potential disability. For the first time since the introduction of the biologics, Baracitinib offers a completely new class of innovative therapy that could, as I understand it from our Chief Medical Advisor be positioned post DMARD failure or post first TNF failure. This is fantastic because it really adds to the therapeutic options available to clinicians and patients. Also the fact that this is an oral therapy means that there are no costs associated with infusions based therapies or those delivered via sub-cut route. All those costs associated with home care delivery companies also disappear. It's really very exciting especially for patients like me who have refractory disease and who have been through all the biologics available. Should my current biologic fail to keep my disease under control, this new drug gives me an option to palliative steroid therapy.

Patients are very likely to prefer an oral (biologic) drug to have a regular infusion or having to inject themselves.

Please explain any advantages that patients or carers think this treatment has over other NHS treatments in England.

I think that what I have said in the above statement summarises why patients would be likely to prefer an oral drug over injection themselves or having to attend hospital (and take time off work) for infusion therapy. Although this may seem a minor point, many people with little fridge space, also may prefer not to have to keep their medicine refrigerated. The potential cost savings by not having to bring people into day case care for infusions or have home heathcare companies delivering drugs must also surely be welcome in a cash-strapped NHS.

If you know of any differences in opinion between patients or carers about the benefits of the treatment being appraised, please tell us about them.

I am not aware of any but should also point out that few patients will be aware of the arrival of these new JAK inhibitors.

5. What do patients and/or carers consider to be the disadvantages of the treatment being appraised?

Disadvantages of a treatment might include:

- aspects of the condition that the treatment cannot help with or might make worse
- difficulties in taking or using the treatment (for example, injection rather than tablets)
- side effects (for example, type or number of problems, how often, for how long, how severe. Please describe which side effects patients might be willing to accept or tolerate and which would be difficult to accept or tolerate)
- where the treatment has to be used (for example, in hospital rather than at home)
- impact on others (for example, family, friends and employers)
- financial impact on the patient and/or their family (for example, the cost of travel to hospital or paying a carer)
- any other issues not listed above

Please list any concerns patients or carers have about current NHS treatments in England.

Current biologics have to be infused or injected. People of working age, and ¾ of people are diagnosed with RA when of working age – (and we also need to bear in mind that age of retirement is extending quite considerably) generally find it problematic to take time off work to visit the hospital for infusions. Often there is more waiting around than they would like and what might have been expected to take half a day can extend into the best part of a whole day. People who self-inject can also find this difficult sometimes and those with major hand deformity or pain have to get someone else to inject for them and family members don't always find this easy. Also if you are living alone and can't self-inject, you may have to get one of the home delivery company nurses to attend or go the hospital. All additional inconveniences. Having said that, many people like myself, have no difficulty injecting themselves.

Please list any concerns patients or carers have about the treatment being appraised.

See my answers above. I am sure that an oral drug would be welcome for all the reasons mentioned.

If you know of any differences in opinion between patients or carers about the disadvantages of the treatment being appraised, please tell us about them.

No.

6. Patient population

Are there any groups of patients who might benefit more from the treatment than others? If so, please describe them and explain why.

I am not sure of the facts here relating to Baracitinib but it may be that this treatment is suitable for patients who might be less suitable for some of the biologic options. For example, we know that Rituximab is generally targeted at those patients who are sero-positive. If Baracitinib is equally effective for both sero-ve and sero+ve, this is an example. Those who are needle phobic would benefit from an oral option as described previously.

Are there any groups of patients who might benefit less from the treatment than others? If so, please describe them and explain why.

I am not aware of any

7. Research evidence on patient or carer views of the treatment

Is your organisation familiar with the published research literature for the treatment?

☐ Yes ☐ No

If you answered 'no', please skip the rest of section 7 and move on to section 8.

Please comment on whether patients' experience of using the treatment as part of their routine NHS care reflects the experiences of patients in the clinical trials.

I am generally aware of the positive trial outcomes and the head to head with Humira but as there are no patients yet using this drug in routine care in the UK, I can't comment further.

Do you think the clinical trials have captured outcomes that are important to patients? Are you aware of any limitations in how the treatment has been assessed in clinical trials?

I don't know what all the end points were but I imagine if reduced pain was one, and reduced das scores, these would be important outcomes to patients.

If the treatment being appraised is already available in the NHS, are there any side effects that were not apparent in the clinical trials but have emerged during routine NHS care?

N/A

Are you aware of any relevant research on patient or carer views of the condition or existing treatments (for example, qualitative studies, surveys and polls)?

☐ Yes ☐ No

If yes, please provide references to the relevant studies.

Our own social research includes

- Family Matters NRAS 2012

- I want to work NRAS 2007
- RA Fatigue Survey and Report 2014
- The Mapping Project, Sue Oliver and Ailsa Bosworth, 2009
- Scotland Work survey, NRAS 2010
- Who Cares Report, Scotand NRAS 2015
- Emotions, Relationships and Sexuality Survey & Report, NRAS 2013
- RA and physiotherapy NRAS 2011
- Wales State of Play Report, BSR and NRAS, 2015

8. Equality

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Protected characteristics are: age; being or becoming a transsexual person; being married or in a civil partnership; being pregnant or having a child; disability; race including colour, nationality, ethnic or national origin; religion, belief or lack of religion/belief; sex; sexual orientation.

Please let us know if you think that recommendations from this appraisal could have an adverse impact on any particular groups of people, such as:

- excluding from full consideration any people protected by the equality legislation who fall within the patient population for which the treatment is/will be licensed;
- having a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the treatment;
- any adverse impact on people with a particular disability or disabilities.

Please let us know if you think that there are any potential equality issues that should be considered in this appraisal.

No

Are there groups of patients who would have difficulties using the treatment or currently available treatments? Please tell us what evidence you think would help the Committee to identify and consider such impacts.

Not that I am aware of

9. *Other issues*

Do you consider the treatment to be innovative?

☐ Yes ☐ No

If yes, please explain what makes it significantly different from other treatments for the condition.

This is a truly innovative drug as it represents the introduction of a new class of medicine which targets the inside of cells involved in the immune system rather than blocking receptors on the outside of cells as per all the other biologic and biosimilar drugs. It is a small molecule drug.

Are there any other issues that you would like the Appraisal Committee to consider?

Not that I can think of

10. *Key messages*

In no more than 5 bullet points, please summarise the key messages of your submission.

- This is a new class of therapy not previously available
- It is truly innovative
- Patients are likely to be more prepared to take an oral medicine than inject themselves or be infused
- It has the potential to save a lot of costs due to the fact that it is oral
- It can be used in different places in the current pathway, ie. post dmard failure and post TNF failure

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Patient/carer expert statement (STA)

Baricitinib for treating moderate to severe rheumatoid arthritis [ID979]

Thank you for agreeing to give us your views on this treatment that is being appraised by NICE and how it could be used in the NHS. Patients, carers and patient organisations can provide a unique perspective on conditions and their treatment that is not typically available from other sources. We are interested in hearing about:

- the experience of having the condition or caring for someone with the condition
- the experience of receiving NHS care for the condition
- the experience of having specific treatments for the condition
- the outcomes of treatment that are important to patients or carers (which might differ from those measured in clinical studies, including health-related quality of life)
- preferences for different treatments and how they are given
- expectations about the risks and benefits of the treatment.

We have already asked your nominating organisation to provide an organisation's view. We are asking you to give your views as an individual whether you are:

- a patient
- a carer (who may be voicing views for a patient who is unable to) or
- somebody who works or volunteers for a patient organisation.

To help you give your views, we have provided a questionnaire. You do not have to answer every question — the questions are there as prompts to guide you. The response area will expand as you type. The length of your response should not normally exceed 10 pages.

1. About you

Your name: Jennie Jones
Name of your nominating organisation:
Do you know if your nominating organisation has submitted a statement?

X Yes ☐ No

Do you wish to agree with your nominating organisation's statement?

X Yes ☐ No

(We would encourage you to complete this form even if you agree with your nominating organisation's statement.)

Are you:

- a patient with the condition?

X Yes ☐ No

- a carer of a patient with the condition?

☐ Yes X No

- a patient organisation employee or volunteer?

X Yes ☐ No

Do you have experience of the treatment being appraised?

☐ Yes X No

If you wrote the organisation submission and do not have anything to add, tick here ☐ (If you tick this box, the rest of this form will be deleted after submission.)

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: No

2. Living with the condition

What is your experience of living with the condition as a patient or carer?

RA has had a devastating impact on my life and life expectations. It has affected my family, my working life and prospects and my emotional well being. When initially diagnosed, my disease was aggressive and getting worse through a number of years while drugs were tried and found not to be effective in controlling my disease. I had to stop my full time work as the travel and hours involved couldn't be sustained with the pain, chronic fatigue and unreliability of my body. It has taken years to get to a steady state of remission and Humira has helped me get a life back (although not as it was before RA). Psychologically getting to a point of acceptance that this disease is here to stay is very difficult. My Mother also had the disease and I cared for her until she passed away, so I am also aware that the effectiveness of drugs and sensitivity to side effects etc can change over time, so no one knows what the future holds. I just have to live as well as possible today. It is particularly hard on your family who have to witness your pain, accommodate your lack of energy and help you when you cannot manage to do even basic tasks. Loss of independence and having to ask for help I have found very hard, even when it is offered freely and with much love.

3. Current practice in treating the condition

Which treatment outcomes are important to you? (That is, what would you like treatment to achieve?) Which of these are most important? If possible, please explain why.

Achieving a target of remission or as close to it as possible with reduced inflammations, pain and less chronic fatigue, in the least possible time to avoid loss of work and independence.

What is your experience of currently available NHS care and of specific treatments? How acceptable are these treatments – which did you prefer

and why?

The quality of care available in the NHS is inconsistent and I had to move hospitals to avoid poor practice. The rationing of Biologic drugs combined with the fact that it usually takes at least 3 months to determine whether a DMARD is working meant for me that my life was on a total downward spiral physically and emotionally for a few years before I got treatments that made a significant difference (steroids work, but they are not sustainable long term without side effects). I have used Methatrexate, Predisolone, Hydroxychloroquine and Sulphasalazine and Humira. It should be noted that even using a biologic is not without problem as normal infections mean the drug has to be stopped and then a period of inflammations etc are likely again until the RA settles down.

The Humira has been the only thing that really works for me, but injecting is not ideal and arranging deliveries, etc takes significant time. Methatrexate I find very unpleasant to take orally with bad effects on my stomach even now.

4. What do you consider to be the advantages of the treatment being appraised?

Benefits of a treatment might include its effect on:

- the course and/or outcome of the condition
- physical symptoms
- pain
- level of disability
- mental health
- quality of life (such as lifestyle and work)
- other people (for example, family, friends and employers)
- ease of use (for example, tablets rather than injection)
- where the treatment has to be used (for example, at home rather than in hospital)
- any other issues not listed above

Please list the benefits that you expect to gain from using the treatment being appraised.

This disease specific treatment should deliver another option for reducing pain and inflammation at any stage of the disease, hopefully with fewer side

effects. I am likely to be living with this disease for the rest of my life, and I am aware drugs can become less effective, other illness or side effects may mean changes in medication has to happen in response. The more effective alternatives, the better for every patient

Please explain any advantages that you think this treatment has over other NHS treatments in England.

Oral dose rather than infusion/injection will be a benefit to many patients who may have difficulty getting to hospital appointments or injecting themselves due to disease/disability in their hands.

If you know of any differences in opinion between you and other patients or carers about the benefits of the treatment being appraised, please tell us about them.

Not aware

5. What do you consider to be the disadvantages of the treatment being appraised?

Disadvantages of a treatment might include:

- aspects of the condition that the treatment cannot help with or might make worse
- difficulties in taking or using the treatment (for example, injection rather than tablets)
- side effects (for example, type or number of problems, how often, for how long, how severe. Please describe which side effects patients might be willing to accept or tolerate and which would be difficult to accept or tolerate)
- where the treatment has to be used (for example, in hospital rather than at home)
- impact on others (for example, family, friends and employers)
- financial impact on the patient and/or their family (for example, the cost of travel to hospital or paying a carer)
- any other issues not listed above

Please list any concerns you have about current NHS treatments in England.

Limiting access to Biologic drugs via strict rationing criteria means many patients endure life changing levels of pain and inflammation and still do not qualify for that therapy. All aspects of their lives are affected – family, work,

relationships and emotional well being. This is a great loss to society as a whole, and devastating for those individuals who have to live with this disease for the rest of their lives.

Please list any concerns you have about the treatment being appraised.

None known

If you know of any differences in opinion between you and other patients or carers about the disadvantages of the treatment being appraised, please tell us about them.

None known

6. Patient population

Do you think some patients might benefit more from the treatment than others? If so, please describe them and explain why.

This drug should benefit patients at all stages of their disease pathway as a first line or alternative treatment.

Do you think some patients might benefit less from the treatment than others? If so, please describe them and explain why.

Not known

7. Research evidence on patient or carer views of the treatment

Are you familiar with the published research literature for the treatment?

☐ Yes X No

If you answered 'no', please skip the rest of section 7 and move on to section 8.

Please comment on whether your experience of using the treatment as part of routine NHS care reflects the experience of patients in the clinical trials.

Do you think the clinical trials have captured outcomes that are important to patients? Are you aware of any limitations in how the treatment has been assessed in clinical trials?

If the treatment being appraised is already available in the NHS, are

there any side effects that were not apparent in the clinical trials but have emerged during routine NHS care?

Are you aware of any relevant research on patient or carer views of the condition or existing treatments?

X Yes ☐ No

If yes, please provide references to the relevant studies.

National Rheumatoid Association

- Family Matters NRAS 2012
- I want to work NRAS 2007
- RA Fatigue Survey and Report 2014
- The Mapping Project, Sue Oliver and Ailsa Bosworth, 2009
- Scotland Work survey, NRAS 2010
- Who Cares Report, Scotand NRAS 2015
- Emotions, Relationships and Sexuality Survey & Report, NRAS 2013
- RA and physiotherapy NRAS 2011
- Wales State of Play Report, BSR and NRAS, 2015

8. Equality

NICE is committed to promoting equality of opportunity and eliminating discrimination. Please let us know if you think that recommendations from this appraisal could have an adverse impact on any particular groups of people, who they are and why.

No

9. Other issues

Do you consider the treatment to be innovative?

X Yes ☐ No

If yes, please explain what makes it significantly different from other treatments for the condition.

Disease specific delivered orally

Is there anything else that you would like the Appraisal Committee to

consider?

No

10. Key messages

In no more than 5 bullet points, please summarise the key messages of your submission.

- This is a new class of therapy not previously available
- It is truly innovative – a targeted drug attacking the disease directly
- Patients are likely to be more prepared to take an oral medicine than inject themselves or be infused
- It has the potential to save a lot of costs due to the fact that it is oral
- It can be used in different places in the current pathway, i.e. post Dmard failure and post TNF failure and adds to the armoury of drugs that may be needed to meet changing requirements during a lifetime of living with this disease.

Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis: A Single Technology Appraisal

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK Iñigo Bermejo, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Emma Simpson, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Shijie Ren, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Ruth Wong, 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
Correspondence Author	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK
Date completed	Date completed (07/04/2017)

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 16/56/23.

Declared competing interests of the authors

David Scott has no conflicts of interest relating to Baricitinib. However, 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.

Also, his department has received a peer-reviewed grant from Pfizer within the last 12 months to undertake academic research on polypharmacy in arthritis. The department has also received free etanercept from Pfizer to use in an NIHR-funded programme grant in rheumatoid arthritis.

None of the other 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 Gill Rooney, 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:

Stevenson M, Bermejo I, Simpson E, Wong R, Scott DL, Young A. Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2016.

Contributions of authors

Emma Simpson summarised and critiqued the clinical effectiveness data reported within the company's submission. Inigo Bermejo and Matt Stevenson critiqued the health economic analysis submitted by the company. Shijie Ren critiqued the company's network meta-analysis and undertook a new analysis. Ruth Wong critiqued the company's search strategy. David Scott and Adam Young

provided clinical advice to the team. All authors were involved in drafting and commenting on the final report.

CONTENTS

List of Tables

List of Figures

ABBREVIATIONS

ABA	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
BARI	Baricitinib
bDMARD	Biologic disease-modifying antirheumatic drug
BIC	Bayesian Information Criterion
BNF	British National Formulary
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
CTZ	Certolizumab pegol
DAS28	Disease Activity Score 28
DAS-CRP	Disease Activity Score C-reactive protein
DES	Discrete event simulation
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
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
IPS	Individual patient simulation
ITT	Intention to treat
IV	Intravenous
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
PALL	Palliative care
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
Q2W	Every two weeks
Q4W	Every four weeks
RA	Rheumatoid arthritis
RCT	Randomised controlled trial
RF	Rheumatoid factor
RR	Rate ratio
RTX	Rituximab
SAE	Serious adverse event
SC	Subcutaneous
SD	Standard deviation
SDAI	Simplified Disease Activity Index
SE	Standard error
SF-36	Short Form (36) Health Survey
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
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 CS generally adhered to the NICE scope. Exceptions related to the exclusion of the subcutaneous (SC) formulation of tocilizumab (TCZ) (TCZ SC) as a comparator, as well as the intravenous (IV) formulation of abatacept (ABA) (ABA IV).

1.2 Summary of clinical effectiveness evidence submitted by the company

The key clinical effectiveness evidence for baricitinib (BARI) was based on three randomised controlled trials (RCTs). Additionally one long-term extension study was included. There were two RCTs in methotrexate (MTX) -or conventional disease-modifying antirheumatic drug (cDMARD)-treated, biologic disease-modifying antirheumatic drug (bDMARD) naïve patients (RA-BEAM, RA-BUILD), both of which included a placebo (PBO) comparator. One RCT also included adalimumab (ADA) as a comparator (RA-BEAM). One PBO-controlled RCT was conducted in bDMARD-treated patients (RA-BEACON).

For the primary endpoint of ACR20 at 12 weeks follow-up, all three RCTs reported that BARI 4mg was statistically significantly superior to PBO ($p \leq 0.001$). At 12 weeks, more patients reached a 20% improvement in the ACR score (ACR20) in the BARI 4mg treated arm than the ADA treated arm ($p=0.01$). There was also an advantage over PBO for BARI 4mg at 24 weeks and for BARI 2mg at 12 weeks and 24 weeks follow-up. At 12 weeks follow-up, all three RCTs reported a significant advantage for BARI 4mg over PBO for EULAR response ($p<0.05$).

The most common adverse events for BARI were low-density lipoprotein cholesterol, upper respiratory tract infections and nausea; other adverse drug reactions included herpes simplex, herpes zoster, acne, increased creatine phosphokinase, increased triglycerides, increased liver function tests (aspartate transaminase, alanine transaminase), neutropenia and thrombocytosis.

Network meta-analyses (NMA) were performed to assess the relative efficacy of BARI compared with the comparators in the inadequate response to cDMARDs (cDMARD-IR) or inadequate response to a tumour necrosis factor inhibitor (TNFi) (TNFi-IR) patients with moderate to severe rheumatoid arthritis (RA).

For the base case analysis at week 24 in the cDMARD-IR population, BARI 4mg was associated with a statistically significant higher odds of an ACR 50 response compared with cDMARD, ADA, PBO,

ETN and SSZ. No statistically significant differences were found versus any other comparators for the ACR50 outcome, with the exception of CTZ + cDMARD, in which odds of ACR50 response was found to be significantly in favour of the comparator. A similar pattern of results was observed for BARI 2mg.

For the base case analysis at week 24 in the TNFi-IR population, BARI demonstrated significantly higher ACR50 response rates than the cDMARD comparator. No statistically significant differences were found versus bDMARDs, with the exception of the comparison of BARI (both 4mg and 2mg) to TCZ, and the comparison of BARI 2mg to RTX, in which statistically significant treatment effects in favour of the comparator were observed.

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

The ERG found the searches for clinical effectiveness evidence reported in the CS to be adequate, and believed that all published RCTs of BARI 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 states that the results presented in NMA should be treated with caution, as a random effects model was assumed for the study-specific baseline treatment effects (pooling non-active and active controls): this was deemed to be inappropriate. In addition, studies that reported European League Against Rheumatism (EULAR) responses were synthesised along with converted EULAR response outcomes from studies that only reported ACR responses. This differs to the approach used in TA375, which performed the conversion having synthesised the ACR data, which ensures that the relative rankings of treatments are maintained.

1.4 Summary of cost effectiveness submitted evidence by the company

The manufacturer supplied a *de novo* discrete event simulation (DES) model constructed in Microsoft Excel®. The model simulates patients' disease progression through the sequences of treatments being compared. For each treatment, patients may achieve good, moderate or no EULAR response; this is assessed at 24 weeks. The EULAR response rates for each treatment are based on 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), adverse event or death. Patients who fail to achieve a moderate or good EULAR response discontinue treatment at 24 weeks and start the next treatment in the sequence. HAQ progression whilst on treatment is assumed to be flat on bDMARDs or BARI, whilst on

cDMARDs and palliative care, HAQ progression is assumed to be non-linear based on latent HAQ trajectory classes. Time to treatment discontinuation for responders is independent of treatment but is dependent on EULAR response category (moderate or good) and is modelled using Weibull curves fitted to British Society for Rheumatology Biologics Register (BSRBR) data. At treatment discontinuation, patients are assumed to suffer a rebound in HAQ equal to that achieved on treatment initiation and 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) over patients' remaining lifetimes. EuroQol 5 Dimensions (EQ-5D) values are calculated 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, and unit costs were taken from the British National Formulary and NHS Reference Costs 2014/15. Serious Adverse Events (SAEs) were excluded from the base case but were included in a scenario analysis.

The analyses presented in the CS relate to four different populations of rheumatoid arthritis patients: (1) patients who have had an inadequate response to cDMARDs (cDMARD-IR) with moderate RA; (2) cDMARD-IR patients with severe RA; (3) patients with severe RA who have had an inadequate response to a tumour necrosis factor inhibitor (TNFi) (TNFi-IR) and who are rituximab (RTX) eligible; and (4) patients who are TNFi-IR with severe RA for whom RTX is contraindicated or not tolerated. The definition of severe RA was a DAS28 > 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 BARI trials.

In the cDMARD-IR population with moderate RA, the deterministic incremental cost-effectiveness ratio (ICER) for BARI + MTX compared with intensive cDMARDs was estimated to be £37,420 per QALY gained. In the cDMARD-IR population with severe RA, BARI + MTX dominated all comparators except for certolizumab pegol (CTZ) + MTX, with the ICER of CTZ + MTZ compared with BARI + MTX estimated to be £18,400 per QALY gained. In the TNFi-IR population with severe RA, when RTX + MTX was an option, BARI + MTX was dominated by RTX + MTX. In the TNFi-IR population with severe RA for whom RTX is contraindicated or not tolerated, BARI + MTX dominated golimumab + MTX and was less effective and less expensive than the remaining comparators. The ICERs for etanercept biosimilars (ETN-b) + MTX, CTZ + MTX and ADA + MTX compared with BARI + MTX were lower than £30,000 per QALY gained. However, the company made a favourable assumption for these interventions (same efficacy as in the severe cDMARD-IR population) in the absence of effectiveness data in this population and therefore, caution is advised when interpreting these results. The ICERs for TCZ IV + MTX and ABA SC + MTX compared with

BARI + MTX were estimated to be higher than £30,000 per QALY gained, but the confidential Patient Access Schemes (PAS) relating to TCZ and ABA were not included.

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) with some minor deviations. The ERG believed that the conceptual model was appropriate but suffered from a series of implementation errors and limitations, as described here.

The company rounded modified HAQ values to the nearest valid HAQ score rather than allowing the valid HAQ score to be sampled based on the continuous HAQ value. The ERG notes that this approach might lead to inaccurate estimations of HAQ scores, as values might be rounded up more often than rounded down or *vice versa*.

The company intended to implement the trajectory of HAQ score whilst on cDMARDs or palliative care based on the latent class approach used by the AG in TA375. However, the company assigned each patient to a single class based on the probability of class membership instead of using an average weighted by the probability of class membership.

The company assumed that patients who achieve a moderate or good EULAR response at 24 weeks experience a reduction in HAQ score instantaneously at treatment initiation. The ERG believes that the company's approach is likely to lead to an overestimation of treatment benefits, as the achievement of response will take at least a few weeks and potentially up to 24 weeks for some patients.

In order to calculate the QALYs and costs produced in the time span between two events, the model uses an area under the curve (AUC) approach for the HAQ score, and then maps this value to the EQ-5D and hospitalisation costs. However, since the relationships between HAQ score and EQ-5D and between HAQ score and hospitalisation costs are not linear, this approach may lead to inaccurate results.

The ABA IV and TCZ SC formulations were not included in the list of comparators, despite ABA SC and TCZ IV being included. The company argued that it had excluded TCZ SC because: (i) the available evidence for TCZ SC was limited; (ii) it provided a lower efficacy estimate than for TCZ IV; and (iii) the cost difference between the two formulations was relatively small. The ERG notes that the difference in costs might be considerable taking into account the administration costs and the confidential PAS. ABA IV was included in the NMA, but was excluded from the analyses. In

response to a clarification request by the ERG, the company presented the results of ABA IV only for the cDMARD-IR population with severe RA, which led to similar results compared with ABA SC (£[redacted] versus £[redacted] and [redacted] versus [redacted] QALYs respectively).

The company used one of the algorithms proposed by Hernández Alava *et al.* to map HAQ scores to EQ-5D. The ERG notes that newer algorithms with a higher accuracy have been since published, such as that in reported in Hernández Alava *et al.* and used in TA375.

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

1.6.1 Strengths

The ERG believes that all available BARI RCTs were included in the CS. The trials were considered by the ERG to be of good quality.

The model used appears conceptually appropriate with relatively few implementation errors, some of which were fixed during the clarification process. The DES approach taken by the company, which was based on the model used in TA375, was deemed appropriate to represent the disease.

1.6.2 Weaknesses and areas of uncertainty

The model contained two programming errors that affected the results of the probabilistic sensitivity analysis (PSA), especially those for the severe TNFi-IR RTX-ineligible population. Similarly, the presented results for a number of scenario analyses lacked face validity.

The only available evidence of BARI monotherapy is in MTX-naïve patients and therefore considerable uncertainty exists about the efficacy of BARI monotherapy in cDMARD-IR and TNFi-IR patients. The company did not present an economic analysis for BARI monotherapy for patients in whom MTX is contraindicated or not tolerated.

The company did not identify effectiveness data for etanercept, infliximab, ADA and CTZ in combination with MTX in the TNFi-IR RTX-ineligible population with severe RA.

The company did not present any evidence relating to the effectiveness of bDMARDs after BARI. The company's economic analysis assumes that the efficacy of bDMARDs after BARI will be equal to their efficacy after another bDMARD. This is a reasonable assumption given the lack of evidence, but the ERG notes that it is possible that the efficacy of bDMARDs after BARI would be better or worse than when following another bDMARD.

No robust evidence was presented to assess the treatment duration of BARI. In their base case, the company assumed that time to treatment discontinuation for BARI was the same as that for bDMARDs.However, BARI is not a bDMARD and it is not clear that time to loss of efficacy would be similar for BARI and bDMARDs meaning that the results are subject to uncertainty.

The company assumed that the flat HAQ progression whilst on treatment assumed for bDMARDs in TA375 also applies to BARI. The scenario analysis in which a linear HAQ increase is assumed instead showed BARI was less effective than its comparators. Whilst uncertain, clinical advice provided to the ERG suggested that it was reasonable to assume the same HAQ progression for BARI as for bDMARDs.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG undertook few exploratory analyses based on the company's submitted model as the ERG believes that [redacted]. The errors affecting the PSA were corrected and the amended results presented. The ERG comments that the added value of any mathematical model for people with severe RA in this Single Technology Appraisal (STA) is debatable given the efficacy and acquisition cost inputs of the bDMARDs.

The ERG highlight the fact that the company drew heavily on TA375 in constructing their mathematical model and that some parts, including latent classes for those on cDMARDs and the HAQ progressions for those on cDMARDs were not implemented correctly. This will affect the ICER for patients with moderate RA, where BARI is compared to cDMARDs. The time required to fix these issues were beyond that available for an STA. The ERG notes that the median ICER of bDMARDs compared with cDMARDs in TA375 was in the region of £50,000 per QALY gained for patients with moderate RA, which is considerably higher than the estimate provided by the company. The ERG believes that the ICER of BARI when used in the moderate, RA population will be closer to that reported in TA375 due to the errors in reconstructing the Assessment Group's model.

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)[1] 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 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, and is manifested with increasing disability and reduced quality of life.[2] The primary symptoms are: pain; morning stiffness; swelling; tenderness; loss of movement; fatigue; and redness of the peripheral joints.[3, 4] RA is associated with substantial costs both directly (associated with drug acquisition and hospitalisation) and indirectly due to reduced productivity.[5] RA has long been reported as being associated with increased mortality,[6, 7] particularly due to cardiovascular events.[8]

Epidemiology

NICE estimates that there are 400,000 people in the UK with RA,[9] based on a prevalence of 0.8% reported by Symmons *et al*.[10] 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).[11] For both genders the peak age of incidence in the UK is in the 70s, but all ages can develop the disease.[11]

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%[12] and a family history of RA is related with a risk ratio of 1.6 compared with the general population.[13] Many genes associated with susceptibility to RA 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 (GLD) 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).[9] 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 inhibitor, abatacept (ABA) 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 (PALL) 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[14] (ACR). NICE Clinical Guideline (CG) 79 provides a summary of the ACR criteria, namely that patients must have at least four of seven criteria: morning stiffness lasting at least 1 hour; swelling in three or more joints; swelling in hand joints; symmetric joint swelling; erosions or decalcification on x-ray of hand; rheumatoid nodules; 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 preferred a clinical diagnosis of RA rather than the ACR criteria because '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' referencing recommendations from the European League Against Rheumatism (EULAR).[15]

In 2010, the ACR and EULAR jointly published RA Classification Criteria, which focussed on the features at earlier stages of disease that are associated with persistent and/or erosive disease rather than defining the disease by its late stage features.[16] The classification criteria allocates 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[17] and EULAR responses.[18]

The initial ACR response was denoted as an ACR20 which 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.

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.[19] 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). This assesses 28 joints in terms of swelling (SW28) and of tenderness to the touch (TEN28) and also 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:[20]

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

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

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.[18] 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.[21] EULAR response has been reported less frequently in RCTs than ACR responses,[22] 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[21]

	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 stated as having inactive disease, those with a DAS28 > 3.2 and ≤ 5.1 are stated as having moderate disease and >5.1 as having very active disease.[20] Within NICE Technology Appraisal (TA) 375, patients with a DAS28 > 3.2 and ≤ 5.1 were denoted as having moderate to severe disease whilst those with a DAS28 > 5.1 were denoted as having severe disease.[23]

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.[24] 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.[22]

2.2 Critique of company's overview of current service provision

The company's overview of current service provision is concise but is 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

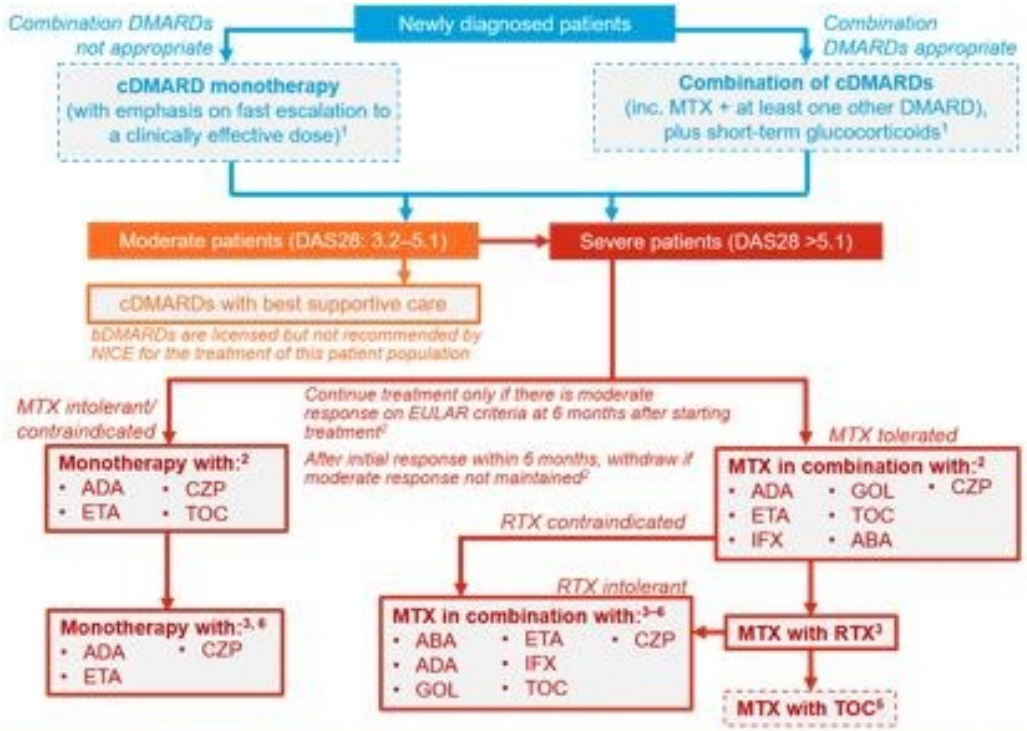
For people with newly diagnosed RA, NICE CG79[9] recommends a combination of cDMARDs (including MTX and at least one other cDMARD plus short-term glucocorticoids) as first-line treatment, ideally beginning within 3 months of the onset of persistent symptoms. Where combination therapies are not appropriate, for example where there are comorbidities or pregnancy, cDMARD monotherapy is recommended. Where cDMARD monotherapy is used, emphasis should be made on increasing 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)[23] recommends the use of ABA, ADA, CTZ, ETN, GOL, IFX, and TCZ in combination with MTX in people with RA after the failure to respond to intensive cDMARDs 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[23] recommends the following bDMARDs as monotherapy options: ADA; CTZ; ETN; or TCZ.

After the failure of the first TNF-inhibitor, TA195[25] 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 ABA, 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[26] recommends TCZ, and TA415[27] 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 replicated in Figure 1. In summary, the typical route for patients with severe disease who could tolerate MTX would be intensive cDMARDs followed by a bDMARD, followed by RTX plus MTX, then TCZ before returning to cDMARDs and potentially PALL.

Figure 1. Treatment pathway presented in the CS[1]



Footnotes: Positions in the treatment pathway where baricitinib might be considered, in accordance with the NICE scope/final license,[3] are demonstrated by thick, solid border lines. Broken border lines indicate positions where baricitinib is not considered in the decision problem.
Abbreviations: RA = rheumatoid arthritis, DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, MTX = methotrexate, DAS28 = Disease Activity Score, bDMARD = biologic DMARD, ADA = adalimumab, CTZ = certolizumab pegol, ETN = etanercept, TCZ = tocilizumab, GOL = golimumab, IFX = infliximab, ABA = abatacept, RTX = rituximab.
Sources: [1]NICE CG79,[38] [2]NICE TA375,[47] [3]NICE TA195,[46] [4]NICE TA225,[44] [5]NICE TA247,[48] [6]TA415[45]

NICE criteria for continuing treatment

NICE TA375[23] states that for patients to continue treatment with their first bDMARD treatment they must maintain at least a moderate EULAR response. TA195,[25] which for all bDMARDs excluding RTX was updated in TA375[23], 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

The final scope issued by NICE defined the population as 'Adults with moderate to severe, active rheumatoid arthritis whose disease has responded inadequately to, or who are intolerant of one or more disease modifying anti-rheumatic drugs (DMARDs), including conventional or biologic DMARDs.' The company have defined four populations for the analyses of the cost-effectiveness of baricitinib. These are:

1. patients with moderate RA who have had an inadequate response to cDMARDs (cDMARD-IR);
2. patients with severe RA who are cDMARD-IR;
3. patients with severe RA who have had an inadequate response to a TNFi (TNFi-IR) and who are RTX eligible; and,
4. patients with severe RA who are TNFi-IR for whom RTX is contraindicated or not tolerated.

Within the categorisation of patients, the company have assumed that severe disease is represented by a DAS28 > 5.1 and that moderate disease is represented by a DAS28 > 3.2 and DAS28 ≤ 5.1.

The ERG believes that such division of the population is appropriate. However, no analyses have been presented for those patients who cannot take MTX and for whom BARI would be used as monotherapy.

3.2 Intervention

The NICE scope defined the population as 'Baricitinib monotherapy or in combination with methotrexate'. This is the intervention assessed by the company.

Baricitinib (BARI), brand name Olumiant®, is a Janus kinase (JAK) inhibitor that is taken orally. BARI has selectivity for JAK1 and JAK2 and is the first JAK1/2 inhibitor licensed for the treatment of moderate to severe RA in the European Union. Inhibition of JAK1 and JAK2 signalling can reduce inflammation, cellular activation and proliferation of key immune cells in patients with RA.[28-30] Further details regarding baricitinib are provided within Chapter 2 of the CS.

BARI is indicated for the treatment of moderately to severely active RA in adult patients who have responded inadequately to, or who are intolerant to one or more DMARDs. BARI can be used as monotherapy or in combination with MTX, with the intervention taken once daily (at any time of the

day) with or without food. BARI treatment should be initiated and supervised by an experienced physician / rheumatologist although the company anticipate that maintenance treatment would be self-administered by the patient at home.

The recommended dose of BARI is 4mg once daily, although a lower dose of 2mg once daily is appropriate for: those aged 75 years and over; for patients with moderate renal impairment, as determined by an estimated glomerular filtration rate (eGFR) between 30 and 60 mL/min/1.73m^2; and for patients taking Organic Anion Transporter 3 inhibitors with strong inhibition potential. A dose of 2mg once daily may also be considered for (i) patients with a history of chronic or recurrent infections and (ii) for patients who have achieved sustained control of disease activity with 4mg once daily and are eligible for dose tapering. BARI is not recommended for patients with an eGFR < 30 mL/min/1.73m^2.

The company state that no additional infrastructure for the NHS would be necessary if BARI was recommended.

The list price of BARI is £805.56 for a pack of 28 tablets and £2416.68 for a pack of 84 tablets (irrespective whether these are 2mg or 4mg packs). The company has made a submission to the Patient Access Schemes Liaison Unit (PASLU) of a simple discount on the cost of baricitinib. The level of the discount is [redacted]. The cost-effectiveness results presented by the company assume that the patient access scheme (PAS) is in place.

3.3 Comparators

The comparators for the assessment of clinical and cost-effectiveness defined in the NICE scope were dependent on the population being analysed. These are discussed in turn.

- For cDMARD-IR patients with severe RA who can receive MTX the following interventions, all in combination with MTX, were defined as comparators: ABA; ADA; CTZ; ETN; IFX; GOL; and TCZ

- For cDMARD-IR patients with severe RA who cannot receive MTX the following interventions were defined as comparators: ADA; CTZ; ETN; and TCZ (each as monotherapy).

- For TNFi-IR patients with severe RA RTX was defined as the primary comparison. If RTX is contraindicated or withdrawn due to AEs the comparators were defined as: ABA; ADA; CTZ;

ETN; IFX; GOL; and TCZ (each in combination with MTX) and: ADA; ETN; and CTZ (each as monotherapy).

- For cDMARD-IR patients with moderate RA the comparators were defined as: intensive cDMARDs; cDMARD monotherapy and PALL.

The company did not consider the intravenous (IV) formulation of ABA: in the clarification response (question B15), the company stated that this was '*a pragmatic decision (…) To attempt to limit the number of sequences included in the submission where it was possible that inclusion of the same intervention with different administration routes was unlikely to be informative.'* The company also did not consider the subcutaneous (SC) formulation of TCZ with three reasons provided in the clarification response (question B9) culminating in the company stating that including 'IV tocilizumab only was felt to be a reasonable choice, with it likely to be representative of the costs and outcomes associated with the S/C version.'

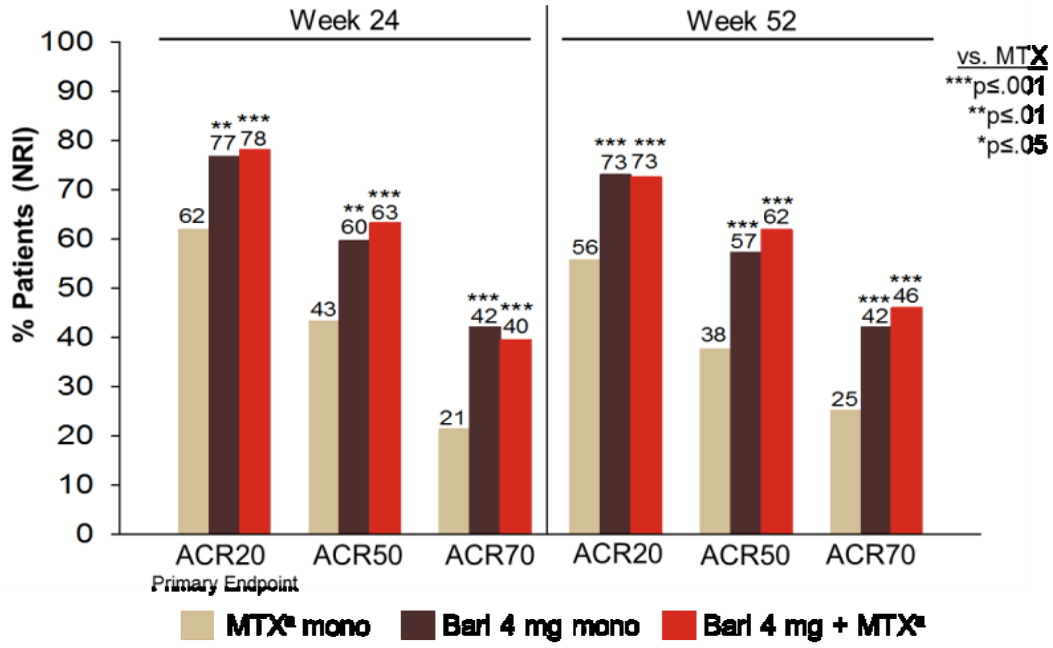
Broadly the ERG believes that the company has evaluated the correct comparators although make two comments.

1) The company have not explicitly modelled BARI used as a monotherapy. The rationale stated by the company for this is 'the paucity of efficacy data in the baricitinib clinical trial programme for patients receiving baricitinib monotherapy, which would be insufficient to form a reliable estimate of efficacy in the modelled populations for baricitinib monotherapy. It should be noted that in the recent MTA regarding the use of biologics in DMARD-naïve and cDMARD-IR patients (TA375), the Appraisal Committee agreed that the minority of (cDMARD-IR) patients with severely active rheumatoid arthritis who could not tolerate methotrexate should not be treated differently from other people with severe disease, as far as possible. The Committee concluded that biologic DMARDs should be recommended as a cost-effective use of NHS resources when used as monotherapy for severely active disease previously treated with DMARDs, where the marketing authorisation of the bDMARD allows for this recommendation to be made. The economic evaluation of baricitinib presented here assumes that a similar rationale will be applied to baricitinib monotherapy.' The lack of data for BARI when used as a monotherapy will increase the uncertainty in its incremental cost effectiveness ratio (ICER) when compared with interventions with a larger evidence base. Clinical advice to the ERG suggests that there is no clearly defined relationship between the efficacy of a bDMARD in combination with MTX and in the bDMARD used as monotherapy. However, data from RA-BEGIN[31] showed that the addition of MTX to BARI 4mg did not produce a marked improvement over BARI

monotherapy in a MTX-naïve population. This therefore provides supportive evidence regarding the efficacy of BARI monotherapy. The data from RA-BEGIN are shown in Figure 2.

2) In all comparisons, the biosimilar prices for IFX and ETN have been used rather than the prices of the original compounds. ABA (both IV and SC), and TCZ (both IV and SC) are subject to commercial-in-confidence (CIC) patient access schemes (PAS). Given this, the company has solely used list prices for these drugs, with the ERG incorporating the discounts for these interventions in a confidential appendix.

Figure 2: ACR responses at weeks 24 and 52 from RA-BEGIN (reproduced from Figure 4 of the CS appendices)



3.4 Outcomes

The outcomes contained within the scope were all addressed within the CS.

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 BARI for treating RA. The clinical evidence provided in the CS comprised a systematic review of BARI RCTs, and RCTs of bDMARDs to populate the network meta-analysis (NMA), and a review of safety evidence of BARI.

4.1.1 Searches

The company performed one clinical effectiveness search to identify all clinical and safety studies of BARI and its comparators (MTX, SSZ, LEF, HCQ, IFX, ADA, CZP, GOL, ETN, ABA, RTX, TCZ and other bDMARDs, cDMARDs and traditional DMARDs).

For the original searches, several electronic bibliographic databases were searched including MEDLINE [via PubMed], MEDLINE in Process [via PubMed], EMBASE [via Elsevier], Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials and the Health Technology Assessment database [via Wiley]) and conference proceedings websites (EULAR, ACR and the British Society for Rheumatology) for the period 1999 to June 2015. For the update searches, similar sources appear to have been searched and covered the period to August 2016.

The company searched several clinical trial registers in August 2016 (clinicaltrials.gov, WHO International Clinical Trials Registry Platform, EU Clinical Trials Register and the German clinical trials database). Supplementary searches by the company included scanning the bibliographies of included studies, reviews, meta-analyses and Scottish Medicines Consortium (SMC) advice, NICE Multiple Technology Appraisals and STA documents, US FDA register and EU EPAR reports (page 68 of the CS).

In Appendix 2 of the CS the company only reported the full literature search strategies for Embase via Elsevier. Medline and Cochrane Library database syntax and MeSH headings are sufficiently different from Embase and therefore these strategies should have been reported for transparency and reproducibility.

The ERG considers that the Embase search strategy is sufficiently comprehensive to retrieve important citations relating to all eligible studies. For the reasons described above, the ERG was unable to assess the adequacy of the searches for Medline and Cochrane Library.

4.1.2 Inclusion criteria

Inclusion criteria are presented in Table 11 of the CS. These are in accordance with the decision problem in the final NICE scope. Inclusion and exclusion criteria are shown in Table 2 of the ERG report.

The population defined in the final NICE scope was adults with moderate to severe RA who are cDMARD-IR or TNFi-IR. The CS also included treatment-naïve patients; however, the BARI trial in this population was recognised as not meeting the scope, and described only in Appendix 1 of the CS. The intervention (technology of interest) was BARI monotherapy or in combination with MTX. Other interventions / comparators of cDMARDs and bDMARDs were included to populate the NMA. Outcome measures included were presented in Tables 1 and 11 of the CS and included: Disease activity (ACR20; ACR50; ACR70; EULAR Response; DAS28 high-sensitivity C-reactive protein (hsCRP); DAS28-ESR; Simplified Disease Activity Index (SDAI); Clinical Disease Activity Index (CDAI)); Physical function (morning joint stiffness (MJS), HAQ-disability index (HAQ-DI)); Joint damage / radiological progression (modified total Sharp score, (mTSS)); Pain (captured as part of the ACR core set); RA-related mortality; Fatigue (Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F)); Worst Tiredness Score); Extra-articular manifestations of the disease (captured under safety reporting); Adverse effects of treatment; Health-related quality of life (EuroQol 5 dimensions 5 levels (EQ-5D-5L); Short Form (36) Health Survey (SF-36) v2); Work Productivity and Activity Index-Rheumatoid Arthritis (WPAI-RA). Study design, for effectiveness data, was restricted to RCTs and their long-term extension studies. This was appropriate given the availability of RCTs meeting the inclusion criteria. Systematic reviews were checked for 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 reviewer was employed to resolve discrepancies between reviewers.

Table 2: Inclusion and exclusion criteria (reproduced from Table 11 of the CS eligibility criteria used in search strategy)

Clinical effectiveness	Inclusion criteria	Exclusion criteria
Population	• Adult (≥ 18 years) patients with moderately to severely active RA (including patients with early and established RA) • Treatment-naïve patients • Patients who had intolerance or inadequate response to prior conventional DMARDs • Patients who had intolerance or inadequate response to previous bDMARDs	• Juvenile idiopathic arthritis • Studies that include only juveniles • Patients with mild RA[a]; if the study population is mixed (i.e., mild to severe), exclude those studies in which data are not reported separately for moderate or severely active RA
Intervention	• Baricitinib o †Licensed treatments: at the labelled doses o †Treatments not yet licensed: in any form or dose • Methotrexate (Trexall, Rheumatrex, amethopterin, Rasuvo, Otrexup) • Sulfasalazine (Azulfidine, Salazopyrin, Sulazine, sulfazine) • Leflunomide (Arabloc, Arava, Lunava, Respo) • Hydroxychloroquine (Plaquenil, Axemal, Dolquine, Quensyl, Quineprox) • Azathioprine (Azasan, Imuran, Azamun, Imurel) • Infliximab (Remicade) • Adalimumab (Humira, Trudexa,ABP 501, BI695501, CHS-1420, GP2017, M923, PF-06410293) • Certolizumab pegol (Cimzia) • Golimumab (Simponi) • Etanercept (Enbrel, Ave*nt, BX2922, CHS-0214, ENIA11, Etacept, Etanar, GP2013, GP2015, HD203, LBEC0101, M923, PRX-106, SB4, TuNEX, Yisaipu) • Abatacept (Orencia) • Anakinra (Kineret) • Rituximab (Rituxan,Mabthera, Zytux, Reditux) • Tocilizumab (Actemra, RoActemra, atlizumab) • Sarilumab • Sirukumab • Tofacitinib (Xeljanz, Jakvinus, tasocitinib)	• Studies that do not have an intervention of interest in at least 1 arm • Non-pharmacological studies, e.g., exercise, Chinese medicine, etc. • *Biosimilars • *Azathioprine (Azasan, Imuran) • *Studies comparing conventional DMARDs to non-DMARD treatments, such as NSAIDs or glucocorticoids

Clinical effectiveness	Inclusion criteria	Exclusion criteria
	○ At the level 1 screening, all therapy versions (i.e., any dose or combination) of the interventions listed above will be included	
Comparators	• Any comparison between any of the listed interventions and each other or placebo	• Studies not reporting on at least one of the interventions of interest
Outcomes	• *Studies reporting efficacy and safety data, HRQOL, WPAI-RA, or health care resource utilisation • *MRI studies that specifically mention the Sharp/Van der Heijde bone erosion score †To be included in the review, a study must report at least 1 of the following outcomes of interest: • †Efficacy measurements: • †ACR criteria • †ACR score • †Proportion of patients achieving an ACR20 response • †Proportion of patients achieving an ACR50 response • †Proportion of patients achieving an ACR70 response • †ACR remission • †Proportion of patients achieving an ACR50 response in the subgroup of patients who are TNF inhibitor naïve, have inadequate response to TNF or other biologics, or who are intolerant to TNF or other biologics (if reported) • †Proportion of patients achieving an ACR20 response in the subgroup of patients who are TNF inhibitor naïve, have inadequate response to TNF or other biologics or who are intolerant to TNF or other biologics (if reported) • †Individual components of the ACR: ○ HAQ-DI ○ Pain VAS ○ Tender joint count ○ Swollen joint count ○ Physician's Global Assessment of Disease Activity ○ Patient's Global Assessment of Disease Activity ○ Modified Total Sharp score ○ Erosion score ○ Joint space narrowing score ○ DAS-28 ESR for RA	• *Studies that report only MRI outcomes and do not specifically mention the Sharp/Van der Heijde bone erosion score • *Studies that report only bone mineral density • *Studies that investigate ultrasound and radiography in assessing bone damage

Clinical effectiveness	Inclusion criteria	Exclusion criteria
	○ DAS-28 CRP for RA ○ SDAI ○ CDAI ○ Physical function assessed by HAQ or HAQ-DI • †Endpoints measuring the following: ○ Morning joint stiffness (severity and duration) and/or joint pain (may be assessed by different instruments) ○ Tiredness or fatigue (may be assessed by different instruments) • †EULAR or ACR remission defined as: ○ CDAI score ≤ 2.8 ○ SDAI score ≤ 3.3 ○ DAS-28 < 2.6 ○ RAPID3 ≤ 1 ○ DAS-44 < 1.6 ○ Boolean definition of remission (EULAR or ACR where all measures must be < 1) • WPAI-RA • Health care resource utilisation • †HRQOL outcomes from the following: ○ EQ-5D ○ SF-36 • †Safety outcomes reported at study endpoint: ○ Overall rate of AEs ○ Overall rate of serious AEs ○ Discontinuations due to ▪ Lack of efficacy ▪ AEs ○ Individual AEs, such as the following: ▪ Specific myelosuppressive events, e.g., anaemia, leukopaenia, neutropaenia, or thrombocytopaenia or lymphopaenia or lymphocytopaenia ▪ Thrombocytosis ▪ Serious infections ▪ Opportunistic infections	

Clinical effectiveness	Inclusion criteria	Exclusion criteria
	▪ Malignancies ▪ Cardiovascular events ▪ Elevations in ALT or AST (> 3 times upper limit of normal) with total bilirubin (> 2 times upper limit of normal) ▪ Injection-related combinations ▪ Intravenous reactions o Death o Initial or prolonged inpatient hospitalisation	
Study design	• Randomised, controlled, prospective clinical trials • Long-term follow-up studies (e.g. open-label follow-up studies with continuation of treatments in their respective randomised group) • Systematic reviews (including meta-analyses)[b]	• *Phase 2, randomised, controlled, prospective clinical trials • Non-randomised clinical trials • Single-arm studies • *Long-term follow-up or extension studies of RCTs in which patients do not remain in their respective randomised group • *Maintenance studies and step-down treatment studies • Preclinical studies • Phase 1 studies • Prognostic studies • Retrospective studies • Prospective observational studies • Case reports • Commentaries and letters (publication type) • Consensus reports • Pooled analyses • *Post hoc analyses • Non-systematic reviews • *Systematic reviews (including meta-analyses) published prior to 2014 • Secondary analyses • Animal models
Language restrictions	• *English-language publications	• *Non–English-language publications
Date restrictions	• 1999 to present	• None

Footnotes: [a]If the disease severity of included patients was not clearly stated in the article, the following approach was used and validated by Lilly: if DAS-28 scores were reported, then DAS-28 scores of > 3.2 were considered to be moderate RA; DAS-28 scores of > 5.1 were considered to be severe RA. If DAS-28 scores were not reported, then swollen and tender joint counts both > 6 was considered to be a good proxy for moderate to severe RA.

[b]Systematic reviews and meta-analyses will be used only for identification of primary studies that may have been missed in the electronic searches.
*Due to the high number of included studies from the abstract/title review, a secondary set of more stringent criteria were used to re-screen included studies.
†Additional criteria used during the full text review process.
Abbreviations: DMARD = disease-modifying antirheumatic drug, cDMARD = conventional DMARD, bDMARD = biologic DMARD, OD = once daily, TNF = tumour necrosis factor inhibitor, ACR = American College of Rheumatology, ACR20/50/70 = 20/50/70% improvement in ACR criteria, EULAR = European League Against Rheumatism, EULAR = EULAR response index, DAS28 = Disease Activity Score modified to include the 28 diarthrodial joint count, hsCRP = high-sensitivity C-reactive protein, ESR = erythrocyte sedimentation rate, SDAI = Simplified Disease Activity Index, CDAI = Clinical Disease Activity Index, MJS = morning joint stiffness, WJP = worst joint pain, FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue, mTSS = modified Total Sharp Score, WPAI-RA = Work Productivity and Activity Index-Rheumatoid Arthritis, EQ-5D-5L = EuroQoL 5 dimensions–5 levels, HAQ-DI = Health Assessment Questionnaire-Disability Index, SF-36v2 = Medical Outcomes Study 36-Item Short Form Health Survey Version 2 Acute, VAS = visual analogue scale, AE = adverse event

4.1.3 Critique of data extraction

Data were extracted by one reviewer and checked by a second (CS clarification response A1[32]) as is good practice. Data extracted for the four included BARI trials by the CS, and reported below, were checked by the ERG against published trial papers, and were found to be accurate. The ERG notes that there was a discrepancy between the EULAR response data reported in the CS and used in NMA. In response to a request for clarification (additional queries regarding data for the NMA), the company stated that the data reported in the CS with respect to EULAR is based on DAS28-hsCRP, but DAS28-ESR EULAR response was used in NMA because the majority of comparator studies in the NMA reported DAS28-ESR EULAR response. The company also noted that noted that 'the response rates for good/moderate responses are lower for DAS28-ESR and were determined via post-hoc analyses'.

The ERG believes that the data provided by the company for those with a moderate EULAR response in van de Putte *et al*.[33] also included those with a good EULAR response.

4.1.4 Quality assessment

Quality assessment of the three included BARI RCTs is presented in Section 4.6 and Appendix 9 of the CS. Quality assessment was conducted by one reviewer (CS clarification response A1). It is considered good practice for two reviewers either to independently assess quality or to check assessed items. Quality items assessed were taken from the Centre for Reviews and Dissemination (CRD) guidelines for undertaking reviews in health care;[34] these are standard and appropriate criteria for assessing the risk of bias in RCTs. Table 3 of this report presents the company's quality assessment of the BARI trials (reproduced from Table 22 of the CS). The ERG checked the company's quality assessment against the publications of the trials, RA-BEAM (Taylor *et al*. [35]), RA-BUILD (Dougados *et al.* [36]), RA-BEACON (Genovese *et al.* [37]), and if not clear from publications, CSRs were also consulted.[38-40]

Details of the generation of random sequences and the concealment of treatment allocation were not provided in the published trial papers, but were listed in Appendix 9 of the CS. Random sequences were generated by computer and allocation was concealed adequately by use of an interactive voice-response system for all three RCTs.[38-40] For the bDMARD-naïve trials RA-BEAM and RA-BUILD, randomisation was stratified by region and joint erosion status. For the bDMARD-experienced trial, RA-BEACON, randomisation was stratified by region and history of bDMARD use.

All three RCTs had blinding of patients, clinicians and outcome assessors. There were no unexpected imbalances in dropouts between treatment groups in any of the three trials. Outcomes that were not

available in published articles at the time of ERG report writing (e.g. EULAR response) were available in the CS as academic-in-confidence (AiC) data.

All three trials employed modified intention to treat (mITT) analyses for effectiveness measures, comprising all randomised patients who received at least one dose of study drug. Of 1307 patients randomised in RA-BEAM, 1305 (99.8%) were included in the mITT analyses; for RA-BUILD and RA-BEACON all randomised patients were included in the mITT analyses. Patients were analysed in their allocated treatment group. All three RCTs employed non-responder imputation analysis for categorical variables. Patients receiving rescue therapy (at, or after, week 16) were considered non-responders at the time of rescue and thereafter.

Table 3: Quality assessment of the BARI trials (reproduced from Table 22 of the CS)

Trial number (acronym)	**RA-BEAM [JADV] NCT01710358**	**RA-BUILD [JADX] NCT01721057**	**RA-BEACON [JADW] NCT01721044**
Was randomisation carried out appropriately?	Yes[38]	Yes [40]	Yes[39]
Was the concealment of treatment allocation adequate?	Yes[38]	Yes[40]	Yes[39]
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes[35]	Yes[36]	Yes[37]
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes[35]	Yes[36]	Yes[37]
Were there any unexpected imbalances in drop-outs between groups?	No[35]	No[36]	No[37]
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	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 (modified intent to treat)[35]	Yes (modified intent to treat)[36]	Yes (modified intent to treat)[37]
Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination)			

Quality assessments of the trials contained in the NMA are presented in Appendix 15 of the CS. The same quality assessment items were used, as is appropriate for RCTs.

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

Five BARI studies were identified by the CS search. These comprised four phase III RCTs (RA-BEAM, RA-BUILD, RA-BEACON), and one long-term extension study (RA-BEYOND). One of the RCTs (RA-BEGIN) had a population of cDMARD-naïve RA patients, and as this is an unlicensed treatment position, this RCT was not of relevance to the decision problem. Details of RA-BEGIN are presented in Appendix 1 of the CS. CS Appendix 19 additionally includes details of three phase II studies of BARI in RA patients.

Details of the three RCTs (RA-BEAM, RA-BUILD, RA-BEACON) and long-term extension study (RA-BEYOND) included in the CS are shown in Table 4 of the ERG (adapted from CS Table 12 and Section 4.3.6.1 of the CS). Both RA-BUILD and RA-BEACON had a 24-week randomised period, whereas RA-BEAM was randomised for 52-weeks, however, all patients randomised to PBO switched to BAR+MTX at week 24. Rescue therapy was permitted in all three RCTs, with patients whose tender joint count and swollen joint count were reduced by less than 20% from baseline at both week 14 and week 16 received open-label rescue treatment (BARI 4 mg) at week 16.[35,36,37] After week 16, rescue therapy was available based on investigator discretion. RA-BEAM and RA-BUILD had populations of cDMARD-experienced, bDMARD-naïve patients, whereas the population for RA-BEACON was bDMARD-experienced.

Table 4: Trial characteristics of included BARI trials (adapted from Table 12 of the CS)

Trial number (acronym)	Population	Number enrolled	Intervention	Comparators	Primary outcome
RA-BEAM (JADV)[35,38]	MTX-inadequate responders, bDMARD-naïve adult patients with moderately to severely RA	1307 randomised (Figure 9 CS) (1305 at least one dose, included in mITT) [35]	Baricitinib 4mg, oral, QD (with background MTX)	Adalimumab 40mg, SC injection, Q2W (with background MTX) Placebo (with background MTX)	proportion of patients achieving an ACR20 response at week 12
RA-BUILD (JADX)[36,40]	cDMARD-inadequate responders, bDMARD-naïve adult patients with moderately to severely active RA	684 randomised (Figure 10 CS)	Baricitinib (2mg, oral, QD) Baricitinib (4 mg, oral, QD). Patients on ≥1 cDMARDs (with or without MTX) continued to take background therapy during study.	Placebo (Patients on ≥1 cDMARDs (with or without MTX) continued to take background therapy during study)	proportion of patients achieving an ACR20 response at week 12
RA-BEACON (JADW)[37,39]	bDMARD inadequate responders adult patients with moderately to severely active RA	527 randomised (Figure 11 CS)	Baricitinib 2mg, oral, QD (with background cDMARDs) Baricitinib 4 mg, oral, QD (with background cDMARDs)	Placebo (with background cDMARDs)	proportion of patients achieving an ACR20 response at week 12
RA-BEYOND[41]	Patients with moderate to severe RA who completed Phase 2b study JADA or Phase 3 studies JADZ, JADV, JADX or JADW	[redacted] (from Table 58 CS)	Baricitinib (2mg, oral, QD) Baricitinib (4 mg, oral, QD)	N/A	long-term safety and tolerability of baricitinib

In the CS, two of the included RCTs (RA-BUILD and RA-BEACON) had published data available at the time of CS writing, whereas data from the other trials were available from Clinical Study Reports (CSRs) provided by the CS. An update search by the ERG identified an additional report of the RA-BEACON trial (Smolen *et al*. 2016),[42] and a full paper publication of the RA-BEAM study (Taylor *et al.* 2017).[35]

An ERG search of the U.S. National Institutes of Health clinical trials registry identified one RCT which was not mentioned in the CS (NCT02265705). This trial, which aims to evaluate the safety and effectiveness of BARI in patients with moderate to severe RA and who had an inadequate response to MTX, was ongoing (at the time of ERG report writing) with an estimated completion date of June 2017. The full list of BARI studies identified by the ERG are provided in Appendix 2.

Details of the three included RCTs

Eligibility criteria for the three included BARI RCTs are shown in Table 5 of the ERG reported (reproduced from Table 13 of the CS). All three RCTs (RA-BEAM, RA-BUILD, RA-BEACON) required a diagnosis of RA by ACR/EULAR 2010 criteria[16] and included patients with moderate to severe, active RA, as defined by the presence of at least 6/68 tender joints and at least 6/66 swollen joints. Further details of eligibility criteria are provided in CS Appendix 5.

For all three included trials, patients received BARI or placebo in combination with MTX (RA-BEAM) or other cDMARDs (RA-BUILD, RA-BEACON). In all three trials, patients were permitted HCQ or SSZ if receiving a stable dose for at least 8 weeks prior to study entry (LEF or AZA were additionally included for RA-BUILD and RA-BEACON); and NSAIDs and/or prednisone (≤10 mg) (or equivalent) were permitted if the patient was on a stable dose for at least 6 weeks prior to randomisation; as were current analgesics at a stable dose (CS Table 14). All three trials excluded live vaccines, non-stable doses of cDMARDs, bDMARD therapy, interferon therapy and parenteral corticosteroids (CS Table 14).

The trials were located in North America, South America, Europe and Asia. Of the 2518 patients randomised to the three RCTs, [redacted] were in the UK (Table 16 of the CS).

Table 5: Eligibility criteria (reproduced from Table 13 of the CS)

	RA-BEAM	RA-BUILD	RA-BEACON
Inclusion criteria	• Adults with a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA • The presence of at least 6/68 tender joints and at least 6/66 swollen joints • HsCRP measurement of ≥6 mg/L • At least 12 weeks of MTX therapy prior to study entry with 8 weeks at a stable dose (7.5–25 mg/week, but if <15 mg/week, documentation of clinical rationale should have been provided) • ≥3 joint erosions in hand, wrist or foot joints based on radiographs or ≥1 joint erosion and be RF or ACPA antibody positive	• Adults with a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA • The presence of at least 6/68 tender joints and at least 6/66 swollen joints • HsCRP measurement of ≥1.2 x ULN • Had failed treatment at an approved dose with more than one cDMARD (experienced insufficient efficacy or were intolerant to treatment)	• Adults with a diagnosis of adult-onset RA as defined by the ACR/EULAR 2010 Criteria for the Classification of RA • The presence of at least 6/68 tender joints and at least 6/66 swollen joints • HsCRP measurement of ≥ 1x ULN • Receiving stable doses of background cDMARD therapy • Had failed treatment at an approved dose with at least one biologic TNFi DMARD (experienced insufficient efficacy or were intolerant to treatment) • Patients who had received other bDMARDs could also participate
Exclusion criteria	• Receiving/had previously received prohibited RA therapies (bDMARDs) • Recent history of infection or tested positive for TB or other serious infections • Immunocompromised or had specific abnormal laboratory tests • Comorbidities that put patients at risk of adverse events when taking study drug	• Receiving/had previously received any bDMARD • Recent history of infection or tested positive for TB or other serious infections • Immunocompromised or had specific abnormal laboratory tests • Comorbidities that put patients at risk of adverse events when taking study drug	• Received bDMARDs within 28 days before randomisation (6 months before for rituximab) • Recent history of infection or tested positive for TB or other serious infections • Immunocompromised or had specific abnormal laboratory tests • Comorbidities that put patients at risk of adverse events when taking study drug

Baseline characteristics of the three RCTs are shown in Table 6 to Table 7 of the ERG report (reproduced from Tables 19-21 of the CS). Baseline characteristics within trials were balanced across trial arms.

Table 6: Baseline characteristics of participants of RA-BEAM (adapted from Table 19 of the CS and Taylor et al 2017[35])

RA-BEAM		PBO + MTX (n=488)	BARI 4 mg QD + MTX (n=487)	ADA 40 mg Q2W + MTX (n=330)
Gender, n (%)	Male	106 (21.7)	112 (23.0)	79 (23.9)
	Female	382 (78.3)	375 (77.0)	251 (76.1)
Age (years) [a]	Mean (SD)	53 (2)	54 (2)	53 (12)
	Median	54.5	55.0	54.5
	Range	19–83	23–80	20–86
Duration of RA symptoms (years) [35]	Mean (SD)	10 (9)	10 (9)	10 (9)
Time from diagnosis of rheumatoid arthritis (years) [b]	Mean	8.9	8.7	8.3
	SD	8.0	8.6	7.9
	Median	6.6	6.2	6.0
	Range	0.05–39.91	0.03–56.42	0.25–34.50
DAS-28(CRP)	Mean (SD)	5.7 (1.0)	5.8 (0.9)	5.8 (0.9)
	Median	5.61	5.75	5.75
	Range	2.91–8.43	3.06–8.04	3.48–7.97
DAS-28(ESR)	Mean (SD)	6.4 (1.0)	6.5 (0.9)	6.4 (1.0)
	Median	6.35	6.45	6.41
	Range	3.29–9.07	3.51–8.81	3.84–8.99
HAQ-DI	Mean	1.55	1.57	1.59
	SD	0.67	0.68	0.70
	Median	1.50	1.56	1.63
	Range	0.0–3.0	0.0–3.0	0.0–3.0
ACPA-positive, n (%)		424 (87)	427 (88)	295 (89)
mTSS	Mean	45	43	44
	SD	50	50	51
	Median	23.25	21.50	25.50
	Range	0.0–300.5	0.0–284.5	0.5–309.5
RF-positive, n (%)		451 (92)	439 (90)	301 (91)
Mean weekly does MTX		14.8 mg across all groups (CS clarification response A3)		

Abbreviations: QD = once daily, Q2W = twice weekly, DAS28 = Disease Activity Score, CRP = C-reactive protein, ESR = erythrocyte sedimentation rate, HAQ-DI = Health Assessment Questionnaire–Disability Index, ACPA = Anti-citrullinated protein antibody, mTSS = modified Total Sharp Score, RF = Rheumatoid factor.
Sources: a = JADV.11.1, b = JADV.11.2.

Table 7: Baseline characteristics of participants of RA-BUILD (reproduced from Table 20 of the CS)[36]

RA-BUILD		PBO QD + MTX (n=228)	BARI 2 mg QD + MTX (n=229)	BARI 4 mg QD + MTX (n=227)
Gender, n (%)	Male	39 (17.1)	45 (19.7)	40 (17.6)
	Female	189 (82.9)	184 (80.3)	187 (82.4)
Age (years)	Mean	51.4	52.2	51.8
	SD	12.5	12.3	12.1
	Median	53.0	52.0	53.0
	Range	21–79	22–82	20–80
Duration of RA symptoms	Mean (SD)	7 (8)	8 (8)	8 (8)
Time from diagnosis of rheumatoid arthritis (years)	Mean	5.9	6.5	6.4
	SD	6.8	7.6	7.5
	Median	3.4	3.6	3.7
	Range	0.07–37.44	0.28–52.76	0.11–41.40
DAS-28(CRP)	Mean	5.53	5.57	5.55
	SD	0.91	0.96	0.87
	Median	5.50	5.49	5.53
	Range	2.27–7.50	3.05–8.03	3.30–7.91
DAS-28(ESR)	Mean	6.19	6.28	6.20
	SD	1.00	0.99	0.91
	Median	6.18	6.25	6.26
	Range	2.90–8.63	3.31–8.52	3.96–8.44
HAQ-DI	Mean	1.5	1.51	1.55
	SD	0.60	0.62	1.60
	Median	1.50	1.50	1.50
	Range	0.0–2.8	0.0–2.9	0.0–3.0
ACPA-positive, n (%)		172 (75.4)	169 (73.8)	163 (71.8)
mTSS	Mean	18.54	25.78	23.71
	SD	31.47	40.26	40.01
	Median	6.00	8.50	6.25
	Range	0.0–241.5	0.0–218.0	0.0–231.0
RF-positive, n (%)		171 (75.0)	177 (77.3)	173 (76.2)
Mean weekly does MTX		16.2 mg across all groups (n=684) (CS clarification response A3)		

Abbreviations: QD = once daily, DAS28 = Disease Activity Score, CRP = C-reactive protein, ESR = erythrocyte sedimentation rate, HAQ-DI = Health Assessment Questionnaire–Disability Index, ACPA = Anti-citrullinated protein antibody, mTSS = modified Total Sharp Score, RF = Rheumatoid factor.

Table 7: Baseline characteristics of participants of RA-BEACON (reproduced from Table 21 of the CS)[37]

RA-BEACON		PBO QD + MTX (n=176)	BARI 2 mg QD + MTX (n=174)	BARI 4 mg QD + MTX (n=177)
Gender, n (%)	Male	31 (17.6)	37 (21.3)	28 (15.8)
	Female	145 (82.4)	137 (78.7)	149 (84.2)
Age (years)	Mean	56.0	55.1	55.9
	SD	10.7	11.1	11.3
	Median	57.0	55.0	58.0
	Range	24–77	21–82	24–82
Duration of RA symptoms years	Mean (SD)	14 (10)	14 (8)	14 (9)
Time from diagnosis of rheumatoid arthritis (years)	Mean	12.8	12.3	12.5
	SD	9.4	7.5	8.7
	Median	10.4	11.1	9.8
	Range	0.62–50.70	1.03–38.04	0.64–37.53
DAS-28(CRP)	Mean	5.89	6.03	5.87
	SD	0.94	0.89	1.00
	Median	5.80	5.99	5.83
	Range	3.64-8.24	3.94-8.07	3.31-8.06
DAS-28(ESR)	Mean	6.59	6.70	6.58
	SD	0.93	0.98	1.06
	Median	6.55	6.74	6.67
	Range	4.58-8.82	4.19-8.74	3.81-8.86
HAQ-DI	Mean	1.78	1.71	1.74
	SD	0.57	0.55	0.59
	Median	1.88	1.75	1.75
	Range	0.48-3.0	0.0-3.0	0.0-3.0
ACPA-positive, n (%)		125 (71.4)	124 (71.3)	119 (67.2)
mTSS	Mean	NR	NR	NR
	SD	NR	NR	NR
	Median	NR	NR	NR
	Range	NR	NR	NR
RF-positive, n (%)		130 (73.9)	128 (73.6)	128 (72.3)
Mean weekly does MTX		16.3 mg across all groups (CS clarification response A3)		

Abbreviations: OD = once daily, DAS28 = Disease Activity Score, CRP = C-reactive protein, ESR = erythrocyte sedimentation rate, HAQ-DI = Health Assessment Questionnaire–Disability Index, ACPA = Anti-citrullinated protein antibody, mTSS = modified Total Sharp Score, RF = Rheumatoid factor.

Effectiveness results of the three included RCTs

ACR response data

ACR response data for the three included RCTs are shown in Table 8, Table 9 and Table 10. The primary outcome for all three RCTs was the proportion of patients achieving an ACR20 response at week 12. For this outcome, all three RCTs found a statistically significant advantage for BARI 4mg over PBO: RA-BEAM (70% vs 40% $p<0.001$), RA-BUILD (66%% vs 40% $p\leq0.001$), RA-BEACON (49% vs 27% $p\leq0.001$). BARI 2mg was also reported to have a significant advantage over PBO in RA-BUILD and RA-BEACON ($p\leq0.001$). RA-BEAM examined non-inferiority of BARI compared to ADA and planned that if non-inferiority was shown, superiority would be assessed. Within this analysis, BARI was considered to have a statistically significant ($p=0.01$) advantage over ADA for ACR20 at 12 weeks (BARI 70%, ADA 61%). ADA was also significantly superior to PBO ($p\leq0.001$).

In RA-BEAM at week 12, ACR20 subgroup analysis of baseline DAS-CRP (≤5.1 versus >5.1) showed no significant interaction with treatment group (BARI and PBO $p=0.967$; BARI and ADA $p=0.249$). In RA-BUILD at week 12, ACR20 subgroup analysis of baseline DAS-CRP (≤5.1 versus >5.1) showed no significant interaction with treatment group (BARI 4mg and PBO $p=0.158$; BARI 2MG and PBO $p=0.080$). In RA-BUILD, subgroup analysis of background cDMARDs showed no statistically significant interaction with treatment group for ACR20 or ACR50 at 12 weeks or 24 weeks follow-up ($p\geq0.271$).

As well as the primary endpoint, at 12 weeks follow-up, BARI 2mg and 4mg also had an advantage over PBO ($p\leq0.01$) for ACR50 and ACR70 responses across the three RCTs. At 24 weeks follow-up, ACR responses for BARI 4mg were significantly better than PBO in all three RCTs ($p\leq0.001$).

Table 8: ACR results RA-BEAM (non-responder imputation) (adapted from CS Tables 23 - Table 25, and CS Table 60 and CS Table 61, and Taylor et al 2017) [35]

Outcome measure	12 weeks			24 weeks			52 weeks		
	PBO (N=488)	BARI 4mg (N=487)	ADA (N=330)	PBO (N=488)	BARI 4mg (N=48	ADA (N=330)	PBO (n=452)	BARI 4mg (N=48	ADA (N=330)
ACR20 (%)[a,b,c]	40	70***+	61***	37	74***+	66***	N/A	71++	62
ACR50 (%)[d]	17	45***++	35***	19	51***	45***	N/A	56+	47
ACR70 (%)	5	19***+	13***	8	30***+	22***	N/A	37	31
Comparison		BARI 4MG vs PBO	BARI 4MG vs ADA						
ACR20 Odds ratio (95% CI)		3.6 (2.7, 4.7) p=0.001	1.5 (1.1, 2.0) p=0.014						
ACR50 Odds ratio (95% CI)		4.2 (3.1, 5.7) p=0.001	1.5 (1.1, 2.1) p=0.005						
ACR70 Odds ratio (95% CI)		4.8 (3.0, 7.8) p=0.001	1.6 (1.1, 2.4) p=0.026						
Subgroup DAS28-hsCRP ≤5.1 ACR20 n/N (%)[a]	[redacted]	[redacted]	[redacted]						
Subgroup DAS28-hsCRP >5.1 ACR20 n/N (%)[a]	[redacted]	[redacted]	[redacted]						

Footnotes: *$p \leq 0.05$, **$p \leq 0.01$, and ***$p \leq 0.001$ versus placebo, and +$p \leq 0.05$, ++$p \leq 0.01$, +++$p \leq 0.001$ versus adalimumab using logistic regression, without control for multiple comparisons.[35]

[a]Note that randomisation was not stratified by DAS28

Table 9: ACR data RA-BUILD (non-responder imputation) (adapted from CS Table 35-37, and CS Table 36 and CS Table 37, and CS Table 62 and CS Table 63)

Outcome measures	12 weeks			24 weeks		
	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)
ACR20 (%)	39.5	65.9***	61.7***	42.1	61.1***	65.2***
ACR50 (%)	12.7	33.6***	33.5***	21.5	41.5***	44.1***
ACR70 (%)	3.1	17.9***	18.1***	7.9	25.3***	24.2***
Comparison		BARI 2MG vs PBO	BARI 4MG vs PBO			
ACR20 Odds ratio (95% CI)		3.0 (2.0, 4.4) p=0.001	2.5 (1.7, 3.7) p=0.001			
ACR50 Odds ratio (95% CI)		3.5 (2.2, 5.6) p=0.001	3.5 (2.2, 5.7) p=0.001			
ACR70 Odds ratio (95% CI)		6.9 (3.0, 15.9) p=0.001	7.2 (3.2, 16.6) p=0.001			
Subgroup ACR20 DAS28-hsCRP ≤5.1 n/N (%)[a]	[redacted]	[redacted]	[redacted]			
Subgroup ACR20 DAS28-hsCRP >5.1 n/N (%)[a]	[redacted]	[redacted]	[redacted]			
Background cDMARD(s) None ACR20 n/N (%)[a] [36]	2/17 (12)	11/18 (61)	7/13 (54)	2/17 (12)	9/18 (50)	7/13 (54)
Background cDMARD(s) MTX only ACR20 n/N (%)[a] [36]	45/109 (41)	78/111 (70)	72/114 (63)	48/109 (44)	72/111 (65)	76/114 (67)
Background cDMARD(s) Non MTX cDMARDs ACR20 n/N (%)[a] [36]	17/44 (39)	26/41 (63)	26/43 (60)	17/44 (39)	24/41 (59)	23/43 (53)
Background cDMARD(s) MTX + other cDMARDs ACR20 n/N (%)[a] [36]	26/58 (45)	36/59 (61)	35/57 (61)	29/58 (50)	35/59 (59)	42/57 (74)
Background cDMARD(s) None	2/17 (12)	4/18 (22)	2/13 (15)	2/17 (12)	7/18 (39)	5/13 (38)

Outcome measures	12 weeks			24 weeks		
	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)
ACR50 n/N (%)[a] [36]						
Background cDMARD(s) MTX only ACR50 n/N (%)[a] [36]	14/109 (13)	40/111 (36)	42/114 (37)	22/109 (20)	47/111 (42)	48/114 (42)
Background cDMARD(s) Non MTX cDMARDs ACR50 n/N (%)[a] [36]	7/44 (16)	18/41 (44)	14/43 (33)	10/44 (23)	15/41 (37)	17/43 (40)
Background cDMARD(s) MTX + other cDMARDs ACR50 n/N (%)[a] [36]	6/58 (10)	15/59 (25)	18/57 (32)	15/58 (26)	26/59 (44)	30/57 (53)

Footnotes: Significance level definitions: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ vs. placebo
[a]Note that randomisation was not stratified by DAS28 or background cDMARDs

Table 10: ACR data RA-BEACON (using non-responder imputation) (adapted from Tables 47 - Table 49 of the CS, and Tables 64 -67 of the CS)

Outcome measures	12 weeks			24 weeks		
	PBO N=176	BARI 2MG N=174	BARI 4MG N=177	PBO N=176	BARI 2MG N=174	BARI 4MG N=177
ACR20 (%)	27.3	48.9***	55.4***	27.3	44.8***	46.3***
ACR50 (%)	8.0	20.1**	28.2***	13.1	23.0*	29.4***
ACR70 (%)	2.3	12.6***	11.3**	3.4	13.2***	16.9***
Comparison		BARI 2MG vs PBO	BARI 4MG vs PBO			
ACR20 Odds ratio (95% CI)		2.7 (1.7, 4.2) p=0.001	3.4 (2.2, 5.4) p=0.001			
ACR50 Odds ratio (95% CI)		3.0 (1.6, 5.9) p=0.002	4.7 (2.5, 8.9) p=0.001			
ACR70 Odds ratio (95% CI)		NR Table 49 p=0.001	NR Table 49 p=0.002			
Subgroup DAS28-hsCRP ≤5.1 ACR20 n/N (%)[a]	[redacted]	[redacted]	[redacted]			
Subgroup DAS28-hsCRP >5.1 ACR20 n/N (%)[a]	[redacted]	[redacted]	[redacted]			
subgroup with number of previous bDMARDs used <3 ACR20 n/N (%)[a]	42/129 (32.6%)	66/124 (53.2%)	74/132 (56.1!)			
subgroup with number of previous bDMARDs used ≥3 ACR20 n/N (%)[a]	6/47 (12.8%)	19/50 (38.0%)	24/45 (53.3%)			

Footnotes: Significance level definitions: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ vs placebo.
[a]Note that randomisation was stratified by number of previous bDMARDs used (<3 or ≥3), but was not stratified by DAS28

EULAR response data

At 12 weeks follow-up, all three RCTs reported a significant advantage for BARI 4MG over PBO for EULAR response ($p \leq 0.001$) (good or moderate EULAR response RA-BEAM [redacted]; RA-BUILD 79.0% vs 53.5%; RA-BEACON 66.1% vs 42.6%). There was an advantage for BARI 4MG over PBO at 24 weeks follow-up for the bDMARD-naïve studies ($p<0.001$), and also for the bDMARD-experienced population ($p<0.05$). BARI 2MG was superior to PBO for EULAR response at 12 weeks ($p<0.001$) for both the bDMARD-naïve and bDMARD-experienced populations. At 24 weeks, BARI 2MG was significantly superior to PBO in the bDMARD-naïve ($p<0.001$) and bDMARD-experienced ($p<0.05$) populations.

Table 11: EULAR results RA-BEAM (non-responder imputation) (adapted from CS Table 23 and CS Table 26)

Outcome measure	12 weeks			24 weeks			52 weeks		
	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	PBO (n=452)	BARI 4MG (N=487)	ADA (N=330)
EULAR (good + moderate) response rate (%)	[redacted]								
EULAR (good) response rate (%)	[redacted]								
Comparison		BARI 4MG vs PBO	BARI 4MG vs ADA						
EULAR good and moderate response Odds ratio (95% CI)		[redacted]							
EULAR good response Odds ratio (95% CI)		[redacted]							

Footnotes: Significance level definitions: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ vs placebo.

Table 12: EULAR data RA-BUILD (using non-responder imputation (adapted from CS Table 35 and CS Table 38)

Outcome measures	12 weeks			24 weeks		
	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)
EULAR (good + moderate) response rate (%)	53.5	79.0***	79.3***	53.5	72.1***	78.0***
EULAR (good) response rate (%)	15.4	34.1***	38.3***	21.9	45.4***	50.7***
Comparison		BARI 2MG vs PBO	BARI 4MG vs PBO			
EULAR good and moderate response Odds ratio (95% CI)		3.3 (2.2, 5.0) p=0.001	3.5 (2.3, 5.4) p=0.001			
EULAR good response Odds ratio (95% CI)		2.9 (1.8, 4.6) p=0.001	3.6 (2.3, 5.7) p=0.001			

Footnotes: Significance level definitions: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ vs. placebo

Table 13: EULAR data RA-BEACON (adapted from CS Table 47 and CS Table 50 of the CS)

Outcome measures	12 weeks			24 weeks		
	PBO N=176	BARI 2MG N=174	BARI 4MG N=177	PBO N=176	BARI 2MG N=174	BARI 4MG N=177
EULAR (good + moderate) response rate (%)	42.6	66.1***	72.3***	37.5	54.0**	59.9***
EULAR (good) response rate (%)	8.5	24.1***	29.9***	11.4	20.1*	31.6***
Comparison		BARI 2MG vs PBO	BARI 4MG vs PBO			
EULAR good and moderate response Odds ratio (95% CI)		2.7 (1.8, 4.2) P=0.001	3.6 (2.3, 5.7) P=0.001			
EULAR good response Odds ratio (95% CI)		3.6 (1.9, 6.8) P=0.001	4.8 (2.6, 9.0) P=0.001			

Footnotes: Significance level definitions: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ vs placebo.

Other efficacy outcomes

At 12 weeks follow-up, BARI 4MG was statistically significantly better than PBO across the three RCTs for DAS28-CRP, HAQ-DI, SDAI low disease activity (LDA), CDAI LDA and MJS duration (see Table 14, Table 15 and Table 16, adapted from CS Tables 23, 35 and 47 and Taylor *et al.*)[35] RA-BEAM planned a statistical comparison of BARI 4MG and ADA at week 12 for DAS28-CRP, and this was found to significantly favour BARI 4MG ($p \leq 0.01$). There was also an advantage for BARI 4MG over PBO at 24 weeks on several measures, and BARI 2MG over PBO at 12 weeks and 24 weeks on several measures.

Table 14: RA-BEAM (adapted from Table 23 of the CS)

Outcome measure	12 weeks			24 weeks			52 weeks	
	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	BARI 4MG (N=487)	ADA (N=330)
DAS28-hsCRP (≤3.2) response rate (%)[g] [35]	14	44***++	35***	19	52***	48***	56+	48
DAS28-hsCRP (<2.6) response rate (%)[g]	4	24***	19***	8	34***	32***	40	39
HAQ-DI CFB LSM (SE)[h]	-0.34 (0.026)	-0.66***++ (0.026)	-0.56*** (0.030)	-0.35 (0.028)	-0.75***++ (0.028)	-0.63*** (0.033)	-0.77***++ (0.031)	-0.66 (0.036)
ΔmTSS CFB LSM (SE)[i]	N/A	N/A	N/A	0.90 (0.10)	0.41*** (0.10)	0.33*** (0.11)	0.71*** (0.18)	0.60*** (0.22)
SDAI LDA (≤11.0) response rate (%)[j] [35]	16	42***+	35***	20	51***	48***	57+	49
SDAI remission (≤3.3) response rate (%)[j] [35]	2	8***	7***	3	16***	14***	23	18
CDAI LDA (≤10.0) response rate (%)[k] [35]	17	40***+	33***	20	50***	48***	57+	49
CDAI remission (≤2.8) response rate (%)[k] [35]	2	8***	7**	4	16***	12***	22	18
FACIT-F (MCID)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Outcome measure	12 weeks			24 weeks			52 weeks	
	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	BARI 4MG (N=487)	ADA (N=330)
improvement ≥ 3.56 (%)[m]								
FACIT-F CFB LSM (SE)[m]	6.7 (0.42)	9.1*** (0.42)	8.7*** (0.49)	6.5 (0.46)	10.0*** (0.45)	9.3*** (0.54)	10.7+ (0.46)	9.3 (0.54)
MJS Duration (min)[n]	60.0	27.1***+	36.6***	N/A	N/A	N/A	N/A	N/A
MJS Severity LSM (SE)[o]	4.1 (0.10)	3.0***++ (0.10)	3.5*** (0.12)	N/A	N/A	N/A	N/A	N/A

*p≤0.05, **p≤0.01, ***p≤0.001 baricitinib-4 mg or adalimumab versus placebo and +P≤0.05, ++P≤0.01, +++P≤0.001 baricitinib 4-mg versus adalimumab by logistic regression, without control for multiple comparisons.

Table 15: RA-BUILD (adapted from CS Table 35)

Outcome measures	12 weeks			24 weeks		
	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)
DAS28-hsCRP (≤3.2) response rate (%)[e]	17.1	35.8***	39.2***	23.7	46.3***	51.5***
DAS28-hsCRP (<2.6) response rate (%)[e]	8.8	25.8***	25.6***	10.5	30.6***	33.0***
HAQ-DI CFB LSM (SE)[f]	-0.34 (0.037)	-0.54*** (0.036)	-0.53*** (0.037)	-0.35 (0.040)	-0.58*** (0.039)	-0.58*** (0.040)
ΔmTSS CFB LSM (SE)[g]	N/A	N/A	N/A	0.70 (0.14)	0.33* (0.14)	0.15** (0.14)
SDAI LDA (≤11.0) response rate (%)[h]	19.7	33.2**	34.8***	28.5	48.0***	52.4***
SDAI remission (≤3.3) response rate (%)[h]	0.9	9.2***	8.8***	3.9	16.6***	15.0***
CDAI LDA (≤10.0) response rate (%)[i]	20.6	34.5**	34.8***	27.6	45.4***	52.0***
CDAI remission (≤2.8) response rate (%)[i]	1.8	10.0***	9.3***	3.9	15.3***	15.4***
FACIT-F (MCID) improvement ≥ 3.56 (%)[k]	58.8	63.3	64.8	42.5	59.0***	59.9***
FACIT-F CFB LSM (SE)[k]	7.5 (0.64)	8.5 (0.61)	9.1 (0.64)	7.9 (0.67)	9.2 (0.64)	10.1* (0.67)
MJS Duration (min)[l]	60.0	44.4**	34.6***	N/A	N/A	N/A

Outcome measures	12 weeks			24 weeks		
	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)
MJS Severity LSM (SE)[m]	4.1 (0.15)	3.5** (0.15)	3.4*** (0.16)	N/A	N/A	N/A

Table 16: RA-BEACON (adapted from CS Table 47)

Outcome measures	12 weeks			24 weeks		
	PBO N=176	BARI 2MG N=174	BARI 4MG N=177	PBO N=176	BARI 2MG N=174	BARI 4MG N=177
DAS28-hsCRP (≤3.2) response rate (%)	9.1	24.1***	31.6***	11.4	20.1*	33.3***
DAS28-hsCRP (<2.6) response rate (%)	4.0	10.9*	16.4***	6.3	10.9	21.5***
HAQ-DI CFB LSM (SE)	-0.17 (0.04)	-0.37*** (0.04)	-0.40*** (0.04)	-0.15 (0.05)	-0.37*** (0.05)	-0.42*** (0.05)
SDAI LDA (≤11.0) response rate (%)	9.1	22.4***	28.2***	14.2	22.4*	31.1***
SDAI remission (≤3.3) response rate (%)	1.7	2.3	5.1	2.3	4.6	9.0**
CDAI LDA (≤10.0) response rate (%)	10.8	23.6**	27.7***	15.3	23.0	31.1***
CDAI remission (≤2.8) response rate (%)	1.7	2.9	5.6	3.4	4.6	9.0*
FACIT-F (MCID) improvement ≥ 3.56 (%)	48.3	63.8**	62.7**	37.5	50.0*	52.5**
FACIT-F CFB LSM (SE)	5.2 (0.9)	8.3 (0.9)**	8.1 (0.9)**	5.7 (0.9)	8.1 (0.9)*	9.2 (0.9)**
MJS Duration (min)	-3.5	-21.0**	-24.0***	-8.0	-25.5**	-27.0**

Health-related quality of life (HRQoL)

All three RCTs reported a significant advantage in EQ-5D-5L of BARI 4MG over PBO ($p \leq 0.001$) at 12 weeks and 24 weeks follow-up (see Table 17, Table 18 and Table 19). BARI 2MG was also statistically significantly superior to PBO ($p \leq 0.01$) at 12 and 24 weeks.

Table 17: RA-BEAM (adapted from CS Table 23)

Outcome measure	12 weeks			24 weeks			52 weeks		
	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	PBO (n=452)	BARI 4MG (N=487)	ADA (N=330)
EQ-5D-5L CFB LSM (SE)	0.102 (0.009)	0.184*** (0.009)	0.167*** (0.011)	0.088 (0.010)	0.199*** (0.010)	0.175*** (0.012)	N/A	0.217+ (0.010)	0.182 (0.012)

EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels
**$p \leq 0.01$, ***$p \leq 0.001$ vs placebo
+ $p \leq 0.05$; vs adalimumab.

Table 18: RA-BUILD (adapted from CS Table 35)

Outcome measures	12 weeks			24 weeks		
	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)
EQ-5D-5L CFB LSM (SE)	0.092 (0.014)	0.165*** (0.013)	0.162***(0.014)	0.091 (0.014)	0.157*** (0.014)	0.186*** (0.014)

EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels
***$p \leq 0.001$ vs placebo

Table 19: RA-BEACON (adapted from CS Table 47)

Outcome measures	12 weeks			24 weeks		
	PBO N=176	BARI 2MG N=174	BARI 4MG N=177	PBO N=176	BARI 2MG N=174	BARI 4MG N=177
EQ-5D-5L CFB LSM (SE)	0.036 (0.019)	0.114*** (0.019)	0.169*** (0.018)	0.038 (0.019)	0.111** (0.019)	0.159*** (0.019)

EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels
**$p \leq 0.01$, ***$p \leq 0.001$ vs placebo

RA-BEYOND study

The RA-BEYOND study (detailed in CS Section 4.7.4) recruited participants from a number of sources: the three RCTs included in the CS (RA-BEAM, RA-BUILD, RA-BEACON); a Phase III study of baricitinib in MTX-naïve patients (RA-BEGIN); and a Phase II study of baricitinib (JADA). Results for the patients from the three included RCTs in the CS who were recruited to RA-BEYOND are shown in Table 20 of the ERG (reproduced from Table 58 of the CS).

Table 20: Effectiveness results of RA-BEYOND (reproduced from CS Table 58)

Outcome measure*		**RA-BEAM BARI 4MG N=[redacted]**	**RA-BUILD**		**RA-BEACON**	
			BARI 2MG N=[redacted]	**BARI 4MG N=[redacted]**	**BARI 2MG N=[redacted]**	**BARI 4MG N=[redacted]**
ACR20 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ACR50 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ACR70 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DAS28 (hsCRP) <2.6 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DAS28(hsCRP) ≤3.2 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CDAI ≤2.8 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CDAI ≤10.0 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SDAI ≤3.3 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SDAI ≤11 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ-DI improvement ≥0.3 (%)	Week 12 of originating study[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	48 weeks after entry into RA-BEYOND[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: JADY CSR.[43]

Footnotes: [a]NRI (rescue not available at week 12), [b]NRI without considering rescue status, *Baseline in the originating study is used in the response rate calculation. The time points are weeks since randomisation in the originating study. Analyses exclude patients who were rescued or switched in the originating studies. RA-BEYOND populations analysed here only include patients who have completed 48 weeks of RA-BEYOND or would have completed 48 weeks if not discontinued. As such, not all patients from the originating study were included in analyses, leading to a different sample size than the patient population from the original study. Therefore, Week 12 results presented in the table above may differ to those presented earlier for the respective originating study due to the difference in sample sizes, which affects the proportion of patients achieving an outcome measure. Data after patients step down to baricitinib 2 mg are imputed based on the model predicted values using data from baricitinib treatment period in the originating and RA-BEYOND studies. NRI without considering rescue is used to impute missing data. Note: Baseline in the originating study is used in the response rate calculation. The time points are weeks since randomisation in the originating study. Data after patients' step-down to baricitinib 2 mg are imputed based on the model predicted values using data from baricitinib treatment period in the originating and RA-BEYOND studies. NRI without considering rescue is used to impute missing data. Note: One year after entry in RA-BEYOND is Week 100 for the 52-week studies (RA-BEGIN and RA-BEAM), Week 72 for the 24-week studies (RA-BUILD and RA-BEACON)
Abbreviations: ACR20 = 20% improvement in American College of Rheumatology Criteria; ACR50 = 50% improvement in American College of Rheumatology Criteria; ACR70 = 70% improvement in American College of Rheumatology Criteria; BAR = baricitinib; CDAI = Clinical Disease Activity Index; DAS28–hsCRP = Disease Activity Score in 28 joints high-sensitivity C-reactive protein; DAS28-ESR = Disease Activity Score in 28 joints-erythrocyte sedimentation rate; HAQ-DI = Health Assessment Questionnaire-Disability Index; NRI = non-responder imputation; PGA = Physician's Global Assessment of Disease Activity; PtGA = Patient's Global Assessment of Disease Activity; SDAI = Simplified Disease Activity Index

Adverse event data

The CS provided AE data for patients who received BARI 4mg from the Phase III trials RA-BEAM, RA-BUILD, RA-BEACON; and the Phase II trials JADA, JADC and JADN (Table 21, reproduced from Table 92 of the CS). The CS provided data on major adverse cardiovascular events (MACE), malignancies, and infections from a combined analysis set including patients on 2mg as well as 4mg BARI (CS Section 4.12). The CS also provided serious adverse events (SAE) data from RA-BEAM (Table 95 of the CS).

Table 21: Overview of tolerability profile up to week 24 (BARI 4 mg from studies JADA, JADC, JADN, RA-BEAM, RA-BUILD, RA-BEACON) (reproduced from CS Table 92)

Adverse event, n (%) [EAIR]	Baricitinib (4 mg QD) (n=997)	Placebo (n=1070)
Overall treatment-emergent adverse events	695 (69.7) [169.8]	659 (61.6) [167.3]
Severe treatment-emergent adverse events	53 (5.3) [12.9]	43 (4.0) [10.9]
Serious adverse events*	53 (5.3) [12.9]	50 (4.7) [12.7]
Permanent discontinuation due to adverse events/death	47 (4.7) [11.5]	35 (3.3) [8.9]
Temporary interruption due to an adverse event	109 (10.9) [27.1]	89 (8.3) [23.0]
Death	3 [0.7]	2 [0.5]

Footnotes: Treatment adverse events were defined as adverse events that either first occurred or worsened in severity after the first dose of study treatment. Patients with multiple occurrences of the same event are counted under the highest severity. *Defined as any AE associated with a patient outcome that met the International Conference on Harmonisation E2A criteria for an SAE.[44]
Sources: Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.34. Page 197).[45] 2016.[45] Eli Lilly Data on File (Clinical Overview. Rheumatoid Arthritis. EMA Submission. Table 2.5.5.2. Page 64). 2016[46, 47]
Abbreviation: EAIR = exposure-adjusted incidence rate, QD = once daily

The summary of product characteristics (SmPC) for BARI from the European Medicines Agency (EMA)[48] states that the most common AEs for BARI were LDL cholesterol (33.6 %), upper respiratory tract infections (14.7 %) and nausea (2.8 %). Table 2 of the SmPC provides the frequency estimates of AEs for BARI based on 3,464 BARI treated patients (see Table 22). Treatment-emergent adverse events by Medical Dictionary for Regulatory Activities (MedDRA) preferred term within system organ class of BARI 4mg and PBO up to week 24 are shown in Table 23 (reproduced from BARI EPAR Table 36). EMA, as reported in the BARI EPAR,[48] considered the following to be adverse drug reactions of BARI: nausea; upper respiratory tract infections; herpes simplex; herpes zoster; acne; increased creatine phosphokinase; increased LDL cholesterol and triglycerides; increased liver function tests (aspartate transaminase (AST), alanine transaminase (ALT)); neutropenia and thrombocytosis[48]. There was a significantly higher rate of infections for BARI 4mg treated patients than for PBO treated patients (36.3% versus 27.9%, $p<0.001$), however when considering only infections defined as serious according to International Conference on Harmonisation criteria, there was no significant difference between groups (BARI 4mg 1.5%, PBO 1.6%) as reported in the BARI EPAR.[48] The most commonly reported infections were upper respiratory tract infections, herpes zoster and herpes simplex.

Upper respiratory tract infections (URTIs) were reported in 14.7% BARI 4mg treated patients, and 11.7% PBO patients (CS Section 4.12.1.5 and EPAR).[48] There were significantly more non-serious herpes simplex (BARI 4mg 1.8%; PBO 0.7%) and herpes zoster infections (BARI 4mg 1.4%; PBO 0.4%) with BARI compared to PBO (CS Section 4.12.1.5, and BARI EPAR [48]. The overall exposure adjusted incidence rate (EAIR) of tuberculosis in RA patients treated with BARI 4mg once daily was 0.20 events per 100 patient years (CS Section 4.12.1.5 and BARI EPAR).[48]

For patients on BARI 2mg or BARI 4mg, the overall incidence rate of positively adjudicated MACE was 0.46 per 100 person-years (see CS Section 4.12.1.3). [redacted] patients developed non-melanoma skin cancer (see CS Section 4.12.1.4). Other malignancies were reported in 38 cases (0.73 per 100 person years(BARI EPAR[48]).

Clinical advice received by the ERG suggested that it would be prudent to make sure arrangements are in place to pick up any safety signal, if one manifested. Whilst there are no data to show major adverse events, the absence of evidence is not definitive evidence of absence.

Table 22: Adverse Reactions Frequency estimate: Very common (≥ 1/10), common (≥ 1/100 to < 1/10), uncommon (≥ 1/1,000 to < 1/100). (Reproduced from BARI SmPC)[48]

System Organ Class	Very common	Common	Uncommon
Infections and infestations	Upper respiratory tract infections[a]	Herpes zoster, Herpes simplex[b] Gastroenteritis Urinary tract infections	
Blood and lymphatic system disorders		Thrombocytosis >600 x 109 cells/L[c]	Neutropaenia <1 x 109 cells/L[c]
Metabolism and nutrition disorders	Hypercholesterolaemia[c]		Hypertriglyceridaemia[c]
Gastrointestinal disorders		Nausea	
Hepatobiliary disorders		ALT increased ≥3 x ULN[c]	AST increased ≥3 x ULN[c]
Skin and subcutaneous tissue disorders			Acne
Investigations			Weight increased Creatine phosphokinase increased >5 x ULN[c]

[a] Combined term (acute sinusitis, epiglottitis, laryngitis, nasopharyngitis, oropharyngeal pain, pharyngitis, pharyngotonsillitis, rhinitis, sinusitis, tonsillitis, tracheitis, upper respiratory tract infection).
[b] Combined term (eczema herpeticum, herpes simplex, ophthalmic herpes simplex, and oral herpes).
[c] Includes changes detected during laboratory monitoring

Table 23: Treatment-Emergent Adverse Events by MedDRA Preferred Term within System Organ Class of BARI 4-mg and PBO up to week 24 (most frequent preferred terms selected by CHMP) (reproduced from BARI EPAR Table 36)[48]

System Organ Class Preferred Term	PBO (N=1070) (Person Years Exposure=393.8) n (%) [Person Years]	BARI 4-mg (N=997) (Person Years Exposure=409.4) n (%)[Person Years]
Patients with >= 1 treatment emergent adverse event	659 (61.6) [167.3]	695 (69.7) [169.8]
Blood and lymphatic system disorders	48 (4.5) [12.2]	54 (5.4) [13.2]
Cardiac Disorders	8 (0.7) [2.0]	13 (1.3) [3.2]
Vascular disorders	35 (3.3) [8.9]	39 (3.9) [9.5]
Congenital, familial and genetic disorders	0	1* (0.1) [0.2]
Ear and labyrinth disorders	15 (1.4) [3.8]	21 (2.1) [5.1]
Endocrine disorders	4 (0.4) [1.0]	2 (0.2) [0.5]
Eye disorders	31 (2.9) [7.9]	33 (3.3) [8.1]
Gastrointestinal disorders	146 (13.6) [37.1]	165 (16.5) [40.3]
General disorders and administration site conditions	71 (6.6) [18.0]	51 (5.1) [12.5]
Hepatobiliary disorders	12 (1.1) [3.0]	18 (1.8) [4.4]
Immune system disorders	8 (0.7) [2.0]	9 (0.9) [2.2]
Infections and infestations	299 (27.9) [75.9]	362 (36.3) [88.4]
Injury, poisoning and procedural complications	50 (4.7) [12.7]	63 (6.3) [15.4]
Investigations	81 (7.6) [20.6]	126 (12.6) [30.8]
Metabolism and nutrition disorders	65 (6.1) [16.5]	91 (9.1) [22.2]
Musculoskeletal and connective tissue disorders	147 (13.7) [37.3]	122 (12.2) [29.8]
Neoplasms benign, malignant and unspecified (including cysts and polyps	7 (0.7) [1.8]	10 (1.0) [2.4]

System Organ Class Preferred Term	PBO (N=1070) (Person Years Exposure=393.8) n (%) [Person Years]	BARI 4-mg (N=997) (Person Years Exposure=409.4) n (%)[Person Years]
Nervous system	77 (7.2) [19.6]	92 (9.2) [22.5]
Psychiatric disorders	31 (2.9) [7.9]	27 (2.7) [6.6]
Renal and urinary disorders	20 (1.9) [5.1]	26 (2.6) [6.4]
Reproductive system and breast disorders	10 (0.9) [2.5]	15 (1.5) [3.7]
Amenorrhoea	1 (0.1) [0.3]	5 (0.6) [1.5]
Respiratory, thoracic and mediastinal disorders	60 (5.6) [15.2]	79 (7.9) [19.3]
Skin and subcutaneous tissue disorders	68 (6.4) [17.3]	66 (6.6) [16.1]

*meningocele

Adverse events from RCTs

Table 24 presents a summary of AEs up to 24 weeks from RA-BEAM, RA-BUILD and RA-BEACON.

In RA-BEAM, throughout 52 weeks of the RCT there were five deaths: one PBO, one PBO switched to BARI, two BARI, and one ADA. From week 0 to week 52, SAEs were experienced by 8% of BARI-treated, and 4% of ADA-treated patients.

Table 24: Adverse events and Exposure adjusted incidence rates Weeks 0-24 in RA-BEAM, RA-BUILD (adapted from CS Table 95 and Taylor et al 2017,[35] Dougados et al,[36] and Genovese et al[37])

	RA-BEAM	RA-BEAM	RA-BEAM	RA-BUILD	RA-BUILD	RA-BUILD	RA-BEACON	RA-BEACON	RA-BEACON
	Placebo (N=488)	Baricitinib (4 mg QD) (N=487)	Adalimumab (N=330)	PBO (n=228)	Baricitinib (2mg QD) (n=229)	Baricitinib (4mg QD) (n=227)	PBO (n=176)	Baricitinib (2mg QD) (n=174)	Baricitinib (4mg QD) (n=177)
Treatment exposure, patient-years (total per group)	197.7	215.0	141.9	89.8	97.7	96.4	65.8	69.9	73.3
Overall AE, n (%) [EAIR]	[redacted]	[redacted]	[redacted]	161 (71)	154 (67)	162 (71)	112 (64)	123 (71)	137 (77)
Serious AE, n (%)[EAIR]	[redacted]	[redacted]	[redacted]	11 (5)	6 (3)	12 (5)	13 (7)	7 (4)	18 (10)
Withdrawal because of AE, n (%) [EAIR]	[redacted]	[redacted]	[redacted]	10 (4)	10 (4)	12 (5)	7 (4)	7 (4)	11 (6)
Temporary interruption due to AE, n [EAIR]	[redacted]	[redacted]	[redacted]	NR	NR	NR	NR	NR	NR
Death, n [EAIR]	[redacted]	[redacted]	[redacted]	2	0	0	0	0	1
Infection, n (%)	134 (27)	176 (36)	110 (33)	79 (35)	70 (31)	96 (42)	55 (31)	76 (44)	70 (40)
Serious infection, n (%)	7 (1)	5 (1)	2 (<1)	4 (2)	2 (<1)	4 (2)	5 (3)	4 (2)	6 (3)
Cancer, n (%)	3 (<1)	2 (<1)	0	0	0	1 (<1)	0	0	2 (1)
MACE	0	1 (<1)	0	2 (<1)	0	0	0	0	2 (1)

RA-BEAM and RA-BUILD Serious adverse events are reported on the basis of conventional International Conference on Harmonisation definitions

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

Trials included in the NMA are listed in the CS Tables 69 and 70 (reproduced below in Table 25 and Table 26 of the ERG report). Trial characteristics of these studies are included in CS Tables 71 and 72 (not reproduced here) and were considered appropriate by the ERG to permit inclusion in the NMA. Quality assessment for these trials is reported in Appendix 15 of the CS.

Trials in the 24-week analysis of the bDMARD-naïve population were largely the same as those in the NMA undertaken by the independent Assessment Group (AG) in TA375. However, there were some exceptions which have been grouped into the following categories: trials in the CS that were not included in TA375; and trials included in TA375 but excluded from the CS. A similar comparison could not be made for the bDMARD-experienced population as this was not the focus of TA375.

Seven trials in the CS that were not included in TA375 for bDMARD-naïve patients:
These were trials published after the cut-off date used within TA375 – Li (2013),[49] BREVACTA,[50] SUMMACTA,[51] RA-SCORE,[52] SERENE,[53] BARI trials – RA-BEAM,[35] and RA-BUILD.[36]

Ten trials included in TA375 but excluded from the CS:
AUGUSTII,[54] IIBCREATE,[55] NCT00254293,[56] and Kremer 2012.[57] These studies were excluded from the CS as they were Phase II trials (see CS Table 11).

ACQUIRE.[58] The company excluded this study because the "*study compared S/C vs IV abatacept. The search strategy specified that studies were to include two different comparators of interest to be included*" (see clarification question A4). This appears to be inconsistent with the inclusion of SUMMACTA, which compares IV TCZ and SC.

ATTRACT.[59] The company excluded this trial as it only provided data relating to ACR20. These data can be used within the NMA and should not be discarded.

CERTAIN.[60] Within the clarification response process (clarification question A4), the company stated that this trial was excluded as it included patients with low to moderate disease activity. The ERG considered baseline DAS28 in the treatment arms of 4.47 and 4.53 to be moderate, to severe, disease activity.

SAMURAI.[61] The company stated that 12 or 24-week data were not identified. However, data at 24 weeks from Nishimoto et al 2007[62] were used in TA375.

Swefot.[63] The company stated that this trial focussed on patients with "*early rheumatoid arthritis (less than a year since diagnosis) and was therefore excluded. Additionally, the infliximab arm, allowed an increase in dose frequency (to every 6 weeks) or a switch to etanercept and it does not appear that reported results take this into account*" (clarification question A4).

TACIT[64] – The company excluded this study as the intervention arm combined bDMARDs (Appendix 4 CS).

In addition, two trials were identified by an ERG update search that would have been considered for inclusion, but were excluded from the CS. These were SURPRISE[65] and NCT01001832[66]. SURPRISE was an open-label study that compared TCZ + MTX to TCZ monotherapy and was therefore excluded (clarification question A4). The study design of NCT01001832 mirrored the design of ACQUIRE in a Japanese-only population and was therefore excluded on the same basis as ACQUIRE (clarification question A4).

Table 25: Summary of trials considered for inclusion in the NMA for the cDMARD-IR population (reproduced from CS Table 69)

Study name	Treatment 1	Treatments 2 / 3	Control	Study design	Endpoints analysed 24 weeks (base case)	Endpoints analysed 12 weeks (sensitivity analysis)
Abe (2006)	IFX 3 mg + MTX (n=49)	---	MTX (n=47)	DB	NA (no MTX arm)	ACR20/50/70
ATTRACT	IFX 3 mg (n=86)		MTX (n=88)	DB	ACR20	ACR20
De Filippis (2006)	IFX 3 mg + MTX (n=16)	ETN 25 mg + MTX (n=16)	---	OL	ACR20/50/70	ACR20/50/70
START	IFX 3 mg + MTX (n=360)	---	MTX (n=363)	DB	ACR20/50/70; EULAR	*No outcomes reported*
ARMADA	ADA 40 mg + MTX (n=67)	---	MTX (n=62)	DB	ACR20/50/70	ACR20/50/70
CHANGE	ADA 40 mg (n=91)	---	Placebo (n=87)	DB	ACR20/50/70	ACR20/50/70
Keystone (2004)	ADA 40 mg + MTX (n=207)	---	MTX (n=200)	DB	ACR20/50/70	ACR20/50/70
Kim (2007)	ADA 40 mg + MTX (n=65)	---	MTX (n=63)	DB	ACR20/50/70	ACR20/50/70
STAR	ADA 40 mg + cDMARD (n=318)	---	cDMARD (n=318)	DB	ACR20/50/70	ACR20/50/70
Van de Putte (2004)	ADA 40 mg (n=113)	---	Placebo (n=110)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR
CNTO 148[a]	GOL 50 mg + MTX (n=35)	---	MTX (n=35)	DB	*[ACR20/50/70][a]*	ACR20/50/70
GO-FORTH	GOL 50 mg + MTX (n=89)	---	MTX (n=90)	DB	ACR20/50/70 EULAR	ACR20/50/70; EULAR

Study name	Treatment 1	Treatments 2 / 3	Control	Study design	Endpoints analysed 24 weeks (base case)	Endpoints analysed 12 weeks (sensitivity analysis)
GO-FORWARD	GOL 50 mg + MTX (n=89)	---	MTX (n=133)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR
Li (2013)	GOL 50 mg + MTX (n=132)	---	MTX (n=132)	DB	ACR20/50/70	ACR20/50/70
J-RAPID[b]	CTZ + MTX (n=82)	---	MTX (n=77)	DB	ACR20/50/70	ACR20/50/70; EULAR
Kang (2013)[b]	CTZ + MTX (n=81)	---	MTX (n=40)	DB	ACR20/50/70	*No outcomes reported*
RAPID1[b]	CTZ + MTX (n=393)	---	MTX (n=199)	DB	ACR20/50/70; EULAR	ACR20/50/70
RAPID2[b]	CTZ + MTX (n=246)	---	MTX (n=127)	DB	ACR20/50/70	ACR20/50/70
REALISTIC[c]	CTZ + cDMARD (n=531/851)	---	cDMARD (n=132 /212)	DB	*NA (12-wk study)*	ACR20/50/70
AIM	ABA 10 mg + MTX (n=433)	---	MTX (n=219)	DB	ACR20/50/70	ACR20/50/70
AMPLE	ABA 125 mg + MTX (n=318)	ADA 40 mg + MTX (n=328)	---	SB	ACR20/50/70	ACR20/50/70
ATTEST	ABA 10 mg + MTX (n=156)	IFX 3 mg + MTX (n=165)	MTX (n=110)	DB	ACR20/50/70	*No outcomes reported*
ACT-RAY	TCZ 8 mg + MTX (n=279)	TCZ 8 mg (n=277)	---	OL & DB	ACR20/50/70; EULAR	ACR20/50/70
ADACTA	TCZ 8 mg (n=163)	ADA 40 mg (n=162)	---	DB	ACR20/50/70; EULAR	*No outcomes reported*
AMBITION[b]	TCZ 8 mg (N=286)	---	MTX (N=284)	DB	ACR20/50/70 EULAR	ACR20/50/70
BREVACTA[d]	TCZ 162 mg + cDMARD (n=348/437)	---	cDMARD (n=172 /219)	DB	ACR20/50/70	*No outcomes reported*
LITHE[b]	TCZ 8 mg + MTX (n=398)	---	MTX (n=393)	DB	ACR20/50/70	ACR20/50/70
Nishimoto (2004)	TCZ 8 mg (n=55)	---	Placebo (n=54)	DB	*NA (12-wk study)*	ACR20/50/70; EULAR
OPTION[b]	TCZ 8 mg + MTX (N=205)	---	MTX (N=204)	DB	ACR20/50/70; EULAR	ACR20/50/70
SATORI	TCZ 8 mg + MTX (n=61)	---	MTX (n=66)	DB	ACR20/50/70; EULAR	ACR20/50/70
SUMMACTA	TCZ 162 mg (n=631)	TCZ 8 mg (n=631)	---	DB	ACR20/50/70	ACR20/50/70
TOWARD[b]	TCZ 8 mg + cDMARD (n=805)	---	cDMARD (n=415)	DB	ACR20/50/70 EULAR	ACR20/50/70
ORAL SCAN[b]	TOF 5 mg + MTX (n=321)	TOF 10 mg + MTX (n=316)	MTX (n=160)	DB	ACR20/50/70	ACR 50/70
ORAL STANDARD[b]	TOF 5 mg + MTX (n=204)	TOF 10 mg + MTX (n=201) ---------- ADA 40 mg + MTX (n=204)	MTX (n=108)	DB	ACR20/50/70	ACR20/50/70
APPEAL	ETN 25 mg +	---	cDMARD	OL	*NA*	ACR20/50/70

Study name	Treatment 1	Treatments 2 / 3	Control	Study design	Endpoints analysed 24 weeks (base case)	Endpoints analysed 12 weeks (sensitivity analysis)
	MTX (n=197)		+ MTX (n=103)		*(16-wk study)*	
Combe (2006)	ETN 25 mg + SSZ (n=101)	SSZ (n=50)	ETN 25 mg (n=103)	DB	ACR20/50/70	*[ACR20/50/70*[e]
JESMR	ETN 25 mg + MTX (n=77)	ETN 25 mg (n=74)	---	OL	ACR20/50/70; EULAR	*[EULAR]*[f]
Lan (2004)	ETN 25 mg + MTX (n=29)	---	MTX (n=29)	DB	*NA (12-wk study)*	ACR20/50/70
Machado (2014)	ETN 50 mg + MTX (n=284)	---	cDMARD + MTX (n=145)	OL	ACR20/50/70; EULAR	ACR20/50/70
Moreland 1999/ Mathias 2000	ETN 25 mg (n=78)	---	Placebo (n=80)	DB	ACR20/50/70	ACR20/50/70
RACAT[b]	ETN 50 mg + MTX (n=175)	SSZ +HCQ + MTX (n=178)	---	DB	ACR20/50/70	*No outcomes reported*
TEMPO	ETN 25 mg + MTX (n=231)	ETN 25 mg (n=223)	MTX (n=228)	DB	ACR20/50/70	ACR20/50/70
Weinblatt (1999)	ETN 25 mg + MTX (n=59)	---	MTX (n=30)	DB	ACR20/50/70	ACR20/50/70
Edwards (2004)	RTX 1000 mg (n=40)	RTX 1000 mg + MTX (n=40)	MTX (n=40)	DB	ACR20/50/70; EULAR	*No outcomes reported*
RA-SCORE	RTX 1000 mg (n=63)	---	MTX (n=60)	DB	ACR20/50/70; EULAR	*No outcomes reported*
SERENE	RTX 1000 mg + MTX (n=168)	RTX 2000 mg + MTX (n=172)	MTX (n=172)	DB	ACR20/50/70; EULAR	*No outcomes reported*
RA-BEAM	BARI 4 mg + MTX (n=487)	ADA 40 mg + MTX (n=330)	MTX (n=488)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR
RA-BUILD	BARI 2 mg + cDMARD (n=229)	BARI 4 mg + cDMARD (n=227)	cDMARD (n=228)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR

Footnotes: Studies in green cells indicate allowance of prior bDMARD treatment up to 20%. [a]CNTO 148: 24-week results were excluded from the analysis due to switch of the placebo group to IFX at week 20. [b]Study includes prior bDMARD use up to 20%. [c]REALISTIC: only results from the subgroup of REALISTIC patients that were cDMARD-IR are used in the analysis. [d]BREVACTA: only results from the subgroup of BREVACTA patients that were cDMARD-IR are used in the analysis. [e]Data not analysed for consistency with the approach taken in the NICE MTA (TA375) [f]No ACR data were available for this time point, thus it was not possible to calculate the EULAR response.

Abbreviations: NMA = Network meta-analysis, ACR20/50/70 = 20/50/70% improvement in ACR disease activity index, EULAR = European League Against Rheumatism, ABTS = subcutaneous abatacept, ADA = adalimumab, BAR = baricitinib, CTZ = certolizumab pegol, ETN = etanercept, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX = infliximab, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab, NA = not available, DB = double-blind, OL = open-label, cDMARD = conventional disease-modifying anti-rheumatic drug, TOF = tofacitinib

Table 26: Summary of trials included in the NMA for the TNFi-IR population

Study name	Treatment 1	Treatment 2	Control	Study design	Endpoints 24 weeks	Endpoints 12 weeks
ATTAIN	ABA 10 mg + cDMARD (n=258)	---	cDMARD (n=133)	DB	ACR20/50/70	ACR20
REALISTIC[a]	CTZ + cDMARD (n=320/851)	---	cDMARD (n=80/212)	DB	NA (12-week study)	ACR20/50/70; *[EULAR]*[b]
GO-AFTER[c]	GOL 50 mg +/- cDMARD (n=153)	---	cDMARD (n=155)	DB	ACR20/50/70	ACR20/50/70
RADIATE	TCZ 8 mg + MTX (n=175)	---	MTX (n=160)	DB	ACR20/50/70	ACR20/50/70
BREVACTA[d]	TCZ 162 mg + cDMARD (n=89/437)	---	cDMARD (n=47/219)	DB	ACR20/50/70	No outcomes reported
ORAL STEP	TOF 5 mg + MTX (n=133)	TOF 10 mg + MTX (n=134)	MTX (n=132)	DB	*[ACR20/50/70]*[e]	ACR20/50/70
REFLEX	RTX 1000 mg + MTX (n=311)	---	MTX (n=209)	DB	ACR20/50/70; EULAR	ACR20/50/70
RA BEACON	BARI 2 mg + cDMARD (n=174)	BARI 4 mg + cDMARD (n=177)	cDMARD (n=176)	DB	ACR20/50/70; EULAR	ACR20/50/70; EULAR

Footnotes: [a]REALISTIC: only results from the subgroup of REALISTIC patients that were TNFi-IR are used in the analysis. [c]GO-AFTER: approx. 30% of patients did not have concomitant cDMARD. [b]Insufficient ACR response data for the TNFi-IR subgroup were available from the REALISTIC study in order to perform the conversion to EULAR response. [d]BREVACTA: only results from the subgroup of BREVACTA patients that were TNFi-IR are used in the analysis. [e]ORAL STEP: results at Week 24 were excluded from the analysis due to a disconnect in the network.
Abbreviations: NMA = Network meta-analysis, ACR20/50/70 = 20/50/70% improvement in ACR disease activity index, EULAR = European League Against Rheumatism, ABTS = subcutaneous abatacept, BAR = baricitinib, CTZ = certolizumab pegol, GOL = golimumab, ICER = incremental cost-effectiveness analysis, IFX = infliximab, LYG = life years gained, MTX = methotrexate, Pall = palliative care, QALY = quality-adjusted life year, RTX = rituximab, TCZ = tocilizumab, NA = not available, DB = double-blind, OL = open-label, cDMARD = conventional disease-modifying anti-rheumatic drug, TOF = tofacitinib

4.4 Critique of the indirect comparison and/or multiple treatment comparison

NMAs were performed separately for the cDMARD-IR and TNFi-IR populations using a Bayesian approach with a probit link function for ACR and EULAR outcome measures. The CS stated that the models were conducted in accordance with NICE Decision Support Unit (DSU) Technical Support Documents (TSDs).[67] Data for ACR were taken directly from reported results in the included studies. For studies that reported ACR data but not EULAR data, EULAR data for these studies were derived using the mapping algorithm derive from the Veteran Affairs Rheumatoid Arthritis database, as reported in Stevenson *et al*.[22] Data were then synthesised using directly reported EULAR data together with ACR-converted EULAR data. The company's clarification response states that ACR

data were converted into EULAR data in 29 out of 45 studies for the cDMARD-IR population and 4 out of 6 studies for the TNFi-IR population.

In TA375,[23] the NMA for EULAR outcome only included the studies which reported EULAR data with a separate NMA analysis performed for ACR data with a subsequent conversion to EULAR undertaken on the convergence diagnostic and output analysis (CODA). It is unclear whether the approach taken by the company would produce substantially different results compared with the method employed by the AG in TA375.[23]

Simultaneous models for baseline and treatment effects were used, with the company citing: "*since the data for both baseline and treatment effects came from the same sources; there were some networks that had zero cells and fitting this type of model increased the stability of the relevant models; and the evidence for TNFi-IR network was sparse.*" A random effects model was assumed for the study-specific baseline treatment effects (pooling non-active and active controls).

The ERG disagrees with the use of simultaneous models for the NMA. Firstly, the control arms of studies were pooled to estimate a baseline rate, irrespective of whether these were cDMARD or non-cDMARD treatments, in a random effects model. For a study with a cDMARD control arm, this would overestimate the treatment effect in that study. For a study with a bDMARD control arm, this would underestimate the treatment effect in that study.

Secondly, using simultaneous models for baseline and treatment effects means that the relative treatment effects are affected by the assumptions made about the baselines, which is not recommended by NICS DSU TSD[67]. The ERG warns that the results presented in NMA should be treated with caution, since an inappropriate pooling of non-active and active controls was carried out for the baselines.

Thirdly, data used in NMAs (no ACR response, ACR 20, ACR 50, no EULAR response and moderate EULAR response) for all the base case analyses at week 24 were without zero cells; zero cells were reported in two studies for ACR 50 category at week 12 in cDMARD-IR population for the sensitivity analysis. The ERG believes that it is unnecessary to use simultaneous models for the base case analyses, and also for the sensitivity analysis with two studies with zero cells since this is unlikely to cause instability problem.

The treatment effect in the probit model was interpreted as "*the pooled effect of the experimental treatment versus the control (in this case, the cDMARD arm of the included studies) is to change the probit (Z) score of the control by δi,bk standard deviation.*" In response to a request for clarification (question A14), the company stated that 'where the control arm was non-cDMARD, its effect was subtracted out so no further adjustment was necessary'. However, the treatment effect δi,bk in the

NICE DSU model is the treatment effect in arm k relative to the control in that study, not to the common reference treatment (in this case cDMARD).

Goodness-of-fit was not assessed in the CS. In response to a request for clarification (question A10), the company provided the total residual deviances which were 427.5 and 717.7 for the ACR and EULAR outcomes in the cDMARD-IR population. Comparing the number of data points 284, this indicates bad model fit, as the total residual deviances should be close to the number of data points. For the TNF-IR population, the same problem was identified with the total residual deviances being 56.6 and 92.3 for the ACR and EULAR outcomes in the TNFi-IR population using 39 data points; this again indicates a bad model fit.

For the TNFi-IR population, different cDMARDs were assumed to be equivalent in order to produce a connected network. For the cDMARD-IR population, unspecified cDMARDs were assumed to be the same as MTX and were labelled cDMARD in the NMA. SSZ and HCQ were assumed to have a differential efficacy to the cDMARD group. In response to a request for clarification (question A5), the company stated that a sensitivity analysis was conducted combining all cDMARDs into one cDMARD node in the NMA, and no notable difference in the results were found.

A random effects model was used for the cDMARD-IR population. In contrast, a fixed effect model was used for the TNFi-IR population since the company stated that random effects models were unstable and did not converge. The choice between the use of fixed effect and random effects models should depend on the objective of the analysis and the conduct of the included studies, rather than on model convergence. When data are sparse, external information should be used to construct the prior distribution for the between-study standard deviation in the random effects model so that it provides plausible posterior uncertainty for the results.

In response to a request for clarification (question A11), the company provided the results using models with a uniform (0, 2) distribution and a log-normal (-2.34, 1.62) distribution for the between-study standard deviation, and concluded that even with informative priors it led to convergence issues. The ERG comments that a uniform (0, 2) distribution is still a reference/vague prior. No conclusion could be made regarding the prior beliefs that were represented by log-normal (-2.34, 1.62) since it was not clear whether 1.62 was the standard deviation, the variance or the precision of that distribution.

Heterogeneity was assessed using Higgins' I^2 and inconsistency was checked using a node-splitting approach. Table 90 in the CS provided results of assessing heterogeneity using Higgins' I^2. Each study was associated with its own I^2 value, however, I^2 is the percentage of variability in estimated treatment effects that is due to heterogeneity and thus it is not clear how this can be calculated for a

single study. In response to a request for clarification (question A19), the company stated that the I^2 value was calculated from a meta-analysis of studies which contained the same treatment arms for example the TEMPO,[68] SATORI[69] and ARMADA[70] studies. However, it was not clear if it meant both treatment arms had to be similar between studies, and the ERG remains unclear on the process of study selection. Furthermore, it is debatable how meaningful the I^2 results are since the model in NMA was probit and the model in meta-analysis used to calculate I^2 is logit. Similar modelling issues also exist in the company's checking of inconsistency. The NMA code for the node-splitting approach given in response to a request for clarification (question A17) used a logit link function instead of a probit link function.

ACR results for both the cDMARD-IR and TNFi-IR populations were presented as the odds ratios of achieving an ACR50 response and the absolute probabilities of achieving each ACR category. EULAR results were presented using the absolute probability of being in each EULAR category for both the cDMARD-IR and TNFi-IR populations.

The CS concluded that for the base case analysis at week 24 in the cDMARD-IR population, BARI 4mg was associated with a statistically significant higher odds of an ACR50 response compared with cDMARD, ADA, PBO, ETN and SSZ. No statistically significant differences were found versus any other comparators for the ACR50 outcome, with the exception of CTZ + cDMARD, in which odds of ACR50 response was found to be significantly in favour of the comparator. A similar pattern of results was observed for BARI 2mg.

For the base case analysis at week 24 in the TNFi-IR population, BARI demonstrated significantly higher ACR50 response rates than the cDMARD comparator. No statistically significant differences were seen versus bDMARDs, with the exception of the comparison of BARI (both 4mg and 2mg) to TCZ, and the comparison of BARI 2mg to RTX, in which statistically significant treatment effects in favour of the comparator were observed.

The ERG notes that as control arms were inappropriately pooled, all results should be interpreted with caution.

The relative treatment effects on the probit scale were presented using the posterior distribution for both populations. However, it was difficult to interpret fully the results due to the high level of overlap between distributions (see Figures 39 and 40 within the CS). The ERG requested that the company presented the treatment effects on the probit scale in forest plots (clarification question A9). However, the treatment effects provided in the company's response were still on the odds ratio scale rather than the probit scale.

4.5 Additional work on clinical effectiveness undertaken by the ERG

The ERG re-analysed both the ACR and EULAR outcomes at week 24 for both the cDMARD-IR and TNFi-IR populations. In the ERG's NMA, all cDMARDs were assumed to have equivalent efficacy and were grouped together.

The company provided data in the format for NMA for the cDMARD-IR population EULAR outcomes. The ERG amended the EULAR data used for van de Putte *et al*.[33] so that the moderate EULAR responders did not include good EULAR responders.

For EULAR outcomes in the TNFi-IR population, and ACR outcome in the cDMARD-IR and TNFi-IR population, the ERG computed the number of responses in each category using the data provided in percentages reported in the CS appendix 14 and in response to clarification request (question A6). The ERG's ACR NMA used the same included studies as those used in the CS. The ERG's EULAR NMA were only included studies that reported EULAR outcomes, rather than introducing EULAR data converted from ACR data.

The model for the relative treatment effect used in ERG's analyses was the same as in NICE DSU TSD[67] which did not assume a random effects model for the baseline for each study. The baseline and relative treatment effect models were run separately to make sure that the information in the baseline model does not propagate to the relative treatment effect model.

A random effects model was used for both ACR and EULAR outcomes in both populations. For the TNFi-IR population, since data were sparse, an informative prior was assumed for the between-study standard deviation. This was a lognormal distribution, with mean -2.56 and variance of 1.74^2 as proposed by Turner et al,[71] which was truncated so that the odds ratio in one study would not be $\geq$50 times than in another. It represented the beliefs that heterogeneity being small is 15%, being moderate is 66%, and being high is19%. Forest plots generated from the ERG's NMA are presented in Appendix 2, and indicate whether the comparators were associated with beneficial treatment effects relative to BARI 4mg. The NMAs conducted by the ERG had total residual deviances that indicated that the model used by the ERG provided a better fit than the NMAs conducted by the company.

For EULAR outcomes in the cDMARD-IR population, BARI 4mg was associated with statistically significant beneficial treatment effects relative to PBO and cDMARD. No statistically significant differences were found versus any other comparator, with the exception of TCZ + cDMARD, which was associated with statistically beneficial treatment effects relative to BARI 4mg.

For ACR outcomes in the cDMARD-IR population, BARI 4mg was associated with statistically significant beneficial treatment effects relative to PBO, cDMARD and ADA monotherapy. No statistically significant differences were found versus any other comparator, with the exception of CTZ + cDMARD, which was associated with a statistically significant beneficial treatment effect relative to BARI 4mg.

For EULAR outcomes in the TNFi-IR population, BARI 4mg was associated with statistically significant beneficial treatment effects relative to cDMARD. No statistically significant differences were found versus RTX 1000mg+MTX with the effect favouring RTX 1000mg+MTX, which was the only other comparator in the network.

For ACR outcomes in the TNFi-IR population, BARI 4mg was associated with a statistically significant beneficial treatment effect relative to cDMARD. No statistically significant differences were found versus any other comparator.

The median values from the ERG's NMAs are presented in Figure 3 and Figure 4 for EULAR outcomes and in Figure 5 and Figure 6 for ACR outcomes.

Figure 3: EULAR response in the ERG's NMA in the cDMARD-IR population

Figure 4: EULAR response in the ERG's NMA in the TNFi-IR population

Figure 5: ACR50 response in the ERG's NMA in the cDMARD-IR population

Figure 6: ACR50 response in the ERG's NMA in the TNFi-IR population

4.6 Conclusions of the clinical effectiveness section

The ERG considers that the company's search approach and strategies is sufficiently comprehensive to retrieve important citations relating to all eligible studies.

Three RCTs (RA-BEAM, RA-BUILD, and RA-BEACON) and one long-term extension study (RA-BEYOND) of BARI in RA were included in the CS. The ERG does not believe any published (at the time of ERG report writing) RCTs of BARI have been omitted from the CS. The company's study selection eligibility criteria were consistent with the decision problem outlined in the final NICE scope. Safety data for BARI were presented from six pooled trials of BARI 4mg and placebo, with additional data from trials of BARI 2mg. Two of the RCTs were in cDMARD-experienced, bDMARD-naïve RA patients (RA-BEAM,[35] RA-BUILD[36]) and one RCT was in bDMARD-experienced patients (RA-BEACON).[37]

The three included BARI RCTs were good quality. The primary endpoint of all three RCTs was ACR20 at 12 weeks. All RCTs allowed rescue treatment from 16 weeks, and patients receiving rescue therapy were considered non-responders from point of rescue in the non-responder analysis. The trials were international, with few patients from the UK. All three RCTs allowed concomitant cDMARDs. All three RCTs included a PBO comparator (with cDMARDs), and RA-BEAM additionally included an ADA comparator.

For the primary endpoint of ACR20 at 12 weeks follow-up, all three RCTs reported than BARI 4mg was statistically significantly superior to PBO ($p \leq 0.001$). At 12 weeks, more patients reached ACR20 in the BARI 4mg treated arm than the ADA treated arm. There was also an advantage over PBO for BARI 4mg at 24 weeks and for BARI 2mg at 12 weeks and 24 weeks follow-up. At 12 weeks follow-up, all three RCTs reported a significant advantage for BARI4mg over PBO for EULAR response.

The most common AEs for BARI were LDL cholesterol, upper respiratory tract infections and nausea. Other adverse drug reactions included: herpes simplex; herpes zoster; acne; increased creatine phosphokinase; increased triglycerides; increased liver function tests (AST, ALT); neutropenia; and thrombocytosis. There was a significantly higher rate in infections for BARI 4mg than PBO ($p<0.001$), but there was no significant difference between groups in the rate of serious infections.

The ERG considers that all of the NMA results presented in CS should be treated with caution because: (1) a random effects model was assumed for the baselines inappropriately pooling non-active and active controls; (2) simultaneous baseline and treatment effect models with an inappropriate assumption for the baselines were conducted; and (3) for EULAR outcomes, studies reported EULAR were synthesised with converted EULAR for studies that only reported ACR data. This differs to the approach used in TA375.[23]

5 COST EFFECTIVENESS

This chapter presents a review of the cost-effectiveness evidence provided in the CS for BARI for treating moderate to severe RA. The cost-effectiveness evidence comprised a systematic review of economic analyses for BARI RCTs and the economic analysis based on the company's *de novo* model.

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 existing economic evaluations of bDMARDs or tofacitinib (TOF) for the treatment of active RA.

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

The company performed two searches to identify economic evaluations of (i) BARI (combined with "RA" and a cost-effectiveness filter) and (ii) bDMARDs or TOF (combined with "RA" and a cost-effectiveness filter) for the treatment of active RA. The following sources were searched: EMBASE [via Ovid]; MEDLINE [via Ovid]; MEDLINE In-Process & Other Non-Indexed Citations [via Ovid]; EconLit [via Ovid]; and NHS EED [via Wiley].

The company carried out supplementary searches of the websites of several international HTA agencies (NICE, SMC, Canadian Agency for Drugs and Technologies in Health, Pharmaceuticals Benefits Advisory Committee). All database and website searches covered the period up to October 2016. The ERG considers that the search for BARI was comprehensive and clearly and fully reported in Appendix 22 of the CS.[1]

In Appendix 20 of the CS,[1] full details of the economic evaluations systematic review for bDMARDs were given for Embase only. The ERG could not assess the adequacy of the other databases searched (Medline, EconLit and NHS EED) because the CS did not provide full search strategies nor did it report which host platforms were used.

The company performed one search to identify the health state utility values and the cost and resource use data for patients with moderate-to-severe RA. The following sources were searched: EMBASE [via Ovid], MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations [via Ovid], EconLit [via Ovid], NHS EED [via Wiley] and HTA [via Wiley]. In addition, the company searched several conference websites (EULAR, ACR and ISPOR) and online databases (CEA Registry, ScHARRHUD and EQ-5D). All the search strategies in both database and website searches for utility values and

costs and resource use were fully reported. The ERG considers that the searches are comprehensive to retrieve all the eligible studies.

5.1.3 Findings of the cost effectiveness review

The company focused on published cost effectiveness reviews that were set within a UK context. The ERG considers this to be a pragmatic way in which to reduce the number of studies to summarise. This resulted in nine relevant studies being identified: eight models that were associated with NICE TAs[23, 25, 26, 72-75] and one independent published review.[76] The backbone of the model submitted by the company was based on the model produced by the independent AG in TA375; this model has been published in a peer-reviewed journal.[22]

5.1.4 Conclusions of the cost effectiveness review

The company stated that none of the models identified included BARI and that they were not aware of any published economic evaluations. As such, the company developed a *de novo* health economic model to assess the cost effectiveness of BARI. The company stated that they had based their model firmly on the model constructed by the AG within TA375[23] due to "*the availability of the details of this model*" and also due "*to the high relevance of this model to NICE decision.*"

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

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[77] is provided in Table 27.

Table 27: Comparison of the company's model with key topics within the NICE reference case.

Element	Reference case	Satisfactorily addressed within the CS	ERG Comments
Defining the decision problem	The scope developed by NICE	Mostly	The evidence for BARI + MTX has been assumed to be representative of BARI used as a monotherapy
Comparators	As listed in the scope developed by NICE	Mostly	ABA IV and TCZ SC have been excluded from the analyses
Perspective on costs	NHS and Personal Social Services (PSS)	Yes	The CS[1] states that an NHS perspective was adopted. No relevant PSS costs were identified.
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	The company's economic evaluation takes the form of a cost-utility analysis. The results of the analysis are presented in terms of the incremental cost per QALY gained for BARI + MTX versus its comparators.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes	The model adopts a lifetime horizon (45 years). A scenario analysis is also presented for a shorter time horizon (15 years).
Synthesis of evidence on health effects	Based on systematic review	Mostly	The probabilities of EULAR response for the intervention and the comparators are based on an NMA performed using data identified through a systematic review. However, the ERG has concerns with the NMA (see Section 4.4). The company performed a systematic review of outcome data. The company transformed ACR data into EULAR data. These have been synthesised in one analysis, rather than EULAR data alone and included ACR data as a scenario analyses
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 effects were expressed in QALYs. HAQ scores were mapped into EQ-5D scores using a mapping algorithm proposed by Hernández-Alava *et al.*[78] Scenario analyses were included using different mapping algorithms.
Evidence on	Costs should relate to	Yes	Resource use estimates associated

Element	Reference case	Satisfactorily addressed within the CS	ERG Comments
resource use and costs	NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS		with HAQ categories were based on data from the Norfolk Arthritis Register database.[79] Cost estimates were based on the BNF[80] and NHS Reference Costs[81].
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	Yes	All costs and QALYs are discounted at a rate of 3.5% per year.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Not Applicable	No additional equity weighting is applied to the estimated QALY gains.

5.2.2 Population

The characteristics of the population included within the modelling base case are shown in Table 28.

Table 28: Population characteristics in the modelling base case.

Population	Female (%)	Baseline age		Baseline HAQ		Source
		Mean	SD	Mean	SD	
Moderate RA, cDMARD-IR	75.01	52.05	12.40	0.98	0.61	Weighted average of RA-BEAM and RA-BUILD (Eli Lilly - Data on File)
Severe RA, cDMARD-IR	79.09	52.89	12.12	1.61	0.63	Weighted average of RA-BEAM and RA-BUILD (Eli Lilly - Data on File)
Severe RA, TNFi-IR	81.70	55.64	11.00	1.78	0.56	RA-BEACON (Eli Lilly - Data on File)

cDMARDs – conventional disease-modifying antirheumatic drugs; HAQ - Health Assessment Questionnaire; RA – rheumatoid arthritis; SD – standard deviation; TNFi - tumour necrosis factors inhibitor; IR – inadequate response

Additionally, the company performed a scenario analysis where baseline characteristics for those who are cDMARD-IR were taken from TA375.[23] In this analysis, 76.3% were assumed to be female with a mean age of 56.10 years.

5.2.3 Interventions and comparators

Descriptions of the intervention and the comparators are provided in Sections 3.2 and 3.3. It should be noted that the IV formulation of ABA and the SC formulation of TCZ have not been included in the analyses and no formal analyses have been presented for patients who could not receive MTX.

The model compares sequences of treatments. In line with TA375,[23] it was assumed that patients who were cDMARD-IR and who could receive RTX + MTX, would receive an initial intervention followed by a fixed sequence of RTX + MTX, TCZ + MTX, MTX then PALL. In both the models presented in TA375 and the CS, it was deemed appropriate that TCZ + MTX would not be later in the sequence if used earlier. In TA375, patients who were initiated on TCZ went straight from RTX+ MTX to MTX, whereas in the CS patients received ADA + MTX after RTX+MTX. In response to clarification question B5,[32] the company stated that the reason why ADA + MTX was chosen in preference to other bDMARDs was that ADA has the largest market share amongst TNFi drugs and provided a scenario analysis whereby CTZ + MTX was used instead.

The base case sequences chosen by the company in the analysis of patients who did not respond adequately to cDMARDs are shown in Table 29. Sequence two is used for all bDMARDs bar TCZ.

Table 29: Treatment sequences for patients who did not adequately respond to cDMARDs

Sequence No	First-line treatment	Second-line treatment	Third-line treatment	Fourth-line treatment	Fifth-line treatment
1	BARI + MTX	RTX + MTX	TCZ + MTX	MTX	PALL
2	bDMARD + MTX *	RTX + MTX	TCZ + MTX	MTX	PALL
3	TCZ + MTZ	RTX + MTX	ADA + MTX	MTX	PALL

* Excluding TCZ.
BARI – Baricitinib; bDMARD – biologic disease-modifying antirheumatic drug; PALL – palliative care; RTX – rituximab; TCZ – tocilizumab; ADA - Adalimumab

5.2.4 Perspective, time horizon and discounting

The model takes the perspective of the NHS. The time horizon is 45 years and assumed to represent a patient's remaining lifetime. All costs and benefits were discounted at 3.5% per annum in line with the NICE Reference Case.[77]

5.2.5 Model structure

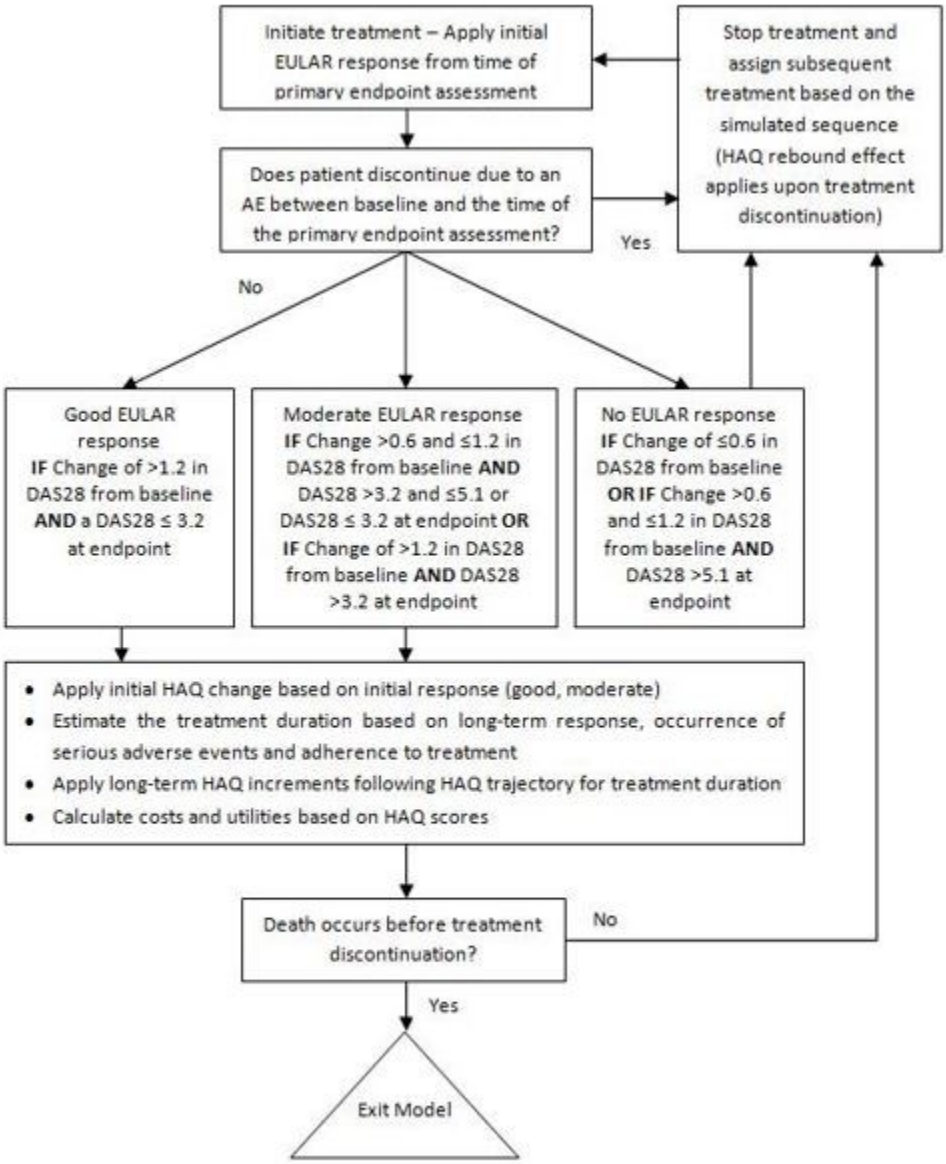
The model presented by the company draws heavily on the model for TA375.[23] A clinical response in terms of EULAR (good; moderate; or none) is estimated at week 24. Patients who experience either a good or a moderate EULAR response remain on treatment; those who experience no response have

their treatment withdrawn and move on to the next treatment in the sequence, unless the patient was already receiving PALL.

For patients experiencing a good or moderate EULAR response, there is an associated HAQ decrease that is assumed to occur at treatment initiation; this contrasts with the AG model whereby the effects of treatment were assumed to be manifested at six months. The CS stated that upon treatment discontinuation HAQ score improvement was lost, '*rebounding immediately to the level prior to initiation of the terminated therapy*'. In TA375, the HAQ score rebound (i.e. increase) was assumed to be equal in size to the HAQ decrease upon response. Whilst these two approaches are equivalent for bDMARDs and BARI, this is not the case for those on MTX, where there is a HAQ increase whilst on treatment. The company misunderstood the ERG's request for confirmation that HAQ score did not rebound to the level previous to initiation of therapy for MTX (question C4),[32] and amended the model so that there was no rebound when a patient discontinued MTX. The ERG notes that the implementation in the original model is in line with TA375 but was discordant with the description in the CS.

Whilst patients are on bDMARD treatment (and BARI in the base case), it is assumed that HAQ progression is zero. For patients on cDMARDs, the company attempted to replicate the latent class analysis used by the AG in TA375[23] and detailed in Stevenson *et al*.[22] although this was not operationalised correctly in the submitted model. After applying changes to HAQ scores, the resulting values were rounded to the nearest valid HAQ score (which is a multiple of 0.125). The ERG notes that this approach can lead to inaccurate results. This contrasts with the approach used in TA375[23] 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). Throughout the model, the costs incurred and the utility of the patient were assumed to be related to HAQ score.
The model for TA375[23] used a discrete event simulation (DES) approach which the company have replicated. The ERG believes that this an appropriate structure which also removes the need for the definition of time cycles and half-cycle correction. The model structure presented by the company is reproduced in Figure 7.

Figure 7: Model structure presented by the company



Abbreviations: EULAR = European League Against Rheumatism, HAQ = Health Assessment Questionnaire, DAS28 = Disease Activity Score 28.

5.2.6 Treatment effectiveness, extrapolation and discontinuation

The estimated treatment effectiveness in terms of EULAR response came from the NMA conducted by the company. These data, which are categorised into patients with good, moderate or no EULAR response, have been marked as academic-in-confidence by the company. These data are presented in Table 30 for CDMARD-IR patients and in Table 31 for TNFi-IR patients. For ADA + MTX, CTZ + MTX, ETN + MTX, and IFX + MTX, the company identified no data relating to patients who had an inadequate response to a TNFi. The company assumed that these values would be the same as for

patients who were cDMARD-IR. Considering the data for other interventions, the ERG does not believe that this assumption is realistic and has removed these data from Table 31.

Table 30: Estimated EULAR response for patients who did not respond adequately to cDMARDs (cDMARD-IR)

Treatment	Good EULAR response	Moderate EULAR response	No EULAR response
BARI 4mg + MTX	[redacted]	[redacted]	[redacted]
BARI 2mg + MTX	[redacted]	[redacted]	[redacted]
ABA SC + MTX	[redacted]	[redacted]	[redacted]
ADA + MTX	[redacted]	[redacted]	[redacted]
CTZ + MTX	[redacted]	[redacted]	[redacted]
ETN + MTX	[redacted]	[redacted]	[redacted]
GOL + MTX	[redacted]	[redacted]	[redacted]
IFX + MTX	[redacted]	[redacted]	[redacted]
RTX + MTX	[redacted]	[redacted]	[redacted]
TCZ IV + MTX	[redacted]	[redacted]	[redacted]
MTX	[redacted]	[redacted]	[redacted]
Intensive cDMARDs	[redacted]	[redacted]	[redacted]
PALL	[redacted]	[redacted]	[redacted]

ABA SC – abatacept subcutaneous; ADA – adalimumab; BARI – baricitinib; cDMARDs – conventional disease-modifying antirheumatic drugs; CTZ – certolizumab pegol; ETN – etanercept; GOL – golimumab; IFX – infliximab; MTX – methotrexate; PALL – palliative care; RTX – rituximab; TCZ – tocilizumab

Table 31: Estimated EULAR response for patients who did not respond adequately to a TNFi (TNFi-IR)

Treatment	**Good EULAR response**	**Moderate EULAR response**	**No EULAR response**
BARI 4mg + MTX	[redacted]	[redacted]	[redacted]
BARI 2mg + MTX	[redacted]	[redacted]	[redacted]
ABA IV + MTX	[redacted]	[redacted]	[redacted]
ABA SC + MTX	[redacted]	[redacted]	[redacted]
GOL + MTX	[redacted]	[redacted]	[redacted]
RTX + MTX	[redacted]	[redacted]	[redacted]
TCZ IV + MTX	[redacted]	[redacted]	[redacted]
MTX	[redacted]	[redacted]	[redacted]
PALL	[redacted]	[redacted]	[redacted]

ABA IV – abatacept intravenous; ABA SC – abatacept subcutaneous; BARI – baricitinib; cDMARDs – conventional disease-modifying antirheumatic drugs; GOL – golimumab; MTX – methotrexate; PALL – palliative care; RTX – rituximab; TCZ – tocilizumab

For the severe cDMARD-IR population, the company used the estimates in Table 30 for all the treatments in the sequence, regardless of their position in the sequence. This implies that, for example, TCZ IV + MTX will have the same efficacy irrespective of whether it is first or third in the sequence. However, Table 30 and Table 31 show that the efficacy of TCZ IV + MTX, as well as all the other treatments, is lower after treatment with a TNFi. The ERG notes that effectiveness estimates from Table 30 should have been used only for the first treatment in the sequence. For the rest of the treatments in the sequence, the estimates from Table 31 should have been used instead.

HAQ improvement upon treatment response

The change in HAQ score was assumed to be conditional on the EULAR response achieved. The reduction in HAQ score was taken from the values reported by the AG in TA375[23]: these were reductions of 0.673 (standard error (SE) 0.012) for patients who experienced a good response, and 0.317 (SE 0.048) for patients who experienced a moderate response. The company assumed that HAQ improvement upon response occurs instantaneously at treatment initiation. In response to a clarification question by the ERG (question B20),[32] the company argued that there is evidence that clinical response to bDMARDs in RA is often rapid, with patients potentially experiencing improvements in symptoms within a few weeks of treatment initiation, perhaps even as early as 48 hours. In addition, it referred to data from RA-BEAM,[38] where change in mean HAQ score at week 12 was similar to that at week 24 for BARI + MTX ([redacted] and [redacted], respectively). The ERG notes that even if that were the case, accounting for the HAQ improvement from time zero is likely to be an overestimation of the treatment benefits.

The company assumed that the HAQ improvement upon response would be lost at treatment discontinuation, and as a consequence the patient would suffer from a rebound in HAQ score equal to the improvement upon treatment response.

HAQ progression

Whilst patients are on treatment with a bDMARD or BARI the company assumed that the HAQ trajectory is flat, that is, that there is no change in the value. This assumption for bDMARDs was also incorporated in the AG model for TA375.[23] The company have assumed that this was also the case for BARI, although test a worsening of HAQ across time in a scenario analysis. Clinical advice to the ERG suggested that the assumption that the HAQ trajectory for BARI is equal to bDMARDs was reasonable.

Treatment duration

Patients who fail to achieve good or moderate EULAR response at 24 weeks discontinue treatment and start the next treatment in the sequence. In contrast, patients who achieve good or moderate EULAR response stay on treatment until loss of efficacy. Time on treatment for these patients is

estimated based on the approach used in TA375.[23] whereby the duration of treatment on the first biologic for adult RA patients was estimated using the British Society for Rheumatology Biologics Register (BSRBR) database, which records the dates on which therapies are initiated and discontinued. Separate curves were fitted to time to treatment discontinuation data for patients with good and moderate EULAR response, independent of treatment. However, instead of using the same survival function as the AG in TA375,[23] the company digitised the Kaplan-Meier curves from the AG's report and fitted their own curves. The AG in TA375[23] fitted several curves to the data and concluded that the generalised gamma provided the best statistical fit in terms of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The company fitted only log logistic, exponential and Weibull curves to the data and selected 'the Weibull distribution as this provided '*the most conservative estimates of long-term extrapolation compared to other models*' although the company noted that the log logistic function provided a better fit in terms of the sum of squared errors.[1] The ERG notes that it is not clear whether a conservative estimate of long-term extrapolation is favourable or unfavourable to BARI. The company did not assess the impact of the choice of the curve in the sensitivity analyses but presented a scenario analysis where a fixed annual discontinuation rate, calculated from the relevant trial for each population, was applied for the BARI arm in the model.

5.2.7 Mortality

The company applied the mortality ratios per HAQ score band reported in TA375[23] to the life tables from the Office for National Statistics.[82] The company[1] used the assumption that only baseline HAQ score, and not changes in HAQ, affected mortality, as was the case in the AG's model in TA375.[23] However, the ERG notes that in the company's model, age of death was recalculated based on the patient's age and HAQ score at every event, which leads to small variations in the life years gained estimated for each intervention. In response to clarification question B1,[32] the company argued that differences in mortality across arms were the result of random variation and would converge to zero as the number of patients run through the model increases. The ERG notes that this would imply that the number of patients simulated in the model was insufficient to minimise the impact of random variation. The ERG also notes that calculating the age at death at different time points can lead to small differences in life expectancy and that therefore sequences of different lengths (or with different treatment efficacies) are likely to produce different life expectancies. Ultimately, the ERG notes that based on the company's assumption, there should be no differences in the estimated life years gained for different interventions.

5.2.8 Health-related quality of life

The company[1] state that it would be preferable to use HRQoL data derived from trials rather than values derived from a database that may not take treatment-specific effects into account. However,

when the company attempted to use data on file from the RA-BEAM study it stated that the EQ-5D score was deemed to be under-estimated. For this reason, the company used an approach similar to that used in TA375.[23] One deviation from the approach conducted in TA375 was that the company used the algorithm published in Hernandez *et al.* in 2012,[78] rather than the four-class mixture model published by Hernandez *et al*. in 2013.[83] Both approaches mapped HAQ scores, amongst other variables to EQ-5D scores. The company stated that the earlier version was chosen as 'the predicted values for selected combinations of covariates were reported in the published article, thus allowing for direct validation of these results'. The coefficients of the mapping algorithm used in the company's model are provided in Table 109 of the CS. The company undertook two scenario analyses using alternative mapping algorithms: a mapping algorithm based on data from the RA-BEAM trial and the quadratic mapping mechanism used by Malottki *et al*. in an earlier evaluation of interventions for the treatment of RA.[84]

5.2.9 Resources and costs

The company's model includes costs associated with drug acquisition, drug administration and monitoring, and hospitalisation. Table 32 summarises the drug acquisition costs for the treatments considered in the economic analysis. The 2mg dosing of BARI is used in a scenario analysis. For weight-dependent dosing calculations, the average dose cost was calculated assuming all patients had the average weight of the population in the relevant BARI trials (see Table 112 of the CS[1]). The ERG notes that this approach to calculating the average dose cost is not appropriate given that the relationship between weight and dosing cost is not linear as explained by Hatswell *et al*.[85] This is because the average cost of a dose is not necessarily equal to the cost of the patient with the average weight, due to drug wastage and differences in cost per mg of some drugs. The ERG notes that the company should have calculated the average cost of a dose using the distribution of the weight of the modelled patient population instead of using the average weight.

The combination of cDMARDs considered by the company was based on TA375[23] and consisted of a combination of HCQ, MTX, prednisolone and SSZ. The cost was calculated by inflating the cost reported in TA375[23] to 2016 prices. The company assumed three loading doses for CTZ and one for IFX. The cost of palliative care was also based on TA375[23] and consisted of a mix of cDMARDs (LEF, GLD, cyclosporine, etc.). In line with TA375,[23] the retreatment interval for RTX was assumed to be 9 months. The company assumed 7 doses of IFX per year, instead of the 6.52 (every 8 weeks) recommended by the BNF[80] and the SmPC.[86] 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 ABA and TCZ were not included, as recommended by NICE.

Drug acquisition costs for each patient were calculated by multiplying the yearly cost of each drug by the time the patient spent on treatment. The ERG notes that this simplification might lead to slight underestimates, especially for drugs with a long retreatment interval, such as IFX and RTX. Loading dose costs and the CTZ PAS were accounted for as one-off costs (or savings).

Table 32: Drug acquisition costs

	Dose per unit	Units per pack	Cost per pack	Units per year (1st year)	Annual cost (1st year)	Source
BARI	4mg	28	£805.56 /[redacted] *	365	£10,501.05 /[redacted]*	CS[1]
	2mg					
ABA SC	125mg	1	£302.40#	52	£15,724.80#	BNF[80]
TCZ IV	80mg	1	102.4	104	£10,649.60	BNF[80]
ADA	40mg	2	£704.28	26	£9,155.64	BNF[80]
CTZ	200mg	2	£715.00	26(29)	£9,295.00 (£6,793.00)	BNF[80]
ETN†	50mg	4	£656.00	52	£8,528.00	BNF[80]
GOL	50mg / 100mg	1	£762.97	12	£9,155.64	BNF[80]
IFX‡	100mg	1	£377.66	21(24)	£7,930.86 (£9,063.84)	Cost: BNF[80], Dose: TA195[87]
RTX	500mg	1	£873.15	5.3	£4,656.80	Cost: BNF[80], Dose: TA375[88]
MTX	10mg	100	£36.78	104	£38.25	MIMS (May 2016)
Intensive cDMARDs	NA			NA	£816.47	TA375[88]
PALL	NA			NA	£747.02	TA375[88]

BARI: baricitinib; ABA: abatacept; TCZ: tocilizumab; ADA: adalimumab; CTZ: certolizumab pegol; ETN: etanercept; GOL: golimumab; IFX: infliximab; RTX: rituximab; MTX: methotrexate; PALL: palliative care: IV: intravenous; SC: subcutaneous
*PAS included.
Confidential PAS in place not included
†Cost of etanercept biosimilar Benepali
‡Cost of infliximab biosimilar Inflectra

Administration costs were based on TA375[88] and were inflated to 2016 prices using the Hospital and Community Health Services Index.[89] The cost per intravenous injection was estimated to be £159.78 and the cost per subcutaneous injection was estimated to be £2.71. Monitoring costs were also based on TA375[88] and included full blood count, erythrocyte sedimentation rate (ESR), biochemical profile, chest x-ray and urine analysis. As in TA375,[88] it was assumed that patients on bDMARDs and cDMARDs incurred the same monitoring cost. This assumption was extended to patients on BARI since the company believed that additional monitoring costs associated with BARI (such as

tuberculosis and hepatitis tests prior to treatment initiation) would be captured by routine patient management in the NHS. The resulting costs from inflating the figures in TA375[88] were £176.38 prior to treatment initiation, £1,763.79 for the first six months of treatment monitoring, and £139.03 for monthly monitoring costs.

Hospitalisation costs were estimated based on the figures reported in TA375,[88] where hospitalisation costs depending on HAQ score band were calculated based on data from the Norfolk Arthritis Register (NOAR) database on impatient days, joint replacements and NHS Reference Costs. The company digitised the graph that showed hospitalisation costs for each HAQ score band in the TA375 AG report,[88] fitted a polynomial to the resulting curve and estimated the hospitalisation cost for each valid HAQ score based on the curve. The ERG notes that it is unclear why the company did not use the costs reported by the AG for each HAQ band, inflated to 2016 prices.

For the scenario analyses where SAEs were incorporated the company assumed that the cost per SAE was £1789 per episode calculated based on the average costs for cellulites and herpes zoster as this were the most prevalent SAEs in RA-BEAM (Eli Lilly data on file). This value is broadly similar to the cost of £1479 used in TA375.[88]

5.2.10 Cost effectiveness results

The company presented cost-effectiveness results for four different populations:

- Moderate, cDMARD-IR
- Severe cDMARD-IR (TNFi naïve)
- Severe TNFi-IR RTX eligible
- Severe TNFi-IR RTX ineligible

The deterministic results in the base case were produced by simulating 27,500 patients. The model generated a pool of random numbers that were used across sequences to alleviate differences stemming from random number bias.

The company presented results of the probabilistic sensitivity analyses (PSA) for the severe cDMARD-IR and the TNFi-IR, RTX-ineligible populations. For the PSA, 500 patients were simulated in each of the 1000 iterations. The company used draws from the joint posterior distribution (i.e. CODA) from the NMA for the probabilities of EULAR response at each iteration. The ERG found two programming errors that affected the results of the PSA, especially the results of the severe TNFi-IR RTX-ineligible populations are. The results of the PSA for the severe cDMARD-IR population and the severe TNFi-IR RTX-ineligible populations have been reproduced in Table 35 and Table 38, respectively.

5.2.10.1 Moderate, cDMARD-IR

Table 33 presents the results for the base-case analysis for patients with moderate RA who are cDMARD-IR. Providing patients with BARI + MTX before current practice results in [redacted] additional QALYs gained at an additional cost of £[redacted] resulting in an ICER of £37,420 per QALY gained compared with current practice. The analysis suggests that the BARI + MTX will result in a small loss of LYG. The ERG notes that this small difference is caused by the calculation of the age at death at every event. It should also be noted that the conceptual analysis in the cDMARD-IR population differs considerably from that in TA375[23] whereby it was assumed that intensive cDMARDs had already been used, as was supported by clinical opinion, rather than assuming that a bDMARD would be used before intensive cDMARDs.

Table 33: Base-case analysis results for the moderate cDMARD-IR population (deterministic)

Interventions	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)
Intensive cDMARDs→MTX→PALL	[redacted]	16.04	[redacted]	[redacted]	[redacted]	-
BARI+MTX→Intensive cDMARDs →MTX→PALL	[redacted]	16.03	[redacted]	[redacted]	[redacted]	**37,420**

BARI: baricitinib; MTX: methotrexate; PALL: palliative care; cDMARDs: conventional disease-modifying antirheumatic drug

However, the ERG highlight the fact that the company drew heavily on TA375 in constructing their mathematical model and that some parts, including latent classes for those on cDMARDs and the HAQ progressions for those on cDMARDs were not implemented correctly. This will affect the ICER of BARI compared with cDMARDs. The time required to fix these issues was beyond that available for this STA. The ERG notes that the median ICER of bDMARDs compared with cDMARDs in TA375 was in the region of £50,000 per QALY, which is considerably higher than the estimate provided by the company.

5.2.10.2 Severe cDMARD-IR

For patients with severe RA who are cDMARD-IR, BARI + MTX dominated all recommended comparators except for CTZ + MTX. CTZ + MTX was estimated to produce [redacted] additional QALYs more compared with BARI + MTX at an additional cost of £[redacted], resulting in an ICER of £18,400 per QALY gained (Table 34). However, the ERG notes that the confidential PASs in place for ABA SC and TCZ IV were not included in the company's analysis.

Table 34: Base case analysis results for the severe cDMARD-IR population (deterministic)

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	Inc. analysis
IFX-b+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
ABA SC+MTX[#]	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated[#]
GOL+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
ADA+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
ETN-b+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated
TCZ IV+MTX[#]	[redacted]	14.73	[redacted]	[redacted]	[redacted]	Dominated[#]
BARI+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	**Baseline**
CTZ+MTX	[redacted]	14.73	[redacted]	[redacted]	[redacted]	**£18,400**

*All treatments followed by sequence RTX+MTX→TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by RTX+MTX→ADA+MTX→MTX→ PALL. Confidential PAS for TCZ IV not included.
[#]Does not include confidential PAS

BARI: baricitinib; ABA: abatacept; TCZ: tocilizumab; ADA: adalimumab; CTZ: certolizumab pegol; ETN-b: etanercept biosimilar; GOL: golimumab; IFX-b: infliximab biosimilar; RTX: rituximab; MTX: methotrexate; PALL: palliative care; IV: intravenous; SC: subcutaneous

The company also presented the results of the PSA for the severe cDMARD-IR population. The results, as shown in Table 35, are very similar to those in the deterministic analysis: BARI + MTX dominates all its comparators except for CTZ+MTX; the ICER for CTZ+MTX compared with BARI + MTX is estimated to be £18,414 per QALY gained.

Table 35: Base case analysis results for the severe cDMARD-IR population (probabilistic)

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	Inc. analysis
IFX-b+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
ABA SC+MTX[#]	[redacted]	14.70	[redacted]	[redacted]	[redacted]	Dominated[#]
ADA+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
GOL+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	Dominated
ETN-b+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	Dominated
TCZ IV+MTX[#]	[redacted]	14.70	[redacted]	[redacted]	[redacted]	Dominated[#]
BARI+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	**Baseline**
CTZ+MTX	[redacted]	14.70	[redacted]	[redacted]	[redacted]	**£18,414**

*All treatments followed by sequence RTX+MTX→TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by RTX+MTX→ADA+MTX→MTX→ PALL. Confidential PAS for TCZ IV not included.
[#]Does not include confidential PAS

BARI: baricitinib; ABA: abatacept; TCZ: tocilizumab; ADA: adalimumab; CTZ: certolizumab pegol; ETN-b: etanercept biosimilar; GOL: golimumab; IFX-b: infliximab biosimilar; RTX: rituximab; MTX: methotrexate; PALL: palliative care; IV: intravenous; SC: subcutaneous

5.2.10.3 Severe TNFi-IR RTX eligible

For patients with severe RA who are TNFi-IR for whom RTX is an option, BARI + MTX was dominated by RTX + MTX (see Table 36), as it was estimated to produce less QALYs at a higher cost. The ERG notes that the company's model suggests that BARI + MTX would result in a loss of LYG compared with RTX +MTX, which is a consequence of recalculating the age at death at every event.

Table 36: Base-case analysis results for the severe TNFi-IR RTX-eligible population (deterministic)

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)
BARI + MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated
RTX +MTX	[redacted]	13.51	[redacted]	[redacted]	[redacted]	-

*All treatments followed by sequence TCZ IV+MTX→MTX→PALL. Confidential PAS for TCZ IV not included.

BARI: baricitinib; MTX: methotrexate; TCZ: tocilizumab; PALL: palliative care; IV: intravenous.

5.2.10.4 Severe TNFi-IR RTX ineligible

For patients with severe RA who are TNFi-IR and for whom RTX is contraindicated or not tolerated, BARI + MTX was the least expensive and the second least effective intervention compared with the comparators included in the company base case (see Table 36). In the full incremental analysis, BARI + MTX was the baseline, as it dominated the only cheaper option (GOL + MTX). All other options were dominated or extendedly dominated except for CTZ + MTX, which was estimated to produce [redacted] additional QALYs compared with BARI + MTX at an additional cost of £[redacted], resulting in an ICER of £16,201 per QALY gained. The ICERs of ETN-b + MTX compared with BARI + MTX and ADA+MTX compared with BARI + MTX are also below £30,000 per QALY gained.

The ERG notes also that the confidential PAS schemes in place for ABA SC and TCZ IV were not included in the analysis. However, the ERG notes that the company did not identify any evidence on the effectiveness of ADA, CTZ, ETN and IFX in combination with MTX in severe TNFi-IR patients. In the absence of such data, the company used the same efficacy estimates of these treatments in severe cDMARD-IR patients instead. As can be seen when comparing Table 30 and Table 31, EULAR responses are lower in the severe TNFi-IR population compared with the severe cDMARD-IR for all treatments. Therefore, the efficacy of ADA, CTZ, ETN and IFX in combination with MTX in severe TNFi-IR patients is likely to be overestimated in the company's base case analysis.

Table 37: Base-case analysis results for the severe TNFi-IR RTX-ineligible population (deterministic)

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)	ICER vs BARI + MTX (£/QALY)
GOL+MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	Dominated
BARI + MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	**Baseline**	
ABA SC+MTX[#]	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	484,782[#]
IFX-b+MTX†	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	34,942†
TCZ IV+MTX[#]	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	36,757[#]
ADA+MTX†	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	27,008†
ETN-b+MTX†	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Extendedly dominated	19,874†
CTZ+MTX†	[redacted]	13.49	[redacted]	[redacted]	[redacted]	**16,201**	16,201†

*All treatments followed by sequence TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by ADA+MTX→MTX→PALL. Confidential PAS for TCZ IV not included.
[#]Does not include confidential PAS
†Efficacy estimates assumed to be equal to those for the severe cDMARD-IR population

BARI: baricitinib; ABA: abatacept; TCZ: tocilizumab; ADA: adalimumab; CTZ: certolizumab pegol; ETN-b: etanercept biosimilar; GOL: golimumab; IFX-b: infliximab biosimilar; RTX: rituximab; MTX: methotrexate; PALL: palliative care; IV: intravenous; SC: subcutaneous

The company also presented the results of the PSA for the severe TNFi-IR RTX-ineligible population. As shown in Table 38, there are important differences between these results and those produced within the deterministic analysis. The ERG notes that such differences are mostly caused by an error in the company's model as explained in Section 5.3 that affects the sequences starting with TCZ IV + MTX, ADA + MTX, ETN-b + MTX, GOL + MTX and IFX-b + MTX. The ERG re-run the PSA after fixing the programming error and these results are presented in Section 5.4.

Table 38: Base-case analysis results for the severe TNFi-IR RTX-ineligible population (probabilistic)

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)	ICER vs BARI + MTX (£/QALY) ‡
GOL+MTX	[redacted]	13.53	[redacted]	[redacted]	[redacted]	Dominated	20,824§¶
TCZ IV+MTX#	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	19,962#§¶
ADA+MTX†	[redacted]	13.53	[redacted]	[redacted]	[redacted]	Dominated	19,947†§¶
ETN-b+MTX†	[redacted]	13.53	[redacted]	[redacted]	[redacted]	**Baseline**	19,457†§¶
IFX-b+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Extendedly dominated	5,367†¶
BARI + MTX	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Extendedly dominated	
ABA SC+MTX#	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	442,044#
CTZ+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	**18,738**	17,149†

*All treatments followed by sequence TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by ADA+MTX→MTX→PALL. Confidential PAS for TCZ IV not included.
#Does not include confidential PAS
†Efficacy estimates assumed to be equal to those for the severe cDMARD-IR population
‡ Approximate ICERs calculated by the ERG based on total costs and QALYs reported by the company
§ These interventions are less effective than BARI + MTX and therefore the ICERs represent savings per QALY lost
¶ These results are affected by a programming error in the PSA

BARI: baricitinib; ABA: abatacept; TCZ: tocilizumab; ADA: adalimumab; CTZ: certolizumab pegol; ETN-b: etanercept biosimilar; GOL: golimumab; IFX-b: infliximab biosimilar; RTX: rituximab; MTX: methotrexate; PALL: palliative care; IV: intravenous; SC: subcutaneous

5.2.11 Sensitivity analyses

The company did not present results of one-way sensitivity analyses due to the computational burden of undertaking them. However, the company presented results of scenario analyses where they explored the impact of changing some of the assumptions made in the base case. The analyses which are most relevant are summarised below. The ERG notes that numerous errors were encountered in the results of the scenario analyses.

The company assumed in a scenario analysis that patients on cDMARDs or palliative care suffered a linear increase in their HAQ score at an annual rate of 0.045 and 0.06, respectively (based on Malottki *et al*.[84]) instead of using the latent class approach. This scenario had a small impact on the severe populations, producing slightly lower ICERs for the most effective drugs. However, for the moderate population, the ICER for BARI + MTX compared with intensive cDMARDs decreased from £37,420 to £20,965 per QALY gained. As detailed in Stevenson *et al*.[22] the ERG do not believe the Malottki *et al.*[84] mapping is as robust as that of Hernandez *et al*.[83]

In the base case, the company assumed that the HAQ score of patients on BARI + MTX would remain constant, as assumed for bDMARDs. The company undertook a scenario analysis where it was assumed that instead of remaining constant, the HAQ score of patients on BARI + MTX would

deteriorate (increase) at an annual rate of 0.025 (approximately half the rate assumed for patients on cDMARDs). As explained by the company in response to clarification question B3,[32] due to a limitation in the model, this scenario analysis can only be run when patients on cDMARDs or palliative care are also assumed to suffer a linear HAQ increase. The results for this scenario analysis are very different to those of the base case. In the severe cDMARD-IR population, BARI + MTX is dominated by two of its comparators and would be less expensive but less efficient than the rest, all of which (except ABA SC, whose confidential PAS has not been included) have an ICER compared with BARI + MTX of less than £20,000 per QALY gained. In the moderate cDMARD-IR population, the ICER for BARI + MTX compared with intensive cDMARDs increases from £20,965 in the previous scenario to £30,280 per QALY gained. The results of this scenario analysis for the severe TNFi-IR populations were not included in the CS.[1] Clinical advice received by the ERG suggested that there was no reason why BARI + MTX would have a different HAQ progression than bDMARDs.

The HAQ score improvements by EULAR response used in TA375[23] which were calculated from the BSRBR database were assumed to be generalisable to BARI. The company presented the results of a scenario analysis where HAQ score improvements for BARI were calculated using the data from the relevant BARI trial for each population. In the severe cDMARD-IR population, BARI + MTX went from being more effective than all of its comparator except for CTZ + MTX in the base case to being less effective than all of them in this scenario analysis. For example, BARI + MTX was estimated to produce [redacted] QALYs more than ETN-b + MTX in the base case but [redacted] QALYs less in this scenario. On the other hand, results in the severe RA TNFi-IR, RTX-ineligible population were similar to those in the base case. These results once again lack face validity, as HAQ improvements by EULAR response observed in the BARI trials were higher than those calculated by the AG of TA375[23] from the BSRBR database. In the case of the moderate RA CDMARD-IR population, the ICER for BARI + MTX compared with intensive cDMARDs decreased from £37,420 to £32,303 per QALY gained as a consequence of a larger difference in incremental QALYs.

The impact of alternative assumptions for time to treatment discontinuation of BARI was also assessed within the company's scenario analyses. In the base case, the company used the time to discontinuation curves by EULAR response category used in TA375,[23] both for BARI and the bDMARDs. In a scenario analysis, the company used a different time to treatment discontinuation for patients on BARI. The CS[1] states that a fixed annual discontinuation rate was applied for the BARI arm in the model, but the ERG noticed that actually a Weibull distribution with an increasing rate was used. The CS[1] explains that the discontinuation rate was derived from the relevant BARI trials for each respective population (CS Section 5.3) but also that it based on data on treatment duration from RA-BEAM[38] (CS Section 5.8.3). The ERG notes that the same Weibull distribution was used in all populations. The CS[1] does not explain how well the Weibull curve fitted the treatment discontinuation

data from the trial, and whether the company assessed the fit of other data. The ERG notes that given that RA-BEAM featured a treatment arm receiving ADA + MTX, applying a curve fitted to the ADA + MTX treatment discontinuation data for bDMARDs would have provided more accurate estimates. The results presented by the company for this scenario analysis for the severe cDMARD-IR population are very different to those presented for the severe TNFi-IR, RTX-ineligible population. In the severe cDMARD-IR population, BARI + MTX was estimated to be less effective and less expensive than all of its comparators. On the other hand, in the severe TNFi-IR, RTX-ineligible population, BARI + MTX was estimated to be more effective and more expensive than all of its comparators except ABA SC + MTX. The results for ABA SC + MTX lack face validity, as this scenario analysis should only affect BARI + MTX. The ERG investigated this issue in the company's model and found that the reason for this inconsistency was a mistake in the model as a result of which the time to treatment discontinuation of BARI was also applied to ABA SC.

The company also explored the impact of using alternative methods to map HAQ scores to the EQ-5D, although this was only undertaken in the severe RA cDMARD-IR population. In the base case, the company used the mixture model proposed by Hernández-Alava *et al*,[78] but undertook scenario analyses using the mapping used by Malottki *et al.*[84] and the company's own mapping algorithm, based on data from RA-BEAM.[38] The results differ only slightly between the base case and these scenario analysis except for two cases based on Malottki *et al*.[84] GOL + MTX, where the incremental QALYs compared with BARI + MTX change from [redacted] to [redacted]; and CTZ + MTX, where the incremental QALYs compared with BARI + MTX change from [redacted] to [redacted]. In the case of GOL + MTX, the result lacks face validity because it has the same EULAR response rate as ADA + MTX and the result for ADA + MTX in this scenario analysis is very different ([redacted] incremental QALYs in the base case, [redacted] in this scenario analysis). In the case of CTZ + MTX, the result has limited credibility given that the incremental costs compared with BARI + MTX also differ from those in the base case (changing from £[redacted] to £[redacted]), which should not happen when only the utility mapping algorithm changes. The ERG believe that this is likely to be a transcription error, and there appear to be two more errors in the results for these scenario analyses. It is also unlikely that the Malottki *et al*.[84] mapping algorithm would produce fewer QALYs for CTZ + MTX than for BARI + MTX, given the former's higher good EULAR response rate.

The company did not include the cost of SAEs in the base case but presented a scenario analysis in which these were accounted for. The company estimated SAE rates of 0.152, 0.049 and 0.095 for BARI, bDMARDs and cDMARDs/PALL, respectively, based on RA-BEAM trial data. In the severe RA populations, the comparative cost-effectiveness of BARI + MTX is slightly reduced, as incremental QALYs remain the same compared to other bDMARDs but the incremental costs are slightly higher. There is an exception in the severe RA, cDMARD-IR population, where the

incremental QALYs produced by ETN-b + MTX compared with BARI + MTX change from [redacted] in the base case to [redacted] when SAE costs are included. This seems to be a transcription error, as the results for ETN-b + MTX are identical to those for CTZ + MTX. In the case of the moderate RA population, the incremental costs of BARI + MTX compared with intensive cDMARDs is slightly higher than in the base case (£[redacted] versus £[redacted]) as it is expected given a higher SAE rate. However, the resulting ICER of BARI + MTX compared with intensive cDMARDs is lower than in the base case (£37,018 versus £37,420), due to an unexplained small increase in the incremental QALYs.

The company presented results of a scenario analysis for the severe RA cDMARD-IR population, that attempted to reflect the cost-effectiveness of tapering down BARI form 4mg QD to 2mg QD. It assumed that patients would be on 4mg until response assessment and then their dose would be tapered down if they had achieved a good EULAR response. The company implemented this by applying a HAQ increment of [redacted] based on data from RA-BEYOND[41] to patients tapered down to the 2mg QD dose. As a result, the cost-effectiveness of BARI + MTX was reduced compared with the bDMARDs, although the conclusion that BARI + MTX dominated the majority of its comparators (when CIC PASs were not incorporated) was not altered. The ERG notes that it is possible that some patients could lose their EULAR response when tapered down from 4mg to 2mg and therefore believes that the results of this scenario analysis could overestimate the cost-effectiveness of tapering down BARI to 2mg QD.

Within their scenario analyses, the company presented an analysis for a different population, the MTX-IR population, based on the RA-BEAM trial. This analysis consists of a head-to-head comparison between BARI + MTX and ADA + MTX based on the data from the RA-BEAM trial, which were used to inform the baseline characteristics of patients, EULAR response rates (from the mITT population) and the HAQ to EQ-5D mapping algorithm. The results show that BARI + MTX dominates ADA + MTX by being £[redacted] less expensive and producing [redacted] additional QALYs.

Additional analyses undertaken by the company in response to the ERG's clarification request

TCZ + MTX is included as the last bDMARD therapy in all the sequences considered for the severe population except for the sequence that starts with TCZ + MTX. In this case, the company assumed that ADA + MTX would be provided as last bDMARD therapy. The ERG requested clarification from the company regarding the reason to do so and requested an analysis where a more effective TNFi (CTZ + MTX) was provided instead (question B5 of the clarification letter[32]). The company explained that ADA was chosen as last bDMARD therapy for the TCZ + MTX sequence because it is the TNFi with the biggest market share. The company also presented the results of the analysis where CTZ + MTX was provided to patients who had been treated with TCZ + MTX as first bDMARD in

the cDMARD-IR population. The TCZ + MTX sequence, which was dominated by BARI + MTX in the base case, was estimated to be more effective than BARI + MTX when CTZ + MTX was used as last bDMARD but with an ICER of £84,106 per QALY gained compared with BARI + MTX. The ERG notes, however, that the confidential PAS for TCZ IV was not accounted for in the analysis. The company did not present results for the severe TNFi-IR, RTX-ineligible population.

In response to another request for clarification by the ERG (question B10[32]), the company presented the results of an analysis whereby patient baseline characteristics (i.e. gender, age, weight and HAQ) were correlated using a multivariate normal distribution (see Table 39). The ERG notes that patients' baseline characteristics should have been correlated in the base case. However, the company only presented the results for the severe cDMARDs-IR population. In this population, all interventions resulted in a higher number of QALYs compared with the base case. In both the base case and this scenario analysis, BARI + MTX dominated all its comparators except CTZ + MTX. The ERG notes that there are important differences in the total LYGs estimated for each intervention, an unintended effect that the company attributed to random variation. There is, therefore, considerable uncertainty regarding the reliability of these results.

Table 39: Scenario analysis including correlation between patient baseline characteristics for the severe, cDMARDs-IR population (deterministic)

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	Inc. analysis
IFX-b+MTX	[redacted]	14.58	[redacted]	[redacted]	[redacted]	Dominated
ABA SC+MTX#	[redacted]	14.59	[redacted]	[redacted]	[redacted]	Dominated#
GOL+MTX	[redacted]	14.59	[redacted]	[redacted]	[redacted]	Dominated
ADA+MTX	[redacted]	14.59	[redacted]	[redacted]	[redacted]	Dominated
ETN-b+MTX	[redacted]	14.60	[redacted]	[redacted]	[redacted]	Dominated
TCZ IV+MTX#	[redacted]	14.61	[redacted]	[redacted]	[redacted]	Dominated#
BARI+MTX	[redacted]	14.67	[redacted]	[redacted]	[redacted]	**Baseline**
CTZ+MTX	[redacted]	14.61	[redacted]	[redacted]	[redacted]	**£26,397**

*All treatments followed by sequence RTX+MTX→TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by RTX+MTX→ADA+MTX→MTX→ PALL. Confidential PAS for TCZ IV not included.
#Does not include confidential PAS

BARI: baricitinib; ABA: abatacept; TCZ: tocilizumab; ADA: adalimumab; CTZ: certolizumab pegol; ETN-b: etanercept biosimilar; GOL: golimumab; IFX-b: infliximab biosimilar; RTX: rituximab; MTX: methotrexate; PALL: palliative care; IV: intravenous; SC: subcutaneous

The company did not identify efficacy estimates of ADA, IFX, CTZ and ETN in combination with MTX in the severe TNFi-IR population. In order to overcome this lack of evidence, the company assumed that these treatments would have the same efficacy in the severe TNFi-IR population as in the severe cDMARD-IR population. Based on the evidence in Table 30 and Table 31, the ERG

considered that such an assumption was favourable to these treatments and requested an analysis excluding the treatments for which there was no evidence. Table 40 shows that BARI + MTX dominates GOL + MTX and that the ICER of ABA SC + MTX and TCZ IV + MTX compared with BARI + MTX is higher than £400,000 per QALY gained.

Table 40: Scenario analysis for the severe TNFi-IR population excluding ADA, IFX, ETN and CTZ (deterministic)

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)	ICER vs BARI + MTX (£/QALY)
GOL+MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	Dominated
BARI + MTX	[redacted]	13.49	[redacted]	[redacted]	[redacted]	**Baseline**	
ABA SC+MTX[#]	[redacted]	13.49	[redacted]	[redacted]	[redacted]	Dominated	£484,782[#]
TCZ IV+MTX[#]	[redacted]	13.50	[redacted]	[redacted]	[redacted]	£430,301	£430,301[#]

*All treatments followed by sequence TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by ABA+MTX→MTX→ PALL. Confidential PAS for TCZ IV not included.
[#]Does not include confidential PAS

BARI: baricitinib; ABA: abatacept; TCZ: tocilizumab; GOL: golimumab; MTX: methotrexate; PALL: palliative care; IV: intravenous; SC: subcutaneous

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.[84, 90, 91]
- 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.[1]
- 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[1] 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 potential limitations identified within the ERG's critical appraisal of the company's economic analysis are described under the following headings:

1. Limitations with the company's NMA
2. Face validity and reproducibility of scenario analyses
3. Limitations of the PSA
4. Using the efficacy of treatments in cDMARD-IR population for all bDMARDs in the sequence
5. Rounding to nearest HAQ score
6. Incorrect implementation on the HAQ trajectory classes
7. HAQ improvement for responders assumed immediate
8. Averaging HAQ across large time periods
9. Exclusion of ABA IV and TCZ SC from the list of comparators
10. Using an older mapping from HAQ score to EQ-5D than the AG
11. Assuming BARI would be inserted before intensive cDMARDs for patients with moderate RA
12. Different life years gained across sequences
13. Lack of consideration of the distribution of weight for interventions where the dosage is weight based
14. Dosage of IFX

1) Limitations with the company's NMA

The results of the company's NMA should be treated with caution owing to a random effects model being assumed for the baselines, which was deemed inappropriate. In addition, simultaneous baseline and treatment effect models were used without ensuring that information in the baseline model did not propagate to the relative treatment effect model. Furthermore, studies that reported EULAR responses were synthesised along with converted EULAR response outcomes from studies that only reported ACR responses.

2) Face validity and reproducibility of scenario analyses

The company presented a comprehensive list of scenario analyses. However, several results presented in Tables 128 to 134 of the CS[1] lack face validity, as explained in Section 5.2.11. Some of these appear to be transcription mistakes whilst at least one (ABA + MTX in the BARI tapering to 2mg scenario analysis) is due to a programming error. There are also inconsistencies between the

description of scenario analyses and the implementation of options in the model. For example, in one of the scenario analyses, the HAQ improvement upon response for patients on BARI is based on data from the BARI trials. However, in the model the only similar option is to apply the HAQ improvement based on the BARI trials to all interventions and not only to BARI. In either case, the ERG was unable to reproduce the results of this scenario analysis, even using the same seed for random number generation and the same number of patients.

3) *Limitations of the PSA*

As explained in Section 5.2.10.4, there were important differences in the severe TNFi-IR, RTX-ineligible population between the results of the deterministic analysis and those of the PSA. For example, in some interventions (TCZ + MTX, ETN + MTX and ADA + MTX), the estimated costs varied by more than 10% whilst in others (BARI + MTX, CTZ + MTX and ABA + MTX) they barely changed. While investigating this, the ERG discovered an error in the model that affected the PSA of the severe TNFi-IR RTX ineligible population. The programming error resulted in patients on GOL + MTX, ETN + MTX, ADA + MTX and IFX + MTX never achieving a good or moderate EULAR response. This affected the sequence starting with TCZ + MTX, given that ADA + MTX is included in the sequence. The error was originated by the way the model copies samples from the CODA to the "Data inputs" sheet, where they are used by the calculations, and given that there are no CODA samples in the model for GOL + MTX, ETN + MTX, ADA + MTX and IFX + MTX. The ERG corrected this error and re-run the PSA, whose results are presented in Section 5.4.

The ERG discovered a further programming error in the model, more precisely in the calculations of the CODA samples for moderate response probability for BARI + MTX in the cDMARD-IR population, that affected the PSA. The error consisted in the probability of moderate response being calculated subtracting the probabilities of good or moderate response by the probability of good response of a different sample, instead of that of the same sample. This error affects row 22 of the 24wks_EULARM worksheet between columns "C" to "ALM" and resulted in negative values. The ERG notes that the results of the probabilistic analysis for the cDMARD-IR population differed only minimally after correcting this programming error.

4) *Using the efficacy of treatments in cDMARD-IR population for all bDMARDs in the sequence*

For the severe cDMARD-IR population, the company used the efficacy estimates of the cDMARD-IR population (see Table 30) for all the treatments in the sequence, regardless of their position in the sequence. This implies that, for example, TCZ IV + MTX will have the same efficacy irrespective of whether it is first or third in the sequence. However, Table 30 and Table 31 show that the efficacy of TCZ IV + MTX, as well as all the other treatments, is lower after the treatment with a TNFi. The ERG considers that after the first treatment, using the efficacy estimates of the TNFi-IR population is more

appropriate. Therefore, for the severe cDMARD-IR population, effectiveness estimates from Table 30 should have been used only for the first treatment in the sequence. For the rest of the treatments in the sequence, the estimates from Table 31 should have been used instead. Similarly, the effectiveness of treatments decays after the second treatment line. In the STA for certolizumab pegol,[27] this was taken into account by using response rates as well as HAQ improvements based on the RADIATE study,[92] which analysed the efficacy of TOC + MTX compared with PBO + MTX and in which approximately half of the patients had received two or more TNFis.

5) *Rounding to nearest 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 company's 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 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. The company approached this issue by rounding the values to the nearest valid HAQ score. The ERG notes that this approach might lead to biased estimations of HAQ scores, as values might be rounded up more often than rounded down or *vice versa*. An example would be that of small changes (lower than 0.0625) that would always be rounded down to zero. In order to avoid this problem, the AG in TA375[23] rounded up with a probability inversely proportional to the distance of the value to the closest valid HAQ score, and rounded down otherwise. For example, a change of 0.4 would have a probability of 0.8 of being rounded down to 0.375 and a probability of 0.2 of being rounded up to 0.5.

6) *Incorrect implementation on the HAQ trajectory classes*

The company implemented the trajectory of HAQ score whilst on cDMARDs or palliative care based on the approach taken by the AG in TA375.[23] This approach was based on a Norton *et al.* study[93] where four latent classes of HAQ trajectories had been identified. The AG in TA375[23] calculated the probability of class membership of each patient based on a series of factors and covariates with the HAQ trajectory of each patient calculated as an average of the trajectories of each class weighted by the probability of class membership. However, the company assigned each patient to a single class based on the probability of class membership and assumed that all patients would follow exactly one of four possible trajectories.

7) *HAQ improvement for responders assumed immediate*

The company assumed that patients who achieved moderate or good EULAR response at 24 weeks would experience a reduction in HAQ score instantaneously at treatment initiation. The ERG believes

that, whilst it is likely that most patients would enjoy of an improvement of HAQ score before week 24, the company's approach is likely to lead to an overestimation of treatment benefits, as achievement of response will take at least a few weeks and up to 24 weeks for some patients. Similarly, the assumption made by the AG of TA375 of no HAQ improvement in the first 6 months is likely to underestimate treatment benefit.

8) Averaging HAQ across large time periods

In order to calculate the QALYs and costs produced in the time span between two events, the model uses and area under the curve approach for the HAQ score, and then maps this value to EQ-5D and hospitalisation costs. However, since the relationships between HAQ score and EQ-5D and between HAQ score and hospitalisation costs are not linear, this approach will lead to inaccurate results.

9) Exclusion of ABA IV and TCZ SC from the list of comparators

The ABA IV and TCZ SC formulations were not included as comparators, although ABA SC and TCZ IV were included. In response to a clarification request by the ERG (question B9),[32] the company stated that it had excluded TCZ SC because the available evidence for TCZ SC was limited, it provided a lower efficacy estimate than for TCZ IV and the cost difference between the two formulations was relatively small. The ERG notes that the difference in costs might be considerable taking into account the administration costs and the CIC PAS. ABA IV was included in the NMA, but excluded from the economic analyses. In response to a clarification request by the ERG (question B15),[32] the company explained that ABA IV was excluded from the analyses to limit the number of sequences considered and because it was unlikely to be informative. The company also presented the results of ABA IV for the severe cDMARD-IR population in their clarification response, which led to similar results compared with ABA SC (£[redacted] versus £[redacted] and [redacted] versus [redacted] QALYs respectively).

10) Using an older mapping from HAQ to EQ-5D than the AG

The company used an algorithm proposed by Hernández Alava *et al.*[78] to map HAQ scores to EQ-5D. The company explained in the CS[1] that the three-class model by Hernández Alava *et al.*[78] was used in the model because the predicted values for selected combinations of covariates were reported in the published article, which allowed the validation of their replicated mapping algorithm. The ERG notes that more recent algorithms have been since published, such as that reported by Hernández Alava *et al.*[83] that purport to have a higher accuracy: this algorithm was used in TA375, which includes all the parameters necessary to implement it. However, the ERG acknowledges that the company's scenario analyses show that the mapping algorithm does not have an important impact on the result of the analyses.

11) Assuming BARI would be inserted before intensive cDMARDs for patients with moderate RA

The ERG notes that in TA375[23] it was assumed that intensive cDMARDs had already been used, as was supported by clinical opinion, instead of assuming that a bDMARD would go in front of intensive cDMARDs. The impact of the change, which was not supported by the clinical advisors to the ERG, is unknown.

12) Different life years gained across sequences

The company assumed,[1] as was the approach used by the AG in TA375,[23] that HAQ scores at baseline affected mortality. However, the company's model re-estimated the age of death at every event, which led to slightly different expected life years gained, as observed in the results. The ERG notes that sequences of different lengths are likely to produce different life expectancies, as age of death is recalculated at different time points. The company also attributed the differences in life years gained to random variation. The ERG notes that using the same random numbers across treatments as the company did in other cases would have eliminated the impact of random variation. Alternatively, more patients should have been run through the model to minimise the impact of random variation.

13) Lack of consideration of the distribution of weight for interventions where the dosage is weight based

The company used the average weight of the population in the relevant BARI trials (see Table 112 of the CS[1]) to calculate the average dose cost. This is equivalent to assuming that all patients had the average weight. The ERG notes that this approach to calculating the average dose cost is not appropriate given that the relationship between weight and dosing cost is not linear as explained by Hatswell *et al*.[85] This is because the average cost of a dose is not necessarily equal to the cost of the patient with the average weight, due to drug wastage and differences in cost per mg of some drugs. The ERG notes that the company should have calculated the average cost of a dose using the distribution of the weight of patients instead of using the average weight.

14) Dosage of IFX

The company assumed patients would receive seven doses of IFX per year and an additional dose in the first year. However, the SmPC[86] and the BNF[80] state that patients should receive IFX doses at 2 and 6 weeks after the first infusion and every 8 weeks thereafter. This implies that the average doses of IFX per year should be 6.52 instead of 7.

5.4 Additional exploratory analyses undertaken by the ERG

The ERG undertook few exploratory analyses based on the company's submitted model as the ERG believes that [redacted]

However, the ERG found two programming errors that affected the company's PSA results, as explained in Section 5.3, and re-run the PSA after fixing them. The ERG's PSA simulated 1000 patients in 1000 iterations. The results of the ERG's PSA are presented in Table 41 and Table 42.

The programming error that affected the PSA of the severe cDMARD-IR population had a minimal impact on the results as can be seen by comparing Table 35 with Table 41. The resulting cost-effectiveness acceptability curve for the severe cDMARD-IR population is presented in Figure 8 with CZP + MTX having the greatest probability of being most cost-effective at ICERs of £20,000 or more, followed by BARI + MTX.

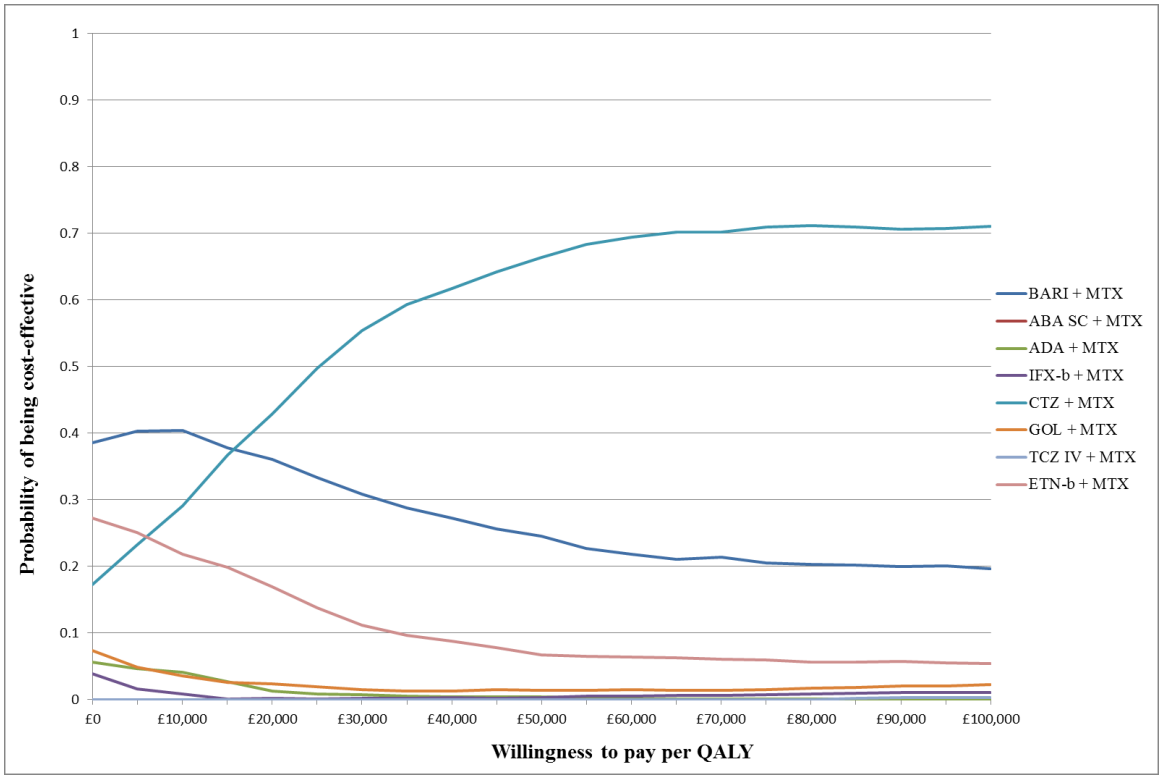
Table 41: ERG-amended base case analysis results for the severe cDMARD-IR population (probabilistic)

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	Inc. analysis
IFX-b+MTX	[redacted]	14.72	[redacted]	[redacted]	[redacted]	Dominated
ABA SC+MTX[#]	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated[#]
ADA+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
GOL+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
TCZ IV+MTX[#]	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated[#]
ETN-b+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	Dominated
BARI+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	**Baseline**
CTZ+MTX	[redacted]	14.71	[redacted]	[redacted]	[redacted]	18,135

*All treatments followed by sequence RTX+MTX→TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by RTX+MTX→ADA+MTX→MTX→ PALL. Confidential PAS for TCZ IV not included.
[#]Does not include confidential PAS

BARI: baricitinib; ABA: abatacept; TCZ: tocilizumab; ADA: adalimumab; CTZ: certolizumab pegol; ETN-b: etanercept biosimilar; GOL: golimumab; IFX-b: infliximab biosimilar; RTX: rituximab; MTX: methotrexate; PALL: palliative care; IV: intravenous; SC: subcutaneous

Figure 8: Cost-effectiveness acceptability curve for the severe cDMARD-IR population (ERG-amended)



In contrast, the programming error that affected the severe TNFi-IR RTX-ineligible population had an important impact in the sequences effected. These differences, which can be seen comparing Table 38 with Table 42, can be summarised as markedly higher costs and QALYs gained for TCZ + MTX, ETN-b + MTX, IFX-b + MTX, GOL + MTX and ADA + MTX. The resulting cost-effectiveness acceptability curve for the severe cDMARD-IR population is presented in Figure 9 with CZP + MTX having the greatest probability of being most cost-effective at ICERs of £20,000 or more. The ERG comments that the results in the TNFi-IR RTX-ineligible population are confounded by the assumption for some interventions that the EULAR responses obtained in a cDMARD-IR population was applicable to the TNFi-IR RTX-ineligible population

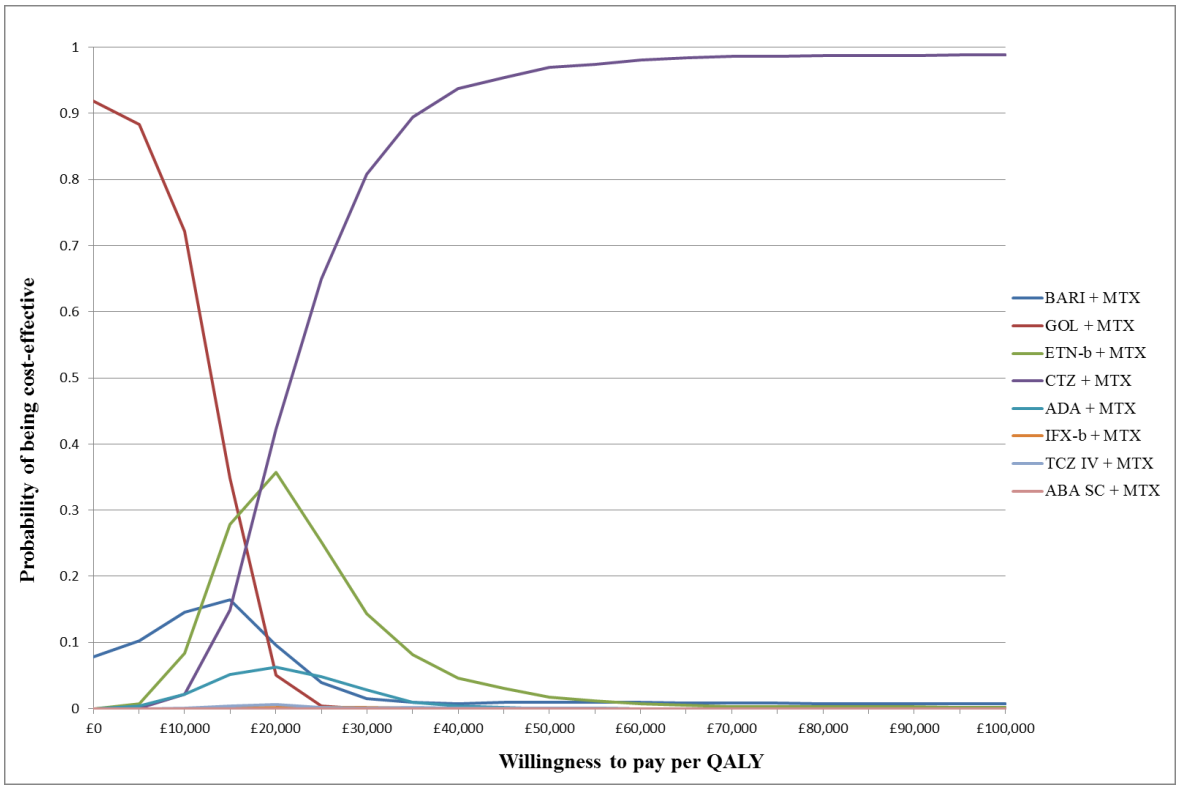
Table 42: ERG-amended base-case analysis results for the severe TNFi-IR RTX-ineligible population (probabilistic)

Interventions*	Total QALYs	Total LYG	Total costs (£)	Inc. QALYs	Inc. costs (£)	ICER (£/QALY)	ICER vs BARI + MTX (£/QALY)
GOL+MTX	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Baseline	18,805§
BARI + MTX	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Extendedly dominated	
ABT SC+MTX#	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	454,225#
TCZ IV+MTX#	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	37,063#
ADA+MTX; †	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	21,494†
ETN-b+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	£15,527	10,197†
IFX-b+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	Dominated	35,045†
CTZ+MTX†	[redacted]	13.52	[redacted]	[redacted]	[redacted]	£20,170	16,962†

*All treatments followed by sequence TCZ IV+MTX→MTX→PALL except TCZ IV+MTX, which is followed by ADA+MTX→MTX→PALL. Confidential PAS for TCZ IV not included.
#Does not include confidential PAS
†Efficacy estimates assumed to be equal to those for the severe cDMARD-IR population
§ This interventions is less effective than BARI + MTX and therefore the ICER represents savings per QALY lost compared with BARI + MTX

BARI: baricitinib; ABA: abatacept; TCZ: tocilizumab; ADA: adalimumab; CTZ: certolizumab pegol; ETN-b: etanercept biosimilar; GOL: golimumab; IFX-b: infliximab biosimilar; RTX: rituximab; MTX: methotrexate; PALL: palliative care; IV: intravenous; SC: subcutaneous

Figure 9: Cost-effectiveness acceptability curve for the severe TNFi-IR RTX-ineligible population (ERG-amended)



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 BARI + 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 BARI + MTX.

The company's *de novo* economic model was largely based on the model developed by the AG in TA375.[23] Costs and health outcomes for BARI + MTX and its comparators were estimated from the perspective of the NHS over a lifetime horizon. The analyses presented in the CS relate to four different populations of RA patients: moderate RA cDMARD-IR; severe RA cDMARD-IR; severe RA TNFi-IR who are RTX eligible; and, severe RA TNFi-IR who are not eligible for RTX.

In the moderate RA cDMARD-IR population, BARI + MTX treatment was estimated to produce [redacted] additional QALYs compared with intensive cDMARDs at an extra cost of £[redacted] resulting in an ICER of £37,420 per QALY gained. In the severe RA cDMARD-IR population, BARI + MTX dominated all of its comparators except CTZ + MTX: the ICER of CTZ + MTZ compared with BARI + MTX was estimated to be £18,400 per QALY gained. In the severe RA TNFi-IR population, when RTX + MTX was an option, BARI + MTX was dominated by RTX + MTX. In the severe RA TNFi-IR population, when RTX + MTX was not an option, BARI + MTX dominated GOL + MTX and less effective and less expensive than the rest of its comparators. The ICERs for ETN-b + MTX, CTZ + MTX and ADA + MTX compared with BARI + MTX were estimated to be lower than £30,000 per QALY. However, the company made a favourable assumption for these interventions in the absence of relevant effectiveness data and therefore, caution is needed when interpreting these results. The ICERs for TCZ IV + MTX and ABA SC + MTX compared with BARI + MTX were estimated to be higher than £30,000 per QALY, but their confidential PAS prices were not included.

For patients with moderate RA, the company's model estimated an ICER (£37,420) which was lower than that estimated by the AG in TA375, which was in the region of £50,000 per QALY for bDMARDs. Given that the company drew heavily on the AG's model but made implementation errors (see Section 5.3) the ERG believes that the ICER for BARI in moderate RA would be nearer to £50,000 per QALY.

The ERG undertook few exploratory analyses based on the company's submitted model, [redacted].

However, the ERG presents the results of the analyses using the company's model whilst incorporating the confidential PAS currently in place for TCZ IV and ABA SC in a confidential appendix.

There remain several potentially important areas of uncertainty:

1. Cost-effectiveness of BARI monotherapy

   Despite BARI being licensed as monotherapy, the company did not provide an economic analysis of the cost-effectiveness of BARI as monotherapy for patients for whom MTX is contraindicated or not tolerated. The ERG asked why evidence from RA-BEGIN was not used to estimate the efficacy of BARI monotherapy (clarification question A2)[32]. The company responded that the population from RA-BEGIN was naïve to MTX and for that reason it should not be combined with data of patients with prior cDMARD and/or TNFi experience. The company referred to the guidance developed in TA375,[23] which recommended bDMARDs as monotherapy within their marketing authorisation on the grounds that people with severe RA who cannot tolerate MTX should not be treated differently and claimed that similar rationale should apply to BARI. However, data from RA-BEGIN[31] showed that the addition of MTX to BARI 4mg did not produce a marked improvement over BARI monotherapy in a MTX-naïve population. This therefore provides supportive evidence regarding the efficacy of BARI monotherapy.

2. Cost-effectiveness of BARI + MTX in the severe TNFi-IR, RTX-ineligible population

   The company only identified evidence on the effectiveness of TCZ, GOL and ABA in combination with MTX for the severe TNFi-IR population. BARI + MTX dominated GOL + MTX, but was less effective than TCZ + MTX and ABA + MTX. The cost-effectiveness of BARI + MTX compared with ETN, IFX, ADA and CTZ in combination with MTX in the severe TNFi-IR population is unknown and it can only be estimated making assumptions on the effectiveness of these comparators.

3. Efficacy of bDMARDs after BARI

   No evidence was presented by the company on the effectiveness of bDMARDs after BARI. In the company's economic analysis, it was assumed that the efficacy of bDMARDs after BARI will be equal to their efficacy after another bDMARD. This is a reasonable assumption to make given the lack of evidence, but the ERG notes that it is possible that the efficacy of bDMARDs after BARI could be better (or worse) than when following another bDMARD.

4. Time to treatment discontinuation of BARI

In their base case, the company assumed that time to treatment discontinuation for BARI is the same as for bDMARDs. The company justified this assumption referring to the similarity of the discontinuation rates from RA-BEAM (6.8% for BARI and 5.6% for ADA). As BARI is an oral treatment, it is plausible that adherence would be higher than for SC or IV treatments. The company presented results of a scenario analysis where time to treatment discontinuation for BARI was based on data from the RA-BEAM trial. However, the ERG notes that adherence is likely to be higher in a trial than in real practice and therefore this comparison may be confounded.

5. Relative HAQ progression whilst on BARI compared with bDMARDs

   In the company's base case it was assumed that, in line with bDMARDs, there would be no HAQ progression whilst on treatment. BARI is not a bDMARD and the relative HAQ progression for BARI compared with bDMARDs is uncertain. Clinical advice provided to the ERG suggested that there was no reason to believe that the progression in HAQ would be worse, or better, than for bDMARDs.

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[1] for BARI 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 BARI would provide an additional 3 months of life compared with its comparators.

7 OVERALL CONCLUSIONS

BARI + MTX treatment was estimated by the company to have an ICER of £37,420 per QALY gained compared with intensive cDMARDs in the moderate RA cDMARD-IR population. In the severe RA cDMARD-IR population, BARI + MTX dominated all its comparators except for CTZ + MTX: the ICER of CTZ + MTZ compared with BARI + MTX was estimated to be £18,400 per QALY gained. In the severe RA TNFi-IR population, when RTX + MTX was an option, BARI + MTX was dominated by RTX + MTX. In patients with severe RA who have had inadequate response to a TNFi and for whom RTX is contraindicated or not tolerated, BARI + MTX dominated GOL + MTX and was less effective and less expensive than the rest of its comparators. The ICERs for ETN-b + MTX, CTZ + MTX and ADA + MTX compared with BARI + MTX were estimated to be lower than £30,000 per QALY gained whilst the ICERs TCZ IV + MTX and ABA SC + MTX compared with BARI + MTX were estimated to be higher than £30,000 per QALY gained. However, the confidential PASs for ABA and TCZ were not included in these analyses.

The ERG's critical appraisal identified a number of issues relating to the company's model and analysis. The ERG believes that: the NMA is subject to potential limitations; some of the scenario analyses as well as the PSA for the severe TNFi-IR, RTX-ineligible population lack face validity; the efficacy estimates for the cDMARD-IR population should only be used for the first line of treatment; rounding to nearest HAQ score might introduce bias; the HAQ trajectory of a patient on cDMARDs or palliative care should be calculated as a weighted average; assuming HAQ improvement upon treatment initiation overestimates treatment benefit; averaging HAQ across large time periods leads to inaccuracies in the calculation of costs and QALYs; ABA IV and TCZ SC should have been included in the list of comparators; newer mapping algorithms from HAQ scores to EQ-5D should have been used; BARI should have not been assumed to be provided before intensive cDMARDs for moderate patients; mortality rates differ between sequences; the distribution of weight for interventions where the dosage is weight based should have been considered; and the dosage of IFX is inaccurate.

There remain several potentially important areas of uncertainty:

1. Cost-effectiveness of BARI monotherapy
2. Cost-effectiveness of BARI + MTX in the severe, TNFi-IR, RTX-ineligible population
3. Efficacy of bDMARDs after BARI
4. Time to treatment discontinuation of BARI
5. Relative HAQ progression whilst on BARI compared with bDMARDs

[redacted] In the moderate RA population, the AG in TA375 estimated that the median ICER of bDMARDs compared with cDMARDs was in the region of £50,000 per QALY gained.

[redacted]

8 REFERENCES

1. Eli Lilly. Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis [ID979]; 2017.
2. Scott DL, Steer S. The course of established rheumatoid arthritis. *Best Practice & Research Clinical Rheumatology* 2007;21:943-67.
3. 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.
4. 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.
5. 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.
6. Naz SM, Symmons DPM. Mortality in established rheumatoid arthritis. *Best Practice & Research Clinical Rheumatology* 2007;21:871-83.
7. 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.
8. 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.
9. National Collaborating Centre for Chronic Conditions. Rheumatoid arthritis: national clinical guideline for management and treatment in adults: Royal College of Physicians (UK); 2009.
10. 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.
11. 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.
12. 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.
13. 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. *Annals of the Rheumatic Diseases* 1996;55:89-93.
14. 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.
15. 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). *Annals of the Rheumatic Diseases* 2007;66:34-45.
16. 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. *Annals of the Rheumatic Diseases* 2010;69:1580-8.

17. 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.

18. 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.

19. Felson DT. Assessing the efficacy and safety of rheumatic disease treatments: Obstacles and proposed solutions. *Arthritis & Rheumatism* 2003;48:1781-7.

20. 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.

21. 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.

22. 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.

23. 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.

24. Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23:S14-S8.

25. 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.

26. National Institute for Health and Clinical Excellence (NICE). Tocilizumab for the treatment of rheumatoid arthritis.; 2012.

27. 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.

28. Shi JG, Chen X, Lee F, Emm T, Scherle PA, Lo Y*, et al.* The pharmacokinetics, pharmacodynamics, and safety of baricitinib, an oral JAK 1/2 inhibitor, in healthy volunteers. *J Clin Pharmacol* 2014;54:1354-61.

29. Shuai K, Liu B. Regulation of JAK-STAT signalling in the immune system. *Nature reviews Immunology* 2003;3:900-11.

30. O'Shea JJ, Plenge R. JAK and STAT signaling molecules in immunoregulation and immune-mediated disease. *Immunity* 2012;36:542-50.

31. Fleischmann R, Schiff M, van der Heijde D, Ramos-Remus C, Spindler A, Stanislav M, *et al.* Baricitinib, Methotrexate, or Combination in Patients with Rheumatoid Arthritis and No or Limited Prior Disease-Modifying Antirheumatic Drug Treatment. *Arthritis & Rheumatology* 2016;09:09.

32. Eli Lilly. Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis [ID979]. Clarification response; 2017.

33. van de Putte LBA, Atkins C, Malaise M, Sany J, Russell A, van Riel PLCM, *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.

34. Centre for Reviews 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.

35. 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. *New England Journal of Medicine* 2017;376:652-62.

36. 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* 2016;76:88-95.

37. Genovese MC, Kremer J, Zamani O, Ludivico C, Krogulec M, Xie L, *et al.* Baricitinib in Patients with Refractory Rheumatoid Arthritis. *The New England Journal of Medicine* 2016;374:1243-52.

38. Eli Lilly and Company. Data on File (JADV Clinical Study Report). 2016.

39. Eli Lilly and Company. Data on File (JADW Clinical Study Report). . 2016.

40. Eli Lilly and Company. Data on File (JADX Clinical Study Report). 2016.

41. Eli Lilly and Company. JADY Abbreviated Clinical Study Report. 2016.

42. Smolen JS, Kremer JM, Gaich CL, DeLozier AM, Schlichting DE, Xie L, *et al.* Patient-reported outcomes from a randomised phase III study of baricitinib in patients with rheumatoid arthritis and an inadequate response to biological agents (RA-BEACON). *Annals of the Rheumatic Diseases* 2016;31:31.

43. Eli Lilly and Company. JADY Abbreviated Clinical Study Report; 2016.

44. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline - Clinical safety data management: definitions and standards for expedited reporting E2A. 1994. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf (Accessed 20-3-17)

45. Eli Lilly and Company. Data on File (Summary of Clinical Safety. Appendix 1. Table APP1.2.7.4.34. Page 197). 2016.

46. Eli Lilly and Company. Clinical Overview Baricitinib Rheumatoid Arthritis. 2016.

47. Eli Lilly and Company. Data on File (Clinical Overview Baricitinib Rheumatoid Arthritis. Table 2.5.5.2. Page 64). 2016.

48. Agency EM. Olumiant (baricitinib). *Science Medicines Health*;http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/004085/human_med_002074.jsp&mid=WC0b01ac058001d124. (Accessed 20-3-17)

49. Li Z, Zhang F, Kay J, Fei K, Han C, Zhuang Y*, et al.* Safety and efficacy of subcutaneous golimumab in chinese patients with active rheumatoid arthritis despite MTX therapy: Results from a randomized, placebo-controlled, phase 3 trial. *Annals of the Rheumatic Diseases Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR* 2013;72.

50. 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.

51. 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.

52. 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. *Annals of the Rheumatic Diseases* 2016;75:170-7.

53. 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(9):1629-35.

54. van Vollenhoven RF, Kinnman N, Vincent E, Wax S, Bathon J. Atacicept in patients with rheumatoid arthritis and an inadequate response to methotrexate: results of a phase II, randomized, placebo-controlled trial. *Arthritis & Rheumatism* 2011;63:1782-92.

55. Keystone EC, Wang MM, Layton M, Hollis S, McInnes IB, Study T. Clinical evaluation of the efficacy of the P2X7 purinergic receptor antagonist AZD9056 on the signs and symptoms of rheumatoid arthritis in patients with active disease despite treatment with methotrexate or sulphasalazine.[Erratum appears in Ann Rheum Dis. 2012 Dec;71(12):2064]. *Annals of the Rheumatic Diseases* 2012;71:1630-5.

56. 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.

57. Kremer J, Zerbini C, Lee EB, Gruben D, Krishnaswami S, Zwillich SH*, et al.* Tofacitinib (CP-690,550), an oral janus kinase inhibitor: Analyses of efficacy endpoints by subgroups in a pooled phase 2 and 3 rheumatoid arthritis study population. *Annals of the Rheumatic Disease Conference: Annual European Congress of Rheumatology of the European League Against Rheumatism, EULAR* 2012;71.

58. Genovese MC, Tena CP, Covarrubias A, Leon G, Mysler E, Keiserman M*, et al.* Subcutaneous abatacept for the treatment of rheumatoid arthritis: longterm data from the ACQUIRE trial. *Journal of Rheumatology* 2014;41:629-39.

59. 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.

60. Smolen J, Emery P, Ferraccioli G. Efficacy and safety of certolizumab pegol after incomplete response to DMARDS in RA patients with low moderate disease activity: Results from certain, a phase IIIb study. *Annals of the Rheumatic Diseases* 2011;70:259.

61. Nishimoto N, Hashimoto J, Miyasaka N, Yamamoto K, Kawai S, Takeuchi T, *et al.* Study of active controlled monotherapy used for rheumatoid arthritis, an IL-6 inhibitor (SAMURAI): Evidence of clinical and radiographic benefit from an x ray reader-blinded randomised controlled trial of tocilizumab. *Annals of the Rheumatic Diseases* 2007;66:1162-7.

62. Nishimoto N, Hashimoto J, Miyasaka N, Yamamoto K, Kawai S, Takeuchi T, *et al.* Study of active controlled monotherapy used for rheumatoid arthritis, an IL-6 inhibitor (SAMURAI): evidence of clinical and radiographic benefit from an x ray reader-blinded randomised controlled trial of tocilizumab. *Annals of the Rheumatic Diseases* 2007;66:1162-7.

63. 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. *Annals of the Rheumatic Diseases* 2013;72:1927-33.

64. Scott DL, Ibrahim F, Farewell V, Go'Keeffe A, Walker D, Kelly C, *et al.* Tumour necrosis factor inhibitors versus combination intensive therapy with conventional disease modifying anti-rheumatic drugs in established rheumatoid arthritis: TACIT non-inferiority randomised controlled trial. *Journal* 2015. http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/304/CN-01068304/frame.html (Accessed 20-3-17)

65. Takeuchi T, Kaneko Y, Atsumi T, Tanaka Y, Inoh M, Kobayashi H, *et al.* 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). *Journal* 2013. http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/554/CN-01011554/frame.html (Accessed 20-3-17)

66. Iwahashi M, Inoue H, Matsubara T, Tanaka T, Amano K, Kanamono T, *et al.* Efficacy, safety, pharmacokinetics and immunogenicity of abatacept administered subcutaneously or intravenously in Japanese patients with rheumatoid arthritis and inadequate response to methotrexate: a Phase II/III, randomized study. *Modern Rheumatology* 2014;24:885-91.

67. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;33:607-17.

68. Klareskog L, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, *et al.* Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. *Lancet (London, England)* 2004;363:675-81.

69. 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.

70. 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 Rheum* 2003;48:35-45.

71. 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.

72. National Institute for Health and Clinical Excellence (NICE). Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis. 2007.

73. National Institute for Health and Clinical Excellence (NICE). Certolizumab pegol for the treatment of rheumatoid arthritis; 2010.

74. National Institute for Health and Clinical Excellence (NICE). Abatacept for treating rheumatoid arthritis after the failure of conventional disease-modifying anti-rheumatic drugs (rapid review of technology appraisal guidance 234); 2013.

75. National Institute for Health and Clinical Excellence (NICE). Golimumab for the treatment of rheumatoid arthritis after the failure of previous disease-modifying anti-rheumatic drugs.; 2011.

76. Diamantopoulos A, Finckh A, Huizinga T, Sungher DK, Sawyer L, Neto D, *et al.* Tocilizumab in the treatment of rheumatoid arthritis: a cost-effectiveness analysis in the UK. *Pharmacoeconomics* 2014;32:775-87.

77. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal; 2013.

78. Hernández Alava M, Wailoo AJ, Ara R. Tails from the peak district: adjusted limited dependent variable mixture models of EQ-5D questionnaire health state utility values. *Value Health* 2012;15.

79. 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.

80. British National Formulary. British National Formulary (BNF) 71. 2016.

81. Deparment of Health. NHS reference Costs 2014 to 2015. 2015. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/480791/2014-15_National_Schedules.xlsx (Accessed 24 March).

82. 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 20-3-17)

83. 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.

84. 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.

85. Hatswell AJ, Porter J, Lee D, Hertel N, Latimer NR. The Cost of Costing Treatments Incorrectly: Errors in the Application of Drug Prices in Economic Evaluation Due to Failing to Account for the Distribution of Patient Weight. *Value in Health*;19:1055-8.

86. European Medicines Agency. Remicade. Summary of product characteristics.; 2016. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000240/WC500050888.pdf (Accessed 24-3-17)

87. National Institute for Health and Care Excellence. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor. 2010. https://www.nice.org.uk/guidance/ta195 (Accessed 24-3-17)

88. 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

89. Curtis L, Burns A. Unit costs of health and social care 2015. ; 2015.

90. 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.

91. 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.

92. 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. *Annals of the Rheumatic Diseases* 2008;67:1516-23.

93. 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.

9 APPENDICES

Appendix 1: Studies of BARI identified by a trial registry search

Table 43: Table of trial registry search

NCT number	Title	Recruitment	Include / exclude
NCT02265705	A Study of Baricitinib (LY3009104) in Participants With Rheumatoid Arthritis (RA)	Active, not recruiting	Not in CS Phase 3 study (JAGS) Estimated study completion date June 2017
NCT01885078	An Extension Study in Participants With Moderate to Severe Rheumatoid Arthritis	Recruiting	Included in CS RA-BEYOND JADY
NCT01710358	A Study in Moderate to Severe Rheumatoid Arthritis	Completed	Included in CS RA-BEAM JADV
NCT01721044	A Moderate to Severe Rheumatoid Arthritis Study	Completed	Included in CS RA-BEACON JADW
NCT01721057	A Study in Moderate to Severe Rheumatoid Arthritis Participants	Completed	Included in CS RA-BUILD JADX
NCT01711359	A Study in Participants With Moderate to Severe Rheumatoid Arthritis	Completed	Included in CS RA-BEGIN JADZ
NCT00902486	INCB028050 Compared to Background Therapy in Patients With Active Rheumatoid Arthritis (RA) With Inadequate Response to Disease Modifying Anti-Rheumatic Drugs	Completed	Included in CS Appendix 19 Phase 2 JADC
NCT01185353	A Study in Participants With Rheumatoid Arthritis on Background Methotrexate	Completed	Included in CS Appendix 19 Phase 2

NCT number	Title	Recruitment	Include / exclude
	Therapy		JADA
NCT01469013	Oral JAK1/JAK2 Selective Inhibitor Treatment in Japanese Participants With Active Rheumatoid Arthritis on Background Methotrexate Therapy	Completed	Included in CS Appendix 19 Phase 2 JADN
NCT02708095	A Study of Baricitinib (LY3009104) in Participants With Systemic Lupus Erythematosus (SLE)	Recruiting	Exclude – population not RA
NCT01398475	A Relative Bioavailability and Food Effect Study of New Formulations	Completed	Exclude – phase 1 study
NCT02759731	Study of Baricitinib, a JAK1/2 Inhibitor, in Chronic Graft-Versus-Host Disease After Allogeneic Hematopoietic Stem Cell Transplantation	Recruiting	Exclude – population not RA

Appendix 2: Forest plots generated from the ERG's NMA

Relative treatment effects are presented on the probit scale, with positive values favouring BARI 4mg + cDMARDs

Figure 10: EULAR results at week 24 for the cDMARD-IR population on probit scale

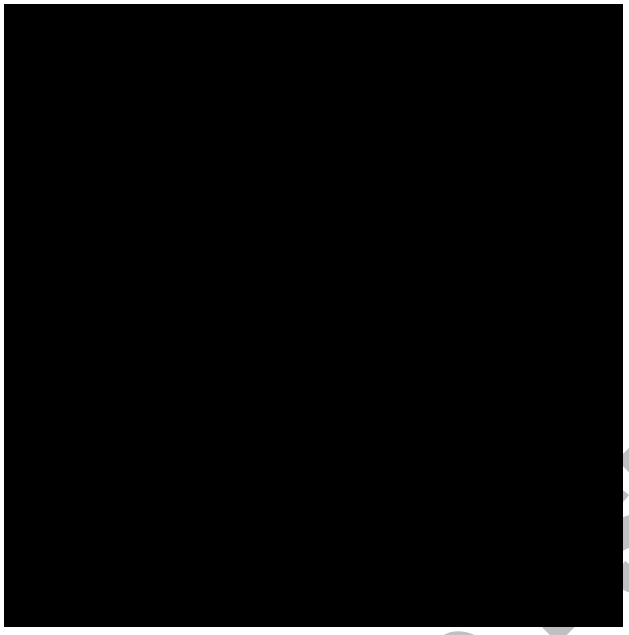


Figure 11: ACR results at week 24 for the cDMARD-IR population on probit scale

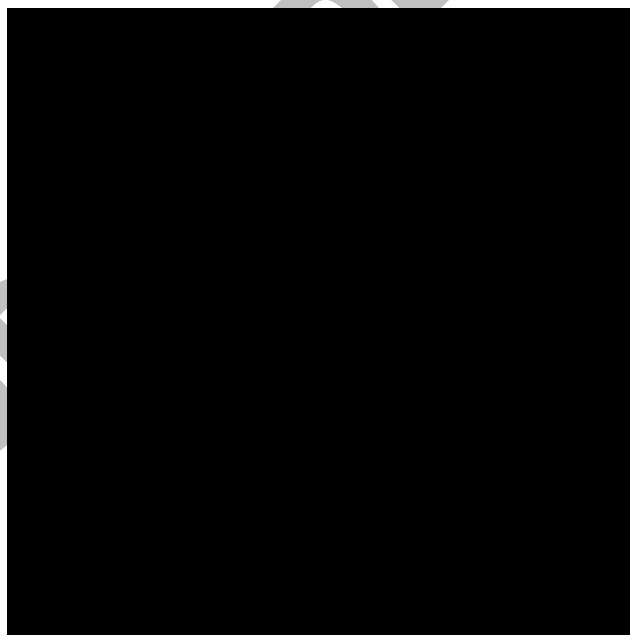


Figure 12: EULAR results at week 24 for the TNFi-IR population on probit scale

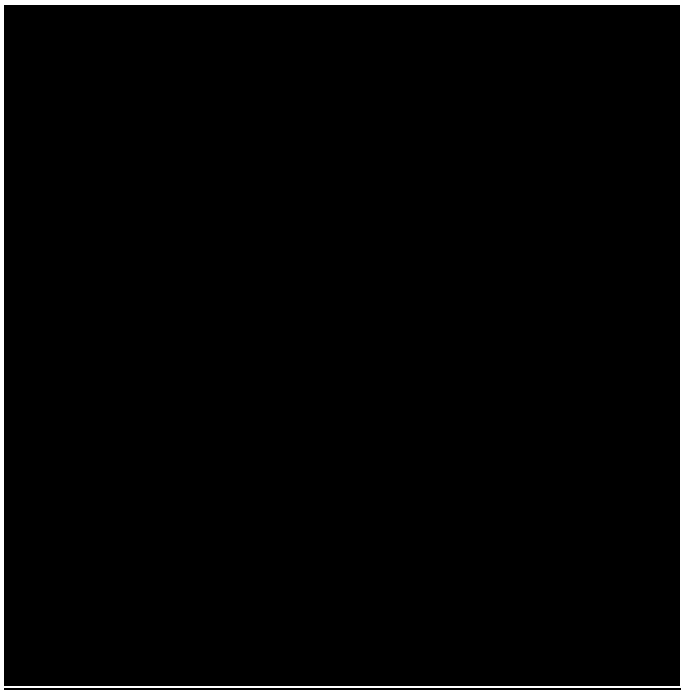
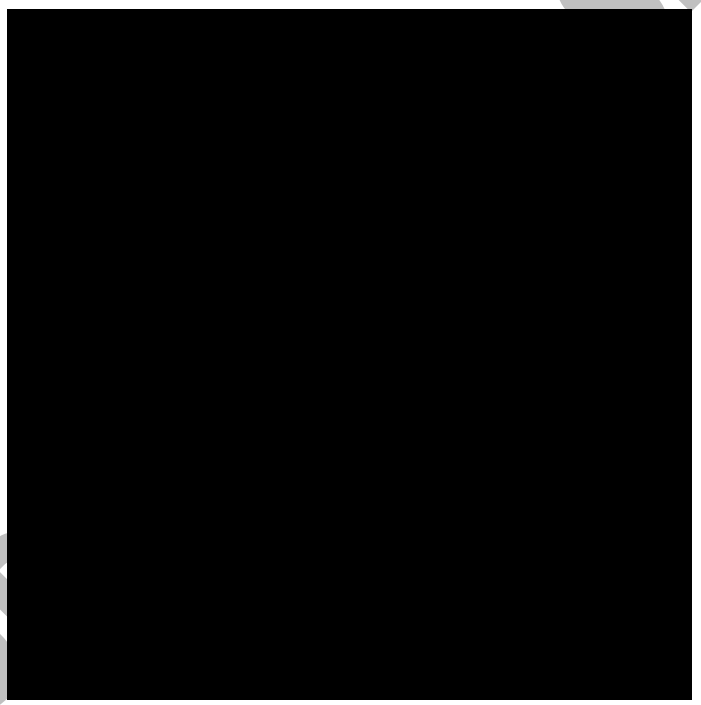


Figure 13: ACR results at week 24 for the TNFi-IR population on probit scale



National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Baricitinib for treating moderate to severe rheumatoid arthritis [ID979]

You are asked to check the ERG report from School of Health & Related Research Sheffield (ScHARR) to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Thursday 20 April 2017** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 Pg 5 Section 1.1

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report states that IV abatacept was excluded as a comparator	IV abatacept was included as a comparator but the base case cost-effectiveness analyses did not present results for abatacept IV. As noted elsewhere in the ERG report, analyses including abatacept IV were provided in response to the clarification request.	To correctly reflect which comparators were included in the company submission.	Text removed

Issue 2 Pg5 Section 1.2

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report states '*The most common adverse events for BARI were low-density lipoprotein cholesterol…'*	The sentence should be amended to state that there were **increases** in LDL cholesterol	To correctly state what changes were observed with respect to LDL cholesterol as this is currently not stated in the report.	Text amended

Issue 3 Pg6 Section 1.3

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In summarising the NMA results, the ERG report notes that baricitinib 4mg showed a significantly higher odds of an ACR 50 response than four other interventions. To be entirely accurate, this should refer to baricitinib in combination with	To state that these results relate to baricitinib 4mg in combination with cDMARDs	To correctly present the results referred to.	Text amended

cDMARDs			

Issue 4 Pg8 Section 1.5

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
As identified in Issue 1, the ERG report states that IV abatacept was not included in the list of comparators	As highlighted in Issue 1, cost-effectiveness results for abatacept IV were provided in response to the clarification request.	As per Issue 1, to correctly reflect the comparators that were included in the company submission	Text amended

Issue 5 Pg 10 Section 1.7

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The first paragraph of section 1.7 contains references to information that is commercial in confidence	The relevant sentences should be redacted.	To protect commercially confidential information.	Text redacted

Issue 6 Pg13 Assessment of response to therapy

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report discusses DAS28 score but neglects to highlight that there are two versions of DAS28 ESR and DAS28 CRP, only detailing the ESR version	The report should be amended to include DAS28 CRP being referenced in this section	To correctly reflect the fact that two versions of the DAS28 measure exist, that there are differences between them and the impact that this might have on comparative effectiveness estimates.	Text has been added referring to the alternative measurement of DAS28 using CRP. Further comment has been made on p42.

Issue 7 Pg19 Section 3.3

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report quotes from the company submission regarding the rationale for not modelling baricitinib monotherapy. The quoted section does not provide the broader context that this refers to the cDMARD-IR population.	The quote should be amended to make clear that the information provided is referencing the cDMARD-IR population given that monotherapy data in the methotrexate naïve population was available.	To provide the correct context for the quote and therefore the necessary clarity for the ERG report.	Text Amended

Issue 8 Pg28 Section 4.1.3

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
It is stated that the ERG believe that the data reported for moderate EULAR response includes that of patients with a good EULAR response with respect to the van de Putte study. Detail on this was provided in the additional clarification request but does not appear to be taken into account.	Further to the information provided in the additional clarification request, we would also like to flag the charts in Fig 3 of the van de putte publication which show moderate and good EULAR response separately and are labelled as such. This information in the charts is supportive of the data provided.	Correctly states the data from the van de Putte publication.	No change made. We believe that a legend in Figure 3 of Van de Putte is incorrect, and that the top right figure should say ‘at least moderate’ rather than moderate. The points in the figure tallies with the data in Table 2 for at least moderate, where for adalimumab 40mg weekly the ‘at least moderate’ value is 63.1% at week 26. If good responses were excluded from the ‘at least moderate’ value than the value for moderate (calculated from Table 2) would be lower than 50%.

Issue 9 Pg30 Section 4.2

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report states that four PIII studies were identified but then only lists 3 studies in parentheses.	Add RA-BEGIN to the list provided in parentheses.	To correctly list the PIII studies being referred to .	Text added

Issue 10 Pg31 Section 4.2

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report notes that their update search identified two other publications related to RA-BEACON and RA-BEAM. These publications were available after the August 2016 update of the company searches	To amend the statement to make clear that these publications were indexed after the August 2016 company search updated	To correctly state why these publications were not part of the company searches.	Text added

Issue 11 Pg36 Table 7

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The table on page 36 is labelled '7' as is the table on pg35. Other tables may also therefore be numbered incorrectly.	Tables should be checked and correctly numbered if necessary	Table are correctly labelled for the reader.	This has been amended to Table 6a and 7 in order to preserve table numbers and not have to include the rest of the report in an erratum.

Issue 12 Pg42 EULAR Response data

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
As highlighted in Issue 6, the reported results are those for EULAR response based on DAS28-hsCRP, not DAS28-ESR	The reported results should be annotated to state that they are DAS28-hsCRP EULAR response.	To make clear which DAS28 measure has been used in the reported results.	Text has been added

Issue 13 Pg47 Paragraph 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report states that 'All three RCTs reported a significant advantage in EQ-5D-5L of BARI 4MG over PBO (p≤0.001) at 12 weeks and 24 weeks follow-up (see Table 17, Table 18 and Table 19). BARI 2MG was also statistically significantly superior to PBO (p≤0.01) at 12 and 24 weeks.'	Should also be stated that a significant advantage over adalimumab was seen EQ-5D-5L responses in RA-BEAM, including at week 52.	To correctly represent the analyses presented in the company submission	Text added

Issue 14 Pg55 Section 4.3

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In the discussion of trials excluded / included in the in the indirect comparison it is stated that the ERG considered the CERTAIN study which by definition included low to	It should be made clear that the inclusion criteria for the CERTAIN study was for low to moderate disease activity. The baseline DAS28 score for patients in the baricitinib phase III studies in a true moderate to severe disease activity population was over	To correctly present the issues with the possible inclusion of CERTAIN in the indirect comparison.	Text amended, although we still believe the majority of patients would have had moderate to severe disease activity in the CERTAIN study.

moderate disease activity patients only to be reflective of a moderate to severe population, citing a mean baseline DAS28 of around 4.5 across the treatment arms.	6.5. From the perspective of the comparability of the population of the baricitinib studies to the population of the CERTAIN, it is questionable as to whether they are comparable and this is also the case for comparability with other studies included in the indirect comparison.		

Issue 15 Pg62 Section 4.4

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
As in Issue 3, the ERG report does not make clear that the baricitinib results stated with respect to ACR50 do not make clear that these are for baricitinib 4mg in combination with cDMARDs. This also applies to the discussion of results in the TNFi-IR population	As per Issue 3	As per Issue 3	Text amended

Issue 16 Pg63 Section 4.5

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report states that EULAR data from the van de Putte study was amended so that moderate EULAR responders did not include good EULAR	See Issue 8	As per Issue 8	No change made. See response to Issue 8

responders			

Issue 17 Pg 65. Figure 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Figure 4 appears to be a replica of figure 3- i.e. not the results for EULAR response in the TNFi-IR population	To correct Figure 4 with the correct data	To present the correct information in the figure.	Correct figure now inserted

Issue 18 Pg 67 Paragraph 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report states that 'At 12 weeks follow-up, all three RCTs reported a significant advantage for BARI4mg over PBO for EULAR response'	Should also be stated that a significant advantage over adalimumab was seen for EULAR responses in RA-BEAM.	To correctly represent the analyses presented in the company submission	Text added, and also on p42

Issue 19 Pg94 point 7

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG states that 'whilst it is likely that most patients would enjoy of an improvement of HAQ score before week 24, the company's approach is likely to lead to an overestimation of treatment benefits, as achievement of response will take	It should be stated that consistent with the BRAM model, it was assumed that start and end effects could be modelled as one-off deductions proportional to the change in QoL score, setting the multiplier to 0.2 years in the base case, sampled from a normal distribution with a standard deviation of 0.02 (separately for start and end), see Malottki	To correctly represent the analyses presented in the company submission	We have amended the text to state that this is only an advantage in relation to costs, where an adjustment to consider the time required to change HAQ was not employed.

at least a few weeks and up to 24 weeks for some patients.'	K, Barton P, Tsourapas A, Uthman AO, Liu Z, Routh K, et al. 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. [Review]. Health Technology Assessment (Winchester, England) 2011 Mar;15(14), p. 156. In this vein, the model does in fact take into consideration the 'gaining' and 'waning' effects of treatment benefit, which is believed to limit the risk of overestimation of treatment benefits. This also applies to Pg 8, section 1.5, Pg 76, paragraph 2, and Pg 103, section 7		A brief summary of the method used by the company has been provided on page 78.

Issue 20 Pg95 point 13

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG notes Hatswell et al put forward that the average cost of a dose is not necessarily equal to the average patient weight due to drug wastage and differences in cost per mg. In the case of infliximab, abatacept and tocilizumab cost per mg is fixed as there is only one vial size available or where different vial sizes are available, the cost per mg is the same.	The ERG report should be amended to make clear that the impact of using weight distribution as opposed to average weight is not a simple choice of one approach being preferred to another.	To correctly represent the issues on this point.	No change made. Not a factual error and the ERG maintains its position on the most appropriate method.

Assuming drug wastage, 3 vials of infliximab cover patients weighing between 67-100kg. Abatacept is dosed by weight band rather than per kg. These issues are not highlighted in the ERG report with respect to interventions dosed by weight			

Issue 21 Pg95 point 14

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
For the dosing of infliximab, the ERG calculates the number of doses as 52.14/8 = 6.52. However, when calculating dosing on a week by week basis, 7 doses will be required per year after Year 1. The difference occurs as the approach adopted by the ERG considers the fact that dosing does not start at week zero in each year, whereas the model adopted an 'on average' approach by using 7 doses per annum.	The ERG report should be amended to acknowledge the adopted approach as an alternative, reasonable approach rather than as inaccurate. This also applies to Pg 103, section 7	To correctly represent the issues on this point.	No change made. A dose every eight weeks is approximately 6.52 doses per year, not seven. The approach used by the company inflates the cost of infliximab across a long-time period. For instance, over 10 years the company would assume 70 doses of infliximab, whereas the ERG's approach would estimate 65.22 (365.25*10/7/8)

Issue 22 Pg95 Section 5.4

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG report alludes to commercial in confidence information	This information should be redacted	To protect commercial in confidence information.	Text redacted

ERRATUM

Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis: A Single Technology Appraisal

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK Iñigo Bermejo, Research Associate, ScHARR, University of Sheffield, Sheffield, UK Emma Simpson, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Shijie Ren, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Ruth Wong, 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
Correspondence Author	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK
Date completed	Date completed (07/04/2017)

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 16/56/23.

Declared competing interests of the authors

David Scott [Text Deleted] 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.

Also, his department has received a peer-reviewed grant from Pfizer within the last 12 months to undertake academic research on polypharmacy in arthritis. The department has also received free etanercept from Pfizer to use in an NIHR-funded programme grant in rheumatoid arthritis.

None of the other 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 Gill Rooney, 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:

Stevenson M, Bermejo I, Simpson E, Wong R, Scott DL, Young A. Baricitinib for Treating Moderate to Severe Rheumatoid Arthritis: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2016.

Contributions of authors

Emma Simpson summarised and critiqued the clinical effectiveness data reported within the company's submission. Inigo Bermejo and Matt Stevenson critiqued the health economic analysis submitted by the company. Shijie Ren critiqued the company's network meta-analysis and undertook a new analysis. Ruth Wong critiqued the company's search strategy. David Scott and Adam Young

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 CS generally adhered to the NICE scope. Exceptions related to the exclusion of the subcutaneous (SC) formulation of tocilizumab (TCZ) (TCZ SC) as a comparator. [Text Deleted].

1.2 Summary of clinical effectiveness evidence submitted by the company

The key clinical effectiveness evidence for baricitinib (BARI) was based on three randomised controlled trials (RCTs). Additionally one long-term extension study was included. There were two RCTs in methotrexate (MTX) -or conventional disease-modifying antirheumatic drug (cDMARD)-treated, biologic disease-modifying antirheumatic drug (bDMARD) naïve patients (RA-BEAM, RA-BUILD), both of which included a placebo (PBO) comparator. One RCT also included adalimumab (ADA) as a comparator (RA-BEAM). One PBO-controlled RCT was conducted in bDMARD-treated patients (RA-BEACON).

For the primary endpoint of ACR20 at 12 weeks follow-up, all three RCTs reported that BARI 4mg was statistically significantly superior to PBO ($p \leq 0.001$). At 12 weeks, more patients reached a 20% improvement in the ACR score (ACR20) in the BARI 4mg treated arm than the ADA treated arm ($p=0.01$). There was also an advantage over PBO for BARI 4mg at 24 weeks and for BARI 2mg at 12 weeks and 24 weeks follow-up. At 12 weeks follow-up, all three RCTs reported a significant advantage for BARI 4mg over PBO for EULAR response ($p<0.05$).

The most common adverse events for BARI were increases in low-density lipoprotein cholesterol, upper respiratory tract infections and nausea; other adverse drug reactions included herpes simplex, herpes zoster, acne, increased creatine phosphokinase, increased triglycerides, increased liver function tests (aspartate transaminase, alanine transaminase), neutropenia and thrombocytosis.

Network meta-analyses (NMA) were performed to assess the relative efficacy of BARI compared with the comparators in the inadequate response to cDMARDs (cDMARD-IR) or inadequate response to a tumour necrosis factor inhibitor (TNFi) (TNFi-IR) patients with moderate to severe rheumatoid arthritis (RA).

For the base case analysis at week 24 in the cDMARD-IR population, BARI 4mg was associated with a statistically significant higher odds of an ACR 50 response compared with cDMARD, ADA, PBO,

ETN and SSZ. No statistically significant differences were found versus any other comparators for the ACR50 outcome, with the exception of CTZ + cDMARD, in which odds of ACR50 response was found to be significantly in favour of the comparator. A similar pattern of results was observed for BARI 2mg.

For the base case analysis at week 24 in the TNFi-IR population, BARI 4mg + cDMARD demonstrated significantly higher ACR50 response rates than the cDMARD comparator. No statistically significant differences were found versus bDMARDs, with the exception of the comparison of BARI (both 4mg and 2mg) to TCZ, and the comparison of BARI 2mg to RTX, in which statistically significant treatment effects in favour of the comparator were observed.

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

The ERG found the searches for clinical effectiveness evidence reported in the CS to be adequate, and believed that all published RCTs of BARI 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 states that the results presented in NMA should be treated with caution, as a random effects model was assumed for the study-specific baseline treatment effects (pooling non-active and active controls): this was deemed to be inappropriate. In addition, studies that reported European League Against Rheumatism (EULAR) responses were synthesised along with converted EULAR response outcomes from studies that only reported ACR responses. This differs to the approach used in TA375, which performed the conversion having synthesised the ACR data, which ensures that the relative rankings of treatments are maintained.

1.4 Summary of cost effectiveness submitted evidence by the company

The manufacturer supplied a *de novo* discrete event simulation (DES) model constructed in Microsoft Excel®. The model simulates patients' disease progression through the sequences of treatments being compared. For each treatment, patients may achieve good, moderate or no EULAR response; this is assessed at 24 weeks. The EULAR response rates for each treatment are based on 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), adverse event or death. Patients who fail to achieve a moderate or good EULAR response discontinue treatment at 24 weeks and start the next treatment in the sequence. HAQ progression whilst on treatment is assumed to be flat on bDMARDs or BARI, whilst on cDMARDs and palliative care, HAQ progression is assumed to be non-linear based on latent HAQ trajectory classes.

Time to treatment discontinuation for responders is independent of treatment but is dependent on EULAR response category (moderate or good) and is modelled using Weibull curves fitted to British Society for Rheumatology Biologics Register (BSRBR) data. At treatment discontinuation, patients are assumed to suffer a rebound in HAQ equal to that achieved on treatment initiation and 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) over patients' remaining lifetimes. EuroQol 5 Dimensions (EQ-5D) values are calculated 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, and unit costs were taken from the British National Formulary and NHS Reference Costs 2014/15. Serious Adverse Events (SAEs) were excluded from the base case but were included in a scenario analysis.

The analyses presented in the CS relate to four different populations of rheumatoid arthritis patients: (1) patients who have had an inadequate response to cDMARDs (cDMARD-IR) with moderate RA; (2) cDMARD-IR patients with severe RA; (3) patients with severe RA who have had an inadequate response to a tumour necrosis factor inhibitor (TNFi) (TNFi-IR) and who are rituximab (RTX) eligible; and (4) patients who are TNFi-IR with severe RA for whom RTX is contraindicated or not tolerated. The definition of severe RA was a DAS28 > 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 BARI trials.

In the cDMARD-IR population with moderate RA, the deterministic incremental cost-effectiveness ratio (ICER) for BARI + MTX compared with intensive cDMARDs was estimated to be £37,420 per QALY gained. In the cDMARD-IR population with severe RA, BARI + MTX dominated all comparators except for certolizumab pegol (CTZ) + MTX, with the ICER of CTZ + MTZ compared with BARI + MTX estimated to be £18,400 per QALY gained. In the TNFi-IR population with severe RA, when RTX + MTX was an option, BARI + MTX was dominated by RTX + MTX. In the TNFi-IR population with severe RA for whom RTX is contraindicated or not tolerated, BARI + MTX dominated golimumab + MTX and was less effective and less expensive than the remaining comparators. The ICERs for etanercept biosimilars (ETN-b) + MTX, CTZ + MTX and ADA + MTX compared with BARI + MTX were lower than £30,000 per QALY gained. However, the company made a favourable assumption for these interventions (same efficacy as in the severe cDMARD-IR population) in the absence of effectiveness data in this population and therefore, caution is advised when interpreting these results. The ICERs for TCZ IV + MTX and abatacept (ABA) SC + MTX compared with BARI + MTX were estimated to be higher than £30,000 per QALY gained, but the confidential Patient Access Schemes (PAS) relating to TCZ and ABA were not included.

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) with some minor deviations. The ERG believed that the conceptual model was appropriate but suffered from a series of implementation errors and limitations, as described here.

The company rounded modified HAQ values to the nearest valid HAQ score rather than allowing the valid HAQ score to be sampled based on the continuous HAQ value. The ERG notes that this approach might lead to inaccurate estimations of HAQ scores, as values might be rounded up more often than rounded down or *vice versa*.

The company intended to implement the trajectory of HAQ score whilst on cDMARDs or palliative care based on the latent class approach used by the AG in TA375. However, the company assigned each patient to a single class based on the probability of class membership instead of using an average weighted by the probability of class membership.

The company assumed that patients who achieve a moderate or good EULAR response at 24 weeks experience a reduction in HAQ score instantaneously at treatment initiation. The ERG believes that the company's approach is likely to lead to an overestimation of treatment benefits in relation to savings in RA-related costs, as the achievement of response will take at least a few weeks and potentially up to 24 weeks for some patients.

In order to calculate the QALYs and costs produced in the time span between two events, the model uses an area under the curve (AUC) approach for the HAQ score, and then maps this value to the EQ-5D and hospitalisation costs. However, since the relationships between HAQ score and EQ-5D and between HAQ score and hospitalisation costs are not linear, this approach may lead to inaccurate results.

The [Text Deleted] TCZ SC formulation were not included in the list of comparators, despite ABA SC and TCZ IV being included. The company argued that it had excluded TCZ SC because: (i) the available evidence for TCZ SC was limited; (ii) it provided a lower efficacy estimate than for TCZ IV; and (iii) the cost difference between the two formulations was relatively small. The ERG notes that the difference in costs might be considerable taking into account the administration costs and the confidential PAS. ABA IV was included in the NMA, but was excluded from the analyses. In

No robust evidence was presented to assess the treatment duration of BARI. In their base case, the company assumed that time to treatment discontinuation for BARI was the same as that for bDMARDs. However, BARI is not a bDMARD and it is not clear that time to loss of efficacy would be similar for BARI and bDMARDs meaning that the results are subject to uncertainty.

The company assumed that the flat HAQ progression whilst on treatment assumed for bDMARDs in TA375 also applies to BARI. The scenario analysis in which a linear HAQ increase is assumed instead showed BARI was less effective than its comparators. Whilst uncertain, clinical advice provided to the ERG suggested that it was reasonable to assume the same HAQ progression for BARI as for bDMARDs.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG undertook few exploratory analyses based on the company's submitted model as the ERG believes that [redacted]. The errors affecting the PSA were corrected and the amended results presented. The ERG comments that the added value of any mathematical model for people with severe RA in this Single Technology Appraisal (STA) is debatable given the efficacy and acquisition cost inputs of the bDMARDs.

The ERG highlight the fact that the company drew heavily on TA375 in constructing their mathematical model and that some parts, including latent classes for those on cDMARDs and the HAQ progressions for those on cDMARDs were not implemented correctly. This will affect the ICER for patients with moderate RA, where BARI is compared to cDMARDs. The time required to fix these issues were beyond that available for an STA. The ERG notes that the median ICER of bDMARDs compared with cDMARDs in TA375 was in the region of £50,000 per QALY gained for patients with moderate RA, which is considerably higher than the estimate provided by the company. The ERG believes that the ICER of BARI when used in the moderate, RA population will be closer to that reported in TA375 due to the errors in reconstructing the Assessment Group's model.

Two classifications have dominated the measurement of improvement in RA symptoms: ACR responses[17] and EULAR responses.[18]

The initial ACR response was denoted as an ACR20 which 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.

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.[19] 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). This assesses 28 joints in terms of swelling (SW28) and of tenderness to the touch (TEN28) and also 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:[20]

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

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.[18] 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.[21] EULAR response has been reported less frequently in RCTs than ACR responses,[22] although EULAR is much more closely aligned to the treatment continuation rules

ETN; IFX; GOL; and TCZ (each in combination with MTX) and: ADA; ETN; and CTZ (each as monotherapy).

- For cDMARD-IR patients with moderate RA the comparators were defined as: intensive cDMARDs; cDMARD monotherapy and PALL.

The company did not consider the intravenous (IV) formulation of ABA: in the clarification response (question B15), the company stated that this was '*a pragmatic decision (…) To attempt to limit the number of sequences included in the submission where it was possible that inclusion of the same intervention with different administration routes was unlikely to be informative.'* The company also did not consider the subcutaneous (SC) formulation of TCZ with three reasons provided in the clarification response (question B9) culminating in the company stating that including 'IV tocilizumab only was felt to be a reasonable choice, with it likely to be representative of the costs and outcomes associated with the S/C version.'

Broadly the ERG believes that the company has evaluated the correct comparators although make two comments.

The company have not explicitly modelled BARI used as a monotherapy in the cDMARD-IR population. The rationale stated by the company for this is 'the paucity of efficacy data in the baricitinib clinical trial programme for patients receiving baricitinib monotherapy, which would be insufficient to form a reliable estimate of efficacy in the modelled populations for baricitinib monotherapy. It should be noted that in the recent MTA regarding the use of biologics in DMARD-naïve and cDMARD-IR patients (TA375), the Appraisal Committee agreed that the minority of (cDMARD-IR) patients with severely active rheumatoid arthritis who could not tolerate methotrexate should not be treated differently from other people with severe disease, as far as possible. The Committee concluded that biologic DMARDs should be recommended as a cost-effective use of NHS resources when used as monotherapy for severely active disease previously treated with DMARDs, where the marketing authorisation of the bDMARD allows for this recommendation to be made. The economic evaluation of baricitinib presented here assumes that a similar rationale will be applied to baricitinib monotherapy.' The lack of data for BARI when used as a monotherapy will increase the uncertainty in its incremental cost effectiveness ratio (ICER) when compared with interventions with a larger evidence base. Clinical advice to the ERG suggests that there is no clearly defined relationship between the efficacy of a bDMARD in combination with MTX and in the bDMARD used as monotherapy. However, data from RA-BEGIN[31] showed that the addition of MTX to BARI 4mg did not produce a marked improvement over BARI

Quality assessments of the trials contained in the NMA are presented in Appendix 15 of the CS. The same quality assessment items were used, as is appropriate for RCTs.

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

Five BARI studies were identified by the CS search. These comprised four phase III RCTs (RA-BEAM, RA-BUILD, RA-BEACON, RA-BEGIN), and one long-term extension study (RA-BEYOND). One of the RCTs (RA-BEGIN) had a population of cDMARD-naïve RA patients, and as this is an unlicensed treatment position, this RCT was not of relevance to the decision problem. Details of RA-BEGIN are presented in Appendix 1 of the CS. CS Appendix 19 additionally includes details of three phase II studies of BARI in RA patients.

Details of the three RCTs (RA-BEAM, RA-BUILD, RA-BEACON) and long-term extension study (RA-BEYOND) included in the CS are shown in **Error! Reference source not found.** of the ERG (adapted from CS Table 12 and Section 4.3.6.1 of the CS). Both RA-BUILD and RA-BEACON had a 24-week randomised period, whereas RA-BEAM was randomised for 52-weeks, however, all patients randomised to PBO switched to BAR+MTX at week 24. Rescue therapy was permitted in all three RCTs, with patients whose tender joint count and swollen joint count were reduced by less than 20% from baseline at both week 14 and week 16 received open-label rescue treatment (BARI 4 mg) at week 16.[35,36,37] After week 16, rescue therapy was available based on investigator discretion. RA-BEAM and RA-BUILD had populations of cDMARD-experienced, bDMARD-naïve patients, whereas the population for RA-BEACON was bDMARD-experienced.

Table 6a: Baseline characteristics of participants of RA-BUILD (reproduced from Table 20 of the CS)[36]

RA-BUILD		PBO QD + MTX (n=228)	BARI 2 mg QD + MTX (n=229)	BARI 4 mg QD + MTX (n=227)
Gender, n (%)	Male	39 (17.1)	45 (19.7)	40 (17.6)
	Female	189 (82.9)	184 (80.3)	187 (82.4)
Age (years)	Mean	51.4	52.2	51.8
	SD	12.5	12.3	12.1
	Median	53.0	52.0	53.0
	Range	21–79	22–82	20–80
Duration of RA symptoms	Mean (SD)	7 (8)	8 (8)	8 (8)
Time from diagnosis of rheumatoid arthritis (years)	Mean	5.9	6.5	6.4
	SD	6.8	7.6	7.5
	Median	3.4	3.6	3.7
	Range	0.07–37.44	0.28–52.76	0.11–41.40
DAS-28(CRP)	Mean	5.53	5.57	5.55
	SD	0.91	0.96	0.87
	Median	5.50	5.49	5.53
	Range	2.27–7.50	3.05–8.03	3.30–7.91
DAS-28(ESR)	Mean	6.19	6.28	6.20
	SD	1.00	0.99	0.91
	Median	6.18	6.25	6.26
	Range	2.90–8.63	3.31–8.52	3.96–8.44
HAQ-DI	Mean	1.5	1.51	1.55
	SD	0.60	0.62	1.60
	Median	1.50	1.50	1.50
	Range	0.0–2.8	0.0–2.9	0.0–3.0
ACPA-positive, n (%)		172 (75.4)	169 (73.8)	163 (71.8)
mTSS	Mean	18.54	25.78	23.71
	SD	31.47	40.26	40.01
	Median	6.00	8.50	6.25
	Range	0.0–241.5	0.0–218.0	0.0–231.0
RF-positive, n (%)		171 (75.0)	177 (77.3)	173 (76.2)
Mean weekly does MTX		16.2 mg across all groups (n=684) (CS clarification response A3)		

Abbreviations: QD = once daily, DAS28 = Disease Activity Score, CRP = C-reactive protein, ESR = erythrocyte sedimentation rate, HAQ-DI = Health Assessment Questionnaire–Disability Index, ACPA = Anti-citrullinated protein antibody, mTSS = modified Total Sharp Score, RF = Rheumatoid factor.

EULAR response data

At 12 weeks follow-up, all three RCTs reported a significant advantage for BARI 4MG over PBO for EULAR response ($p \leq 0.001$) (good or moderate EULAR response RA-BEAM [redacted]; RA-BUILD 79.0% vs 53.5%; RA-BEACON 66.1% vs 42.6%). A significant advantage was also seen in RA-BEAM in the comparison with ADA in relation to good and moderate response (p=0.002), and a good response (p=0.010). There was an advantage for BARI 4MG over PBO at 24 weeks follow-up for the bDMARD-naïve studies ($p<0.001$), and also for the bDMARD-experienced population ($p<0.05$). BARI 2MG was superior to PBO for EULAR response at 12 weeks ($p<0.001$) for both the bDMARD-naïve and bDMARD-experienced populations. At 24 weeks, BARI 2MG was significantly superior to PBO in the bDMARD-naïve ($p<0.001$) and bDMARD-experienced ($p<0.05$) populations. The results presented refer to EULAR responses calculated using DAS28-hsCRP which are more favourable than the DAS28-ESR values determined via a post-hoc analysis and used in the cost-effectiveness modelling.

Table 1: EULAR response (based on DAS28-hsCRP) results RA-BEAM (non-responder imputation) (adapted from CS Table 23 and CS Table 26)

Outcome measure	12 weeks			24 weeks			52 weeks		
	PBO (N=488)	BARI 4MG (N=487	ADA (N=330)	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	PBO (n=452)	BARI 4MG (N=487)	ADA (N=330)
EULAR (good + moderate) response rate (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EULAR (good) response rate (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Comparison		BARI 4MG vs PBO	BARI 4MG vs ADA						
EULAR good and moderate response Odds ratio (95% CI)		[redacted]	[redacted]						
EULAR good response Odds ratio (95% CI)		[redacted]	[redacted]						

Footnotes: Significance level definitions: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ vs placebo.

Health-related quality of life (HRQoL)

All three RCTs reported a significant advantage in EQ-5D-5L of BARI 4MG over PBO ($p \leq 0.001$) at 12 weeks and 24 weeks follow-up (see Table 2, Table 3 and Table 4). A statistically significant benefit of BARI 4mg compared with ADA, both in addition to cDMARD in EQ-5D-5L was observed at week 52. (see Table 2) BARI 2mg was also statistically significantly superior to PBO ($p \leq 0.01$) at 12 and 24 weeks.

Table 2: RA-BEAM (adapted from CS Table 23)

Outcome measure	12 weeks			24 weeks			52 weeks		
	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	PBO (N=488)	BARI 4MG (N=487)	ADA (N=330)	PBO (n=452)	BARI 4MG (N=487	ADA (N=330)
EQ-5D-5L CFB LSM (SE)	0.102 (0.009)	0.184*** (0.009)	0.167*** (0.011)	0.088 (0.010)	0.199*** (0.010)	0.175*** (0.012)	N/A	0.217+ (0.010)	0.182 (0.012)

EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels
**$p \leq 0.01$, ***$p \leq 0.001$ vs placebo
+ $p \leq 0.05$; vs adalimumab.

Table 3: RA-BUILD (adapted from CS Table 35)

Outcome measures	12 weeks			24 weeks		
	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)	PBO (N=228)	BARI 2MG (N=229)	BARI 4MG (N=227)
EQ-5D-5L CFB LSM (SE)	0.092 (0.014)	0.165*** (0.013)	0.162***(0.014)	0.091 (0.014)	0.157*** (0.014)	0.186*** (0.014)

EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels
***$p \leq 0.001$ vs placebo

Table 4: RA-BEACON (adapted from CS Table 47)

Outcome measures	12 weeks			24 weeks		
	PBO N=176	BARI 2MG N=174	BARI 4MG N=177	PBO N=176	BARI 2MG N=174	BARI 4MG N=177
EQ-5D-5L CFB LSM (SE)	0.036 (0.019)	0.114*** (0.019)	0.169*** (0.018)	0.038 (0.019)	0.111** (0.019)	0.159*** (0.019)

EQ-5D-5L = European Quality of Life-5 Dimensions-5 levels
**$p \leq 0.01$, ***$p \leq 0.001$ vs placebo

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

Trials included in the NMA are listed in the CS Tables 69 and 70 (reproduced below in **Error! Reference source not found.** and **Error! Reference source not found.** of the ERG report). Trial characteristics of these studies are included in CS Tables 71 and 72 (not reproduced here) and were considered appropriate by the ERG to permit inclusion in the NMA. Quality assessment for these trials is reported in Appendix 15 of the CS.

Trials in the 24-week analysis of the bDMARD-naïve population were largely the same as those in the NMA undertaken by the independent Assessment Group (AG) in TA375. However, there were some exceptions which have been grouped into the following categories: trials in the CS that were not included in TA375; and trials included in TA375 but excluded from the CS. A similar comparison could not be made for the bDMARD-experienced population as this was not the focus of TA375.

Seven trials in the CS that were not included in TA375 for bDMARD-naïve patients:
These were trials published after the cut-off date used within TA375 – Li (2013),[49] BREVACTA,[50] SUMMACTA,[51] RA-SCORE,[52] SERENE,[53] BARI trials – RA-BEAM,[35] and RA-BUILD.[36]

Ten trials included in TA375 but excluded from the CS:
AUGUSTII,[54] IIBCREATE,[55] NCT00254293,[56] and Kremer 2012.[57] These studies were excluded from the CS as they were Phase II trials (see CS Table 11).

ACQUIRE.[58] The company excluded this study because the "*study compared S/C vs IV abatacept. The search strategy specified that studies were to include two different comparators of interest to be included*" (see clarification question A4). This appears to be inconsistent with the inclusion of SUMMACTA, which compares IV TCZ and SC.

ATTRACT.[59] The company excluded this trial as it only provided data relating to ACR20. These data can be used within the NMA and should not be discarded.

CERTAIN.[60] Within the clarification response process (clarification question A4), the company stated that this trial was excluded as it included patients with low to moderate disease activity. As the baseline DAS28 in the treatment arms of this trial were 4.47 and 4.53 the ERG believes that the majority of patients would have had moderate, to severe, disease activity.

SAMURAI.[61] The company stated that 12 or 24-week data were not identified. However, data at 24 weeks from Nishimoto et al 2007[62] were used in TA375.

single study. In response to a request for clarification (question A19), the company stated that the I^2 value was calculated from a meta-analysis of studies which contained the same treatment arms for example the TEMPO,[68] SATORI[69] and ARMADA[70] studies. However, it was not clear if it meant both treatment arms had to be similar between studies, and the ERG remains unclear on the process of study selection. Furthermore, it is debatable how meaningful the I^2 results are since the model in NMA was probit and the model in meta-analysis used to calculate I^2 is logit. Similar modelling issues also exist in the company's checking of inconsistency. The NMA code for the node-splitting approach given in response to a request for clarification (question A17) used a logit link function instead of a probit link function.

ACR results for both the cDMARD-IR and TNFi-IR populations were presented as the odds ratios of achieving an ACR50 response and the absolute probabilities of achieving each ACR category. EULAR results were presented using the absolute probability of being in each EULAR category for both the cDMARD-IR and TNFi-IR populations.

The CS concluded that for the base case analysis at week 24 in the cDMARD-IR population, BARI 4mg + cDMARD was associated with a statistically significant higher odds of an ACR50 response compared with cDMARD, ADA, PBO, ETN and SSZ. No statistically significant differences were found versus any other comparators for the ACR50 outcome, with the exception of CTZ + cDMARD, in which odds of ACR50 response was found to be significantly in favour of the comparator. A similar pattern of results was observed for BARI 2mg.

For the base case analysis at week 24 in the TNFi-IR population, BARI 4mg + cDMARD demonstrated significantly higher ACR50 response rates than the cDMARD comparator. No statistically significant differences were seen versus bDMARDs + cDMARD, with the exception of the comparison of BARI (both 4mg and 2mg both + cDMARD) to TCZ + cDMARD, and the comparison of BARI 2mg + cDMARD to RTX + cDMARD, in which statistically significant treatment effects in favour of the comparator were observed.

The ERG notes that as control arms were inappropriately pooled, all results should be interpreted with caution.

The relative treatment effects on the probit scale were presented using the posterior distribution for both populations. However, it was difficult to interpret fully the results due to the high level of overlap between distributions (see Figures 39 and 40 within the CS). The ERG requested that the company presented the treatment effects on the probit scale in forest plots (clarification question A9). However,

the treatment effects provided in the company's response were still on the odds ratio scale rather than the probit scale

Figure 1: EULAR response in the ERG's NMA in the TNFi-IR population

Figure 2: ACR50 response in the ERG's NMA in the cDMARD-IR population

For the primary endpoint of ACR20 at 12 weeks follow-up, all three RCTs reported than BARI 4mg was statistically significantly superior to PBO ($p \leq 0.001$). At 12 weeks, more patients reached ACR20 in the BARI 4mg treated arm than the ADA treated arm. There was also an advantage over PBO for BARI 4mg at 24 weeks and for BARI 2mg at 12 weeks and 24 weeks follow-up. At 12 weeks follow-up, all three RCTs reported a significant advantage for BARI 4mg over PBO for EULAR response. A significant advantage was also seen in RA-BEAM in the comparison of BARI 4mg with ADA in relation to good and moderate response ($p=0.002$), and a good response ($p=0.010$).

The most common AEs for BARI were LDL cholesterol, upper respiratory tract infections and nausea. Other adverse drug reactions included: herpes simplex; herpes zoster; acne; increased creatine phosphokinase; increased triglycerides; increased liver function tests (AST, ALT); neutropenia; and thrombocytosis. There was a significantly higher rate in infections for BARI 4mg than PBO ($p<0.001$), but there was no significant difference between groups in the rate of serious infections.

The ERG considers that all of the NMA results presented in CS should be treated with caution because: (1) a random effects model was assumed for the baselines inappropriately pooling non-active and active controls; (2) simultaneous baseline and treatment effect models with an inappropriate assumption for the baselines were conducted; and (3) for EULAR outcomes, studies reported EULAR were synthesised with converted EULAR for studies that only reported ACR data. This differs to the approach used in TA375.[23]

For the severe cDMARD-IR population, the company used the estimates in **Error! Reference source not found.** for all the treatments in the sequence, regardless of their position in the sequence. This implies that, for example, TCZ IV + MTX will have the same efficacy irrespective of whether it is first or third in the sequence. However, **Error! Reference source not found.** and **Error! Reference source not found.** show that the efficacy of TCZ IV + MTX, as well as all the other treatments, is lower after treatment with a TNFi. The ERG notes that effectiveness estimates from **Error! Reference source not found.** should have been used only for the first treatment in the sequence. For the rest of the treatments in the sequence, the estimates from **Error! Reference source not found.** should have been used instead.

HAQ improvement upon treatment response

The change in HAQ score was assumed to be conditional on the EULAR response achieved. The reduction in HAQ score was taken from the values reported by the AG in TA375[23]: these were reductions of 0.673 (standard error (SE) 0.012) for patients who experienced a good response, and 0.317 (SE 0.048) for patients who experienced a moderate response. The company assumed that HAQ improvement upon response occurs instantaneously at treatment initiation. In response to a clarification question by the ERG (question B20),[32] the company argued that there is evidence that clinical response to bDMARDs in RA is often rapid, with patients potentially experiencing improvements in symptoms within a few weeks of treatment initiation, perhaps even as early as 48 hours. In addition, it referred to data from RA-BEAM,[38] where change in mean HAQ score at week 12 was similar to that at week 24 for BARI + MTX ([redacted] and [redacted], respectively). The ERG notes that even if that were the case, accounting for the HAQ improvement from time zero is likely to be an overestimation of the treatment benefits in relation to savings in RA-related costs.

The company assumed that the HAQ improvement upon response would be lost at treatment discontinuation, and as a consequence the patient would suffer from a rebound in HAQ score equal to the improvement upon treatment response.

HAQ progression

Whilst patients are on treatment with a bDMARD or BARI the company assumed that the HAQ trajectory is flat, that is, that there is no change in the value. This assumption for bDMARDs was also incorporated in the AG model for TA375.[23] The company have assumed that this was also the case for BARI, although test a worsening of HAQ across time in a scenario analysis. Clinical advice to the ERG suggested that the assumption that the HAQ trajectory for BARI is equal to bDMARDs was reasonable.

Treatment duration

Patients who fail to achieve good or moderate EULAR response at 24 weeks discontinue treatment and start the next treatment in the sequence. In contrast, patients who achieve good or moderate EULAR response stay on treatment until loss of efficacy. Time on treatment for these patients is

that, whilst it is likely that most patients would enjoy of an improvement of HAQ score before week 24, the company's approach is likely to lead to an overestimation of treatment benefits in relation to savings in RA-related costs, as achievement of response will take at least a few weeks and up to 24 weeks for some patients. Similarly, the assumption made by the AG of TA375 of no HAQ improvement in the first 6 months is likely to underestimate treatment benefit in relation to in relation to savings in RA-related costs and QALY gains.

8) Averaging HAQ across large time periods

In order to calculate the QALYs and costs produced in the time span between two events, the model uses and area under the curve approach for the HAQ score, and then maps this value to EQ-5D and hospitalisation costs. However, since the relationships between HAQ score and EQ-5D and between HAQ score and hospitalisation costs are not linear, this approach will lead to inaccurate results.

9) Exclusion of ABA IV and TCZ SC from the list of comparators

The ABA IV and TCZ SC formulations were not included as comparators, although ABA SC and TCZ IV were included. In response to a clarification request by the ERG (question B9),[32] the company stated that it had excluded TCZ SC because the available evidence for TCZ SC was limited, it provided a lower efficacy estimate than for TCZ IV and the cost difference between the two formulations was relatively small. The ERG notes that the difference in costs might be considerable taking into account the administration costs and the CIC PAS. ABA IV was included in the NMA, but excluded from the economic analyses. In response to a clarification request by the ERG (question B15),[32] the company explained that ABA IV was excluded from the analyses to limit the number of sequences considered and because it was unlikely to be informative. The company also presented the results of ABA IV for the severe cDMARD-IR population in their clarification response, which led to similar results compared with ABA SC (£[redacted] versus £[redacted] and [redacted] versus [redacted] QALYs respectively).

10) Using an older mapping from HAQ to EQ-5D than the AG

The company used an algorithm proposed by Hernández Alava *et al.*[78] to map HAQ scores to EQ-5D. The company explained in the CS[1] that the three-class model by Hernández Alava *et al.*[78] was used in the model because the predicted values for selected combinations of covariates were reported in the published article, which allowed the validation of their replicated mapping algorithm. The ERG notes that more recent algorithms have been since published, such as that reported by Hernández Alava *et al.*[83] that purport to have a higher accuracy: this algorithm was used in TA375, which includes all the parameters necessary to implement it. However, the ERG acknowledges that the company's scenario analyses show that the mapping algorithm does not have an important impact on the result of the analyses.

11) Assuming BARI would be inserted before intensive cDMARDs for patients with moderate RA

The ERG notes that in TA375[23] it was assumed that intensive cDMARDs had already been used, as was supported by clinical opinion, instead of assuming that a bDMARD would go in front of intensive cDMARDs. The impact of the change, which was not supported by the clinical advisors to the ERG, is unknown.

12) Different life years gained across sequences

The company assumed,[1] as was the approach used by the AG in TA375,[23] that HAQ scores at baseline affected mortality. However, the company's model re-estimated the age of death at every event, which led to slightly different expected life years gained, as observed in the results. The ERG notes that sequences of different lengths are likely to produce different life expectancies, as age of death is recalculated at different time points. The company also attributed the differences in life years gained to random variation. The ERG notes that using the same random numbers across treatments as the company did in other cases would have eliminated the impact of random variation. Alternatively, more patients should have been run through the model to minimise the impact of random variation.

13) Lack of consideration of the distribution of weight for interventions where the dosage is weight based

The company used the average weight of the population in the relevant BARI trials (see Table 112 of the CS[1]) to calculate the average dose cost. This is equivalent to assuming that all patients had the average weight. The ERG notes that this approach to calculating the average dose cost is not appropriate given that the relationship between weight and dosing cost is not linear as explained by Hatswell *et al*.[85] This is because the average cost of a dose is not necessarily equal to the cost of the patient with the average weight, due to drug wastage and differences in cost per mg of some drugs. The ERG notes that the company should have calculated the average cost of a dose using the distribution of the weight of patients instead of using the average weight.

14) Dosage of IFX

The company assumed patients would receive seven doses of IFX per year and an additional dose in the first year. However, the SmPC[86] and the BNF[80] state that patients should receive IFX doses at 2 and 6 weeks after the first infusion and every 8 weeks thereafter. This implies that the average doses of IFX per year should be 6.52 instead of 7.

7. OVERALL CONCLUSIONS

BARI + MTX treatment was estimated by the company to have an ICER of £37,420 per QALY gained compared with intensive cDMARDs in the moderate RA cDMARD-IR population. In the severe RA cDMARD-IR population, BARI + MTX dominated all its comparators except for CTZ + MTX: the ICER of CTZ + MTZ compared with BARI + MTX was estimated to be £18,400 per QALY gained. In the severe RA TNFi-IR population, when RTX + MTX was an option, BARI + MTX was dominated by RTX + MTX. In patients with severe RA who have had inadequate response to a TNFi and for whom RTX is contraindicated or not tolerated, BARI + MTX dominated GOL + MTX and was less effective and less expensive than the rest of its comparators. The ICERs for ETN-b + MTX, CTZ + MTX and ADA + MTX compared with BARI + MTX were estimated to be lower than £30,000 per QALY gained whilst the ICERs TCZ IV + MTX and ABA SC + MTX compared with BARI + MTX were estimated to be higher than £30,000 per QALY gained. However, the confidential PASs for ABA and TCZ were not included in these analyses.

The ERG's critical appraisal identified a number of issues relating to the company's model and analysis. The ERG believes that: the NMA is subject to potential limitations; some of the scenario analyses as well as the PSA for the severe TNFi-IR, RTX-ineligible population lack face validity; the efficacy estimates for the cDMARD-IR population should only be used for the first line of treatment; rounding to nearest HAQ score might introduce bias; the HAQ trajectory of a patient on cDMARDs or palliative care should be calculated as a weighted average; assuming HAQ improvement upon treatment initiation overestimates treatment benefit in relation to savings in RA-related costs; averaging HAQ across large time periods leads to inaccuracies in the calculation of costs and QALYs; [Text Deleted] TCZ SC should have been included in the list of comparators; newer mapping algorithms from HAQ scores to EQ-5D should have been used; BARI should have not been assumed to be provided before intensive cDMARDs for moderate patients; mortality rates differ between sequences; the distribution of weight for interventions where the dosage is weight based should have been considered; and the dosage of IFX is inaccurate.

There remain several potentially important areas of uncertainty:

1. Cost-effectiveness of BARI monotherapy
2. Cost-effectiveness of BARI + MTX in the severe, TNFi-IR, RTX-ineligible population
3. Efficacy of bDMARDs after BARI
4. Time to treatment discontinuation of BARI
5. Relative HAQ progression whilst on BARI compared with bDMARDs