


# Certolizumab Pegol for Treating Rheumatoid Arthritis Following Inadequate Response to a TNF- $\alpha$ Inhibitor: An Evidence Review Group Perspective of a NICE Single Technology Appraisal

Iñigo Bermejo<sup>1</sup>  · Matt Stevenson<sup>1</sup> · Rachel Archer<sup>1</sup> · John W. Stevens<sup>1</sup> · Edward Goka<sup>1</sup> · Mark Clowes<sup>1</sup> · David L. Scott<sup>2</sup> · Adam Young<sup>3</sup>

Published online: 26 May 2017  
© Springer International Publishing Switzerland 2017

**Abstract** As part of its single technology appraisal (STA) process, the National Institute for Health and Care Excellence (NICE) invited the manufacturer (UCB Pharma) of certolizumab pegol (CZP; Cimzia<sup>®</sup>) to submit evidence of its clinical and cost effectiveness for the treatment of rheumatoid arthritis (RA) following inadequate response to a tumour necrosis factor- $\alpha$  inhibitor (TNFi). The School of Health and Related Research Technology Appraisal Group at the University of Sheffield was commissioned to act as the independent Evidence Review Group (ERG). The ERG produced a detailed review of the evidence for the clinical and cost effectiveness of the technology, based upon the company's submission to NICE. The clinical effectiveness evidence in the company's submission for CZP was based predominantly on six randomised controlled trials (RCTs) comparing the efficacy of CZP against placebo. The clinical effectiveness review identified no head-to-head evidence on the efficacy of CZP against the comparators within the scope; therefore, the company performed a network meta-analysis (NMA). The company's NMA concluded that CZP had a similar efficacy to that of its comparators. The company submitted a Markov model that assessed the incremental cost effectiveness of CZP versus comparator biologic disease-modifying antirheumatic drugs (bDMARDs) for the

treatment of RA from the perspective of the National Health Service for three decision problems, each of which followed an inadequate response to a TNFi. These were (1) a comparison against rituximab (RTX) in combination with methotrexate (MTX); (2) a comparison against bDMARDs when RTX was contraindicated or withdrawn due to an adverse event; and (3) a comparison against bDMARDs when MTX was contraindicated or withdrawn due to an adverse event. Results from the company's economic evaluation showed that CZP resulted in a similar number of quality-adjusted life years (QALYs) produced at similar or lower costs compared with comparator bDMARDs. The commercial-in-confidence patient access schemes for abatacept and tocilizumab could not be incorporated by the company, but were incorporated by the ERG in a confidential appendix for the NICE Appraisal Committee (AC). The company estimated that the addition of CZP before RTX in a sequence for patients who could receive MTX produced more QALYs at an increased cost, with a cost per QALY of £33,222. Following a critique of the model, the ERG undertook exploratory analyses that did not change the conclusions reached based on the company's economic evaluation in relation to the comparison with bDMARDs. The ERG estimated that where CZP replaced RTX, CZP was dominated, as it produced fewer QALYs at an increased cost. The AC concluded that there was little difference in effectiveness between CZP and comparator bDMARDs and that equivalence among bDMARDs could be accepted. The AC consequently recommended CZP plus MTX for people for whom RTX is contraindicated or not tolerated and CZP monotherapy for people for whom MTX is contraindicated or not tolerated. The AC concluded that CZP plus MTX could not be considered a cost-effective use of National Health Service resources when RTX plus MTX is a treatment option.

✉ Iñigo Bermejo  
i.bermejo@sheffield.ac.uk

<sup>1</sup> School of Health and Related Research (ScHARR),  
University of Sheffield, Regent Court, 30 Regent Street,  
Sheffield S1 4DA, UK  
<sup>2</sup> Department of Rheumatology, King's College Hospital NHS  
Foundation Trust, London, UK  
<sup>3</sup> Department of Rheumatology, West Hertfordshire Hospital  
NHS Trust, Hertfordshire, UK

### Key Points for Decision Makers

Certolizumab pegol (CZP) has shown similar clinical efficacy to other recommended biologic disease-modifying antirheumatic drugs (bDMARDs) in patients who had an inadequate response to tumour necrosis factor- $\alpha$  inhibitors (TNFi). The lack of published evidence on the effectiveness of some comparators following inadequate response to a TNFi adds considerable uncertainty to the incremental cost effectiveness of CZP.

In the population eligible for rituximab (RTX) in combination with methotrexate (MTX), RTX is of similar clinical efficacy to CZP but has a significantly lower cost. Therefore, RTX in combination with MTX should be preferred to CZP with MTX.

In the population for whom RTX is contraindicated or withdrawn, CZP in combination with MTX has a similar efficacy and comparable costs to other bDMARDs in combination with MTX recommended by the National Institute for Health and Care Excellence (NICE).

In the population for whom MTX is contraindicated or withdrawn, CZP monotherapy has a similar efficacy and comparable cost to some of the other bDMARD monotherapies recommended by NICE.

The relative simplicity of the decision when bDMARDs were the main comparator provides supportive evidence that abbreviated appraisals which have been proposed by NICE where efficacy and costs are comparable can be delivered.

## 1 Introduction

The National Institute for Health and Care Excellence (NICE) is an independent organisation responsible for providing national guidance on promoting good health and preventing and treating ill health in priority areas with significant impact. Health technologies must be shown to be clinically effective and to represent a cost-effective use of National Health Service (NHS) resources in order for NICE to recommend their use within the NHS in England. The NICE Single Technology Appraisal (STA) process usually covers new single health technologies within a single indication, soon after their UK market authorisation [1]. Within the STA process, the company provides NICE

with a written submission, alongside a mathematical model that summarises the company's estimates of the clinical and cost effectiveness of the technology. This submission is reviewed by an external organisation independent of NICE [the Evidence Review Group (ERG)], which consults with clinical specialists and produces a report. After consideration of the company's submission, the ERG report and testimony from experts and other stakeholders, the NICE Appraisal Committee (AC) formulates preliminary guidance, the Appraisal Consultation Document (ACD), which indicates the initial decision of the AC regarding the recommendation (or not) of the technology. Stakeholders are then invited to comment on the submitted evidence and the ACD, after which a further ACD may be produced or a Final Appraisal Determination (FAD) issued, which is open to appeal. An ACD is not produced when the technology is recommended within its full marketing authorisation; in this case, a FAD is produced directly.

This paper presents a summary of the ERG report [2] for the STA of certolizumab pegol (CZP) for treating rheumatoid arthritis (RA) following inadequate response to a tumour necrosis factor- $\alpha$  inhibitor (TNFi) and a summary of the subsequent development of the NICE guidance for the use of this technology in England. Full details of all relevant appraisal documents (including the appraisal scope, ERG report, company and consultee submissions, FAD and comments from consultees) can be found on the NICE website [3].

## 2 The Decision Problem

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 [4]. The primary symptoms are pain, morning stiffness, swelling, tenderness, loss of movement, fatigue, and redness of the peripheral joints [5, 6]. RA is associated with substantial costs both directly (due to drug acquisition and hospitalisation) and indirectly (due to reduced productivity) [7]. RA has long been reported as being associated with increased mortality [8, 9], particularly due to cardiovascular events [10]. There are an estimated 580,000 people in England and Wales with RA, with approximately 26,000 incident cases per year [11]. RA is more prevalent in females (1.16%) than in males (0.44%) [12], with the majority of cases being diagnosed between the ages of 40 and 80 years [13].

Two classifications have dominated the measurement of improvement in RA symptoms: American College of Rheumatology (ACR) responses [14] and European League Against Rheumatism (EULAR) responses [15]. ACR

response has been widely adopted in randomised controlled trials (RCTs), although studies have shown that the value of the measure can vary between studies due to the timing of the response [16]. In the UK, monitoring the progression of RA is often undertaken using the disease activity score of 28 joints (DAS28). 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 [15]. EULAR responses have been reported less frequently in RCTs than ACR responses [2]. However, a EULAR response is much more closely aligned to the treatment continuation rules stipulated by NICE, which require either a moderate or good EULAR response or a DAS28 improvement of more than 1.2 points to continue treatment with biologic disease-modifying antirheumatic drugs (bDMARDs).

## 2.1 Current Treatment

For people with newly diagnosed RA, NICE recommends considering a combination of conventional disease-modifying antirheumatic drugs (cDMARDs) including methotrexate (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 [17]. NICE guidance [Technology Appraisal (TA) 375] [18] recommends the use of the following bDMARDs: abatacept (ABA), adalimumab (ADA), CZP, etanercept (ETA), golimumab (GOL), infliximab (IFX) and tocilizumab (TOC), each in combination with MTX, for patients who have severe active RA (defined as a DAS28 score >5.1) after the failure to respond to cDMARD treatment. For people who meet these criteria but for whom MTX is contraindicated or has been withdrawn, NICE recommends the use of ADA, CZP, ETA and TOC as monotherapy [18]. Most of these bDMARDs (all except ABA and TOC) are TNFis. After the failure of the first TNFi, NICE recommends rituximab (RTX) in combination with MTX for the treatment of severe active RA [19]. If RTX is contraindicated or withdrawn because of an adverse event (AE), NICE recommends ABA, ADA, ETA, GOL, IFX or TOC in combination with MTX [19–21]. If MTX is contraindicated or withdrawn because of an AE, NICE recommends ADA or ETA [19] as monotherapy. NICE also recommends TOC in combination with MTX as a third-line biologic after inadequate response to RTX in combination with MTX [20].

Treatment continuation criteria vary across TAs: TA375 [18] states that for patients to continue treatment with their first bDMARD treatment they must achieve and maintain at least a moderate EULAR response. For RTX, TA195

[22] states that treatment should be continued only if there is an improvement in the DAS28 score of at least 1.2 points at initiation of treatment and whilst this response is maintained. If the relevant continuation criterion is not met, then the treatment should be stopped and the next treatment in the sequence initiated.

## 3 The Independent Evidence Review Group (ERG) Review

In accordance with the process for STAs, the ERG and NICE had the opportunity to seek clarification on specific points in the company's submission (CS) [23], in response to which the company provided additional information. The ERG also modified the company's decision analytic model to produce an ERG base case and to assess the impact of alternative parameter values and assumptions on the model results. The evidence presented in the company's submission and the ERG's review of that evidence is summarised here.

### 3.1 Clinical Evidence Provided by the Company

Evidence was presented in the CS [23] for the efficacy of CZP in combination with MTX and other cDMARDs or as monotherapy in the treatment of moderate to severe RA in patients with a previously inadequate response or intolerance to TNFi therapy. This evidence was based on six RCTs (REALISTIC [24], DOSEFLEX [25], PREDICT [26], SWITCH [27], J-RAPID [28] and HIKARI [29]; see Table 1 for full study names). All of these trials recruited both TNFi-naïve and TNFi-experienced patients, with the exception of the SWITCH study, which was performed solely in a TNFi-experienced population. Five RCTs were placebo-controlled (PREDICT did not have a non-CZP comparator arm). The durations of the randomised controlled phases in the RCTs were 12 weeks (REALISTIC and SWITCH); 16 weeks (DOSEFLEX); 24 weeks (J-RAPID and HIKARI) and 52 weeks (PREDICT). The primary outcome in four of the RCTs (REALISTIC, SWITCH, J-RAPID and HIKARI) was ACR20 response at week 12. The primary endpoint of DOSEFLEX was ACR20 response at 34 weeks in patients randomised at week 18, whilst the primary endpoints in PREDICT were clinical disease activity scale (CDAI) and RAPID-3 scores (routine assessment of patient index data 3) at 12 and 52 weeks. J-RAPID and HIKARI were undertaken exclusively in Japan. The company also included supplementary observational evidence from the Swedish registry-based study ARTIS (Antirheumatic Therapies in Sweden) [30]. Disease activity was reported in the CS [23] as ACR and EULAR responses, DAS28 and CDAI. The clinical

**Table 1** Full study names for randomised controlled trials

Acronym	Study
ATTAIN [35]	Abatacept Trial in Treatment of Anti-TNF INadequate responders
DOSEFLEX [25]	Dosing Flexibility Study in Patients with Rheumatoid Arthritis
GO-AFTER [36]	Golimulab After Former anti-tumour necrosis factor $\alpha$ Therapy Evaluated in Rheumatoid arthritis
HIKARI [29]	A multicenter, Double-blind, Randomized, Placebo-Controlled, Parallel-Group Trial to Assess the Efficacy, Pharmacokinetics, and Safety of CDP870 Without Coadministration of MTX in Japanese Active RA Patients in Whom MTX Cannot be Administrated
J-RAPID [28]	Japan Rheumatoid Arthritis Prevention of structural Damage
PREDICT [26]	Phase 4, Randomized, 52-Week Study to Evaluate Two Assessment Tools to Predict Treatment Success at 52 Weeks Based on a Treatment Decision at Week 12 in Subjects with Moderate to Severe Rheumatoid Arthritis Receiving Cimzia
RADIATE [31]	Rheumatoid Arthritis Study in Anti-TNF Failures
REALISTIC [24]	RA Evaluation in Subjects Receiving TNF Inhibitor CZP
REFLEX [33]	Randomized Evaluation of Long-term Efficacy of Rituximab in RA
SWITCH [27]	Randomised Controlled Trial of Switching to Alternative Tumour Necrosis Factor (TNF)-Blocking Drugs or Abatacept or Rituximab in Patients with Rheumatoid Arthritis Who Have Failed an Initial TNF-Blocking Drug

effectiveness results of the described trials were confidential and therefore cannot be reported here.

No head-to-head evidence evaluating CZP against comparator bDMARDs was available and therefore the company performed a Bayesian network meta-analysis (NMA) to assess the effectiveness of CZP compared with other recommended bDMARDs. The results of nine relevant RCTs were included in the NMA: three trials were included for CZP + MTX (REALISTIC [24], J-RAPID [28] and SWITCH [27]); two for TOC + MTX (RADIATE [31], Genovese et al. [32]); two for RTX + MTX (REFLEX [33], Combe et al. [34]); one for ETN + MTX (Combe et al. [34]); one RCT for ABA + MTX (ATTAIN [35]); and one for GOL + MTX (GO-AFTER [36]). The company only considered fixed-effect models and justified its decision based on the limited number of studies.

### 3.1.1 Critique of the Clinical Evidence and Interpretation

The eligibility criteria applied in the selection of evidence for the clinical effectiveness review were considered by the ERG to be reasonable and generally consistent with the decision problem as outlined in the final NICE scope. The ERG was satisfied that the searches for clinical effectiveness evidence reported in the CS [23] were likely to have identified all relevant published RCT evidence. However, an RCT by Kang et al. [37] that included CZP was identified by the ERG and clarification was sought from the company as to why it was not included in the CS [23]. The

company responded that the Kang et al. study was not included because the number of patients in the study who were TNFi experienced was small. However, the ERG noted that two CZP RCTs were included even though they also had low numbers of TNFi-experienced patients (J-RAPID and HIKARI) and therefore the ERG considered that the justification provided by the company to support their decision to exclude the Kang et al. study [37] was not applied consistently.

The quality of the included RCTs including CZP and the ARTIS non-randomised study were assessed using well established and recognised criteria. Data for radiological progression and joint damage were not presented in the CS [23]; however, data on inhibition of joint structural damage were available in the published articles for both J-RAPID and HIKARI. Extra-articular manifestations of disease were not included in the CS [23]. Study and patient characteristics for included CZP trials were clearly described in a narrative summary alongside clinical and safety data. However, *p* values were frequently unreported and therefore the ERG requested that these be provided by the company where available. Classical meta-analyses were performed for CZP used in combination with MTX and for CZP as monotherapy. Classical meta-analyses were performed separately for the outcomes of ACR20/50/70, EULAR response and DAS28 (Erythrocyte Sedimentation Rate [ESR]) remission at 3 months. No meta-analysis was performed for outcomes at 6 months due to data unavailability. Both fixed-effects (Mantel–Haenszel) and random-



effects (DerSimonian and Laird) models were used. Heterogeneity between trials was investigated using  $I^2$  values. The ERG noted that it is generally recommended that at least five studies should be available for a frequentist meta-analysis, whereas the analyses in the CS [23] included, at most, only three studies. A Bayesian NMA was performed to assess CZP against comparator interventions, which had several limitations. The ERG believed that several changes would have been required to the analyses conducted and to the reporting of the results in order for them to represent the genuine uncertainty and be useful for decision-making purposes. These changes included incorporating weakly informative prior information for the between-study standard deviation; generating predictive distributions of the effects of treatments in a new study; using the evidence from the REALISTIC study to generate the probabilities of being in each ACR and EULAR category for the reference treatment; and taking draws from the joint posterior distribution of treatment effects rather than assuming univariate normal distributions for them. It was not possible for the ERG in the time available to make the required changes to produce robust results and therefore the ERG did not amend the NMA presented in the CS [23].

### 3.2 Cost-Effectiveness Evidence Provided by the Company

The company supplied a de novo cohort Markov model constructed in Microsoft Excel©. The perspective was that of the NHS and a 6-month cycle length and a time horizon of 45 years (assumed to be lifetime) was used. A discount rate of 3.5% per annum was used both for costs and for utilities. Patients entered the model after inadequate response to a TNFi and transitioned to one of three health states depending on their EULAR response: none, moderate or good. Non-responders discontinued treatment after a cycle and transitioned to a state representing the first 6 months of the first follow-up treatment. Good and moderate EULAR responders remained in their states until treatment discontinuation, after which they transitioned to the state representing the first 6 months of the next treatment in the sequence. Patients achieving good or moderate EULAR response in follow-up treatments transitioned to a state representing the rest of the duration of the treatment whereas non-responders transitioned to the state representing the first 6 months of the next follow-up treatment in the sequence. During any cycle, patients could transition from any of the alive states to death.

The company considered three different populations: population A, formed by patients eligible for RTX in combination with MTX (RTX + MTX); population B, formed by patients for whom RTX is contraindicated or withdrawn due to an adverse event; and population C,

formed by patients for whom MTX is contraindicated or withdrawn due to an adverse event. For population A, the company compared a sequence that it believed to reflect currently recommended clinical practice (consisting of RTX + MTX, TOC + MTX, ABA + MTX, MTX + hydroxychloroquine + sulfasalazine, non-biologic treatment mixture and palliative care) with a sequence consisting of CZP in combination with MTX (CZP + MTX) inserted at the start of the comparator sequence. For population B, the company compared a sequence starting with a treatment of CZP + MTX with the sequences starting with treatments of ABA, ADA, GOL, ETA, IFX and TOC each in combination with MTX. For population C, the company compared a sequence starting with a treatment of CZP monotherapy with sequences starting with treatments of ADA, ETA and TOC monotherapies.

The company modelled treatment efficacy for the first treatment in the sequence differently from subsequent treatments. The NMA conducted by the company was used to estimate the probabilities of no, moderate and good EULAR responses of CZP and comparators when the interventions were used in combination with MTX. The probabilities of EULAR responses for CZP and comparators when used as monotherapy were estimated based on the relative efficacy compared with CZP + MTX.

Changes in Health Assessment Questionnaire (HAQ) score for each of the EULAR response categories were estimated using a linear regression fitted to data from the REALISTIC trial. Changes in EQ-5D from baseline was conditional on EULAR response to the first therapy and were estimated through a series of linear regression analyses with patient-level data from the PREDICT study [26]. Treatment discontinuation rate for patients with a good or moderate EULAR response was modelled with a Weibull distribution based on the Assessment Group's approach in NICE TA195 [19], but assuming instead that all bDMARDs had the same discontinuation rate.

For subsequent treatments, both the probabilities of EULAR response and the changes in HAQ scores conditioned on response were estimated based on the RADIATE study [31], which analysed the efficacy of TOC + MTX compared with placebo + MTX in patients who had failed to respond to one or more TNFis. Treatment discontinuation rate following response was assumed to be constant and equal to that of the first treatment between the 6th month and a year.

Patients' utilities were assumed to depend on the HAQ score in each cycle. Patients achieving good or moderate EULAR response experienced a decrease (improvement) in HAQ score, the value of which was added at treatment discontinuation. Whilst on bDMARD treatment, the HAQ score was assumed to remain constant. Contrastingly, for patients on cDMARDs or palliative care, the HAQ score

was assumed to increase linearly at a rate of 0.045 and 0.06 per year, respectively. Changes in EQ-5D were estimated following a linear mapping algorithm from changes in HAQ scores reported by Brennan et al. [38]. Mortality was assumed to be affected by HAQ score, with a hazard ratio of 1.43 per HAQ score point applied following Norton et al. [39].

Unit costs were taken from the Personal Social Services Research Unit [40], British National Formulary (BNF) [41] and NHS Reference Costs [42]. The cost of CZP and GOL used in the model included the public Patient Access Scheme (PAS) in place. For CZP, this results in the first ten syringes of CZP being provided to the NHS free of charge. The list prices reported in the BNF were used for the rest of the drugs, as directed by NICE. Costs were valued in 2015 Great British pounds.

In their base-case analysis, the company estimated that for population A, the probabilistic incremental cost-effectiveness ratio (ICER) of adding CZP + MTX before the currently recommended treatment sequence was £33,222 per quality-adjusted life-year (QALY) gained (0.290 QALYs gained at a cost of £9842). For population B, the estimated probabilistic ICER of CZP + MTX versus GOL + MTX was £3461 (0.256 QALYs gained at a cost of £884), whilst the estimated probabilistic ICER of TOC (intravenous [IV]) + MTX versus CZP + MTX was £132,783 (0.201 QALYs gained at a cost of £26,659). For population C, the estimated probabilistic ICER of CZP monotherapy versus ADA monotherapy was £3461 (0.260 QALYs gained at a cost of £1336), whilst the estimated probabilistic ICER of TOC (IV) monotherapy versus CZP monotherapy was £133,655 (0.196 QALYs gained at a cost of £26,179). One-way sensitivity analyses undertaken by the company, where the mean values were replaced with values from the relevant 95% confidence intervals, showed that the net monetary benefit of CZP, assuming a threshold of £30,000 per QALY gained, was most sensitive to the efficacies of RTX + MTX, CZP (as monotherapy or in combination with MTX), and TOC (as monotherapy or in combination with MTX). Scenario analyses undertaken by the company showed that assuming the efficacy of CZP is equal to the other TNFi has the biggest impact on the ICER, followed by the treatment duration of RTX + MTX and assuming a flat HAQ score progression for cDMARDs and palliative care. All of these changes produced ICERs less favourable to CZP, with the exception of setting the efficacy of CZP equal to other bDMARDs in population C.

### 3.2.1 Critique of the Cost-Effectiveness Evidence and Interpretation

The ERG had concerns regarding the NMAs used to estimate the efficacy of CZP and its comparators, which were

used to characterise uncertainty in the economic model. The company expected heterogeneity but assumed that a fixed-effects model was appropriate. The evidence for the reference treatment from the REALISTIC study was assumed by the company to represent the evidence for the target population; however, the company only used the 'no EULAR response' rates from the REALISTIC study and used evidence from all other studies to estimate the response rates for other ACR and EULAR response categories. The company generated estimates of absolute probabilities of being in each ACR and EULAR response category using mean and standard deviations extracted from the NMA and assuming univariate normal. However, this approach fails to preserve the underlying joint distribution between parameters, and using draws from the joint posterior distribution would have been preferred. The ERG also believed that the exclusion of J-RAPID from the NMA was not justified.

The ERG noted that the company used a simplistic approach to map changes in HAQ score to changes in EQ-5D utility and that better approaches exist to capture the non-linearity of the relationship between HAQ score and EQ-5D [43, 44].

The ERG believed that the treatment sequences considered by the company for population A were inappropriate because they include TOC + MTX followed by ABA + MTX after RTX + MTX. Clinical experts consulted by the ERG claimed that usually TOC + MTX or ABA + MTX were provided, but not both.

Due to the lack of published evidence of the efficacy of IFX, ADA and ETA, each in combination with MTX in patients with an inadequate response to a TNFi, the company assumed the efficacy of these drugs to be equal to that of GOL + MTX. Similarly, the company made assumptions on the efficacy of TOC, ADA and ETA, each used in monotherapy in patients with inadequate response to a TNFi due to the lack of published evidence. It was assumed that the relative efficacy of each intervention when used in combination with MTX compared with CZP + MTX was generalisable to when the treatment was used as a monotherapy. The ERG believes that these assumptions introduce considerable uncertainty which is not fully captured and that therefore the results of the base-case analysis should be interpreted with caution.

The company assumed the same treatment duration for all bDMARDs for its base-case analysis, despite evidence suggesting different treatment durations for different bDMARDs [19]. The ERG notes that the company identified treatment duration as a parameter with a large impact on the ICER (especially in population A) in one of their scenario analyses.

The ERG had concerns regarding the modelling of the efficacy of subsequent treatments due to the lack of

evidence on treatment efficacy in patients with an inadequate response to a previous TNFi. In addition, the ERG believed that the difference in the modelling of the first and subsequent treatments meant that the model was not properly suited to comparing sequences of different lengths.

### 3.3 Additional Work Undertaken by the ERG

The ERG applied a series of modifications to the company's base-case analysis. The most relevant were (1) adding biosimilars of IFX and ETA and subcutaneous (SC) formulations of TOC and ABA as comparators; (2) comparing four possible sequences (CZP before RTX, CZP after RTX, no RTX, no CZP) for population A; (3) removing ABA + MTX treatment after TOC + MTX from the sequences in population A; (4) using different durations for different treatments based on the data provided in TA195 [19]; (5) setting the RTX retreatment interval to 7.35 months; (6) using the results of the NMA including J-RAPID; (7) amending the cost of TOC by considering the 80-mg formulation and setting the 800-mg limit per administration recommended in TOC's summary of product characteristics; and (vii) adjusting the mean HAQ improvements reported in RADIATE to be more appropriate for responders.

These modifications resulted in the sequence including CZP + MTX being dominated in population A in the ERG's base-case analysis. For population B, the estimated probabilistic ICER of CZP + MTX versus GOL + MTX was £13,155 (0.287 QALYs gained at a cost of £3774) whilst the estimated probabilistic ICER of TOC (SC) + MTX versus CZP + MTX was £43,994 (0.544 QALYs gained at a cost of £23,954). For population C, the estimated probabilistic ICER of CZP monotherapy versus ADA was £14,437 (0.291 QALYs gained at a cost of £4206), whilst the estimated probabilistic ICER of TOC (SC) monotherapy versus CZP monotherapy was £45,090 (0.525 QALYs gained at a cost of £23,690).

The ERG also undertook two scenario analyses using the results from the NMA excluding J-RAPID, as the company did for its base case, and assuming ADA, ETA and IFX had the same efficacy as CZP (instead of assuming their efficacy was equal to that of GOL). The first scenario analysis had little impact on the results; contrastingly, the second scenario analysis showed very different results in which biosimilar ETA dominated CZP in populations B and C (population A was unaffected). However, there remained treatments currently recommended by NICE that were estimated to be less cost effective than CZP.

Estimates of the cost effectiveness of CZP when the ABA and TOC PASs were taken into consideration were provided to the NICE AC in a confidential appendix.

### 3.4 Conclusions of the ERG Report

The ERG's critical appraisal identified a number of issues relating to the company's model and analysis. The most pertinent of these relate to (1) the weaknesses of the NMA; (2) inclusion of two lines of bDMARDs after RTX + MTX; (3) exclusion from the base case of biosimilars for IFX and ETA; (4) exclusion from the base case of SC formulations of TOC and ABA; (5) assuming the same treatment duration for all bDMARDs; (6) assuming a retreatment interval of RTX that was deemed too short by the NICE AC in TA195 [19]; (7) ignoring the 80-mg formulation of TOC and the 800-mg limit per administration; and (8) assuming that the mean HAQ improvements reported in RADIATE apply to responders. The ERG undertook a series of exploratory analyses based on the company's submitted model in order to address the limitations listed above; however, no additional work was undertaken correcting the NMA and as such, the level of uncertainty in all presented results is underestimated.

The ERG's base-case analysis suggests that for population A, CZP + MTX should not be used before RTX + MTX. Limitations of the company's model in the methods for modelling subsequent treatments mean that the results of a fully incremental analysis comparing sequences of different lengths was deemed unreliable. However, when comparing sequences of equal length, the use of RTX + MTX before CZP + MTX, or the use of RTX + MTX rather than CZP + MTX was dominant. This result is not unexpected given the similar efficacies of RTX + MTX and CZP + MTX and the lower acquisition price associated with RTX compared with CZP.

For population B, the probabilistic ICER of CZP + MTX versus biosimilar ETA + MTX is expected to be £12,116 per QALY gained and the probabilistic ICER of TOC (SC) + MTX versus CZP + MTX is expected to be £45,414 per QALY gained. These ICERs are less favourable to CZP + MTX than the company's base-case ICERs. However, the probability that CZP + MTX produces more net benefit than its comparators assuming a willingness-to-pay (WTP) threshold of £30,000 per QALY gained remains essentially unchanged at 0.96. However, the PAS for TOC has not been included in these calculations.

For population C, the probabilistic ICER of CZP monotherapy versus biosimilar ETA monotherapy is estimated to be £13,784 per QALY gained and the probabilistic ICER of TOC (SC) monotherapy versus CZP monotherapy is expected to be £46,501 per QALY gained. These ICERs are less favourable to CZP monotherapy than the company's base-case ICERs. However, the probability that CZP monotherapy produces more net benefit than its

comparators assuming a WTP threshold of £30,000 per QALY gained is reduced slightly to 0.96. However, the PAS for TOC has not been included in these calculations.

Additional analyses undertaken by the ERG using this revised base-case model indicate that excluding J-RAPID from the NMA has little impact on the results of the analyses. In contrast, assuming that ADA, IFX and ETA in combination with MTX have the same efficacy as CZP + MTX (rather than GOL + MTX) leads to biosimilar ETA + MTX dominating CZP + MTX; similarly, assuming ADA and ETA monotherapy have the same efficacy as CZP monotherapy leads to biosimilar ETA monotherapy dominating CZP monotherapy. The ERG notes that even if CZP + MTX were dominated by biosimilar ETA + MTX, there remain comparators for which it is estimated that CZP + MTX is dominant, such as IFX + MTX and ADA + MTX. The latter two interventions will remain options for treatment in population B as they were recommended in TA195.

With respect to the company's economic analysis and the ERG's additional exploratory analyses, there remain several potentially important areas of uncertainty:

1. The lack of data on the efficacy of ETA, ADA and IFX in combination with MTX in patients who have not responded adequately to a TNFi; there is a similar lack of data on the efficacy of ETA and ADA monotherapy in these patients. Alternative assumptions for the efficacy of these drugs to those used by the company produced markedly different results. This limitation had already been highlighted by the AC of TA195 [19] and was acknowledged during the scoping meeting for the current appraisal.
2. The scarcity of data on the efficacy of bDMARDs in general, and TNFis in particular, in patients who have had an inadequate response to two or more bDMARDs. There is also the possibility that there could be reduced efficacy of TNFis following inadequate response to a previous TNFi.
3. The relative efficacies of the bDMARDs are uncertain given the limitations of the NMA within the CS [23], namely (1) not incorporating weakly informative prior information for the between-study standard deviation; (2) not using predictive distributions of the effects of treatments in a new study; (3) calculating the 'no response' rates based only on the evidence from the REALISTIC study and using the evidence from other sources only to estimate other response category rates instead of directly generating the probabilities of being in each response category; and (4) assuming univariate normal distribution treatment effects instead of taking draws from their joint posterior distribution.

## 4 Key Methodological Issues

The ERG considered that the company's model was not appropriate to compare sequences of different lengths due to the difference in the implementation of the first and subsequent treatments and in the assumptions made when modelling subsequent treatments. Furthermore, the NMA used in the economic model had several shortcomings that prevented a genuine representation of uncertainty and limited its usefulness for decision-making purposes. Finally, the choice of a cohort model as modelling approach proved inappropriate to represent the nature of the disease. For example, the company acknowledged that due to the inability of cohort models to handle non-linear functions, they had to use a linear HAQ progression for patients on cDMARDs or palliative care and their mapping of HAQ to EQ-5D was restricted to linear models. Additionally, the inability of cohort models to track the time a patient has spent in such treatments resulted in the treatment discontinuation rate being assumed to be constant for subsequent treatments instead of time-dependent. An individual patient model would have resolved this methodological issue.

The ERG noted that the conclusions of the company's analyses tally with the expectations before constructing a mathematical model, given the comparable efficacy and costs of the intervention and its comparators. The relative simplicity of this decision provides supportive evidence that abbreviated appraisals, which have been proposed by NICE [45], can be delivered under conditions such as those in the CZP STA.

## 5 National Institute for Health and Care Excellence Guidance

In September 2016, on the basis of the evidence available (including verbal testimony of invited clinical experts and patient representatives), the AC produced guidance that CZP in combination with MTX was recommended as an option following an inadequate response to a TNFi for treating severe RA if RTX is contraindicated or not tolerated. The AC also produced guidance that CZP monotherapy was recommended if MTX was contraindicated or not tolerated. Both recommendations were conditional on the company providing CZP with the agreed PAS.

### 5.1 Consideration of Clinical and Cost-Effectiveness Issues Included in the Final Appraisal Determination

This section summarises the key issues considered by the AC. The full list of the issues considered by the AC can be found in the FAD [46].



### 5.1.1 Current Clinical Management

The AC considered the current clinical management of severe active RA following inadequate response to a TNFi in England and noted that the NICE guidance recommends ADA, ETA, IFX, ABA, TOC and GOL (each with MTX) as options, when RTX (plus MTX) is contraindicated or not tolerated, and ADA and ETA monotherapy as alternative options if RTX therapy cannot be given because MTX is contraindicated or not tolerated. The AC heard from clinical experts that responses to bDMARDs differ between people and therefore it is important to have a range of options for bDMARD treatments. The AC was aware that the marketing authorisation covers the use of CZP in moderate to severe disease but that TA375 [18] recommends that treatment with a bDMARD should only be started when disease is severe—that is, a disease activity (DAS28) score of more than 5.1.

### 5.1.2 Uncertainties in the Clinical Evidence

The AC considered the company's clinical evidence and accepted that the results showed that CZP was more clinically effective than placebo. It understood that the only evidence available on the comparative effectiveness of CZP and the comparator bDMARDs was from the company's NMA. The committee concluded that there are uncertainties from the methods used and it could not reliably conclude whether CZP was more clinically effective than the comparator bDMARDs on the basis of the evidence presented by the company. The AC heard from the clinical experts that CZP, which is already in use in clinical practice, is not considered to be better or worse than other TNFis. The AC concluded that CZP has a similar efficacy to other available bDMARDs.

### 5.1.3 Uncertainties in the Economic Modelling

The AC had concerns about the company's approach to evaluating the cost effectiveness of CZP plus MTX for patients for whom RTX plus MTX was an option. Specifically, the AC was not persuaded that a treatment sequence containing CZP and six other treatments should be compared with the same sequence without CZP. The AC was aware that using different sequence lengths can increase modelling uncertainties and concluded that treatment sequences of the same length are preferable. After consultation, the AC expressed uncertainties about the assumptions used in the company's model. The AC preferred the ERG's values for the retreatment interval for RTX and using different treatment durations for TNFis and non-TNFis based on the REFLEX study and its

extension. Based on the ERG's exploratory analysis, the AC concluded that CZP plus MTX was not a cost-effective treatment in patients for whom RTX plus MTX was an option.

For people for whom RTX or MTX are contraindicated or not tolerated, the AC noted the similarities in costs between bDMARDs and its conclusions on comparative efficacy, and therefore that equivalence among bDMARDs could be accepted. Therefore, the AC concluded that CZP plus MTX or CZP monotherapy can be considered a cost-effective use of NHS resources for people for whom RTX or MTX are contraindicated or not tolerated.

## 6 Conclusions

The evidence suggests that CZP plus MTX or CZP as monotherapy has a similar efficacy for treating severe active RA following inadequate response to a TNFi to that of other bDMARDs already recommended by NICE. Therefore, CZP plus MTX or as monotherapy was considered by NICE to be a cost-effective use of NHS resources for people for whom RTX or MTX are contraindicated or not tolerated. However, the cost of RTX treatment is significantly lower than that of CZP, with comparable efficacy, so CZP was not considered by NICE to be a cost-effective use of NHS resources when RTX plus MTX is a treatment option for a patient.

**Acknowledgements** This summary of the ERG report was compiled after NICE issued the FAD. All authors have commented on the submitted manuscript and have given their approval for the final version to be published. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of NICE or the Department of Health. Any errors are the responsibility of the authors.

**Author Contributions** IB and MS critiqued the mathematical model provided and the cost-effectiveness analyses submitted by the company. RA and EG critiqued the clinical effectiveness data reported by the company. JWS critiqued the NMA performed by the company. MC critiqued the literature searches undertaken by the company. DLS and AY provided clinical advice to the ERG throughout the project. All authors were involved in drafting and commenting on the final document. IB acts as the guarantor of the manuscript. This summary has not been externally reviewed by PharmacoEconomics.

### Compliance with Ethical Standards

**Funding** This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment Programme (Project Number 15/06/07). Visit the HTA programme website for further project information (<http://www.hta.ac.uk>).

**Conflicts of interest** John W. Stevens undertook methodological work for BMS, the manufacturer of abatacept, for SchARR, although this work was not related to rheumatoid arthritis. The remaining authors have no conflicts of interest to declare.

## References

- National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. 2013. <https://www.nice.org.uk/article/pmg9/>. Accessed 27 Apr 2017.
- Stevenson M, 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 Technol Assess*. 2016;20(35):1–610.
- National Institute for Health and Clinical Excellence (NICE). Certolizumab pegol for treating rheumatoid arthritis after inadequate response to a TNF-alpha inhibitor. 2016. <https://www.nice.org.uk/guidance/ta415>. Accessed 27 Apr 2017.
- Scott DL, Steer S. The course of established rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2007;21(5):943–67.
- Pincus T, et al. Early radiographic joint space narrowing and erosion and later malalignment in rheumatoid arthritis: a longitudinal analysis. *J Rheumatol*. 1998;25(4):636–40.
- Drossaers-Bakker KW, et al. Radiographic damage of large joints in long-term rheumatoid arthritis and its relation to function. *Rheumatology*. 2000;39(9):998–1003.
- Allaire S, et al. Current risk factors for work disability associated with rheumatoid arthritis: recent data from a US national cohort. *Arthritis Care Res*. 2009;61(3):321–8.
- Naz SM, Symmons DPM. Mortality in established rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2007;21(5):871–83.
- Dadoun S, et al. Mortality in rheumatoid arthritis over the last fifty years: Systematic review and meta-analysis. *Joint Bone Spine*. 2013;80(1):29–33.
- Meune C, et al. Trends in cardiovascular mortality in patients with rheumatoid arthritis over 50 years: a systematic review and meta-analysis of cohort studies. *Rheumatology*. 2009;48(10):1309–13.
- The National Audit Office. Services for people with rheumatoid arthritis. 2009 [cited 2016 24 March]. <https://www.nao.org.uk/wp-content/uploads/2009/07/0809823.pdf>. Accessed 27 Apr 2017.
- Symmons DPM, et al. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *Rheumatology*. 1994;33(8):735–9.
- Alamanos Y, Drosos AA. Epidemiology of adult rheumatoid arthritis. *Autoimmun Rev*. 2005;4(3):130–6.
- Felson DT, et al. American college of rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum*. 1995;38(6):727–35.
- van Gestel AM, et al. 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 Rheum*. 1996;39(1):34–40.
- Felson DT. Assessing the efficacy and safety of rheumatic disease treatments: Obstacles and proposed solutions. *Arthritis Rheum*. 2003;48(7):1781–7.
- National Collaborating Centre for Chronic Conditions. Rheumatoid arthritis: national clinical guideline for management and treatment in adults. 2009. <https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0009576/>. Accessed 27 Apr 2017.
- 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. <https://www.nice.org.uk/guidance/ta375>. Accessed 27 Apr 2017.
- 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. 2010. <http://www.nice.org.uk/guidance/ta195>. Accessed 27 Apr 2017.
- National Institute for Health and Clinical Excellence (NICE). Tocilizumab for the treatment of rheumatoid arthritis. 2012. <https://www.nice.org.uk/guidance/ta247>. Accessed 27 Apr 2017.
- 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. <http://www.nice.org.uk/guidance/ta225>. Accessed 27 Apr 2017.
- Wu EQ, et al. Mapping FACT-P and EORTC QLQ-C30 to patient health status measured by EQ-5D in metastatic hormone-refractory prostate cancer patients. *Value Health*. 2007;10:408–14.
- UCB Pharma. Certolizumab Pegol for treating rheumatoid arthritis after inadequate response to a TNF inhibitor. 2016. <https://www.nice.org.uk/guidance/ta415/documents/committee-papers>. Accessed 27 Apr 2017.
- Weinblatt ME, et al. Efficacy and safety of certolizumab pegol in a broad population of patients with active rheumatoid arthritis: results from the REALISTIC phase IIIb study. *Rheumatology*. 2012;51(12):2204–14.
- Furst DE, et al. Two dosing regimens of certolizumab pegol in patients with active rheumatoid arthritis. *Arthritis Care Res (Hoboken)*. 2015;67(2):151–60.
- Curtis JR, et al. A randomized trial comparing disease activity measures for the assessment and prediction of response in rheumatoid arthritis patients initiating certolizumab pegol. *Arthritis Rheumatol*. 2015;67(12):3104–12.
- Schiff MH, et al. Rheumatoid arthritis secondary non-responders to TNF can attain an efficacious and safe response by switching to certolizumab pegol: a phase IV, randomised, multicentre, double-blind, 12-week study, followed by a 12-week open-label phase. *Ann Rheum Dis*. 2014;73(12):2174–7.
- Yamamoto K, et al. Efficacy and safety of certolizumab pegol plus methotrexate in Japanese rheumatoid arthritis patients with an inadequate response to methotrexate: the J-RAPID randomized, placebo-controlled trial. *Mod Rheumatol*. 2014;24(5):715–24.
- Yamamoto K, et al. Efficacy and safety of certolizumab pegol without methotrexate co-administration in Japanese patients with active rheumatoid arthritis: the HIKARI randomized, placebo-controlled trial. *Mod Rheumatol*. 2014;24(4):552–60.
- Chatzidionysiou K, et al. Effectiveness and survival-on-drug of certolizumab pegol in rheumatoid arthritis in clinical practice: results from the national Swedish register. *Scand J Rheumatol*. 2015;44(6):431–7.
- Emery P, et al. IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial. *Ann Rheum Dis*. 2008;67(11):1516–23.
- Genovese MC, et al. Efficacy and safety of olokizumab in patients with rheumatoid arthritis with an inadequate response to TNF inhibitor therapy: outcomes of a randomised Phase IIb study. *Ann Rheum Dis*. 2014;73(9):1607–15.
- Cohen SB, 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*. 2006;54(9):2793–806.

34. Combe B, et al. AB0516 An open-label randomized controlled study to evaluate the efficacy of etanercept (ETN) versus rituximab (RTX), in patients with active rheumatoid arthritis previously treated with RTX and TNF blockers. *Ann Rheum Dis*. 2013;71(Suppl 3):667.
35. Genovese MC, et al. Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. *N Engl J Med*. 2005;353(11):1114–23.
36. Smolen JS, et al. Golimumab in patients with active rheumatoid arthritis after treatment with tumour necrosis factor alpha inhibitors (GO-AFTER study): a multicentre, randomised, double-blind, placebo-controlled, phase III trial. *Lancet*. 2009;374(9685):210–21.
37. Kang Y, et al. Efficacy and safety of certolizumab pegol with concomitant methotrexate in Korean rheumatoid arthritis (RA) patients with an inadequate response to MTX. *Ann Rheum Dis*. 2012;71(3):666.
38. Brennan A, et al. Modelling the cost effectiveness of TNF- $\alpha$  antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry. *Rheumatology*. 2007;46(8):1345–54.
39. Norton S, et al. Trajectories of functional limitation in early rheumatoid arthritis and their association with mortality. *Rheumatology*. 2013;52(11):2016–24.
40. Personal Social Services Research Unit. Unit costs of health and social care. 2015. <http://www.pssru.ac.uk/project-pages/unit-costs/2015/>. Accessed 27 Apr 2017.
41. British National Formulary. BNF June 2015. 2015. [www.bnf.org/](http://www.bnf.org/). Accessed 27 Apr 2017.
42. Department of Health. NHS reference Costs 2014 to 2015. 2015 [cited 2016 24 March]. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/480791/2014-15\\_National\\_Schedules.xlsx](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/480791/2014-15_National_Schedules.xlsx). Accessed 27 Apr 2017.
43. Hernández Alava M, et al. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. *Rheumatology*. 2013;52(5):944–50.
44. 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;15:550–61.
45. National Institute for Health and Care Excellence. Abbreviated technology appraisal process consultation. 2016. <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/abbreviated-technology-appraisal-process-consultation>. Accessed 27 Apr 2017.
46. National Institute for Health and Clinical Excellence (NICE). Certolizumab pegol for treating rheumatoid arthritis after inadequate response to a TNF-alpha inhibitor—final appraisal determination. 2016. <https://www.nice.org.uk/guidance/ta415/documents/final-appraisal-determination-document>. Accessed 27 Apr 2017.