Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE

Certolizumab pegol (CIMZIA®) for the treatment of Rheumatoid Arthritis

Produced by

West Midlands Health Technology Assessment Collaboration
Public Health, Epidemiology and Biostatistics
University of Birmingham
Edgbaston
Birmingham B15 2TT

Authors

Martin Connock[1]
Sandy Tubeuf[2]
Kinga Malottki[1]
Abdulrahman Uthman[1]
Jeff Round[2]
Sue Bayliss[1]
Catherine Meads[1]
David Moore[1]

[1] Department of Public Health and Epidemiology, University of Birmingham
[2] Leeds Institute of Health Sciences, University of Leeds

Acknowledgements: Professor George Kitas, Consultant Rheumatologist for his clinical advice on the background and clinical effectiveness sections.

Correspondence to

David Moore
Public Health, Epidemiology and Biostatistics,
University of Birmingham
Edgbaston
Birmingham B15 2TT

Date Completed:

August 2009

This report was commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Clinical Excellence as STA report number 07/13/01. The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme or the National Institute for Health and Clinical Excellence.

Table of Contents

Appendices

Table of Tables

Table of Figures

1 SUMMARY

1.1 Scope of the submission

The submission considers the effectiveness and cost-effectiveness of certolizumab pegol (CZP) in the treatment of patients with moderate to severe rheumatoid arthritis (RA) who have had an inadequate response to disease-modifying antirheumatic drugs (DMARDs), including methotrexate (MTX).

CZP is a recombinant humanised antibody conjugated to polyethylene glycol with inhibitory action against tumour necrosis factor α (TNF–α).

CZP was considered as mono-therapy or as combined therapy with MTX. For combined therapy in the economic analysis four comparators were considered: etanercept, infliximab, adalimumab and rituximab; however for clinical effectiveness tocilizumab represented an additional comparator. For monotherapy economic analysis and clinical effectiveness there were two comparators, adalimumab and etanercept. Two scenarios were adopted for the cost-effectiveness analyses; one scenario included a patient access scheme (PAS) for CZP in which patients received the first ten syringes of CZP treatment at no cost to the NHS and in the other scenario there was no PAS.

1.2 Summary of submitted clinical effectiveness evidence

The manufacturer's submitted evidence consisted of:

- a meta-analysis of outcomes from two triple-arm RCTs comparing CZP plus MTX at two different CZP dose regimens versus MTX + placebo (these trials were RAPID 1 lasting 52 weeks, RAPID 2 lasting 24 weeks)

- results from a 24-week RCT comparing a single dose regimen of CZP versus placebo (the FAST4WARD trial)

- an indirect comparison meta-analysis (ICM) that included thirteen studies of combined therapy of a biological DMARD (CZP, etanercept, infliximab, adalimumab, rituximab or tocilizumab) plus MTX versus MTX + placebo
- an ICM that included five studies of three biological DMARDs (adalimumab, etanercept and CZP) in monotherapy versus placebo
- an analysis of the safety of CZP using data from various RCTs and their extension studies

Results from the RCTs and from direct meta-analyses demonstrated that CZP + MTX was statistically significantly more effective than MTX + placebo for ACR20, ACR50, and ACR70 outcome measures and for quality of life measures at 3 and 6 months. Similarly CZP monotherapy was more effective than placebo for these outcomes.

On the basis of results from the ICMs, the manufacturer suggested that CZP may be at least as effective as other bDMARD comparators, and in a few ACR measures at 3 and 6 months, more effective. These ICM estimates were associated with considerable uncertainty. There is some evidence that CZP inhibits progression of structural damage to joints.

1.3 Summary of submitted cost effectiveness evidence

The model supplied by the manufacturers was internally consistent and robust.

Clinical and cost data used to populate the model were taken from appropriate sources and clearly indicated in the written portion of the submission.

Data on quality of life and health utilities was not as robust as the effectiveness and cost data used. The inclusion of the PAS has a significant impact on the model results.

The scenario with the PAS is the option that would be favoured by the NHS as it is less costly yet delivers the same benefit. Therefore the results submitted by the manufacturer for this scenario are summarised in the table below:

From the submission:

	Mean costs	Difference	Mean life years	Difference	Mean QALYs	Difference	ICER	PSA results £20 000 threshold	PSA results £30 000 threshold
Combination therapies									
Certolizumab pegol + MTX	85 583		14.63		6.654				
Etanercept + MTX	86 165	-582	14.62	0.01	6.589	0.065	CZP dominates	70.1	65.7
Adalimumab + MTX	86 034	-451	14.59	0.04	6.412	0.242	CZP dominates	71.0	67.6
Rituximab + MTX	82 940	2643	14.58	0.05	6.362	0.292	9072	60.2	61.8
Infliximab + MTX	95 599	-10 016	14.55	0.08	6.196	0.458	CZP dominates	98.8	95.5
Monotherapies									
Certolizumab pegol	81 849		14.57		6.305				
Etanercept	85 941	-4092	14.60	-0.03	6.435	-0.130	31 582*	56.3	51.8
Adalimumab	84 201	-2352	14.54	0.03	6.090	0.215	CZP dominates	82.9	78.2

*Comparator was more costly and more effective than CZP + MTX

1.4 Commentary on the robustness of submitted evidence

The NICE Specification for Manufacturer/Sponsor Submission states “A submission should be as succinct and informative as possible. It is expected that the main body of the submission will not usually exceed 75 pages.” The submission for this STA was 190 pages long. Although the submission was detailed, a large number of elements required clarification. The clarification document subsequently received was 84 pages long and included revised economic models and analyses. The modified analyses encompassed some changes implemented in response to clarification issues, whilst some others followed from reassessments by the manufacturer independent of any clarification issue. After consultation with NICE it was agreed that the ERG would only appraise the revised economic analyses. In part, it was difficult to follow the narrative of the submitted evidence because the logic of argument was split between the two documents.

1.4.1 Strengths

It was clearly demonstrated in two adequately powered double blind placebo controlled RCTs of acceptable quality that combined therapy of CZP + MTX was more effective than MTX + placebo. The effectiveness of CZP in monotherapy relative to placebo was similarly demonstrated in a double blind RCT of reasonable quality.

The economic model supplied by the manufacturers was internally consistent and robust.

Clinical and cost data used to populate the model were taken from appropriate sources and clearly indicated in the written portion of the submission.

1.4.2 Weaknesses

In relation to the decision problem, previous NICE guidance[1] on the use of anti-TNF bDMARDs (infliximab, etanercept and adalimiumab) states that patients should have previously undergone trial of two DMARDs including MTX (unless contraindicated). Recruitment to the CZP trials did not necessarily only include patients satisfying this criterion. Although the average number of DMARDs previously experienced by patients entering the three CZP trials was greater than 2 it was unclear from the submission what proportion of participants fell short of two DMARD exposures. The influence of this on the estimates of effectiveness is difficult to judge.

A completed unpublished industry sponsored RCT of CZP + MTX versus MTX + placebo (C87014) was not included in the review of clinical effectiveness. This maybe because unpublished studies were excluded. However data from this study was encompassed in the evidence presented about safety of CZP. The specific findings of this trial are unknown but probably relevant to the decision problem.

There were no RCTs of CZP against any other bDMARDs. Due to this lack of relevant clinical evidence, indirect comparisons were conducted for CZP versus TNF-α inhibitors (etanercept, inflximab and adalimumab), rituximab (anti-CD20 monoclonal-antibody) and tocilizumab (IL6-receptor binding monoclonal-antibody). The indirect comparison meta-analyses (ICMs) were conducted so that only multiple pair-wise comparisons could be made. The results of the ICM did not directly inform the effectiveness inputs into the submitted economic model.

There were several concerns regarding the methodology adopted for the ICM and how evidence was subsequently incorporated into the economic analysis:

- The inclusion and exclusion of studies for the ICM did not appear to be systematic.
- The inclusion of data from the included studies lacked consistency.
- There was a possibility that relevant information from several excluded studies, including one of CZP + MTX versus MTX + placebo, could have been used in the ICM.
- There was insufficient consideration and exploration of underlying heterogeneity amongst the studies included for ICM.
- The development of effectiveness input for the economic analysis included data for a bDMARD comparator omitted from the subsequent economic analysis, raising the issue of whether data for other omitted bDMARDs should not also have been included.
- The development of effectiveness input for the economic analysis sacrificed the strengths of randomisation and underestimated the associated uncertainty.

It was noted that, of the 10 trials used in manufacturer's indirect comparison meta-analyses for combination therapy at six months, the two CZP trials had a low mean number of previous DMARDs exposures compared to trials of

comparator bDMARDs. Furthermore, the mean MTX dosage at entry in the trials reporting 6 month follow up was lowest in the CZP trials.

Regarding the economic model data on quality of life and health utilities was not as robust as the effectiveness and cost data used. The exact details of how many patients were given health related quality of life questionnaires and response rates to the questionnaires were not provided in the submission. There is a degree of uncertainty in the final EQ-5D estimates and utility after first line treatment. It is not clear how this uncertainty might affect the estimates of cost-effectiveness that are generated.

The adverse event costs as well as related health outcomes were not included in the revised model though they were included in the original submission. It is possible that further information could have been provided to justify this revision (initially differential adverse event effects between comparator therapies were considered). It is unclear what sources of information were used in this exercise and if the assumption is reliable.

The working model supplied to the ERG lacked some functionalities which precluded validation of some aspects of the submitted economic results.

1.4.3 Areas of uncertainty

There is uncertainty about the clinical effectiveness and cost-effectiveness of CZP relative to other bDMARDs. This stems from:

- The lack of head to head trials
- The potential lack of exchangeability amongst the trials for the indirect comparisons of combination therapies (bDMARD + MTX versus MTX + placebo).
- Uncertainty about the appropriateness of the studies included in the indirect comparisons for combination therapy.
- The underestimate of uncertainty in parameters of clinical effectiveness that fed into the manufacturer’s economic model

- The lack of RCT evidence about CZP monotherapy used at the dose anticipated in the pending licence (licensed maintenance dose likely 200mg every 2 weeks, trial dose 400mg every 4 weeks).

1.5 Key issues

Where available the ERG usually consult the scientific discussion document produced by the EMEA. This document is a detailed source of evidence underpinning the marketing approval. At the time of writing this report, the document pertaining to CZP was not yet publically available.

Previous NICE guidance (130 www.nice.org.uk/TA130) states:

"TNF-α inhibitors should normally be used in combination with methotrexate. Where a patient is intolerant of methotrexate or where methotrexate treatment is considered to be inappropriate, adalimumab and etanercept may be given as monotherapy."

Results published in 2007 (Hyrich et al[2]) for a large UK National Cohort may reflect current UK practice regarding combination therapy; this indicates that 3,770 (56%) of 6,739 patients that started anti-TNF treatment received concomitant MTX; of those administered adalimumab etanercept and infliximab 30%, 40% and 85% respectively received co-therapy with MTX.

From the submission:

> The tumour necrosis factor alpha (TNF-α) inhibitors adalimumab, etanercept and infliximab are recommended as options for the treatment of adults who have both of the following characteristics.
> • Active rheumatoid arthritis as measured by disease activity score (DAS28) greater than 5.1 confirmed on at least two occasions, 1 month apart.
> • Have undergone trials of two disease-modifying anti-rheumatic drugs (DMARDs), including methotrexate (unless contraindicated). A trial of a DMARD is defined as being normally of 6 months, with 2 months at standard dose, unless significant toxicity has limited the dose or duration of treatment.

The following issues are important for the decision problem

- Was the manufacturer's selection of comparators too narrow?

- Does the low rate of ACR response observed in the control arms of the CZP combination trials represent an issue of concern?

- The superior ACR effectiveness of CZP relative to comparators that was fed into the economic model was not well supported by submitted evidence and may have resulted in an overestimate of CZP cost-effectiveness.

- Eligibility criteria for recruitment in the FAST4WARD mono-therapy CZP trial specified that participants should have failed ≥ 1 DMARD but did not specify contra-indication or intolerance to MTX; the population therefore is not necessarily consistent with NICE guidance 130. The CZP maintenance dose in the trial (400mg every four weeks) did not correspond to the anticipated EMEA licensed maintenance dose (200mg every two weeks). However, a maintenance dose of 400mg every four weeks can be considered according to USA prescribing information.

- There are concerns regarding underestimation of uncertainty surrounding utility values entered into the economic model. These stem from the use of two mapping processes each associated with uncertainty that were not fully taken into account.

- Only pair-wise comparisons were available in the working model supplied to the ERG so that confirmation of the submitted cost-effectiveness frontiers was not possible.

- The revised model submitted by the manufacturer excluded any costs associated with adverse events; this departure from the original model structure was not implemented in response to any ERG request for clarification and might have favoured CZP relative to competitors.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The manufacturer's description was appropriate.

2.2 Critique of manufacturer's overview of current service provision

The manufacturer's description was appropriate.

3 Critique of manufacturer's definition of decision problem

3.1 The indicated scope.

The manufacturer's definition of the decision problem taken from the submission is shown below.

	Decision problem addressed in the submission
Population	Adults who have active, moderate-to-severe rheumatoid arthritis that has not responded adequately to DMARDs, including methotrexate.
Intervention	Certolizumab pegol
Comparator(s)	Comparisons to both conventional and biological DMARDs will be presented in the submission. The clinical effectiveness of certolizumab pegol will be directly compared to methotrexate based on the pivotal clinical trials. The clinical effectiveness of certolizumab pegol will be indirectly compared to adalimumab, etanercept, infliximab, rituximab and tocilizumab using methotrexate as a common comparator for combination therapies and placebo as a common comparator for monotherapies.
Outcomes	All outcomes as defined in the final scope will be presented, with the exception of extra-articular manifestations of disease, as the pivotal clinical trials did not routinely report this particular outcome. The outcome measures considered include: ▪ disease activity ▪ physical function ▪ joint damage ▪ pain ▪ mortality ▪ fatigue ▪ radiological progression ▪ adverse effects of treatment ▪ health related quality of life
Economic Analysis	The economic analysis will present a cost-utility analysis with cost-effectiveness expressed in terms of incremental cost per quality-adjusted life year. The cost utility of certolizumab pegol will be compared to the UK licensed and reimbursed interventions indicated for patients with active, moderate-to-severe RA that has not responded adequately or is intolerant to conventional DMARDs (adalimumab, etanercept, infliximab and rituximab).

	Decision problem addressed in the submission
	The time horizon for the economic evaluation will reflect the chronic nature of RA and be presented for a lifetime horizon. The analyses will be conducted in accordance with the NICE reference case for economic evaluation. Costs will be considered from a National Health Service and Personal Social Services perspective.
Subgroups to be considered	No subgroups were studied in the pivotal trials.
Special considerations, including issues related to equity or equality	Frequency of joint replacement or hospital admissions are not considered as the pivotal clinical trials did not routinely report these two outcomes. Evidence from RCTs suggests significant inhibition of structural damage to joints by 16 weeks. This may lead to fewer referrals for surgery, although this has yet to be proven in studies.

3.2 Population

The population of interest in the manufacturer's submission is consistent with the NICE scope and with the anticipated marketing authorisation as described in the EMEA Pre-Authorisation statement 25 June 2009.

3.3 Intervention

The submission is consistent with the NICE scope. The proposed course of treatment (MS page 4) is described as follows:

> The recommended starting dose for adults with RA is 400mg subcutaneously (s.c.) at weeks 0, 2 and 4, followed by a maintenance dose of 200mg s.c. every 2 weeks.

It is worth noting that the maintenance dose above (200mg every two weeks) is somewhat different to that in the only RCT providing evidence of clinical effectiveness of mono-therapy where maintenance was at 400mg every 4 weeks. In the USA the prescribing information states a maintenance dose of 400mg every four weeks can be considered.

3.4 Comparators

The comparators itemised in NICE and manufacturer's scope, taken from the submission, are shown below.

NICE	Manufacturer
Management strategies involving DMARDs without certolizumab pegol, including treatment with: ▪ conventional DMARDs (for example sulfasalazine, leflunomide) ▪ biological agents (including adalimumab, etanercept, infliximab, rituximab, tocilizumab)	Comparisons to both conventional and biological DMARDs will be presented in the submission. The clinical effectiveness of certolizumab pegol will be directly compared to methotrexate based on the pivotal clinical trials. The clinical effectiveness of certolizumab pegol will be indirectly compared to adalimumab, etanercept, infliximab, rituximab and tocilizumab using methotrexate as a common comparator for combination therapies and placebo as a common comparator for monotherapies.

In contrast to the manufacturer, the NICE scope allows for additional conventional DMARD comparators beyond MTX, and does not necessarily restrict bDMARD comparators to adalimumab, etanercept, infliximab, rituximab and tocilizumab; in other words other bDMARDs such as abatacept, anakinra, and ustekinumab might be comparators. No clear justification was presented in the submission for the apparent restriction of comparator bDMARDs; the MS (pg 6) comments that anakinra and abatacept are not recommended for combination therapy employing two anti-cytokine drugs of different mechanistic action, that abatacept is not recommended on health economic grounds (NICE), and that tocilizumab is not currently a treatment option in the UK (UK guideline[3]). Currently abatacept and anakinra are not approved by NICE for use in RA.

Since the scope specified that indirect comparison methodology was to be employed to estimate effectiveness of CZP relative to other bDMARDs the inclusion of additional bDMARDS could have been considered, however the narrower list presented might reasonably reflect current UK practice.

Further, the manufacturer considers combination therapy only in terms of bDMARD + MTX versus MTX whereas the NICE scope would allow for bDMARD + a conventional DMARD other than MTX. Of 6,739 UK patients that completed 6 months treatment with a bDMARD 69% received concomitant DMARD and 56% received concomitant MTX (Hyrich et al 2007[2]).

The manufacturer's comparators list for analysis of clinical effectiveness by indirect comparison to CZP includes tocilizumab and is not wholly consistent

with the comparisons listed for economic analysis which do not include tocilizumab (see below)

The economic analysis will present a cost-utility analysis with cost-effectiveness expressed in terms of incremental cost per quality-adjusted life year. The cost utility of certolizumab pegol will be compared to the UK licensed and reimbursed interventions indicated for patients with active, moderate-to-severe RA that has not responded adequately or is intolerant to conventional DMARDs (adalimumab, etanercept, infliximab and rituximab).

As effectiveness data from tocilizumab contributed input for the economic analysis while tocilzumab was not an economic comparator, an issue arises regarding data from studies of other bDMARDs that could have contributed input for the economic analysis but like tocilizumab were also not part of the economic decision problem.

3.5 Outcomes

Outcomes listed by the manufacturer coincide with those itemised in the NICE scope but with the addition of mortality.

The clinical effectiveness outcomes that fed the economic modelling were: health related quality of life and disease activity as measured according to ACR 20, ACR50 and ACR70 criteria. The EQ-5D results in the MS effectiveness section were presented using VAS scores, however EQ-5D elicited utilities were used in the cost-effectiveness section.

A description of the ACR outcome measure is given in Appendix 1 along with details of the Health Assessment Questionnaire Disability Index (HAQ-DI), the Disease Activity Score (DAS) and the Modified Total Sharp Score (mTSS) outcome measures.

3.6 Time frame

The time frame adopted is consistent with the NICE scope and attempts to extrapolate to a life time horizon.

3.7 Other relevant factors

From the submission:

Frequency of joint replacement or hospital admissions are not considered as the pivotal clinical trials did not routinely report these two outcomes. Evidence from RCTs suggests significant inhibition of structural damage to joints by 16 weeks. This may lead to fewer referrals for surgery, although this has yet to be proven in studies.

4 CLINICAL EFFECTIVENESS

4.1 Critique of manufacturer's approach

4.1.1 Description and critique of clinical effectiveness search strategy

The manufacturer conducted an initial search to July 2007; this was subsequently broadened and updated to April 2009 (see Appendix 2).

The stated aim of the search was to identify studies allowing systematic review of the efficacy of bDMARD treatments for rheumatoid arthritis (MS page 22).

> The identification of relevant studies was performed in a two step process. An initial systematic review of the efficacy of biological DMARD treatments for rheumatoid arthritis was conducted with databases searched up to July 2007.

Comment:

The above is a broader aim than implied in manufacturer's definition of the decision problem.

In effect the manufacturer's search strategy and selection of studies served three distinct purposes:

- The identification of appropriate studies to allow assessment of the clinical effectiveness of CZP + MTX versus MTX + placebo and of CZP versus placebo in patients intolerant of MTX or for who MTX was contra-indicated.
- The identification of appropriate studies to allow assessment of the effectiveness of CZP combination therapy (CZP + MTX), and of CZP mono-therapy relative to combination therapy and mono-therapy with other bDMARDs using indirect comparison methodology
- The identification of appropriate studies to inform the decision problem addressed in the economic analysis

The search strategy specifically named the bDMARDs that the manufacturer identified as comparators for the decision problem (i.e. etanercept, adalimumab, infliximab, rituximab and tocilizumab); it also included terms for anti-TNF agents and for monoclonal antibodies. Studies of some bDMARDs such as fusion protein anti-cytokine agents (e.g. abatacept and anakinra) may not be efficiently retrieved with this search strategy and therefore this search may not fulfil the broad aim stated in the MS (see above).

It is difficult to determine if the studies were retrieved because the literature search tree presented page 27 of the submission is confusing and an appropriate list of excluded studies was not provided (e.g. the “100 references excluded at second screening stage” were not itemised).

In summary the database strategies appear comprehensive and it is unlikely that studies of CZP and the specified comparators will have been missed.

Omission of foreign language papers

Although it was stated that searches were limited to articles published in English this does not appear to have been implemented in the search strategies (see submission section 10.2.4, pg 181). However the submission does state that non-English language papers were excluded (see submission section 10.2.6, pg 184). This procedure may exclude relevant studies. The ERG re-ran the MEDLINE and EMBASE strategies to test how many papers in the time period 2007-2009 would have been disregarded as a result of this exclusion criterion (see Appendix 3). The manufacturer’s response to ERG’s request for clarification on this issue is shown in Appendix 4.

144 foreign language papers (45 MEDLINE, 108 EMBASE, 9 duplicates) were identified and their publication languages were: 38% German, 18% Japanese and 15% Spanish, the remaining 27% relatively evenly spread between 10 other languages. The ERG did not assess whether any of these would have met all of the manufacturers’ other inclusion criteria.

Ongoing trials

The manufacturer's strategy did not include any details of any searches for ongoing trials, thus there is a possibility that studies which are ongoing or which are completed but not yet published may have been overlooked.

The ERG conducted searches of the following publicly available trials registers: ClinicalTrials.gov, Current Controlled Trials meta Register and the UKCRN Portfolio database, using terms cimzia or certolizumab combined with rheumatoid. Table 1 lists the ongoing studies of CZP identified. Of the ongoing studies, only one trial (C87051) was mentioned in MS ongoing studies section (section 6.2.5, pg 26).

Table 1 Studies identified by ERG searches for ongoing studies

Study Number	Status*	Intervention v Control	Design	Completion Date	Comments
C87011	completed	CZP v PL	RCT	Jul-04	FAST4WARD
C87027	completed	CZP+MTX v PL+MTX	RCT	Oct-06	RAPID 1
C87050	completed	CZP+MTX v PL+MTX	RCT	Sep-06	RAPID 2
C87014	completed	CZP+MTX v PL+MTX	RCT	Jan-04	not provided by UCB
C87015	ongoing, not recruiting	CZP	open-label extension of -014 or -011	Jun-11	FAST4WARD and ??? Extension; mentioned in MS
C87051	ongoing, not recruiting	CZP+MTX	open-label extension of -050	Jun-11	RAPID 2 extension; mentioned in MS
C87028	ongoing, not recruiting	CZP+MTX	open-label extension of -027	Jun-11	RAPID 1 extension; not mentioned in MS
C97076	recruiting	CZP v PL	RCT	Jun-11	CERTAIN; patients with moderate to low DAS; not mentioned in MS
C87077	recruiting	CZP+MTX v PL+MTX	RCT	Sep-10	study only in responders after a run-in on CZP; not mentioned in MS
C87080	enrolling patients only by invitation	CZP	open-label extension of -076	Apr-10	CERTAIN 2; completion date earlier than in main study; not mentioned in MS
C87094	recruiting	CZP v PL	RCT	Nov-10	not mentioned in MS
C87084	enrolling patients only by invitation	CZP+MTX	open-label extension of -077	Apr-11	not mentioned in MS
CDP970-275-08-001	recruiting	CZP+MTX v PL+MTX	RCT	Mar-11	Japanese; not mentioned in MS
CDP970-275-08-003	recruiting	CZP v PL	RCT	Mar-11	Japanese; not mentioned in MS
CDP970-275-08-002	not yet open to recruitment	CZP+MTX	open-label extension of -001	Mar-12	Japanese; not mentioned in MS
CDP970-275-08-004	not yet open to recruitment	CZP	open-label extension of -003	Mar-12	Japanese; not mentioned in MS
*Searches for ongoing studies will also identify some studies which are completed.					

Unpublished Completed RCT

Searches for CZP trials using ClinicalTrials.gov resulted in identification of a completed study (C87014) comparing CZP+MTX with Placebo+MTX. No published report appears to exist for this study and as a consequence it was not included in the manufacturer's effectiveness review as this only included published studies. The ERG conducted further searches for details of this trial (see Appendix 8). A summary of information on study C87014 obtained from these searches is presented in Table 2 below.

Table 2 Details available regarding study C87014

Study: C87014

Design: RCT
Double-blind (no further details provided)

Location: NA

Duration: 24 week follow-up; study conducted October 2002 – January 2004

Sponsor: UCB (according to www.clinicaltrials.gov)

Participant numbers:

	CZP 400mg + MTX	**MTX**
Eligible	NA	
Randomised	247	
Completed	NA	
Analysed	NA	

Inclusion criteria:
- age 18-75 years
- diagnosis of adult-onset RA
- active disease (moderate to severe)
- had received MTX for at least 6 months and have been on a stable dose of 15-25mg per week for at least 8 weeks prior to the first dose of CZP
- on a stable dose of folic acid

Exclusion criteria:
- contraindication for MTX or anti-TNF inhibitor
- other inflammatory arthritis outside RA
- evidence of latent TB (chest x-ray or PPD test)
- prior treatment with another TNF alpha inhibitor
- prior treatment with other experimental therapies, including biological therapies within 6 months prior to the first screening visit

Intervention: CZP 400mg every four weeks + MTX

Comparator: PL every four weeks + MTX

Other drugs:
- other DMARDs washed out at least 28 days prior to the first study date
- stable dose of prednisone up to 10mg/day allowed
- patients allowed to have concomitant insets, including COX-II; must have been stable on these drugs for at least 4 weeks prior to the study

Primary outcome:
- ACR 20 response rate at week 24

Secondary outcomes:
- Safety and tolerability
- Health outcomes measures
- Immunogenic profile of CZP plus MTX
- Systemic exposure of CZP

Assessment time:
- Responder rates measured at week 1, 2, 4 and every 4 weeks thereafter up to week 24

Sample size calculation: NA

Analysis methods: NA

Results:
the study met its primary endpoint (proportion of patients achieving ACR 20 response)

Study C87014 is mentioned in the submission on page 33 where it is stated to be a previous study upon which the power calculation for RAPID1 and RAPID2 trials was based. From the submission:

> In a previous study (C87014), the observed placebo response at week 24 was 26.9%. A placebo + MTX ACR response of 30% was therefore assumed,

In the CZP USA Prescribing Information this trial is listed as one of the four CZP RCTs conducted in RA patients. As such it was presumably submitted to the FDA during the USA marketing approval process (see Table 3). As the EMEA scientific discussion on CZP for RA is not yet public it is unclear if the trial informed these discussions.

Table 3 Map of CZP trials contained within various documents relevant to the decision problem

Trial / Document	RAPID 1 (combination)	RAPID 2 (combination)	C87014 (combination)	FAST4WARD (monotherapy)
Manufacturer submission				
SmPC (as supplied to NICE/ERG)				
FDA approval			?	
EMEA	?	?	?	?

This study was also mentioned in the adverse events section of the submission.

Finally, with 247 randomised patients, C87014 is equivalent to about 15% of patients in RAPID 1 & 2 (also on CZP+MTX vs MTX+placebo).

4.1.2 Statement of the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate.

The inclusion and exclusion criteria specified in the submission are shown below.

Inclusion criteria were as follows:

- Studies must be published randomised controlled trials.
- Studies must be conducted in human adult patients (≥18 years) with active RA who have had an inadequate response to prior DMARD therapy, including MTX.
- Studies must contain one of the following interventions administered with or without MTX: adalimumab, CZP, etanercept, infliximab, rituximab, tocilizumab.
- The treatment comparison must be to another biological DMARD, a conventional DMARD or placebo.
- The study must report results for an outcome of interest, namely: patient's global assessment of disease activity; DAS28 scores; ACR responses; EULAR responses; HAQ-DI; swollen joint count; tender joint count; patient's global assessment of pain; patient's assessment of fatigue; mTSS (van der Heijde); EQ-5D scores; mortality; hospitalisation; joint replacement surgery; rheumatoid nodules; vasculitis; neuropathy; any adverse event; any serious adverse event; infusion reactions/acute hypersensitivity reactions; injection site reactions; tuberculosis; serious infections; malignancies/lymphoma; withdrawals due to adverse events.
- The interventions of interest must be studied at a licensed dose (EMEA or FDA). If a study included more than one treatment arm of the intervention of interest, one of them must be a licensed dose.

The following studies were excluded:

- Observational studies, reviews and open label extensions of RCTs
- Non-English language papers
- Dose ranging studies and studies where the comparator was an untreated control group

Comment:

1. It is not clear if these criteria apply for the direct comparisons undertaken or the indirect comparisons or both. Preferably explicit separate sets of criteria are desirable.

2. The bDMARDs for inclusion are consistent with the manufacturer's definition of the decision problem, and those listed in the NICE scope. However, the NICE scope could be considered not to limit comparators to these named bDMARDs.

3. The treatment comparator may be “a conventional DMARD” (by implication not necessarily MTX) yet the intervention appears limited to studies in which a named bDMARD is “administered with or without MTX”. This would rule out studies of bDMARDs in combination with other “conventional DMARDs”. Thus, in the context of the decision problem, there appears to be a slight misbalance between the intervention and comparator arm options.

4.1.3 Table of identified studies. What studies were included in the submission and what were excluded.

NICE requests that the submission: “provide a list of all RCTs that compare intervention with other therapies (including placebo)”

The submission lists 37 studies (table 1 page 23); three of these are CZP (± MTX) versus other therapies (placebo ± MTX). The others are studies of other bDMARDs and are presumably relevant to the indirect comparisons. However some of these included studies were subsequently excluded from the indirect comparisons. Therefore, there is some confusion in the submission about what studies are “included” and for which purpose.

As previously mentioned the flow diagram in the submission (MS figure 3 pg27) lacks clarity; it is difficult to understand the status of the 135 references designated as “suitable for data extraction” or the relevance of “37 primary RCTs” to the direct and indirect comparisons undertaken.

It is clear that three CZP RCTs met the inclusion criteria.

4.1.4 Details of any relevant studies that were not included in the submission?

The ERG identified an unpublished RCT on CZP 400mg every four weeks in combination with MTX (study C87014). See section 4.1.1 for details of this study.

Non-English language papers and unpublished studies were excluded.

Other than the trial mentioned above it is probably unlikely any CZP studies were omitted.

Although unlikely, it is not clear if studies relevant to the ICM as conducted by the manufacturer were omitted, as a list of all potentially relevant papers that were excluded (with reasons for exclusion) was not supplied (e.g. in Fig3 in the submission pg 27 details on the “100 references excluded at the second screening stage” are not given).

4.1.5 Description and critique of manufacturers approach to validity assessment

Only RCTs were included in the manufacturers review. The manufacturer applied clear and appropriate criteria to assess the validity of the three included studies on CZP. These criteria were:

- Method of concealing treatment allocation
- Randomisation technique
- Sample size justified?
- Observers blinded?
- Study design
- Location
- Comparability of patients with UK RA population
- CZP dosage regimen
- Dosage regimen within SmPC?
- Study groups comparable?
- Statistical analyses appropriate?
- ITT analyses performed?
- Confounding factors affecting interpretation present?

The results of the quality assessment were tabulated and not used in any analyses. The ERG considers the quality assessment to be fair and reasonable. However, the ERG does have some concern about a feature of the study design rules for early withdrawal of patients (see section 4.2).

Quality assessment was not undertaken for trials on comparator drugs included in the indirect comparisons.

4.1.6 Description and critique of manufacturers outcome selection

In the inclusion/exclusion criteria a list of outcomes of interest was provided. The choice of outcomes seems to be appropriate.

The submission concentrated mainly on ACR responses which were later used for the economic model. The choice of this outcome seems justified as it measures clinical improvement and data from trials is available.

Given the importance of utility measures for economic modelling, it is surprising that only EQ-5D VAS results were presented in the effectiveness section of the submission. A comparison of EQ-5D scores between different bDMARDs was not undertaken.

4.1.7 Describe and critique the statistical approach used

For the direct and indirect comparisons the approach appears appropriate.

However the approach for obtaining effectiveness inputs for the economic modelling for CZP compared to other bDMARDs had some shortcomings. See section 4.2.3.1 for further details.

4.1.8 Summary statement

The manufacturer adopted a generally systematic and robust approach however there were concerns regarding:

- The lack of information about the completed RCT of CZP + MTX versus MTX + placebo (study C87104)
- The exclusion of published studies reported in foreign languages

- Some lack of clarity in the flow of studies from the searches to inclusion for the direct and indirect comparisons that were undertaken

- The statistical approach to obtaining effectiveness inputs for the economic modelling

4.2 Summary of submitted evidence

The manufacturer submitted evidence from three adequately powered double blind CZP-RCTs of acceptable quality. There were two RCTs of combination-therapy that directly compared CZP + MTX versus MTX + placebo (RAPID1 and RAPID2) and one mono-therapy RCT (FAST4WARD) that directly compared CZP versus placebo. Meta-analyses of the two combination therapy RCTs were undertaken for several outcomes.

Details of the CZP trials are summarised in Table 6, Table 7 and Table 8. Given the size of these tables they have been placed in Appendix 5.

An important aspect of study design was the withdrawal at week 16 of patients not achieving ACR20 (week 12 in FAST4WARD). This resulted in more than half the patients in the control arms of the trials being withdrawn at this time. In effect all intervention versus control comparisons by ITT made at times after 16 weeks required some imputation for missing data. Withdrawal rates were strikingly greater in control arms (63-81%) than intervention arms (17-21%) of the trials so that any bias resulting from imputation would have a greater effect on control results.

The submission presented results of indirect meta-analyses of combination therapy and of mono-therapy in which CZP was compared to several bDMARDs with respect to ACR outcomes. These ICMs encompassed results taken from 18 RCTs of which three were the CZP trials mentioned above.

The sections below summarise the results of the direct and indirect comparisons.

4.2.1 Summary of direct comparison results

Many direct comparison results were presented in the submission extending from pages 36 to 72 (with safety outcomes reported mainly in an additional section of the MS).

The primary outcome(s) in each of the three CZP RCTs were stated in the submission. These were (from MS table 6 page 32):

	RAPID 1	RAPID 2	FAST4WARD
Primary outcome(s)	• ACR20 at wk 24 • Prevention of structural damage (mTSS at week 52)	• ACR20 at wk 24	• ACR20 at wk 24

Secondary outcomes were as follows (from table 6 page 32 of the submission):

	RAPID 1	RAPID 2	FAST4WARD
Secondary outcomes	• ACR20 at wk 52 • ACR50/70 at wks 24/52 • DAS28 at wks 24/52 • mTSS at wk 24 • Physical function • HRQoL, fatigue and productivity • PK profile and immunogenicity of CZP • Safety and tolerability of CZP	• ACR50/70 at wk 24 • mTSS at wk 24 • DAS28 • HAQ-DI at wk 24 • SF-36 • HRQoL, fatigue and productivity • PK profile and immunogenicity of CZP • Safety and tolerability of CZP	• ACR50/70 • DAS-ESR-3 • Pain • Physical function • HR-QoL, fatigue and productivity • Systemic exposure and immunogenicity of CZP • Safety and tolerability of CZP

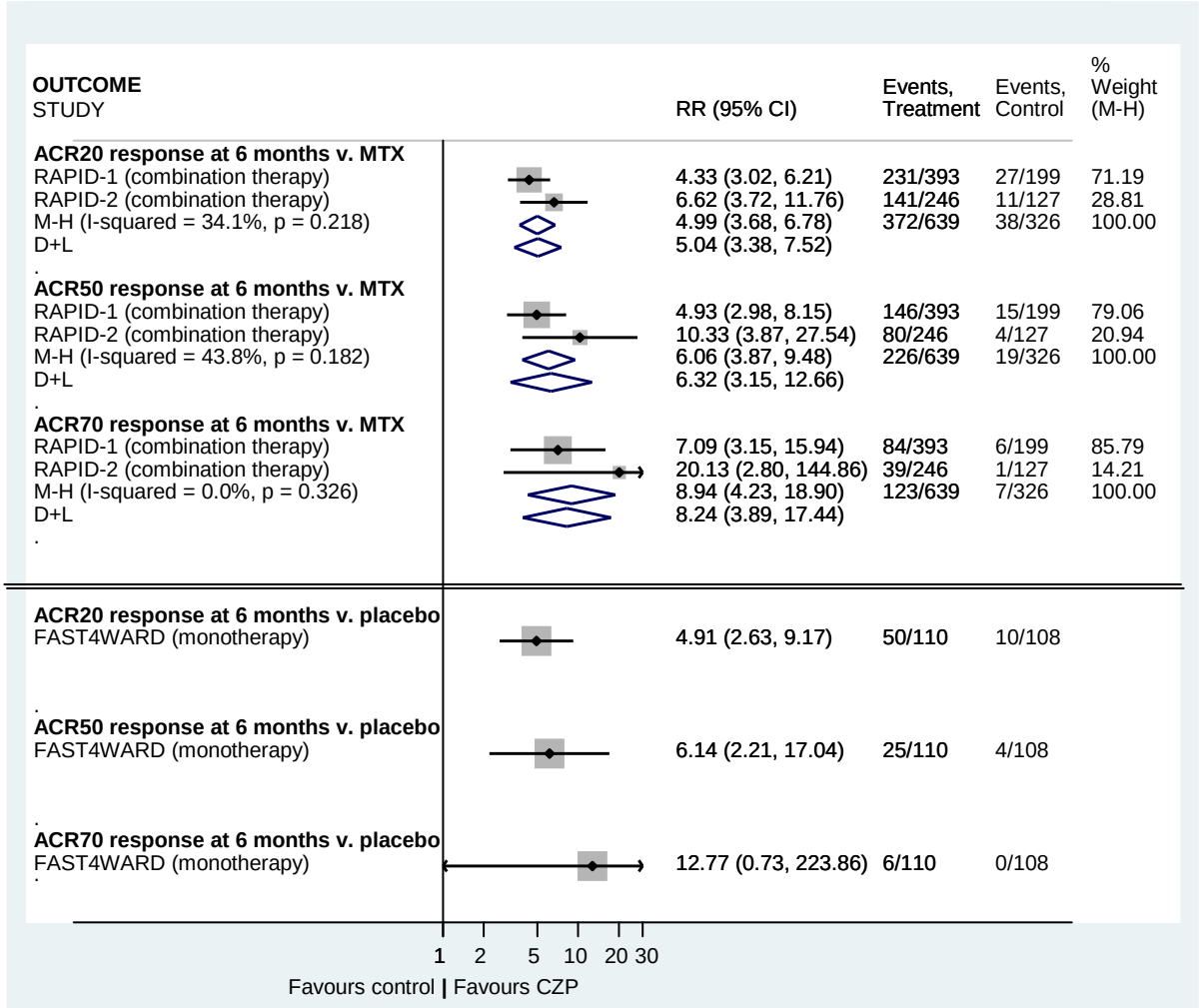
<u>ACR measures</u>

A main emphasis in the MS was on trial results according to ACR measures and ACR20 ACR50 and ACR70 criteria. Table 9 in Appendix 6 is an itinerary of the multiple ACR measures.

Overall, the ACR results demonstrated a superiority of CZP versus control for both combination therapy and monotherapy and in most instances the comparison reached statistical significance in favour of CZP. In addition the results indicated that CZP elicits a rapid response that is sustained reasonably well to week 52. There was evidence of improvement in all the core ACR components.

For combination therapy the submission provided forest plots of meta-analyses for ACR outcomes at six months in RAPID1 and RAPID2 (MS Figures 27 to 29). These were sometimes confusing (for example in figure 27 the event column for the MTX arm was labelled “Treatment” and the event column for the CZP + MTX arm was labelled “Control”). It was difficult to see if the results supported greater effectiveness for CZP + MTX or for MTX and the accompanying table of results (MS table 20 page 66) compared CZP + MTX versus MTX while the forest plots compared to MTX versus CZP + MTX. To provide a summary and improve clarity the ERG prepared a single forest plot (Figure 1) to summarise ACR results for both combination and mono-therapy.

Figure 1 ACR outcomes at 6 months: CZP combination therapy (meta-analysis) & mono-therapy



Disease remission according to DAS-28 (ESR) score.

Remission was defined as a DAS28 (ESR) score of < 2.6. At baseline the mean DAS28 (ESR) score was 6.3 (FAST4WARD) 6.8 to 6.85 (RAPID1, depending on trial arm) and in RAPID2 median score ranged 6.9 to 7 depending on treatment arm.

A small proportion of patients achieved remission in RAPID1 and RAPID2 trials; a statistically significant greater proportion achieved remission in the CZP + MTX arms than in the MTX + placebo arms. The results from the submission (table 12, pg 43) are shown below.

	RAPID 1			RAPID 2		
	Placebo + MTX	**CZP 200 mg + MTX**	**CZP 400 mg + MTX**	**Placebo + MTX**	**CZP 200 mg + MTX**	**CZP 400 mg + MTX**
DAS28 remission, wk 24						
No. of patients	196	391	387	125	245	246
No. achieving remission (%)	3 (1.5%)	45 (11.5%)	50 (12.9%)	1 (0.8%)	23 (9.4%)	21 (8.5%)
Odds ratio (vs. placebo)		8.7	9.8		12.550	11.807
95% CI		2.6, 28.4	3.0, 32.1		1.671, 94.265	1.566, 89.043
P value		<0.001	<0.001		0.014	0.017
DAS remission, wk 52						
No. achieving remission (%)	3 (1.5%)	62 (15.9%)	74 (19.1%)			
Odds ratio (vs. placebo)		12.5	15.6			
95% CI		3.8, 40.3	4.8, 50.4			
P value		<0.001	<0.001			

ITT population, LOCF

EULAR responses were also reported (see pg 44 MS).

Radiographic progression of joint damage

In RAPID1 and RAPID2 radiographs of hands and feet were collected at baseline, at withdrawal (for early withdrawals) and at week 24 and week 52 (RAPID1 only). For each radiograph the mean score from two independent observers was used; a Kappa score for extent of agreement between observers was not published. The mean change from baseline in erosion score (ES, maximum score 280), joint space narrowing (JSN) score (maximum score 168), and modified Total Sharp Score (mTSS, maximum score 448) were the outcome statistics. Since more than half of control patients were withdrawn at 16 weeks, the estimates for these patients (and for

other withdrawals) were obtained by linear extrapolation to weeks 24 and 52. Sensitivity analyses were undertaken.

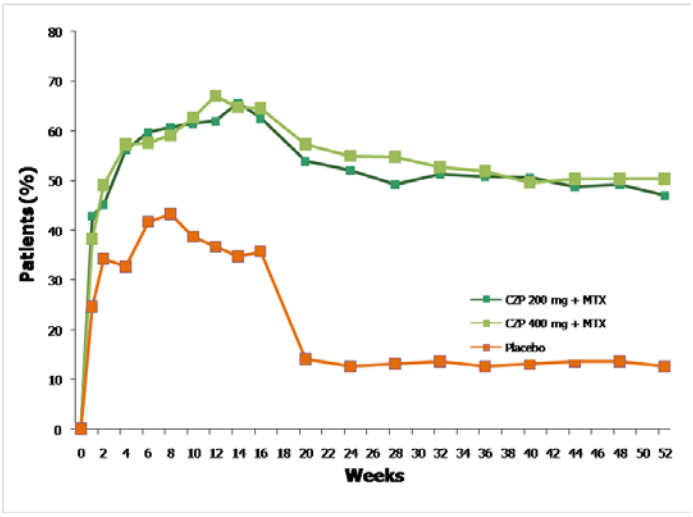
Statistically significant less progression was reported for intervention groups (at both dose regimens) according to all three scores in both trials at 24 weeks and for RAPID1at 52 weeks. The results were reasonably robust in sensitivity analysis. However, the difference in change (CZP v. control) were relatively small; for example in RAPID1 over a period of 24 weeks the difference in mean change from baseline between intervention (200mg regimen) and control was -1 for mTSS, -0.6 for ES, and -0.4 for JSN; these were from baseline scores of 39 and 38.4 (mTSS), 14.3 and 14.9 (ES), and 24 and 24.6 (JSN) for control and intervention (200mg regimen) respectively.

Physical function, pain and fatigue (HAQ-DI and PAAP scores)

In the three trials all arms showed a reduction (improvement) from baseline in mean HAQ score (total range 0 to 3, lower score worse condition). The reduction was statistically significantly greater for the intervention than the control groups; e.g. the difference from control at week 24 was 0.41 and 0.43 for 200mg and 400mg groups in RAPID1 and correspondingly by 0.35 and 0.36 in RAPID2. Essentially similar results were reported at week 52 for RAPID1 and at week 24 for the monotherapy FAST4WARD trial.

The submission also presented results for the proportion of patients achieving a minimum clinically important difference (MCID) from baseline, defined as a score improvement of ≥ 0.22. At 24 weeks in RAPID1 13%, 47% and 50% of patients in control, 200mg and 400mg arms achieved MCID. Similar results were reported at 24 weeks for RAPID2 and FAST4WARD and for 52 weeks in RAPID1. Figure 15 in the MS (reproduced below) is a plot of % MCID against week of study; in this graph a very considerable drop in the control arm occurs at the time of withdrawal according to ACR20 response. This is difficult to interpret.

Figure 15 RAPID 1 HAQ-DI MCID responders over time (ITT population)



Patient's Assessment of Arthritis Pain – VAS mean change in score from baseline results indicated superiority of CZP to comparator in all three trials. Similar beneficial results were reported for fatigue measures.

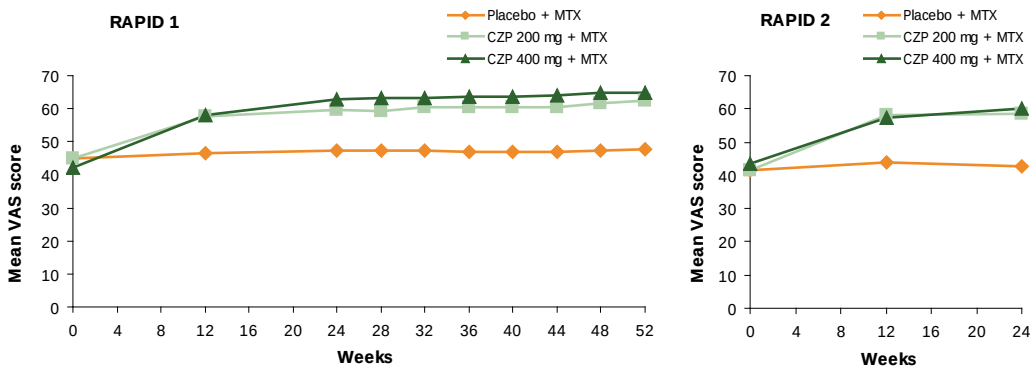
Quality of life

For RAPID1 the submission included results by eight SF36 domains for the adjusted mean change at week 52 relative to baseline. These indicated a greater favourable response for both intervention arms than for MTX+ placebo, and the difference reached statistical significance. Similar results were obtained at week 24 in RAPID2.

The % of patients achieving a MCID relative to baseline was also reported. Again CZP was clearly superior to control. At week 24, compared to control, CZP in combination or monotherapy was statistically significantly superior for the proportion of patients achieving a MCID improvement from baseline. The number of patients contributing SF36 data was not clear in the submission.

At European sites in RAPID1 and RAPID2 EQ-5D was assessed. The submission presented the improvements observed according to VAS results by visit. The number of patients contributing data was not indicated. The graphs taken from the submission are shown below.

Figure 21 Summary of EQ-5D health state (VAS) by visit – LOCF (ITT population)



Safety

For assessment of safety, the submission used evidence from a wider selection of CZP studies that the three included RCTs. Thirty one clinical studies were used.

The submission presented data derived from a total duration of exposure to CZP of > 4000 patient years with median duration of exposure approximately 1.7 years. Events were categorised as: adverse events, serious adverse events, infections, malignancies, deaths, and injection site reactions.

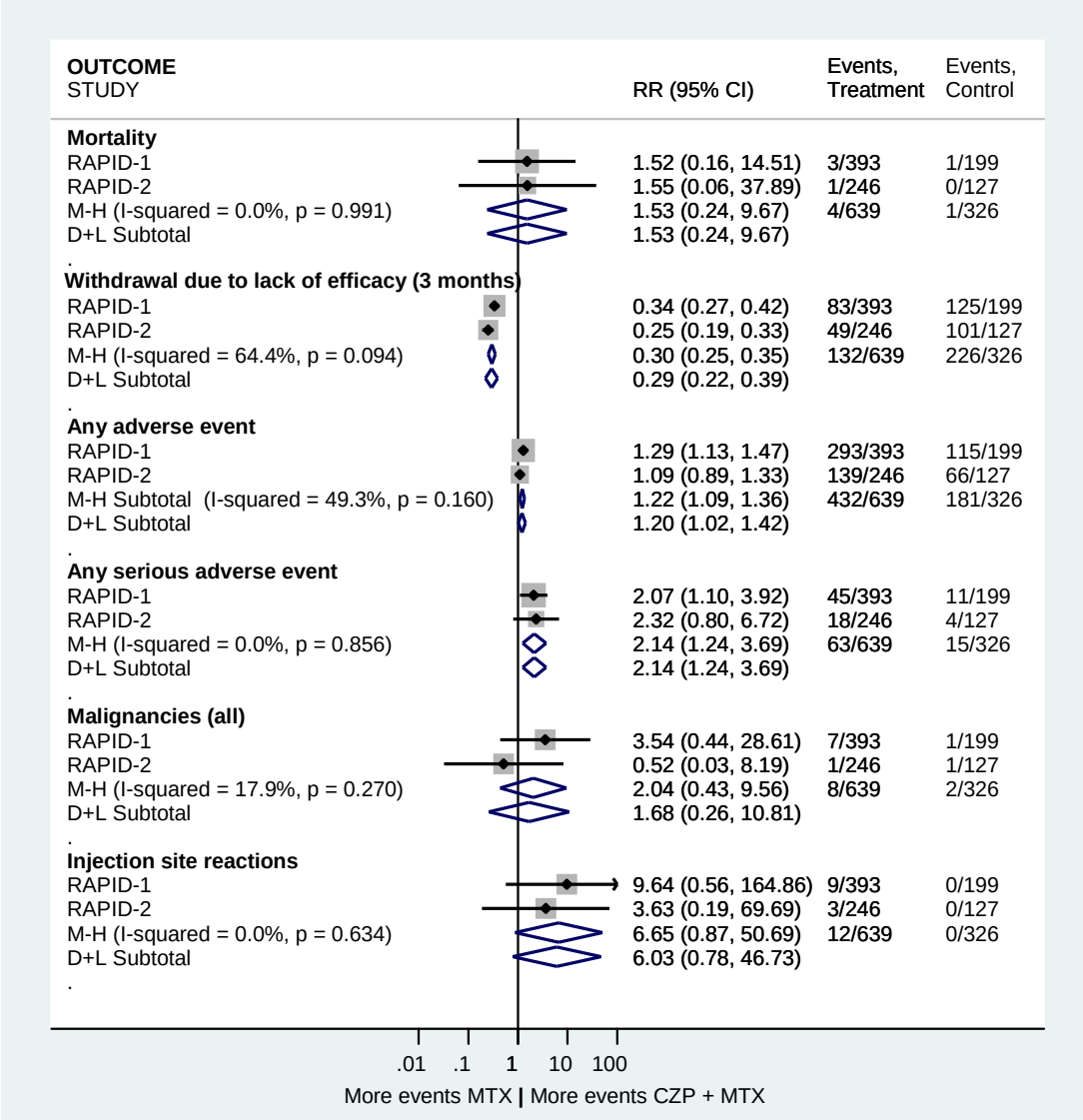
The manufacturer commented that the pattern of malignancies and deaths was “consistent with that seen in anti-TNF-α-treated population”.

Adverse events (occurring in ≥ 5% patients, data from placebo controlled trials) were various and generally more common in intervention than placebo arms. The most common amongst many AEs were infections and infestations (37% of CZP-treated patients compared with 29% of control patients) and musculoskeletal and connective tissue disorders (equally common in CZP and placebo patients). These were also the most common serious adverse events (infections and infestations: 3.5% of CZP patients and 0.6% for placebo; musculoskeletal and connective tissue disorders 1.9% and 1.7% for CZP and

placebo patients respectively). The incidence rate (cases per 100 patient years) for infections and infestations was 6.34 for CZP and 1.65 for placebo. The rate for tuberculosis was 0.9 for CZP and 0 for placebo.

The incidence of adverse events was quite consistent between placebo-controlled trials and open label trials (MS Table 42 pg 96). The submission contained meta-analyses of safety outcomes observed in the RAPID1 and RAPID2 trials (MS figures 30 to 35). For convenience the ERG has condensed these into a single forest plot (Figure 2).

Figure 2 Meta-analysis of relative risk of outcomes at 6 months CZP+MTX vs MTX



Serious adverse events associated with anti-TNF therapies.

The submission's summary of effectiveness evidence (section 6.9) does not contain a statement referring to safety / adverse events but concentrates on clinical benefit only.

Anti-TNF therapies have now been licensed for multiple indications and data about rare serious adverse events have gradually accumulated. In this section the relevant safety issues following from this data are briefly reviewed.

Bongertz et al 2006[4] meta-analysed rates of malignancy and of serious infection reported in placebo controlled RCTs of infliximab and adalimumab in rheumatoid arthritis. Information in published papers and from the US FDA website was used for the analysis. Odds ratios (anti-TNF versus placebo) for malignancy and for infection were 3.3 (95%CI: 1.2 to 9.1) and 2 (95% CI: 1.2 to 3.1) respectively and the numbers needed to harm were 154 and 59 respectively over a treatment period of 3 to 12 months. Higher drug doses were associated with greater risk. Similar results were reported in a meta-analysis conducted by Shoor 2006.[5] TNF-α has an important role in the host immune response to *Mycobacterium tuberculosis* and in the immunopathology of tuberculosis. Patients to be treated with anti-TNF agents should be screened for tuberculosis before starting anti-TNF therapy, they should be monitored for tuberculosis during therapy and those with latent tuberculosis should be appropriately treated prior to initiation of anti- TNF therapy. A recent publication (Raval 2007[6]) detailed 130 infliximab-associated cases of tuberculosis spontaneously reported to the US FDA between November 1 2001 and May 30 2006. In 45% of cases there was extrapulmonary disease. In a subset of 67 cases notified after the addition of a tuberculosis warning to the boxed medication it was noted that in six instances no test had been performed and that of 47 tuberculin skin tests performed 34 gave a negative result. The false negative rate was unknown. These results emphasise the requirement for vigilance by physicians caring for patients treated or about to be treated with anti-TNF therapies. Ramos-Casals et al 2007 identified 233

cases of autoimmune disease apparently associated with anti-TNF therapies.[7] Of these 17 occurred in CD patients. Anti-TNF agents infliximab, adalimumab and etanercept were associated with various autoimmune manifestations including lupus, vasculitis, and interstitial lung diseases. Overall incidence rates or rates for individual anti-TNF agents are unknown. Elevated TNF-α is associated with heart failure and its level is correlated with severity of heart failure. Case reports (n=47) reviewed by Kwon et al 2003[8] indicate that anti-TNF therapy might trigger new onset heart failure in a subset of patients and might exacerbate the condition of some patients. The SPCs carry warning of this potential risk. Treatment with monoclonal antibodies has been associated with potentially fatal induction of progressive multifocal leuko-encephalopathy. A recent systematic review of primary data by Socal et al (2008) identified 29 cases most of which (n=23) were associated with Rituximab therapy which depletes the B cell population.[9] The single instance associated with anti-TNF treatment was reported for a 74 year old woman given etanercept for rheumatoid arthritis.

4.2.2 Summary of indirect comparison results

A total of 18 RCTs contributed data for the ICs (13 for combination therapy and 5 for mono-therapy). Three outcomes were analysed (ACR20, ACR50 and ACR70) at each of 3 months and 6 months. A random effects meta-analysis was used that generated odds ratios and 95% confidence intervals for pair-wise comparisons between bDMARDs (between CZP, adalimumab, infliximab, etanercept, rituximab and tocilizumab for combination therapy and between CZP, adalimumab and etanercept for monotherapy).

The results are presented in Appendix 7:

- For combination therapy at ACR response at 6 months in Table 12, Table 13 and Table 14

- For combination therapy at ACR response at 6 months in Table 15 Table 16 and Table 17

- Corresponding results for mono-therapy are in Table 18 to Table 23

Brief summary of manufacturer's IC results

The submission reported that point estimates for pair-wise indirect comparisons between different bDMARDs were associated with wide confidence intervals. For example, amongst 42 comparisons at 6 months for combination therapy (bDMARD + MTX) only four comparisons provided statistically significant results (CZP better than infliximab and tocilizumab for ACR20, and CZP and tocilizumab superior to infliximab for ACR70). No combination therapy ICs at 3 months and none of the mono-therapy ICs reached statistical significance.

4.2.3 Critique of submitted evidence syntheses

The ERG undertook a formal critical appraisal of the manufacturer's systematic review of the evidence underpinning the effectiveness and safety of CZP used in combination therapy with MTX or as mono-therapy (see Appendix 7).

4.2.3.1 Indirect comparisons

Description of manufacturer's submission

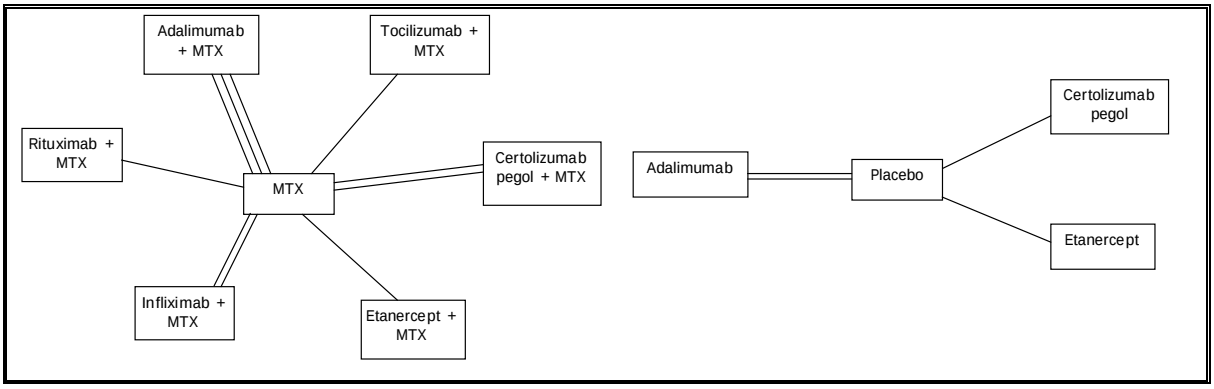
The submission used the indirect comparison (IC) method of Bucher et al[10] rather than a mixed treatment comparison (MTC) to compare efficacy of CTZ versus other bDMARDs. This choice was justified as follows:

> An indirect analysis method was chosen rather than a mixed treatment comparison due to the fact that most data came from trials of biological DMARDs versus MTX or placebo, forming star shaped network diagrams (figures 37 to 40). There were two direct trials of etanercept plus MTX versus infliximab plus MTX identified in the systematic review, however these were excluded from the indirect analysis as they contained small patient populations.

Comment

A potential disadvantage of this approach is that only multiple two way comparisons between bDMARDs is possible. The alternative MTC approach has the potential advantage of allowing direct probability statements about which treatment is the most effective, even when standard methods might determine no significant difference between treatments.[11]

The manufacturer's network diagrams for outcomes at 6 months for combination therapy (left) and for mono-therapy (right) are shown below:



Reproducibility of manufacturers IC results

The ERG used the manufacturer's input data and was able to reproduce the results reported in the submission.

Identification and selection of studies

The method for selecting the studies for inclusion in the IC was not totally explicitly reported. The literature search tree illustrating the identification and selection of studies included in the systematic review did not document the 18 included studies (submission Fig 3 page 27).

As such it was not possible for the ERG to determine if other potentially relevant studies were excluded.

The submission excluded two studies (De Filippis 2006 and Furst 2007[12,13]) from the IC due to small sample sizes. The sample size cut-off for excluding studies was not stated.

The ERG became aware of one study (LITHE, a trial of tocilizumab in patients with inadequate response to MTX by Kremer et al was presented at EULAR 2009) that apparently satisfied inclusion criteria used for the IC but which did not appear in either the submission inclusion or exclusion lists; availability of this conference proceeding probably post-dated the end of the search.

Data selection from included studies

The IC included unpublished data from the CZP studies but no unpublished results were included from studies of the comparator bDMARDs. It was unclear whether such unpublished data was not sought or was assumed unavailable.

Several of the publications of studies of combination therapy provided graphs showing ACR20 ACR50 and ACR70 outcome results at 3 months but no numerical statement of this data. These studies included the CZP trials (RAPID1 and RAPID2), Kim 2007[14], Keystone 2004[15], ARMADA, ATTEST and OPTION. The RAPID1 and RAPID2 trial data was included for the manufacturers IC analysis of effectiveness at three months, but the other studies were not included. The CZP trial results at three months were obtained from the unpublished full trial reports. Inclusion of CZP results but omission of data for five of the trials represents an inconsistency in data selection/ extraction. Of 12 trials with available data only 7 were used in the IC introducing a potential for bias into the analysis.

Similarly for the manufacturer's mono-therapy IC the publication of the van de Putte 2004 study (adalimumab versus placebo) included graphical data for ACR20 ACR50 and ACR70 results at both 12 and 24 weeks; however the 12 week data was not included in the IC for 3 months. In contrast, 12 week data from the CZP study (FAST4WARD, referenced as abstract Fleischmann et al 2008[16]) was included, and was obtained from the unpublished full study report.

Study details and quality assessment

Except for the three CZP trials the MS did not present any quality assessment of the studies included for ICs but did present some study details covering the following aspects:

- Trial design – all studies included were RCTs
- Intervention and control treatments were listed
- Patient numbers (eligible, randomised, allocated, drop-outs) – numbers of participants in each trial arm were provided, but no further details such as number of drop out were listed.
- Outcome – a summary of ACR20, ACR50, ACR70 were provided

A summary of characteristics of trial participants was not provided.

Although not documented in the submission all of the studies in the IC were sponsored by manufacturers of bDMARDs.

Heterogeneity of included studies

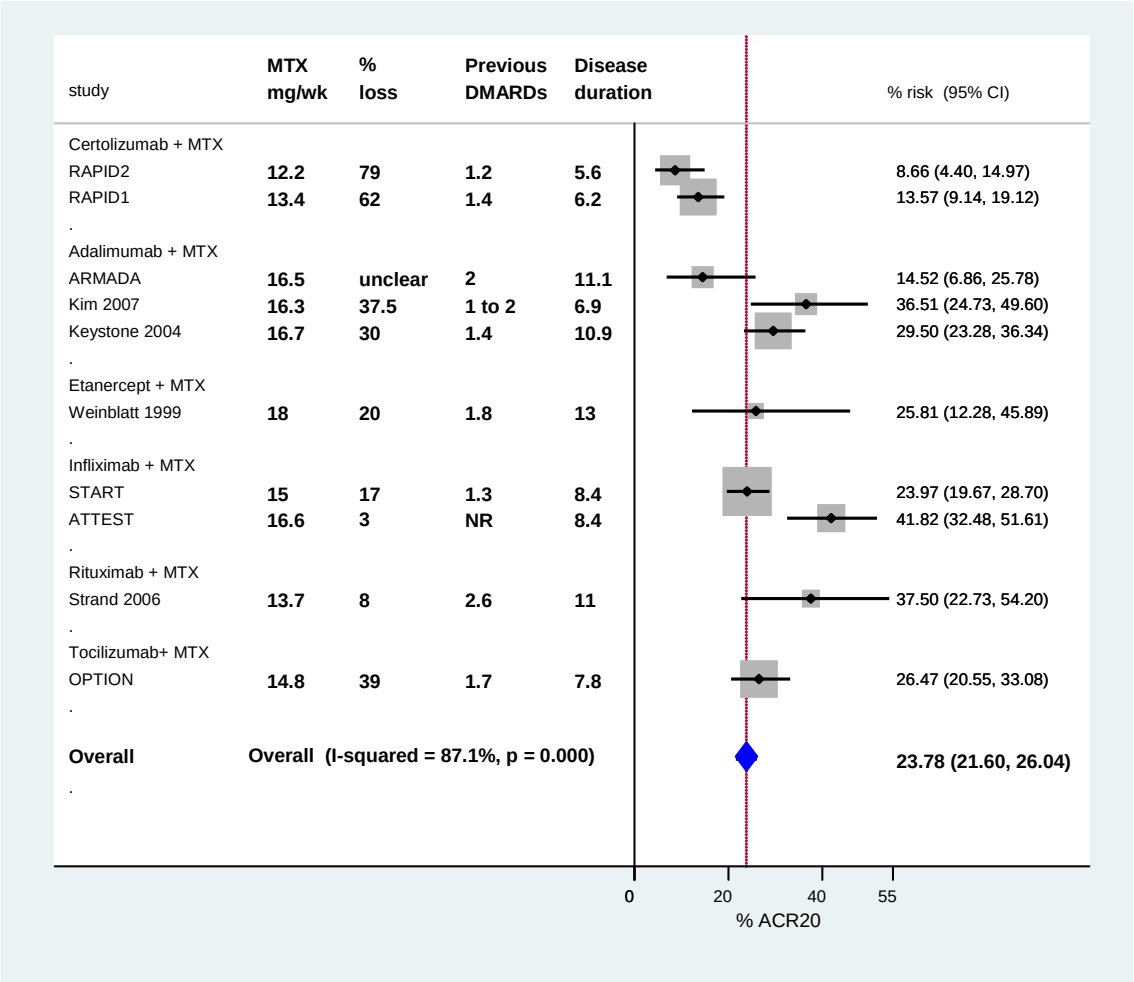
The MS did not comment on whether baseline characteristics of participants were similar across the RCTs contributing to the ICs and whether this might represent a potential source of bias. Similarly, an indication or discussion of clinical or statistical heterogeneity has not been found with the MS.

The submission did not explore the heterogeneity that might exist amongst the studies included for the ICs. As IC makes an assumption of exchangeability between trials,(Glenny et al[17]) the ERG undertook an analysis of the heterogeneity amongst the control arms of the studies used in the estimation of effectiveness for the ACR20 outcome at six months. This choice was made because it involved the largest number of studies and the largest number of events.

The control rates were extracted from the submission and meta-analysed. It should be emphasised that this meta-analysis was undertaken to explore heterogeneity not for the purpose of comparison with active intervention arms.

The results are shown in Figure 3. Data for four study level variables (chosen by the ERG) are also included in the figure.

Figure 3 Risk of ACR 20 in MTX arms of studies in the ICM
Inverse variance fixed effects model (SE of studies for pooling was calculated using Freeman-Tukey arcsine transformation)



Different withdrawal rules in the studies mean that the % loss data may not be strictly comparable between studies.

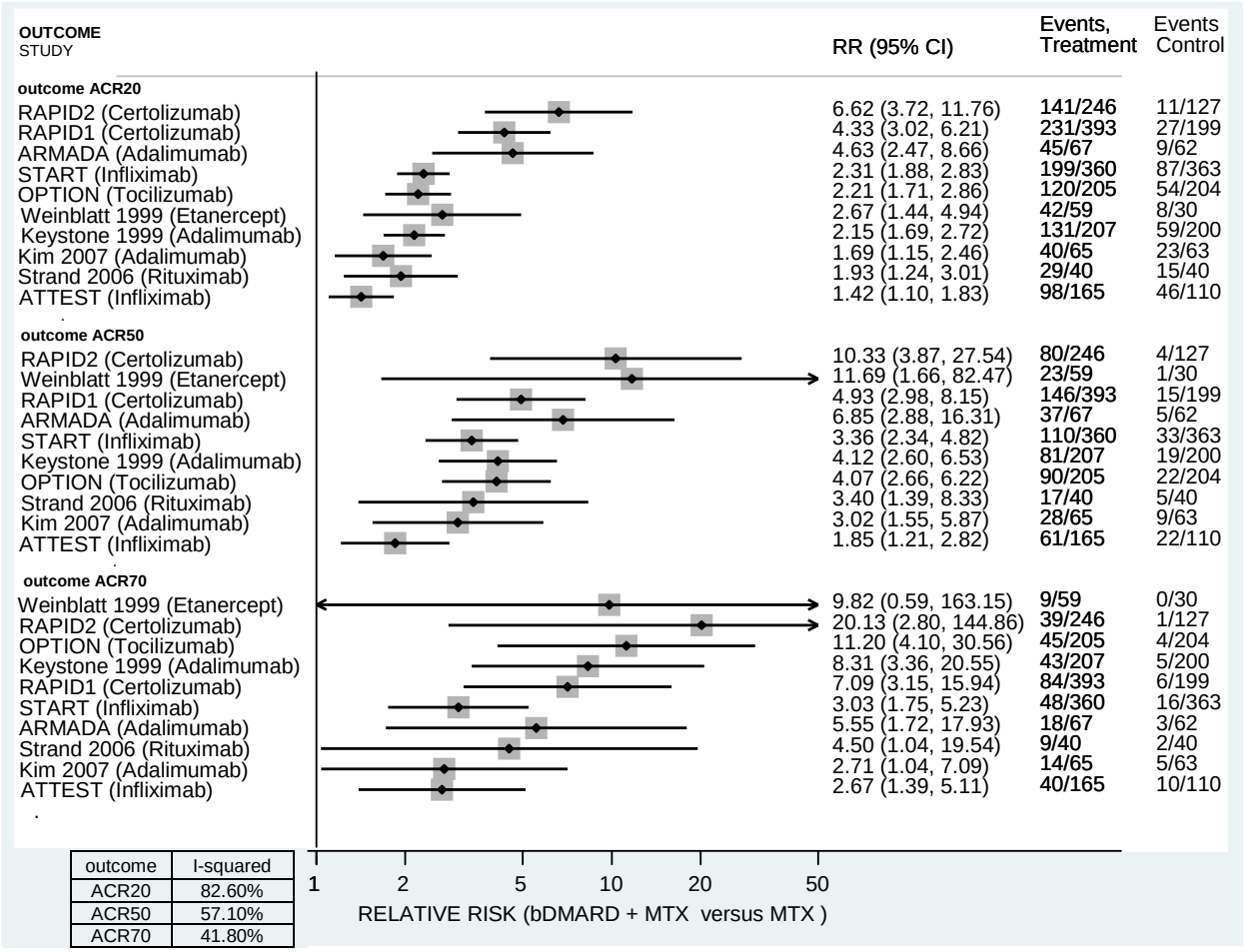
It is evident that the control rate in the two CZP RCTs was the lowest amongst the ten trials and the I squared statistic indicated considerable heterogeneity. When these two studies were omitted from the analysis the I squared statistic was reduced to 70% and pooled estimate increased to 28%.

The four study level variables were looked at as potential contributors to the observed heterogeneity, these were: entry level MTX dose as a potential indicator of treatment intensity and population differences, % withdrawals for the ACR20 outcome as indicator of completeness of data, and duration of RA and number of previous DMARDs trialled as indicators of possible population

differences. For each of these variables the two CZP RCTs were at the extreme of the distributions. These results should be judged in a clinical context and it should be cautioned that when sufficient study level variables are examined chance is likely to throw up apparent associations that can be spurious.

To further explore heterogeneity in the submission the ERG calculated the I squared statistic and prepared forest plots for the ACR20 ACR50 and ACR70 relative risk at six months for combination therapies (bDMARD + MTX versus MTX + placebo). The results are shown in Figure 4.

Figure 4 Relative risk at 6 months for ACR outcomes in 10 studies used for indirect comparison meta-analyses



outcome	I-squared
ACR20	82.60%
ACR50	57.10%
ACR70	41.80%

Considerable heterogeneity is evident especially for the ACR20 outcome. In the figure the studies are arranged in ascending order according to ACR response risk in the MTX control group. Superficial analysis indicates that low

control rates are apparently associated with high relative risk. It should be born in mind however that control risk itself contributes to relative risk.

This examination of heterogeneity amongst the studies used for indirect comparisons indicates that a mixed treatment analysis with methods that allow for differences in control rate or baseline risk (similar in principle to the analyses undertaken by Nixon et al[18]) probably represents the best choice of methodology for the decision problem. Undertaking this analysis was beyond the remit of the ERG.

Response rates estimated for input to the economic model

For use in the economic model the response risks for ACR20 ACR50 and ACR70 for bDMARD arms were calculated using data extracted from the studies identified for the ICs. The method used was as follows:

Response risk across the control arms of all the trials was aggregated as n/N; (where n is the total number achieving response (e.g. ACR20), and N is the total number of participants in the control arms) (submission table 51 page 120). The odds for control response were then calculated (O_c)

The odds for response for a given bDMARD (O_{bDMARD}) was then derived by multiplying the observed odds ratio ($OR_{obs:}$ bDMARD vs Control) by the odds for the control response: The resulting odds were then converted to a risks.

$$O_{bDMARD} = OR_{obs} * O_c$$

To obtain lower and upper confidence intervals for O_{bDMARD} the procedure was repeated using lower and upper confidence intervals for the OR_{obs}.

Thus the control risk was calculated as though the data came from a single trial. Two drawbacks are associated with this procedure: Aggregating control risk across different trials for multiple interventions breaks randomisation and forfeits the strength derived from randomisation making results susceptible to confounders (e.g. prognostic factors that differ between trials). Ignoring between trial variance in the aggregated control risk will underestimate the

uncertainty in the derived values for O_{bDMARD}. The submission acknowledges this as follows (MS page 140):

> It is appreciated that 95% ranges calculated in this way do not account for uncertainty in the effect of the comparator treatment (in this case MTX or placebo). However, it is differences between treatments to which the model is sensitive rather than absolute values and it is believed that the derivation of confidence intervals for absolute treatment effects is an area of continuing methodological development.

According to Glenny et al[17] when there is between trial heterogeneity the underestimate of uncertainty with this procedure will be even greater.

In the manufacturer's clarification document (pg 16) the adopted procedure is claimed to derive risk values from the indirect analyses presented in the submission. Thus from the clarification document:

> All values are obtained from the indirect analysis results reported within section 6.6 of the submission document.

In fact the results from the indirect comparison meta-analyses are not used in these calculations except in the sense that the control risk (MTX arm or placebo arm) is aggregated across the same set of studies that were included for the indirect comparisons and that the observed estimates of bDMARD versus control were used for both the ICs and for the calculation of response risk for the bDMARD arms.

The response risk for the bDMARD arms at 6 months calculated by the manufacturer for ACR20 (clarification document table 18 pg 26; reproduced in Appendix 9) indicate that CZP is superior to the other bDMARDs (e.g. 77.2% risk CZP, 48.2% risk infliximab), and superior to the other bDMARDs except for etanercept for ACR50 and ACR70.

If it is assumed: a] the bDMARDS are equally beneficial (in ACR outcomes); b] that the effect of combining a bDMARD with MTX is similarly additive or multiplicative (multiplier >1) for all bDMARDs; c] that the combination therapy trials deliver accurate estimates (ORs); then arithmetically the manufacturer's

method would generate bDMARD + MTX arm response risks simply in inverse proportion to the trial control risk and in a similar ranking to that actually reported. This could represent a different interpretation to that of superiority of CZP.

Reproducibility of bDMARD response rates in the submission

The ERG attempted to reproduce the bDMARD response rates presented in the submission using the submission procedure. Several errors were identified for the response rates in mono-therapy. These were corrected by the manufacturer in the clarification document (e.g. compare adalimumab table 64 submission with table 18 and table11 in clarification document). The revised model however appeared to still employ the original values.

4.2.4 Summary

A reasonable interpretation of the results is that there is little convincing evidence that CZP is more or less effective than the comparators examined.

DIRECT COMPARISONS

Three RCTs provided evidence; two of combination therapy (CZP + MTX versus placebo + MTX) and one of monotherapy (CZP versus placebo). These demonstrated superiority of CZP relative to control for outcomes specified relevant in the scope.

Of particular importance was data confirming the effectiveness of CZP with respect to ACR20, ACR50 and ACR70 criteria, to quality of life (EQ-5D VAS and SF-36), and to progression of joint destruction. Only EQ-5D VAS results and not EQ-5D derived utility data were presented in the effectiveness section, even though utility data from trials was incorporated into the economic analysis. With regard to the progression of joint damage the

robustness of radiographic estimates at 24 and 52 weeks was somewhat compromised because more than half of control patients had withdrawn by week 16 necessitating linear extrapolation to determine their radiographic progression.

INDIRECT COMPARISONS

The manufacturer undertook indirect comparisons to investigate relative effectiveness of bDMARDs for ACR outcomes. The selection criteria for inclusion of studies for IC were not totally explicit. There was no exploration of heterogeneity amongst the included studies.

The manufacturer's summary statement based on the IC was:

> In summary, CZP administered as a monotherapy or in combination with MTX is at least as effective, and in several cases more effective than other biological DMARDs in terms of ACR response at three and 6 months.

This underplays the considerable and acknowledged uncertainty in the IC results. A more defendable conclusion is that there is currently no convincing evidence that any one of the bDMARDs examined is superior or inferior to another. A further problem arises from some inconsistency in the data extracted for ICs at 3 months. The inclusion of studies should be informed by the decision problem.[11] Confusingly the submission offers two slightly different decisions; one excluding tocilizumab (economic decision) and the other including tocilizumab (effectiveness analysis). This inconsistency undermines the appropriateness of the data input to the economic model; thus it raises the issue of whether data from studies of other bDMARDs should have been included as well as that from tocilizumab trials. A further problem with model input derives from the aggregation of control risk across all trials and the acknowledged underestimate of uncertainty associated with this procedure. This leads to over-precise estimates of risk for ACR outcomes in the bDMARD arms and the susceptibility of results to confounding factors. Of particular importance might be differences in baseline risk between the CZP studies and the other studies included for analysis. A superior approach to that adopted in

the submission would be a mixed treatment comparison that allowed for differences between trial populations that might derive from different prognosis.

5 ECONOMIC EVALUATION

5.1 Description of manufacturer's search strategy and comments

From the submission:

> Details of the strategies including the full search strategies run on MEDLINE EMBASE and the Cochrane Library are detailed in section 6.1 and 10.3 appendix 3 of the manufacturer's submission.
>
> Databases searched were :
>
> MEDLINE (Ovid) and MEDLINE In Process (Ovid) 1950 to date
> EMBASE (Ovid) 1980 to date
> Cochrane Library (Wiley) (NHS EED) 1982 to date
> OHE Health Economic Evaluation Database to date
>
> The searches were conducted on 18 May 2009
>
> The searches were limited to English language papers only (10.2.6).
>
> The searches were based on those in an NHS R&D HTA study by Chen et al which included a systematic review of economic evaluations of anti TNFs (adalimumab, etanercept, infliximab) in the treatment of rheumatoid arthritis in adults. The Chen review was updated from 2005 onwards , the newer biological DMARDs/TNF inhibitors not covered in the original review were searched for without date limit.

Comments

- A combination of index terms and textwords are used to create a comprehensive strategy which is unlikely to have missed many relevant references.
- No searches relating to quality of life appear to have been conducted.
- The strategy has been translated successfully between the databases, however the results of the Cochrane Library search across NHS EED will have been restricted unnecessarily by the incorrect use of the Boolean "AND" operator where "OR" has been used in the MEDLINE and EMBASE searches (line#29). Therefore, the ERG re-ran the search using both the "AND" and "OR" operators in the final line – the "AND" retrieved 3 economic evaluations but the broader search using

the “OR” operator and which correctly replicated the strategy which the manufacturer ran on MEDLINE and EMBASE retrieved 50 studies. The ERG specifically examined five of these studies which were deemed to be relevant because they involved the comparators (adalimumab, etanercept, infliximab, rituximab, or abatacept). One study did consider both the quality of life and economic impact of infliximab by comparison to adalimumab and could have been quoted (Walsh et al. 2007).

5.2 Overview of manufacturer’s economic evaluation

The submitted economic evaluation is a cost–utility analysis using a lifetime Markov model built in Microsoft Excel® and comparing the cost-effectiveness of CZP to the other bDMARDs. The CZP dose regimen used in the economic analysis was an initial loading of 400mg at weeks 0, 2 and 4, followed by 200mg every 2 weeks (see page 111 of submission).

The comparators for the cost-effectiveness analysis have been chosen based on existing NICE clinical guidelines which state that following an inadequate response to at least two conventional DMARDs, the patient can progress to an appropriate bDMARD. The comparator therapies included in the model for RA patients that have responded inadequately or are intolerant of MTX are:

- Adalimumab (40 mg EOW) + MTX
- Adalimumab (40 mg EOW)
- Infliximab (3 mg/kg) + MTX
- Etanercept (25 mg twice weekly) + MTX
- Etanercept (25 mg twice weekly)
- Rituximab (1000 mg every other week)

Tocilizumab whose technology appraisal is in progress, is not considered as a comparator to CZP.

The manufacturer adopted mean estimates from the CZP trials population (RAPID 1, RAPID 2 and FAST4WARD) as the baseline characteristics of the model.

The time frame for the analysis is 45 years in order to capture the lifetime of the average patient up to 100 years of age (the average age of the model population is 52.2 years, see Table 48 of the submission).

Patients enter the model after an inadequate response, or intolerance, to MTX and proceed on treatment with CZP or a bDMARD comparator. Treatment may be discontinued due to the following events:

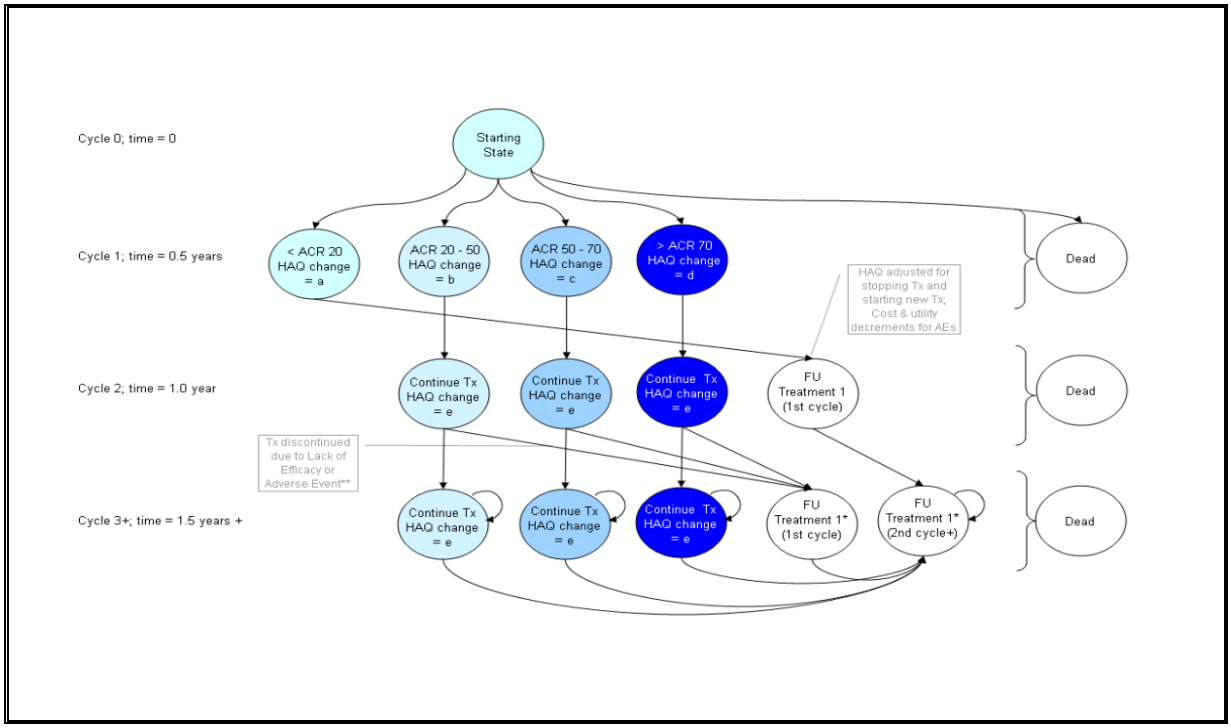
- At the end of the first model cycle (after six months of treatment) for all non-responders (defined as patients who obtain inadequate response, ACR <20);
- During any subsequent model cycle due to lack of efficacy or adverse events (at the mid-point of the cycle);
- Due to death.

The first cycle lasts for 6 months, at the end of that cycle patients are assigned to one of four response groups based on assumed risks of responding to the relevant treatment, as measured on the ACR scale:

- ACR of <20
- ACR of 20-50
- ACR of 50-70
- ACR of >70

Only patients who obtain an adequate response in the first timestep continue on the modelled initial therapy. At the end of the next and following cycles, patients may remain in the same Markov treatment health state, discontinue treatment due to lack of efficacy or die. There are no state transitions other than discontinuation of treatment or death.

The model schema from the submission (Fig 41 page 119) is shown below:



The manufacturer submitted two different economic models during the appraisal process. The first model (“original” model) was included as part of the initial submission. The second model (“revised” model) was received during the clarification stage of the appraisal process. The manufacturer made changes to the model during this time that went beyond those prompted by requests for clarification by the ERG. As such the ERG consider these as two separate models. The original model included:

- Costs of drugs and monitoring medications and routine follow-up (Tables 59 and 60, pg 135 in the submission) and assumed that there is no wastage of drugs
- Differential adverse event effects on QALY between treatment comparators, but no associated costs of adverse event.
- A Patient Access Scheme (PAS) that gives away the initial 10 syringes free of charge to all new patients treated with CZP (Appendix 4 of the submission).
- Two model scenarios: scenario A with the PAS, and scenario B without PAS.

- Health outcomes (quality-adjusted life-years) and incremental cost-effectiveness (Tables 65 and 66, pg 144). Utilities for patients are based on ACR response and other factors and are quantified via regression analysis on EQ-5D values or HAQ-DI linked to utility weights via a simple linear relation (Tables 52 and 53 pg 129).

- A probabilistic sensitivity analysis (PSA) to assess the overall effect of the uncertainty in the model (Figures 42-57 – 7.7c).

Details and results of the revised model provided by the manufacturer were included in Section C of the response to clarifications. This revised model incorporates three key changes to the original model.

1. Utility assumptions

The utility assumptions have been revised. The originally submitted model assumed that utility as measured by EQ-5D decreased by 0.037 per year if response had been assessed at 6 months (or 0.014 per year at 3 months). The new model adopts an assumption of a utility decrease of 0.0025 per year if response had been assessed at both 6 months and 3 months (see table 41 pg 68 of the clarification document in replacement of table 53 pg 129 of the submission).

2. Adverse events

The assumption of differential adverse event effects between comparator therapies is considered inappropriate and inconsistent with the accounting of costs, and the revised model does not directly include any adverse events.

3. Discontinuation probabilities

Treatment discontinuation probabilities are revised with exact values rather than approximate values using an exponential survival distribution (see table 21 pg 29 in the clarification document in replacement of table 49 pg 116 in the submission).

The differences between the models are summarised by the ERG in Table 4.

Table 4 Summary of results from the original and the revised economic models

	Original model				Revised model			
Response definition	**HAQ-DI**		**EQ-5D**		**HAQ-DI**		**EQ-5D**	
	3 months	6 months	3 months	6 months	3 months	6 months	3 months	6 months
CZP or comparator	0.066	0.177	-0.014	-0.037	-0.1913	-0.0963	0.0402	0.0202
Follow-up bDMARDs	0.066	0.177	-0.014	-0.037	0.012	0.012	-0.0025	-0.0025
Follow-up cDMARDs	0.066	0.177	-0.014	-0.037	0.012	0.012	-0.0025	-0.0025
Palliation	0.066	0.177	-0.014	-0.037	0.012	0.012	-0.0025	-0.0025
Base case deterministic with PAS	**Mean costs**	**Mean life years**	**Mean QALys**	**ICER**	**Mean costs**	**Mean life years**	**Mean QALys**	**ICER**
CZP + MTX	£96,417	13.547	2.903		£85,583	14.63	6.654	
Etanercept + MTX	£97,317	13.554	2.908	N/A (£197,037)*	£86,165	14.62	6.589	CZP dominates
Adalimumab + MTX	£96,428	13.555	2.801	CZP dominates	£86,034	14.59	6.412	CZP dominates
Rituximab + MTX	£92,936	13.539	2.77	£26,157	£82,940	14.58	6.362	£9,072
Infliximab + MTX	£104,460	13.534	2.692	CZP dominates	£95,599	14.55	6.196	CZP dominates
CZP	£91,820	13.535	2.736		£81,849	14.57	6.305	
Etanercept	£95,691	13.489	2.782	N/A (£82,695)*	£85,941	14.60	6.435	£31,582*
Adalimumab	£90,048	13.365	2.609	£13,982	£84,201	14.54	6.090	CZP dominates
Base case deterministic without PAS	**Mean costs**	**Mean life years**	**Mean QALys**	**ICER**	**Mean costs**	**Mean life years**	**Mean QALys**	**ICER**
CZP + MTX	£99,992	13.547	2.903		£89,158	14.63	6.654	
Etanercept + MTX	£97,317	13.554	2.908	ETA dominates	£86,165	14.62	6.589	£46,192
Adalimumab + MTX	£96,428	13.555	2.801	£34,930	£86,034	14.59	6.412	£12,937
Rituximab + MTX	£92,936	13.539	2.77	£53,023	£82,940	14.58	6.362	£21,345
Infliximab + MTX	£104,460	13.534	2.692	CZP dominates	£95,599	14.55	6.196	CZP dominates
CZP	£95,395	13.535	2.736		£85,424	14.57	6.305	
Etanercept	£95,691	13.489	2.782	N/A (£6.341)*	£85,941	14.60	6.435	£3,991*
Adalimumab	£90,048	13.365	2.609	£42,197	£84,201	14.54	6.090	£5,687

* Comparator was more costly and more effective than CZP+MTX

The results of the revised economic analysis were summarised in tables 44 and 42 of the clarification document, reproduced below:

Table 42: Base case deterministic results Scenario A2

	Mean costs	Difference	Mean life years	Difference	Mean QALYs	Difference	ICER	PSA results £20 000 threshold	PSA results £30 000 threshold
Combination therapies									
Certolizumab pegol + MTX	85 583		14.63		6.654				
Etanercept + MTX	86 165	-582	14.62	0.01	6.589	0.065	CZP dominates	70.1	65.7
Adalimumab + MTX	86 034	-451	14.59	0.04	6.412	0.242	CZP dominates	71.0	67.6
Rituximab + MTX	82 940	2643	14.58	0.05	6.362	0.292	9072	60.2	61.8
Infliximab + MTX	95 599	-10 016	14.55	0.08	6.196	0.458	CZP dominates	98.8	95.5
Monotherapies									
Certolizumab pegol	81 849		14.57		6.305				
Etanercept	85 941	-4092	14.60	-0.03	6.435	-0.130	31 582*	56.3	51.8
Adalimumab	84 201	-2352	14.54	0.03	6.090	0.215	CZP dominates	82.9	78.2

Table 44: Base case deterministic results for Scenario B2

	Mean costs	Difference	Mean life years	Difference	Mean QALYs	Difference	ICERs	PSA results £20 000 threshold	PSA results £30 000 threshold
Combination therapies									
Certolizumab pegol + MTX	89 158		14.63		6.654				
Etanercept + MTX	86 165	2993	14.62	0.01	6.589	0.065	46 192	55.0	56.4
Adalimumab + MTX	86 034	3124	14.59	0.04	6.412	0.242	12 937	56.7	58.5
Rituximab + MTX	82 940	6218	14.58	0.05	6.362	0.292	21 345	45.4	51.0
Infliximab + MTX	95 599	-6441	14.55	0.08	6.196	0.458	CZP dominates	97.0	92.6
Monotherapies									
Certolizumab pegol	85 424		14.57		6.305				
Etanercept	85 941	-517	14.60	-0.03	6.435	-0.130	3991*	43.5	43.7
Adalimumab	84 201	1223	14.54	0.03	6.090	0.215	5687	66.6	67.8

* Comparator was more costly and more effective than CZP+MTX

The manufacturer concludes:

> The revised model continues to demonstrate the cost-effectiveness of CZP within the NHS in England and Wales, with a probability greater than 50% of being within NICE's £30,000 per QALY threshold for CZP + MTX compared to combination therapies, and greater than 40% for CZP compared to monotherapies. By accounting for the proposed Patient Access Scheme, further savings are provided to the NHS and the degree of cost-effectiveness is even greater

Comments from the ERG:

- The revised model has been considered by the ERG as a replacement to the original model presented in the submission. The ERG would like to underline that some of the manufacturer's model changes did not stem from the ERG's clarification requests but derived from the fact that the manufacturer used the clarification period to run further research into their modelling.
- The use of a Markov model is appropriate. The ERG was concerned by the differing cycle lengths applied in the first three model cycles. The manufacturer's response to clarifications stated that the differing cycle lengths was to allow for assessment of response at either three or six months. Although it is more common for cycle lengths within a Markov model to be the same length of time, in this case the difference in cycle length was driven by clinical considerations. However, the manufacturers clearly state (section 7.2.6.7) no half-cycle correction was applied after the first state. The half-cycle correction is a generally accepted component of Markov of models, though given that the length of each cycle is relatively short compared to the duration of the model, the results in this case are unlikely to be meaningfully different if the half-cycle correction were applied.
- The time horizon is appropriate to the decision problem and as requested for clarification, the manufacturer had provided sensitivity analyses with five and ten years (see tables 25-28 pg 45-46 in the clarification document).

5.2.1 Natural history

The manufacturer does not describe the natural history of RA in the cost-effectiveness section but this is described elsewhere in the submission, particularly in terms of disease progression and response to treatment.

5.2.2 Treatment effectiveness within the submission

Treatment effectiveness within the submission is measured using ACR response as the primary clinical method for monitoring disease activity and response to therapy. This is mentioned in previous sections of this report.

No extrapolation of survival was performed as part of the economic model. Long term differences between CZP and the comparators are driven solely by the relative differences in ACR response rates applied in the model. The subsequent HAQ-DI progression rate and time on treatment for responding patients is assumed to be equivalent.

As mentioned above the costs or effects of adverse events are not incorporate in the revised model.

Comments:

- The assumption of no difference in adverse effects between drugs (certolizumab pegol, infliximab, adalimumab, etanercept, rituximab) may on average be a reasonable assumption, but cannot be considered evidence based. For example, infusion reactions can obviously only apply to the IV drugs but not the S/C, and vice versa, local reactions to S/C but not IV. It is possible that patients are at greater risk of TB with infliximab than etanercept.

5.2.3 Health related quality of life

The model measures patients' health states in terms of EQ-5D directly collected by the manufacturer as part of the RAPID 1 & 2 trials at trial centres

in Europe. Patients are initially assigned a quality of life equal to 0.38 as measured by EQ-5D utility.

Patient quality of life depends on ACR response category and the magnitude of changes in utility were estimated using coefficients from an ANCOVA regression analysis controlling for age, gender, baseline EQ-5D, disease duration, number of previous conventional DMARDs and anti-CCP antibody status as covariates.

Patients continuing on CZP or the comparator treatment are subject to a decline in quality of life of an annual amount derived from repeated measures analysis of CZP trial data.

Assumptions related to change in quality of life are presented in the table 41 (pg 68) of the clarification document.

Treatment benefit while on first line bDMARD treatment

To model the continuing benefit of treatment following on from the initial response, the manufacturer has included an estimate of utility based on a mapped estimate of EQ-5D utility scores derived from HAQ-DI scores. The algorithm used to derive these scores is based on a model published by Bansback et al (see below for a description of this model).

When a patient enters this phase of the model, their initial utility score is based on the data collected as part of the clinical trials. As patients progress through this phase of the model, utility scores are then revised based on the mapped estimates.

Utility after first line bDMARD treatment

Utility values used in the model following the discontinuation of first line treatment are determined in the same manner as those for estimating benefit while on first line treatment.

Comments:

- The exact details of how many patients were given health related quality of life questionnaires and response rates to the questionnaires were not provided in the submission. From the trial reports it appears that data was collected for only European patients in RAPID 1&2.
- Elicitation of utility values using an indirect method such as the EQ-5D is recommended in the NICE methods guidance. The manufacturers have followed this approach and have included an appropriate estimate of baseline utility value for this portion of the model. However, reporting of the health related quality of life data in the submission was poor, making it difficult to evaluate. Only one table with EQ-5D data was presented and this was based on the results of the VAS component of the instrument. Even this did not report how many patients had responded or any other details. Full responses to the EQ-5D were provided as part of the trial reports, although without aggregated data provided these are of little use in evaluating the manufacturer's economic model.
- Reasons for the change in using the observed data as collected in the trial to the mapped estimates are not clearly stated in the report. It is likely however that the reason was to allow estimation of utility scores over time. EQ-5D data was collected only during the trial period and accurate extrapolation would be difficult. By linking EQ-5D scores to HAQ-DI scores and ACR response, the manufacturer can provide an estimate of changes in utility over time whilst conforming as closely as possible to the NICE methodology guidance on utility values used in economic models. By effectively using two mapping algorithms (ACR results to HAQ-DI; HAQ-DI to EQ-5D data) the final estimates of utility will contain a further element of uncertainty arising from using one instrument to predict results for another. This additional uncertainty is not captured in the model, although the current mapping literature does not provide adequate methodologies for addressing the problem in this context. However, this does make it more difficult to draw firm conclusions about the reliability of the final estimates and by extension

the reliability of the final results of the model where these are based on the mapped estimates.

- Regarding utility after first line treatment, the same comments and concerns apply to this approach as described in the previous section.

Comment on the Bansback et al model

The mapping model used to estimate utility scores for treatment benefit effect and for post first-line treatment are based on a previously published model from Bansback et al (2007).Using two samples of patients, one from Canada (n=319) and one from the UK (n=151) the authors tested five different linear regression models to explore the relationship between HAQ-DI scores and responses to the EQ-5D and SF-6D health related quality of life instruments. Validity of the models was tested according to their predictive ability – that is, how well the algorithm did predict EQ-5D or SF-6D scores from HAQ-DI scores compared with the observed study results. Predictive ability of the models was determined by the root of the mean squared error statistic (RMSE): the smaller the RMSE, the better the assumed predictive ability of the model.

The authors conclude that a model that best predicted EQ-5D results was one where the EQ-5D index score was regressed on the individual levels of the HAQ-DI item scores. This model also included the most covariates of all models tested. The paper does not list the covariates used in the model, a significant weakness in determining whether or not this is the most appropriate model for use in estimating EQ-5D scores in the manufacturers' economic model for CZP. In other respects, the model by Bansback et al is consistent with other mapping papers in the literature. The use of linear regression techniques to model results and the RMSE statistic to evaluate model fit and are common approaches. Though these methods have potential weaknesses, there is currently no commonly agreed framework for mapping between non-preference and preference based measures. The use of this model can therefore be considered appropriate within the context of the state of the art of mapping methods and the fact that the NICE methods guidance permits such an approach.

5.2.4 Resources and costs

The manufacturer assumed a price per 200 mg of CZP equal to £357.50 using the yearly cost of adalimumab (£357.50 per 40mg).

Resources were measured using published literature and indexed to 2009 prices using the NHS Pay and Prices index. Costs were considered from an NHS and Personal Social Services perspective (in practice NHS only), consistent with the NICE reference case.

Resources and unit costs used in the model are described in tables 55-61 (pg 132-136) of the manufacturer submission. In the economic evaluation three different types of costs associated with the interventions were included. These are the costs of drug administration, costs of patient monitoring, and drug costs. The costs were obtained in US dollars year 2001, and subsequently inflated (Table 61 pg 136 in the submission) and converted to pounds using an applied exchange rate of £0.67 to $1.00.

The registration of ACR response is assumed to coincide with regular monitoring of the disease by a rheumatologist who should see the patient every six months. Hence, monitoring of CZP does not incur any extra costs.

The impact of drug-related adverse events on medical resources was assumed to be negligible. Costs of adverse events were not incorporated in the model.

Also, UCB has decided to introduce a Patient Access Scheme (PAS) that includes providing the initial 10 syringes (i.e. treatment for three months) free of charge to all new patients treated with CZP and two separate scenarios with and without the PAS were tested.

Comments:

- Costs related to adverse events were included in the original model but not in the revised model. This had consequences on the mean costs of each drug which were not negligible (for example, the mean cost of CZP in combination with MTX reduces from £96,417 with adverse

events to £85,583 without adverse events in scenario A (PAS)). Therefore, this change is potentially influential and the ERG could not understand how this approach was consistent with the NICE clinical guidelines on the management of rheumatoid arthritis in adults.

- The manufacturer estimated that uptake of the PAS would be in the region of 95% of eligible patients. It is not clear what would happen to the remaining 5%, The PAS proposal claims no additional costs to the NHS of managing the scheme.
- The difference between scenarios A and B is simply the exclusion of the cost of the ten free syringes that would be made available under the PAS.
- The ERG are not aware of any PASs for the comparator bDMARDs. If any exist or become approved by the Department of Health then the scenario A analyses will require revision.

5.2.5 Discounting

An annual rate of 3.5% on both costs and health effects, which accords with NICE reference case.

In the sensitivity analyses, the manufacturer tested the implications of an annual rate of 1.5% and then 6% for either costs or health effects, or both (see table 67 pg 146 of the manufacturer submission / table 43 pg 71 of the clarification response for scenario A and table 68 pg 152 of the manufacturer submission / table 45 pg 78 of the clarification response for scenario B).

5.2.6 Sensitivity analyses

Probabilistic sensitivity analysis (PSA) was the primary approach to deal with uncertainty in the results input to the model. This included parameters relating to the following: clinical effectiveness, the association between mortality and HAQ-DI scores, age, gender, weight, baseline HAQ-DI scores, number of previous DMARD, disease duration, anti-CCP antibody positive, anti-CCP antibody negative and utility weight. The manufacturer applied estimates of

uncertainty around the parameters. Uncertainty may have been under-estimated for the ACR response.

In addition to the PSA, a number of one way sensitivity analyses were conducted for scenarios A & B. Within each scenario the manufacturer altered the time horizon, discount rate, ACR response, treatment duration, baseline HAQ score, utility progression for first line and subsequent treatments, the perspective and the costing method.

The ERG attempted to run a number of additional one-way scenario analyses. Not all parameters varied by the manufacturer in the sensitivity could be altered within the model due to the design of the model interface. The ERG was only able to modify the time horizon, discount rate, ACR response and costing method. Therefore, it has not been possible to verify all the results.

An apparent anomaly was noted in figure 31 p 75 of the manufacturers' responses to clarification questions. The scatter plot for the comparison of CZP to adalimumab takes an unusual "banana" shape. It is not clear to the ERG at present what parameters may be driving this result. All other comparators produced a more traditional approximate ellipsoid scatter plot. It is unclear if this result is data driven or an error in programming in the model.

5.2.7 Model validation

The ERG was concerned about some technical aspects of the model provided.

The ERG noticed a flaw in the programming for the probabilistic sensitivity analysis in the model. When each new simulation is run, only the first 1000 results from the immediate previous simulation are deleted. If a previous PSA using more than 1000 iterations is run then all iterations from 1001 upwards would still be reported in the PSA results tab. It does not appear that this programming flaw has affected the accuracy of the model results as presented. However the model could not be described as working as claimed if more than 1000 iterations are required for a simulation. If as the

manufacturers state in their submission p.139 “1000 simulations were run using base case assumptions and parameter variability” then the PSA results reported into the submission are correct. Nevertheless, if any results were based on greater than 1000 iterations (which should be possible as the manufacturers claim that 32,000 iterations should be possible) then the flaw could have important consequences on the cost-effectiveness results as the PSA summary results would rely on all the PSA results in the sheet, potentially including results from previous simulations. .

Other issues relating to the model were noted, including:

- The probabilistic sensitivity analysis was run regardless of whether the deterministic or probabilistic was selected on the settings tab.
- Despite the range of comparators, only pair-wise comparisons were possible in the model provided. A comparison between all drugs based on the results of the deterministic analysis is provided in section 5.5 of this report. Whereas the manufacturer provided cost-effectiveness acceptability frontier graphs in the clarification document (Figures 32-33, 24-25), the revised model could not be used to generate similar graphs. It was thus not possible for the ERG to reproduce in full the manufacturer’s results or to confirm the accuracy of the results reported in Table 39.
- The Y-axis of PSA scatter plot was inaccurately labelled “Probability of cost-effectiveness” rather than “Incremental costs” within the Excel model
- The model was not user-friendly for one-way sensitivity analyses: no facility was provided for changes to certain parameters varied in the manufacturers reported one-way analyses, including: discount rates for costs and benefits, ACR response, treatment duration for CZP and comparators, utility progression rates, rebound assumption, or perspective (NHS and Personal Social Services vs Societal perspective)

- Results in life years are available in the Excel model but the manufacturer has not reported these in the submission. Several results when measured in terms of LYs exhibited a cost-effectiveness dominance of comparators over CZP (e.g. etanercept+MTX).
- Some variables and parameter values in the executable model did not correspond to those contained in the submission document (e.g. table 21 in the clarification document or table 49 of the submission; table 50 of the submission)
- In the manufacturer's responses to the clarification they note that the model has been designed for use in markets other than the UK. The resulting additional complexity has hampered the ability of the ERG to evaluate in full the reliability of the model. These limitations have been noted above.

5.3 Critique of approach used

With respect to the economic evaluation report and the economic model provided by the manufacturer, the ERG have four key concerns. These are related to the way that health related quality of life and utility values have been estimated, the way in which adverse events have been treated as part of the model, the fact that only pair-wise comparisons were possible and finally the degree to which the original model and the revised model differed.

The ERG concerns over the way health related quality of life has been measured and the way in which utilities have been estimated for use in the model have been described in Section 5.2.3 of this report.

The ERG concerns over the fact that despite the range of comparators, only pair-wise comparisons were possible. The PSA could not be conducted for all the technologies together. The manufacturer provided cost-effectiveness acceptability frontier graphs in the clarification document, but they only showed the results of pair-wise comparisons. The decision problem in this case is not to choose CZP relative to individual alternatives, but whether or

not CZP is cost-effective relative to all alternative therapies available to the NHS.

The ERG has concerns about the change in the manufacturers' models regarding adverse events, the costs of which were taken into account in the original model but not in the revised model. This change was not motivated by the ERG's clarification requests about adverse events. The ERG could not determine how this approach was consistent with the NICE Clinical Guideline 79 but noticed significant differences in the mean costs of each drug when costs related to adverse events were not included. Except in scenario A where Etanercept +MTX experiences the highest reduction in mean costs (£11,152 less), it always was the mean costs of CZP monotherapy or in combination with MTX which reduces the most (£10,402 less on average). As a result, the non –inclusion of the costs of adverse events could be seen to favour the manufacturers' technology.

The model that the ERG has appraised in detail is the revised model provided to us by the manufacturer following our requests for clarification. As has been stated, this model differed significantly from that included in the original submission. The results of these changes have been summarised in section 5.2 of this report. Due to the submission of a revised model, the length of time available to the ERG to assess the cost-effectiveness analyses was curtailed.

5.4 Results included in manufacturer's submission

The base case results were given with the following assumptions: ACR response at six months, life time horizon, discounting of costs and QALYs at 3.5%, inflation rate of 3 %, no wastage of drugs and sulfasalazine as the first follow-up treatment. The results included in the manufacturers' clarification document are presented below (table 42 pg 70). Whereas in the original submission, the PSA results as well the sensitivity analyses were commented on by the manufacturer, only the deterministic results were commented with the revised model.

SCENARIO A2 – Revised CE results (incorporating PAS)
CZP has favourable cost-effectiveness against each of the combination therapy and monotherapy comparators. Rituximab + MTX is cost-effective below a willingness to pay of £9051 per QALY; certolizumab pegol + MTX is cost-effective above this threshold.
Among monotherapies, certolizumab pegol is cost-effective below a willingness to pay of £31 582 per QALY; above this threshold etanercept is cost-effective. However, at higher willingness to pay thresholds, the probabilities that certolizumab pegol and etanercept are cost-effective are similar, leading to a large probability of error if any one treatment is preferentially recommended.

SCENARIO A2 – Revised CE results (incorporating PAS) table of results:

	Mean costs	Difference	Mean life years	Difference	Mean QALYs	Difference	ICER	PSA results £20 000 threshold	PSA results £30 000 threshold
Combination therapies									
Certolizumab pegol + MTX	85 583		14.63		6.654				
Etanercept + MTX	86 165	-582	14.62	0.01	6.589	0.065	CZP dominates	70.1	65.7
Adalimumab + MTX	86 034	-451	14.59	0.04	6.412	0.242	CZP dominates	71.0	67.6
Rituximab + MTX	82 940	2643	14.58	0.05	6.362	0.292	9072	60.2	61.8
Infliximab + MTX	95 599	-10 016	14.55	0.08	6.196	0.458	CZP dominates	98.8	95.5
Monotherapies									
Certolizumab pegol	81 849		14.57		6.305				
Etanercept	85 941	-4092	14.60	-0.03	6.435	-0.130	31 582*	56.3	51.8
Adalimumab	84 201	-2352	14.54	0.03	6.090	0.215	CZP dominates	82.9	78.2

SCENARIO B2 – Revised CE results (not incorporating PAS)
Certolizumab pegol + MTX has favourable cost-effectiveness against each of the combination therapy comparators. Rituximab + MTX is cost-effective below a willingness to pay of £14 207 per QALY. Above this threshold, probabilities of cost-effectiveness are similar amongst the comparators. Deterministic results suggest etanercept + MTX would be cost-effective between thresholds of £14 207 and £46 046 per QALY, and certolizumab pegol + MTX cost-effective above this. That the probabilities of cost-effectiveness are similar leads to a large probability of error if any one treatment is preferentially recommended.

Deterministic results suggest that certolizumab pegol is cost-effective below a willingness to pay threshold of £3991 per QALY and that etanercept would be cost-effective above this. However, as with scenario A2, the probability that certolizumab pegol is cost-effective at higher willingness to pay reflects the uncertainty around the relative effectiveness of certolizumab pegol vs etanercept. Again there would be a large probability of error if any one treatment is preferentially recommended, given a higher willingness to pay threshold.

SCENARIO B2 – Revised CE results (not incorporating PAS) table of results:

	Mean costs	Difference	Mean life years	Difference	Mean QALYs	Difference	ICERs	PSA results £20 000 threshold	PSA results £30 000 threshold
Combination therapies									
Certolizumab pegol + MTX	89 158		14.63		6.654				
Etanercept + MTX	86 165	2993	14.62	0.01	6.589	0.065	46 192	55.0	56.4
Adalimumab + MTX	86 034	3124	14.59	0.04	6.412	0.242	12 937	56.7	58.5
Rituximab + MTX	82 940	6218	14.58	0.05	6.362	0.292	21 345	45.4	51.0
Infliximab + MTX	95 599	-6441	14.55	0.08	6.196	0.458	CZP dominates	97.0	92.6
Monotherapies									
Certolizumab pegol	85 424		14.57		6.305				
Etanercept	85 941	-517	14.60	-0.03	6.435	-0.130	3991*	43.5	43.7
Adalimumab	84 201	1223	14.54	0.03	6.090	0.215	5687	66.6	67.8

Cost-effectiveness acceptability frontiers for each scenario are presented in Appendix 10**.**

The results generated by the manufacturer's economic model were mainly driven by the effectiveness inputs for ACR20 ACR50 and ACR70 responses estimated for the different bDMARDs and by the costs associated with each of these therapies. Uncertainty around the ACR responses as highlighted elsewhere in this report and further uncertainties associated with utility values suggest that the PSA findings in the tables above should be viewed with some caution.

5.5 Comment on validity of results presented with reference to methodology used

Quality assessment of the economic model using the ScHARR-TAG check list can be found in Appendix 11.

Setting aside all the concerns above, the results of the analysis provided in the submission can be considered an approximate estimate of the likelihood that CZP has favourable cost-effectiveness against each of the combination therapy and mono-therapy comparators.

The ERG has compared the deterministic results of all the drugs together and that comparative analysis is summarised in the Table 5 below.

Table 5 Comparison of the deterministic results for combination and mono-therapy for the senarios with and without PAS

Scenario A - Combination therapies					
Comparators	Mean costs	Mean QALYs	dCosts	dQALY	ICER
Infliximab	95,599	6.196			
Rituximab	**82,940**	6.362			
Adalimumab	86,034	6.412	3,094	0.05	61880
Etanercept	86,165	6.589	3,225	0.227	14207
Certoluzimab	85,583	**6.654**	-582	0.065	-8954
Scenario B - Combination therapies					
Comparators	Mean costs	Mean QALYs	dCosts	dQALY	ICER
Infliximab	95,599	6.196			
Rituximab	82,940	6.362			
Adalimumab	**86,034**	6.412			
Etanercept	86,165	6.589	131	0.177	740
Certoluzimab	89,158	**6.654**	3,124	0.242	12909
Scenario A - Monotherapies					
Comparators	Mean costs	Mean QALYs	dCosts	dQALY	ICER
Adalimumab	84201	6.09			
Certoluzimab	**81849**	6.305			
Etanercept	85941	**6.435**	4092	0.13	31477
Scenario B - Monotherapies					
Comparators	Mean costs	Mean QALYs	dCosts	dQALY	ICER
Adalimumab	**84201**	6.09			
Certoluzimab	85424	6.305	1223	0.215	5688
Etanercept	85941	**6.435**	517	0.13	3977
Scenario A - Monotherapies with response rate corrected by ERG					
Comparators	Mean costs	Mean QALYs	dCosts	dQALY	ICER
Adalimumab	83390	6.058			
Certoluzimab	**81849**	6.305			
Etanercept	85394	**6.403**	3545	0.098	36173
Scenario B - Monotherapies with response rate corrected by ERG					
Comparators	Mean costs	Mean QALYs	dCosts	dQALY	ICER
Adalimumab	**83390**	6.058			

Certoluzimab	85424	6.305	2034	0.247	8235
Etanercept	85394	**6.403**	-30	0.098	-306

In the table are reported the base-case analysis deterministic results from the clarification documents as well as the corrected results of adalimumab and Etanercept monotherapy response rates (ACR20, ACR50 and ACR70). We ranked the comparators by effectiveness. Whereas Certoluzimab is the most effective in combination therapies, etanercept appears more effective in monotherapies.

Regarding the costs, the table makes clear that the manufacturers' technology is cost-effective when the patient access scheme is considered.

5.6 Summary of uncertainties and issues

The original and revised models submitted by the manufacturer for this appraisal were generally of a good quality. The model was internally consistent, despite the programming errors identified by the ERG and noted above, and produced results consistent with those claimed by the manufacturer in their submission. The description of the model provided by the manufacturer in the original submission was of poorer quality and in places and generated a number of questions. However, the manufacturers' responses to our clarification questions addressed these concerns adequately. Despite the overall quality of the model, the ERG had a number of key concerns.

Utilities for the initial response to treatment were based on EQ-5D data collected as part of the trial, as recommended in the NICE methodology guidance. However, no summary information was provided in the submission relating to this data, apart from a single table indicating VAS responses from this instrument. The ERG were not able to assess the validity of these results. Full results were included in the trial reports, however without access to the data in aggregated form it is not possible for the ERG to make meaningful comment on the data.

Utilities for later phases of the model were based on mapped estimates. First, ACR response was mapped to HAQ-DI data using trial data. Then HAQ-DI data was mapped to EQ-5D data using a previously published mapping algorithm. This method is acceptable under the NICE methodology guidance, The use of mapping approaches to estimate utility scores does generate a significant degree of uncertainty in the results, and it was not clear in the manufacturers written submission how this uncertainty was captured in the model. It is noted though that utility estimates were used based on the disease and not the treatment, so all patients regardless of treatment attracted the same utility scores except where treatment response was greater.

The ERG also had concerns about the model functionality and usability when running the one-way and probalistic sensitivity analyses for verification. These have been noted in detail above, but centre around two key points. The first is the fact that not all variables included in the one-way sensitivity analyses could be modified by the ERG to assess the accuracy and robustness of the reported results. The second issue relates to the limited functionality of the model in generating comparisons between all relevant treatment options. Only pairwise comparisons were possible and there was no function for generating CEACs comparing all options at once. This omission is significant, as CZP should be compared with that treatment considered to be the best option on grounds of cost-effectiveness. Without the ability to evaluate all five treatments at once, it is more difficult to ascertain whether or not CZP is indeed a viable, cost-effective treatment for RA.

Another key issue relates to the changes in the model between the original submission and the responses to the ERGs requests for clarification; again, these changes have been noted above in detail. The manufacturers not only modified the model in response to our points for clarification, they also made changes beyond those requested by the ERG. This creates two difficulties. The first is that the model then becomes a de-facto new model and requires full appraisal but within a reduced span of time. The second is that there is no further opportunity to ask for clarification from the manufacturers relating to the changes in the model.

6 Additional work undertaken by the ERG

The ERG conducted a number of additional pieces of work which are documented throughout this report.

7 Discussion

It is clear that CZP is an effective treatment for RA. Relative to a passive control in mono-therapy or in combination therapy with MTX, like other anti TNF agents it improves signs and symptoms of disease, quality of life and slows the progressive destruction of joints. The major difficulties for the decision problem is the complete lack of head to head trials of CZP with any other bDMARD and the identification of the appropriate range of comparators. The manufacturer's approach was to implement indirect meta-analyses involving CZP, adalimumab, infliximab, etanercept, rituximab and tocilizumab. Amongst a total of 98 indirect comparisons undertaken only 4 produced statistically significant results. In the ERG's opinion no convincing evidence has been produced that demonstrates that any of these bDMARDs is superior or inferior to any other. The lack of power in the ICs meant that the manufacturer's approach has not advance a solution to the decision problem. Furthermore the manufacturer's method of incorporating effectiveness of the bDMARDs into the economic model may have been biased in favour of CZP, underestimated uncertainty and was susceptible to the influence of confounders. A more powerful approach that would advance a more reliable solution to the decision problem would be a mixed treatment analysis. This would best take into account prognostic factors of populations participating in the contributory trials used in the analysis.

The results generated by the manufacturer's economic model were driven by the effectiveness inputs for ACR20 ACR50 and ACR70 responses for the different bDMARDs and by the costs associated with each of these therapies. The manufacturer assumed equal harms for all bDMARDs also equal bDMARD-treatment duration. As the evidence for differences in bDMARD

effectiveness was limited, if equal effectiveness is assumed between bDMARDs and all other assumptions that entered the revised model hold then the model outputs are predominantly driven by treatment costs which are predominantly composed of drug costs.

7.1 Implications for research

There is a need for head to head RCTs comparing the effectiveness of bDMARDs. These are unlikely to be sponsored by industry and therefore they are unlikely to be readily undertaken. In the absence of such research a full mixed treatment analysis using a broad spectrum of randomised evidence would be useful in highlighting any differences in effectiveness between these drugs. Linking the findings of such analyses to investigation and exploration of cost-effectiveness of competing bDMARDs would serve a useful purpose.

8 APPENDICES

Appendix 1 Description of Common RA Outcome Measures

ACR outcome and its components (Felson 1995[19])

The American College of Rheumatology (ACR) have developed criteria that are commonly known as ACR scores (ACR20, ACR50, ACR70). The ACR20 is a change score which counts the percentage of patients who have improved by 20% in a combination of measures – (i) tender and swollen joint counts and (ii), improvement in three of the following five parameters:

- Acute phase reactant (C-RP, ESR)
- Patient assessment
- Physician assessment
- Pain scale
- Disability/functional questionnaire

The selection of three out of five parameters in each patient will vary over time because at each visit the best three are selected. As the ACR20 uses the best three from five parameters, the HAQ score (see below) may not contribute to the ACR20 at each follow up point.

The Health Assessment Questionnaire Disability Index (HAQ-DI)

The Health Assessment Questionnaire – Disease Index (HAQ or HAQ-DI). consists of 20 questions about physical functioning referring to 8 component sets: dressing/grooming, arising, eating, walking, hygiene, reach, grip, and activities. For each question, people indicate whether, over the past week, they were able to undertake the task without any difficulty, with some difficulty, with much difficulty or were unable to. HAQ-DI scores range from 0 to 3. Scores of 0 to 1 are considered to be mild to moderate difficulty, 1 to 2 are moderate to severe disability and 2 to 3 are severe to very severe disability.

An average HAQ score reported in a population based study was 0.49 and in RA was 1.2. In another study of 1109 French RA patients, the distribution was roughly normal with a median of 1.25 and mean of 1.32 (SD 0.77). The HAQ in RA patients and the general population tend to increase by age. However, the HAQ score rises dramatically in the general population after the age of 70 but the HAQ score changes for RA patients in this age group are even higher.

A minimum clinically important difference (MCIOD) in HAQ score in RAPID1 and RAPID2 was defined as ≥0.22 point change from baseline.

The DAS-28

The Disease Activity Score (DAS-28) is a continuous measure of disease activity in RA rather than the categorical response criteria of the ACR. It includes the proportion of tender and swollen joints (out of 28 joints), the ESR and general health or global disease activity from a 100mm VAS. These are fed into the following formula (usually using a computer package)

DAS28 = 0.56xsqrt(tender28) + 0.28xsqrt(swollen28) + 0.7xln(ESR) + 0.014xGH

The result is a number from 1 to10 where a higher number is worse disease. A DAS-28 score of 2.6 or less is considered to represent remission, 3.2 or less is low disease activity and 5.1 or more is high disease activity. However, in clinical practice, the DAS score may not be as useful tool for decision making in individual patient encounters.

Modified Total Sharp Score (mTSS)

Total Modified Sharp Score (mTSS) sums a joint erosion score (ES) and a joint space narrowing (JSN) score.

From the Trial report for RAPID1.

> This methodology quantifies the extent of bone erosions and joint space narrowing for 44 and 42 joints, respectively, with higher scores representing greater damage. Each radiograph was read centrally and independently by 2 of 3 experienced readers who were blinded for treatment, visit, and patient identification. The mean score of the readers was used for analysis.

The (ES) ranges from 0 –160 in the hands + wrists (2 x 16 joints scored 0 to 5) and 0 –120 in the feet (2 x 6 joints scored 0 to 10), giving a total of 280.

Joint space narrowing (JSN) scores range from 0–120 (2 x 15 joints scored 0 to 4 in hands/wrists) and 0–48 (2 x 6 joints scored 0 to 4 in feet), giving a total of 168. The mTSS is computed as the sum of the ES and JSN scores (range, 0–448), with higher scores representing greater damage.

Appendix 2 Details and comment on Manufacturer's search strategy for clinical effectiveness

Details of the strategies including the full search strategies run on MEDLINE EMBASE and the Cochrane Library are detailed in MS sections 6.1 and 10.2 MS appendix 2. A summary of the manufacturer's search strategy is presented in the box below

Databases searched for initial systematic review were :
MEDLINE and MEDLINE In Process 1966 to date
EMBASE 1980 to date
Cochrane Library (CENTRAL, CDSR, DARE) to date
These searches were conducted up to July 2007.
Databases searched to update the initial systematic review were:

MEDLINE and MEDLINE In Process (Ovid) 1950 to present
EMBASE (Ovid) 1980 to 2009 week 14
Cochrane Library (CENTRAL) (Wiley) to present
These searches were limited to the period between 2007 and 6 April 2009 when the update searches were conducted.
In addition to the above, conference abstracts from the European League against Rheumatism (EULAR) and American College of Rheumatology (ACR) were searched from 2004-2009.

The searches were limited to English language papers only.

Comments:

- The issue of the Cochrane Library searched is not stated but it may be assumed to be 2009 issue 2 as the search was conducted on 6 April 2009
- The strategy used on MEDLINE and MEDLINE In Process uses a combination of textwords and MeSH terms to describe the population (RA) and intervention/comparators (monoclonal antibodies and anti-TNF agents).
- Although a separate strategy might be expected for use with MEDLINE In Process, where only textwords are applicable, the one displayed should nonetheless have performed adequately across the two databases
- A methodological filter picks up clinical trials and excludes other study types.

- The strategy displayed for EMBASE is equally comprehensive and appropriate EMTREE equivalents to the MeSH terms in the MEDLINE search have been selected. Again studies are unlikely to have been missed
- The strategy for use on the Cochrane Library also expresses the population and intervention in terms consistent with the searches on MEDLINE and EMBASE.

Appendix 3 Additional searches conducted by the ERG

Ongoing trials search

Registers searched:
ClinicalTrials.gov, MetaRegister, UKCRN Portfolio

Terms used:
Cimzia or certolizumab and rheumatoid arthritis

Search for foreign language papers

Database: Ovid MEDLINE(R) <1950 to June Week 4 2009>
Search Strategy:
---
1 Randomized Controlled Trials as Topic/ (61577)
2 randomized controlled trial/ (275071)
3 random allocation/ (65034)
4 double blind method/ (102499)
5 single blind method/ (13081)
6 clinical trial/ (454772)
7 exp clinical trials as topic/ (218328)
8 or/1-7 (703312)
9 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).mp. [mp=title, original title, abstract, name of substance word, subject heading word] (138350)
10 placebos/ (28161)
11 placebo$.tw. (117578)
12 randomly allocated.tw. (11135)
13 (allocated adj2 random).tw. (630)
14 (clinic$ adj trial$1).tw. (132255)
15 or/9-14 (314369)
16 8 or 15 (787583)
17 case report.tw. (144153)
18 letter/ (657630)
19 historical article/ (274061)
20 review of reported cases.pt. (0)
21 review,multicase.pt. (0)
22 or/17-21 (1067015)
23 16 not 22 (765050)
24 TNFR-Fc fusion protein$.mp. (2402)
25 Interleukin 1 Receptor Antagonist Protein/ (2932)
26 tumor necrosis factors/ (1117)
27 tnf receptor fusion protein$.mp. (43)
28 anti tumo?r necrosis factor.mp. (1563)
29 anti TNF.mp. (3711)
30 anti interleukin.mp. (814)
31 certolizumab.mp. (80)
32 cimzia.mp. (4)
33 CDP870.mp. (84)
34 adalimumab.mp. (1193)
35 humira.mp. (66)
36 etanercept.mp. (2104)
37 enbrel.mp. (131)
38 infliximab.mp. (4808)
39 remicade.mp. (154)
40 monoclonal antibod$ ca2.mp. (23)
41 Mab ca2.mp. (3)
42 rituximab.mp. (5038)
43 rituxan.mp. (169)
44 mabthera.mp. (81)
45 tocilizumab.mp. (124)
46 actemra.mp. (4)
47 golimumab.mp. (23)
48 CNTO 148.mp. (1)
49 or/24-48 (19612)
50 rheumatoid arthritis.mp. or Arthritis, Rheumatoid/ (83434)
51 felty's syndrome.mp. or Felty's Syndrome/ (742)
52 caplan syndrome.mp. or Caplan Syndrome/ (145)

53 rheumatoid nodule.mp. or Rheumatoid Nodule/ (849)
54 Still's Disease, Adult-Onset/ or still$ disease.mp. (1253)
55 or/50-54 (85153)
56 23 and 49 and 55 (1161)
57 limit 56 to yr="2007 - 2009" (366)
58 limit 57 to english language (321)
59 57 not 58 (45)

Database: EMBASE <1980 to 2009 Week 27>
Search Strategy:
---
1 clinical trial/ (546535)
2 randomized controlled trial/ (170607)
3 randomization/ (26914)
4 single blind procedure/ (8290)
5 double blind procedure/ (73001)
6 crossover procedure/ (21495)
7 placebo/ (128352)
8 randomi?ed controlled trial$.tw. (33960)
9 Rct.tw. (2825)
10 random allocation.tw. (641)
11 Randomly allocated.tw. (10349)
12 Allocated randomly.tw. (1359)
13 (allocated adj2 random).tw. (562)
14 single blind$.tw. (7581)
15 Double blind$.tw. (85786)
16 ((triple or treble) adj blind$).tw. (141)
17 placebo$.tw. (111617)
18 Prospective study/ (83374)
19 or/1-18 (717903)
20 case study/ (6178)
21 case report.tw. (121063)
22 abstract report/ or letter/ (503495)
23 or/20-22 (628355)
24 19 not 23 (692867)
25 TNFR-Fc fusion protein$.mp. (9)
26 interleukin 1 receptor blocking agent.mp. or interleukin 1 receptor blocking agent/ (7016)
27 recombinant interleukin 1 receptor blocking agent/ (1843)
28 Tumor necrosis factor/ (20124)
29 tumor necrosis factor antibody/ (2644)
30 tnf receptor fusion protein$.mp. (41)
31 anti tumo?r necrosis factor.mp. (1527)
32 anti tnf.mp. (3447)
33 anti interleukin.mp. (789)
34 certolizumab.mp. or certolizumab pegol/ (632)
35 cimzia.mp. (108)
36 CDP870.mp. (22)
37 adalimumab.mp. or adalimumab/ (4110)
38 humira.mp. (1082)
39 etanercept.mp. or etanercept/ (8343)
40 enbrel.mp. (1826)
41 infliximab.mp. or infliximab/ (12093)
42 remicade.mp. (2341)
43 monoclonal antibod$ ca2.mp. (35)
44 Mab ca2.mp. (3)
45 rituximab.mp. or rituximab/ (12605)
46 rituxan.mp. (1443)
47 mabthera.mp. (994)
48 tocilizumab.mp. (125)
49 actemra.mp. (53)
50 golimumab.mp. or golimumab/ (96)
51 CNTO 148.mp. (29)
52 or/25-51 (54454)
53 rheumatoid arthritis.mp. or rheumatoid arthritis/ (71601)
54 felty's syndrome.mp. or felty syndrome/ (497)
55 caplan syndrome.mp. or pneumoconiosis/ (2135)
56 rheumatoid nodule.mp. or rheumatoid nodule/ (663)
57 still$ disease.mp. or Adult onset still disease/ (1012)
58 or/53-57 (74307)
59 24 and 52 and 58 (3289)
60 limit 59 to yr="2007 - 2009" (1109)
61 limit 60 to english language (1001)
62 60 not 61 (108)

Cost-effectiveness search

Cochrane Library (Wiley) NHS EED 2009 Issue 3

#1 rheumatoid next arthritis
#2 MeSH descriptor Arthritis, Rheumatoid, this term only
#3 (#1 OR #2)
#4 "tumor necrosis factor*"
#5 "tumour necrosis factor*"
#6 MeSH descriptor Receptors, Tumor Necrosis Factor, this term only
#7 "anti tnf"
#8 antitnf
#9 infliximab
#10 remicade
#11 enbrel
#12 etanercept
#13 adalimumab
#14 humira
#15 certolizumab
#16 "certolizumab pegol"
#17 cimzia
#18 rituximab
#19 rituxan
#20 tocilizumab
#21 actemra
#22 (#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21)
#23 (#3 AND #22)
#24 (#23), from 2005 to 2009
#25 (#15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21)
#26 (#3 AND #25)
#27 (#24 AND #26)
#28 (#24 OR #26)

Appendix 4 Manufacturer's response to query about non-English language papers

a. Please state if searches for ongoing studies were undertaken. Please provide details.

From the clarification document:

Searches for ongoing studies were not undertaken, as these would be unlikely to provide finalised results of similar quality to the fully published studies included in the systematic review. However hand searching was undertaken in order to capture results that were finalised but not yet fully published. There are no ongoing trials that UCB are aware of, funded by UCB or others, involving CZP in RA.

Appendix 5 Details of CZP trials

Table 6 Details of RAPID 1

Study: RAPID 1 (Keystone et al 2008)

Design: RCT
Patients randomised 2:2:1 (randomisation not described) to one of CZP+MTX arms or PL+MTX
Double-blind (each radiograph was read by two independent, blind readers)

Location: 147 centres worldwide

Duration: 52 weeks (February 2005 – October 2006)

Sponsor: UCB

Participant numbers:

	CZP 200mg + MTX	CZP 400mg + MTX	PL + MTX
Screened	1262		
Randomised*	992 (based on confidential report)		
Completed (week 52) (% relative to analysed population)	255 (64.9%)	274 (70.3%)	43 (21.6%)
Analysed – ITT	393	390	199

* 10 patients were excluded from analysis due to major abuses found during audit

At week 16 patients withdrew because of lack of efficacy: 62.8% PL, 21.1% CZP 200mg, 17.4% CZP 400 mg
One patient in each group was lost to follow-up

Table 10:1 Summary of Patients Withdrawn by Reason for Withdrawal - ITT Population

Reason for Withdrawal	Placebo + MTX (N = 199)	CZP 200 mg + MTX (N = 393)	CZP 400 mg + MTX (N = 390)	Overall (N = 982)
Total number of patients withdrawn	156 (78.4%)	138 (35.1%)	116 (29.7%)	410 (41.8%)
Lack of efficacy	141 (70.9%)	98 (24.9%)	74 (19.0%)	313 (31.9%)
Adverse event	3 (1.5%)	17 (4.3%)	22 (5.6%)	42 (4.3%)
Protocol non-compliance	0	4 (1.0%)	3 (0.8%)	7 (0.7%)
Patient decision	10 (5.0%)	15 (3.8%)	11 (2.8%)	36 (3.7%)
Lost to follow-up	1 (0.5%)	1 (0.3%)	1 (0.3%)	3 (0.3%)
Other	3 (1.5%)	5 (1.3%)	6 (1.5%)	14 (1.4%)

Note: More than 1 reason for withdrawal may be recorded for a patient.
Source: Table 14.1.1:2

Inclusion criteria:
- age ≥18 years
- RA defined by ACR criteria
- RA duration ≥6 months and < 15 years
- failed ≥1 prior DMARD due to lack of efficacy or intolerance
- have received MTX for ≥6 months, with a stable dosage of ≥10mg/week for ≥2 months prior to baseline
- active disease at screening and baseline:
  1. ≥9 (out of 68) tender joints
  2. ≥9 (out of 66) swollen joints
  3. and ≥1 of:
     - ESR ≥30mm/h
     - CRP > 15mg/l

Exclusion criteria:
- any other inflammatory arthritis or secondary noninflamatory arthritis that could have interfered with evaluation of the effects of CZP
- high risk of infection in the opinion of the investigator
- history of malignancy, dymelyelinating disease, blood dyscarsias, or severe, progressive, and/or uncontrolled renal, hepatic, hematologic, gastrointestinal, endocrine, pulmonary, cardiac, neurologic, or cerebral disease
- history of or a chest x-ray showing active or latent TB
- positive purified protein derivative skin test (by local standard) unless this was associated with Bacille Calmette-Guérin vaccination and there was no clinical or radiographic suspicion of TB (at discretion of investigator)
- biologic therapies for RA within 6 months (or etanercept and/or anakinra within 3 months) of

<table>
<tr><td colspan="2">Study: RAPID 1 (Keystone et al 2008)</td></tr>
<tr><td></td><td>baseline
o any previous biologic therapy that resulted in a severe hypersensitivity or anaphylactic reaction
o prior failure to respond to treatment with a TNF-α inhibitor</td></tr>
<tr><td>Intervention: lyophilised subcutaneous CZP 400mg at weeks 0, 2, and 4 followed by 200mg or 400mg every 2 weeks thereafter + MTX (continued at the same dosage they were taking at study entry)</td><td>Comparator: placebo (saline) + MTX (continued at the same dosage they were taking at study entry)</td></tr>
<tr><td colspan="2">Other drugs:
o DMARDs (exluding MTX) had to be discontinued 28 days prior to baseline except for LEF which had to be discontinued 6 months prior to baseline unless a cholestyramine washout was performed
o concomitant oral corticosteroids (prednisone ≤10mg/day or equivalent, stable for 4 weeks prior to baseline and throughout the study) were allowed,
o non-steroidal anti-inflammatory drugs/ cyclooxygenase 2 inhibitors and analgesics were allowed
o parenteral corticosteroids were prohibited
o concomitant medications were monitored at each visit according to protocol</td></tr>
<tr><td colspan="2">Primary outcome:
o ACR 20 response rate at week 24
o change from baseline in the modified total Sharp score at week 52</td></tr>
<tr><td colspan="2">Secondary outcomes included (more outcomes might have been measured):
o change from baseline in the modified total Sharp score at week 24
o change from baseline in the disability index (DI)
o HAQ at weeks 24 and 54
o ACR 20 responder rate at week 52
o ACR 50/70 responder rate at week 24 and 52
o Mean change from baseline in erosion and joint space narrowing scores
o Swollen (66 joints) and tender (68 joints) joint counts
o Physician’s and patient’s global assessments of disease activity
o Patient’s assessment of arthritis pain
o Physical function (according to HAQ DI)
o DAS 28
o ESR
o CRP
o Proportion of patients achieving clinically meaningful improvements in physical function (minimum clinically important difference defined as a decrease in HAQ DI of ≥0.22 points from baseline)
o Safety assessment included: measurement of vital signs, physical examination, hematologic analysis, serum biochemical analysis and urinanalysis;
o Systolic and diastolic blood pressure was measured at each visit before and after the injection of the study drug
o Adverse events were monitored at every study visit; treatment-emergent adverse events were defined as occurring after the first administration of the study drug and up to 12 weeks after the last dose was administered; adverse events were classified according to Medical Dictionary for Regulatory Activities (version 9.0) by primary system organ class and preferred term
o Plasma concentrations of anti-CZP antibodies were measured by an enzyme-linked immunosorbent assay; levels >2.4 units/ml were considered positive</td></tr>
<tr><td colspan="2">Assessment time:
o NA</td></tr>
<tr><td colspan="2">Sample size calculation:
o Based on both primary outcomes; larger required sample size used
o For ACR20: 590 patients to have 90% power to detect a difference ≥ 20% between CZP groups and PL with a 2-sided significance level of 2.5%
o For modified total Sharp score: 950 patients to have 90% power to detect a difference ≥ 2.2 Sharp units between CZP groups and PL with SD assumed to be 7 Sharp units</td></tr>
<tr><td colspan="2">Analysis:
o efficacy: ITT (all patients randomised to the study)
o patients who received rescue medication or withdrew for any reason, including safety were considered nonresponders (ACR 20?) from that time point onwards
o for patients who withdrew before week 52 and who had their radiographs taken at the withdrawal visit, the modified total Sharp score was estimated by linear extrapolation of the scores on the radiographs taken at their withdrawal visit or, if this was not performed, at week 24
o sensitivity analyses were also carried out (including LOCF)
o safety: all patients who received at least 1 dose of medication</td></tr>
<tr><td colspan="2">Comments
o Patients who failed to achieve ACR 20 at weeks 12 and 14 were designated treatment failures and withdrawn from the study at week 16</td></tr>
</table>

Study: RAPID 1 (Keystone et al 2008)

- Patients who withdrew early due to reasons other than consent withdrawal, underwent mandatory radiographic assessment at the time of withdrawal and at week 52

Table 7 Details of RAPID 2

Study: RAPID 2 (Smollen et al 2008)

Design: RCT
Patients randomised 2:2:1 (randomisation not described) to one of CZP+MTX arms or PL+MTX
Double-blind (each radiograph was read by two independent, blinded readers)

Location: 76 international sites

Duration: 24 weeks (June 2005 to September 2006)

Sponsor: UCB

Participant numbers:

	CZP 200mg + MTX	CZP 400mg + MTX	PL + MTX
Screened	733		
Randomised	619		
Completed (week 24) (% relative to analysed population)	174	181	17
Analysed – ITT	246	246	127

Patients who did not demonstrate an ACR20 response at both weeks 12 and 14 were withdrawn from the study and designated nonresponders

Table10:1 Summary of Patients Withdrawing by Reason for Withdrawal - ITT Population

Reason for Withdrawal [a]	Placebo + MTX (N=127)	CZP 200 mg + MTX (N=246)	CZP 400 mg + MTX (N=246)	Overall (N=619)
Total number of patients withdrawn	110 (86.6%)	72 (29.3%)	65 (26.4%)	247 (39.9%)
Adverse event	2 (1.6%)	11 (4.5%)	6 (2.4%)	19 (3.1%)
Lack of efficacy	107 (84.3%)	54 (22.0%)	53 (21.5%)	214 (34.6%)
Lost to Follow-up	0	0	0	0
Protocol violation	1 (0.8%)	1 (0.4%)	2 (0.8%)	4 (0.6%)
Withdrawal of consent	0	5 (2.0%)	3 (1.2%)	8 (1.3%)
Other reason	0	1 (0.4%)	1 (0.4%)	2 (0.3%)

[a] More than 1 reason for withdrawal may be recorded for a patient.
Note: Percentages are based on the ITT Population.
Source: Table 14.1.1:2

Most of the patients withdrawn due to lack of efficacy were withdrawn at week 16: 103 in the placebo arm and 52 in each CZP arm

Inclusion criteria:
- age ≥18 years
- adult onset RA defined by ACR criteria
- RA duration ≥6 months and < 15 years
- failed ≥1 prior DMARD due to lack of efficacy or intolerance
- have received MTX for ≥6 months, with a stable dosage of ≥10mg/week for ≥2 months prior to baseline
- active disease at screening and baseline:
  4. ≥9 (out of 68) tender joints
  5. ≥9 (out of 66) swollen joints
  6. and ≥1 of:
     - ESR ≥30mm/h
     - CRP > 15mg/l

Exclusion criteria:
- any other inflammatory arthritis or secondary noninflamatory arthritis that could have interfered with evaluation of the effects of CZP
- high risk of infection in the opinion of the investigator
- history of malignancy, dymelyelinating disease, blood dyscarsias, or severe, progressive, and/or uncontrolled renal, hepatic, hematologic, gastrointestinal, endocrine, pulmonary, cardiac, neurologic, or cerebral disease
- history of or a chest x-ray showing active or latent TB
- positive purified protein derivative skin test (by local standard) unless this was associated with Bacille Calmette-Guérin vaccination and there was no clinical or radiographic suspicion of TB (at discretion of investigator)
- biologic therapies for RA within 6 months (or

<table>
<tr><td colspan="2">Study: RAPID 2 (Smollen et al 2008)</td></tr>
<tr><td></td><td>etanercept and/or anakinra within 3 months) of baseline
o any previous biologic therapy that resulted in a severe hypersensitivity or anaphylactic reaction
o had not initially responded to a prior TNF-α inhibitor</td></tr>
<tr><td>Intervention: subcutaneous liquid CZP 400mg at weeks 0, 2, and 4 followed by 200mg or 400mg every 2 weeks thereafter + MTX</td><td>Comparator: placebo (saline) + MTX</td></tr>
<tr><td colspan="2">Other drugs:

o DMARDs (excluding MTX) had to be discontinued 28 days prior to baseline except for LEF which had to be discontinued 6 months prior to baseline unless a cholestyramine washout was performed
o concomitant oral corticosteroids (prednisone ≤10mg/day or equivalent prior to baseline and throughout the study) were allowed provided the doses were stable within 28 days of baseline and remained stable during the study,
o non-steroidal anti-inflammatory drugs/ cyclooxygenase 2 inhibitors were allowed provided the doses were stable within 14 days of baseline and remained stable during the study
o parenteral corticosteroids were prohibited
o concomitant medications were monitored at each visit according to protocol</td></tr>
<tr><td colspan="2">Primary outcome:
o ACR 20 response rate at week 24</td></tr>
<tr><td colspan="2">Secondary outcomes included (more outcomes might have been measured):
o ACR 50/70 responder rate
o change from baseline in the modified total Sharp score
o SF-36 Health Survey
o Individual ACR core set variables
o DAS 28
o HAQ-DI
o safety</td></tr>
<tr><td colspan="2">Assessment time:
o Baseline, weeks 1, 2, 4, 6, 12, 14, 10 and 24, or withdrawal</td></tr>
<tr><td colspan="2">Sample size calculation:
o Based on ACR20
o For ACR20: 590 patients to have 90% power to detect a difference ≥ 20% between CZP groups and PL with a 2-sided significance level of 2.5%, assuming a placebo rate of 30%;</td></tr>
<tr><td colspan="2">Analysis:
o efficacy: ITT (all patients randomised to the study)
o in primary analyses “nonresponder imputation” was used: patients who received rescue medication or were withdrawn from the study were considered nonresponders from that time point onward;
o “the number of subjects in the summaries varies slightly from the ITT numbers due to nonimputable missing data”
o sensitivity analyses were also carried out (including LOCF)
o safety: all patients who received at least 1 dose of medication</td></tr>
<tr><td colspan="2">Comments
Patients who did not achieve ACR 20 response at both weeks 12 and 14 were designated treatment failures and withdrawn from the study</td></tr>
</table>

Table 8 Details of FAST4WARD

Study: FAST4WARD (Fleischmann et al 2008)

Design: RCT
Randomisation 1:1 using an interactive voice randomisation service
Study personnel distributing the drugs was blinded

Location: 36 sites in Austria, Czech Republic and US

Duration: 24 weeks; June 2003-2004

Sponsor: UCB

Participant numbers:

	CZP	PL
Eligible	NA	
Randomised	111	109
Completed (week 24)	76	28
Analysed – mITT	111 (110 for ACR 20)	109 (108 for ACR 20)

21.6% CZP patients and 68.8% PL patients withdrew due to lack of efficacy (p<0.001)

Based on the confidential trial report, the reasons for discontinuation were:

	Discontinuation before week 12	Discontinuation after week 12 and before week 24
CZP	23 (20.72%)	12 (10.81%)
Lack of efficacy	15	9
Adverse events	4	1
Consent withdrawal	2	0
Protocol violation	2	2
PL	59 (54.13%)	22 (20.18%)
Lack of efficacy	55	21
Adverse events	2	0
Protocol violation	0	1
Lost to follow-up	2	0

Table 10:2 Summary of Patient Accountability (All Randomized Patients)

	Placebo (N = 109) n (%)	CDP870 400 mg (N = 111) n (%)	Overall (N = 220) n (%)
Randomized but did not take study medication	0	0	0
Randomized and received at least one dose of study medication (mITT)	109 (100.0%)	111 (100.0%)	220 (100.0%)
Completed study up to Week 12	50 (45.9%)	88 (79.3%)	138 (62.7%)
Completed study at Week 24	28 (25.7%)	76 (68.5%)	104 (47.3%)
Withdrawn from study	81 (74.3%)	35 (31.5%)	116 (52.7%)
Reason for withdrawal from study			
AE	2 (1.8%)	5 (4.5%)	7 (3.2%)
Protocol violation	1 (0.9%)	4 (3.6%)	5 (2.3%)
Lost to follow-up	3 (2.8%)	0	3 (1.4%)
Consent withdrawn	0	2 (1.8%)	2 (0.9%)
Lack of efficacy	75 (68.8%)	24 (21.6%)	99 (45.0%)
Sponsor's Decision	0	0	0
Patient completed study per the protocol	81 (74.3%)	89 (80.2%)	170 (77.3%)

Source: Table 14.1.1:2, Listing 16.2.1:1 and Listing 16.2.3:1

Inclusion criteria:
- age 18-75 years
- adult-onset RA defined by ACR criteria
- RA duration ≥6 months
- failed ≥1 prior DMARD due to lack of efficacy or intolerance

Exclusion criteria:
- inflammatory arthritis other than RA
- history of serious or life-threatening infection
- any current infection
- history of or a chest x-ray suggesting TB
- positive (defined by local practice) purified protein

<table>
<tr><td colspan="2">Study: FAST4WARD (Fleischmann et al 2008)</td></tr>
<tr><td>o active disease at screening and baseline:
7. ≥9 (out of 68) tender joints
8. ≥9 (out of 66) swollen joints
9. and ≥1 of:
o ≥45 minutes of morning stiffness
o ESR ≥28mm/h
o CRP > 10mg/l</td><td>derivative skin test (unless they had received Bacille Calmette-Guérin vaccination and had a chest x-ray suggesting no TB and no clinical symptoms)
o biologic therapies for RA within 6 months
o prior treatment with TNF-α inhibitors</td></tr>
<tr><td>Intervention: lyophilised subcutaneous CZP 400mg every 4 weeks from baseline to week 20</td><td>Comparator: placebo (sorbitol solution) every 4 weeks from baseline to week 20</td></tr>
<tr><td colspan="2">Other drugs:
o DMARDs discontinued for ≥28 days or 5 half-lives of the drug (whichever was longer) prior to first study dose except for LEF which was eliminated using cholestyramine administration and a 28 day washout
o concurrent oral corticosteroids (prednisone equivalent ≤10mg/day, stable for ≥4 weeks prior to enrolment and during the study), non-steroidal anti-inflammatory drugs and analgesics were allowed
o intra-articular, peri-articular, intramuscular and intravenous corticosteroids were prohibited</td></tr>
<tr><td colspan="2">Primary outcome: ACR20 response at week 24</td></tr>
<tr><td colspan="2">Secondary outcomes included (it seems more outcomes might have been measured):
o ACR 50/70 response
o ACR component scores
o DAS
o ESR-3
o HAQ-DI
o SF-36
o pain (measured with 100mm VAS)
o modified Brief Pain Inventory (mBPI)
o 11-point Fatigue Assessment Scale (FAS)
o Safety</td></tr>
<tr><td colspan="2">Assessment time:
o efficacy and safety at baseline and weeks: 1, 2, 4, 8, 12, 16, 20 and 24
o additional safety assessment 4 and 12 weeks after final dose – week 24 and 32</td></tr>
<tr><td colspan="2">Sample size calculation:
o expected 25% ACR20 responders in PL and 50% in CZP arm at week 24
o ≥90% power at 5% 2-sided significance level
o sample of 100 per group</td></tr>
<tr><td colspan="2">Analysis:
o efficacy: modified ITT: all randomised patients who have taken ≥1 dose of study medication
o safety population: all randomised patients who have taken ≥1 dose of study medication
o it seems that patients with missing data were sometimes excluded from analysis
o patients who withdrew for any reason were considered ACR20/50/70 non-responders
o where possible, last observations were carried forward in a sensitivity analysis</td></tr>
<tr><td colspan="2">Comments
Minimal clinically important differences at week 24 used in a post-hoc analysis were defined as:
o HAQ-DI: ≥0.22 point decrease from baseline
o arthritis pain: ≥10 point decrease
o SF-36 domain: ≥5 point increase in individual domains
o Physical and Mental Component Summary scores: ≥5 point increase
o FAS ≥1 point decrease

Although authors claim that 100% of participants were included in the modified ITT population, data from 108 (PL) and 110 (CZP) patients was used to calculate ACR20/50 response (based on confidential submission); possibly little impact on results</td></tr>
</table>

Appendix 6 Further Trial Details

Table 9 Results presented in submission deriving from ACR data

OUTCOME	Time point(s) and Measure reported	MS reference	TRIAL (dose)	Comment
ACR20 ACR50 ACR70	Wk 24 & wk 52; Odds Ratio: CZP v. control [by dose]	Table 8	RAPID1 (200mg & 400mg) RAPID 2 (200mg & 400mg)	
ACR20 ACR50 ACR70	Wk 24; % response in CZP & placebo arms. P values for comparison	Table 9	FAST4WARD	Modified ITT analysis
ACR20 ACR50 ACR70	Wk 24 & wk 52; Meta-analysis, Relative Risk CZP v. control	Table 20	RAPID 1 + RAPID2 pooled (200 mg)	
ACR20 ACR50 ACR70	Wk 24 & wk 52; Meta-analysis, Risk Difference CZP v. control	Table 21	RAPID 1 + RAPID2 pooled (200 mg)	
ACR20 ACR50 ACR70	Wk 24; forest plots of Relative Risk	Figs: 27, 28 & 29	RAPID 1 + RAPID2 pooled (200 mg)	
ACR20	Every 2 wks to wk 24; % response in CZP + MTX & MTX + placebo arms	Fig: 9	RAPID 2 (200mg & 400mg)	Indicates rapid response to CZP
ACR20	Every 2 wks to wk 52; % response CZP + MTX &MTX + placebo arms	Fig: 8	RAPID 1 (200mg & 400mg)	Indicates rapid response to CZP
ACR20	Wk 1 & 2 then every 4 wks from wk 4 to wk 52; % response CZP & placebo arms	Fig: 8	FAST4WARD	Indicates rapid response to CZP
ACR core components	Early wks 1, 4 to 12 Wk 24 & wk 52: mean change from baseline in each trial arm	Text narrative Table 10	RAPID 1 RAPID 1	Indicates rapid response to CZP
ACR core components	Wk 1, Wk 24: mean change from baseline in each trial arm	Text narrative Table 11	RAPID 2 RAPID 2	Indicates rapid response to CZP

Appendix 7 ERG formal critical appraisal of review underpinning Clinical evidence section of submission

PURPOSE OF THIS APPRAISAL:
Critical appraisals attempt to identify the strengths and weaknesses of pieces of information, often research literature, so that readers may apply that information within the limits identified. There are two important sets of limits:

a) the INTERNAL validity of the information ie how the information was collected and/or summarised
b) the EXTERNAL validity of the information ie how relevant the information is to any specific question posed by a reader

This appraisal checklist is specifically designed for reviews of research information. It is based on: Oxman AD. Checklists for review articles. BMJ, 1994;309:648-51. Updated version in: Chalmers I & Altman DG (eds). *Systematic reviews.* London: BMJ Publishing, 1995. This has in turn been modified on the basis of ARIF's experience reviewing many different types of reviews of research retrieved in its responses to requests for research information on the effects/effectiveness of health care interventions.

Implicit in the checklist is our belief that the following elements of a review are particularly important:

- Clear, explicit statement of method (in sufficient detail that another person undertaking the same review might be able to repeat the processes and arrive at the same conclusion AND that a reader can make an assessment of any bias that the reviewer has introduced in the way that the research was identified and summarised).
- Comprehensive ascertainment of all the available research literature relevant to the question the reviewer sets out to answer.
- Processing the ascertained literature in a way which reduces bias or makes explicit any bias which has been introduced, so that the reviewers or the reader can make allowance for this in their conclusions.
- An appropriate numerical summary of the size of any effect (or equivalent), including its confidence intervals.

If a review meets the first three general criteria a review would be a "systematic review"; if a review met all four criteria it would be a "systematic review with meta-analysis".

ASSESSOR'S SCREENING QUESTIONS

ON FIRST READING IS THERE SUFFICIENT INFORMATION TO MAKE A DETAILED APPRAISAL?

Yes

IN RELATION TO WHAT QUESTION IS THIS REVIEW BEING APPRAISED (TARGET QUESTION)?

This appraisal relates to the "Clinical Evidence" component of the report. Other sections of the report, particularly the "Cost effectiveness" section are not dealt with here.

State question, in terms of:

Question type - effects/effectiveness

Population/condition - adults who have active RA that has not responded adequately to DMARDs, including MTX

Intervention – certolizumab pegol

Comparator – management strategies involving DMARDs without certolizumab pegol, including treatment with: conventional DMARDs (for example sulfasalazine, leflunomide), biological agents (including adalimumab, etanercept, infliximab, rituximab, tocilizumab)

Outcomes - disease activity, physical function, joint damage, pain, mortality, fatigue, radiological progression, extra-articular manifestations of disease, adverse effects of treatment, health related quality of life

All as specified in the final scope as set out by NICE

HAS A CLEAR QUESTION BEEN DEFINED (REVIEW QUESTION)?

State specific question, to which the further assessment of this review relates, in terms of:

Question type - effects/effectiveness
Population/condition - adults who have active moderate to severe RA that has not responded adequately to DMARDs, including MTX

Intervention – certolizumab pegol

Comparator – comparison to both conventional and biological DMARDs; direct comparison to methotrexate and indirect to other biological DMARDs (adalimumab, etanercept, infliximab, rituximab, tocilizumab)

Outcomes – disease activity, physical function, joint damage, pain, mortality, fatigue, radiological progression, adverse effects of treatment, health related quality of life

****

Comments relating to internal validity:
Review question was defined based on the “decision problem addressed in the submission” (section 2, p. 7).

Comments relating to external validity:
Some of the drugs that would have been included based on the final scope (conventional DMARDs other than MTX, other biological DMARDS) were not considered in this review. The

submission did not evaluate extra-articular manifestations of disease as the *"pivotal clinical trials did not routinely report this particular outcome".*

****

WHAT ARE THE IMPLICATIONS FOR THE VALIDITY OF THE REVIEW OF THE TYPE AND RANGE OF STUDY DESIGNS INCLUDED?

State type/types of study designs which were included:

Randomised controlled trials

****

Comments:

The general validity of the included study designs to answer the review question was high.

****

WERE INCLUSION/EXCLUSION CRITERIA CLEARLY STATED?

From 6.2.2 p 24

List any INCLUSION criteria:

- Studies must be published randomised controlled trials
- Studies must be conducted in human adult patients (≥18 years) with active RA who have had an inadequate response to prior DMARD therapy, including MTX.
- Studies must contain one of the following interventions administered with or without MTX: adalimumab, CZP, etanercept, infliximab, rituximab, tocilizumab.
- The treatment comparison must be to another biological DMARD, a conventional DMARD or placebo.
- The study must report results for an outcome of interest, namely: patient's global assessment of disease activity; DAS28 scores; ACR responses; EULAR responses; HAQ-DI; swollen joint count; tender joint count; patient's global assessment of pain; patient's assessment of fatigue; mTSS (van der Heijde); EQ-5D scores; mortality; hospitalisation; joint replacement surgery; rheumatoid nodules; vasculitis; neuropathy; any adverse event; any serious adverse event; infusion reactions/acute hypersensitivity reactions; injection site reactions; tuberculosis; serious infections; malignancies/lymphoma; withdrawals due to adverse events.
- The interventions of interest must be studied at a licensed dose (EMEA or FDA). If a study included more than one treatment arm of the intervention of interest, one of them must be a licensed dose.

List any EXCLUSION criteria:

- Observational studies, reviews and open label extensions of RCTs
- Non-English language papers
- Dose ranging studies and studies where the comparator was an untreated control group

****

Comments:

The inclusion/exclusion criteria were consistent with the implicit review question and were stated with sufficient clarity.

It is not clear why authors did not search for or include unpublished studies and studies published in languages other than English.

****

WAS THE SEARCH STRATEGY ADOPTED LIKELY TO HAVE MISSED MANY POTENTIALLY RELEVANT STUDIES?

State the search strategy:

Detailed in Appx 2, p 180

The searches were conducted in two steps. First databases searched up to July 2007:

- Medline and Medline in progress (1966 to date)
- EMBASE (1980 to date)
- Cochrane Library (CENTRAL, CDSR, DARE to date)

Later an update to these searches was performed from 2007 to 6th April 2009:

- Medline, Medline in-process and other non-indexed citations (1950 to present) via Ovid
- Embase 1980 to 2009 week 14 via Ovid
- Cochrane Central Trials Register via Wiley Interscience

Abstracts from conferences were also searched:

- European League against Rheumatism (EULAR), 2004 to 2008, via the website http://www.eular.org/
- American College of Rheumatology (ACR), 2006 to 2008, via the website http://www.rheumatology.org/

"If a study was found published as a conference abstract and had no link to a primary publication from a journal, a full publication of the study was searched via PubMed."

****

Comments:

The search strategy was generally consistent with the review question.

It is unclear if a study was considered for inclusion/exclusion if it was identified only as a conference abstract, but no full publication was found.

British Society for Rheumatology website was not searched.

It appears that in the second step searches in CDSR and DARE were not updated. It is also unclear why these databases were searched if the review did not aim at identifying secondary research.

The limited steps to identify unpublished studies were a concern. On-going trials were not searched for. Results of hand searching were included in the QUOROM diagram (3 references). However hand searching was not mentioned in the description of searches.

****

HOW WERE INCLUSION/EXCLUSION CRITERIA APPLIED?

QUOROM flow diagram provided p. 27

In both stages of searching 1618 citations were identified. 263 duplicates and 1220 citations were excluded (review did not state if this was done based on title and abstract or full publication). This resulted in 135 references considered *"suitable for data extraction"*: 37 primary RCTs (3 relating to CZP) and 98 linked studies (23 relating to CZP).

It is unclear why the FAST4WARD study was included, as authors state that it investigated an unlicensed dose of certolizumab pegol (p. 36).

State whether a list of excluded studies was available and whether any excluded articles were examined:

A list of excluded studies together with reasons for exclusion was provided (table 22, p.75).

WAS THE VALIDITY OF INCLUDED STUDIES ASSESSED?

a) *Validity implicit in inclusion/exclusion criteria*
b) *Validity of all included studies re-examined*
c) *Both*
d) *Apparently not assessed at all*

b) of the above (but it seems only for trials on certolizumab pegol)

State whether the criteria used to assess validity were reasonable OR

Whether a recognised validity checklist was employed (that is one which has had its validity assessed):

Only RCTs were included in this review. Quality of the included studies was assessed using explicit and appropriate criteria. It however seems that quality assessment was carried out only for trials on certolizumab pegol.

State how information on the validity of the included studies was used:

a) *To provide narrative or tabulated information on the strengths and weaknesses of the included studies*
b) *As a check on the nature of the included studies to identify "late exclusions" (potentially inappropriate)*
c) *As a check on the nature of the included studies to identify wide variation in characteristics, suggesting that meta-analysis was not appropriate*
d) *Where meta-analysis was employed, to conduct sensitivity analyses to check robustness of findings*
e) *Other (please state)*
f) *Apparently not at all (potentially inappropriate if variation in important characteristics of included studies was likely)*

a) of the above

****

Comments:

Included studies were restricted to RCTs.

Quality assessment was undertaken probably only for certolizumab pegol trials.

****

WAS THE PROCESS OF DATA ABSTRACTION ADEQUATE?

State how the relevant data items were extracted:

a) Reference to pre-determined list
b) Use of data abstraction sheet
c) Other (please state)
d) No detail

d) of the above

****

Comments:

The data abstracted was consistent with the review question.

****

WERE THE IMPORTANT STEPS IN THE REVIEW REPRODUCIBLE & BIAS FREE?

State whether the repeatability of the following steps was examined, reported and acted upon:

Searching for all potentially relevant studies – search strategies were provided, but no information on repeatability was given.

Applying study inclusion/exclusion criteria – No, it was only reported that two reviewers applied inclusion/exclusion criteria and a third reviewer resolved any disagreements, no indication if any forms or checklists were used was provided. The level of agreement between the two reviewers was not stated.

Assessment of validity of included studies – No information was provided on how many reviewers assessed quality or what methods were used. Only the results of validity assessment (only for CZP trials) were provided.

Data abstraction – No detail of data abstraction forms was provided. Two reviewers independently extracted data from trials. Disagreements were resolved by a third reviewer.

****

Comments:

The repeatability of important steps in the review was not tested and reported. Number of reviewers involved in each stage of the review was unclear.

****

WHAT WAS(WERE) THE RELEVANT AND JUSTIFIABLE REVIEW BOTTOM LINE(S) - AS STATED IN THE REVIEW ?

The review inclusion criteria refer to six drugs, but results were presented in sufficient detail only for certolizumab pegol.

State whether meta-analysis was used:

Meta-analysis was used to combine results of trials on all included drugs when there was more than one trial. Forest plots were presented only for certolizumab pegol trials. Drugs were also compared using indirect analysis.

The main results are reported below.

Table 10 Summary of relative risk ratio meta-analysis results (as in the submission)

Outcome	Treatment	Control	N	Effect measure	Statistical heterogeneity (I^2)	Fixed effects			Random effects		
						Effect size	CI	P value	Effect size	CI	P value
ACR20 response at 6 months	CZP + MTX	MTX + placebo	965	Risk ratio	34	5.00	(3.7 to 6.67)	<0.001	5.00	(3.33 to 7.69)	<0.001
ACR50 response at 6 months	CZP + MTX	MTX + placebo	965	Risk ratio	44	5.88	(3.85 to 9.09)	<0.001	6.25	(3.13 to 12.5)	<0.001
ACR70 response at 6 months	CZP + MTX	MTX + placebo	965	Risk ratio	0	9.09	(4.17 to 20)	<0.001	8.33	(3.85 to 16.67)	<0.001
Mortality	CZP + MTX	MTX + placebo	965	Risk ratio	0	1.54	(0.24 to 10)	0.65	1.54	(0.24 to 10)	0.65
Withdrawal due to lack of efficacy at 3 months	CZP + MTX	MTX + placebo	965	Risk ratio	64	0.30	(0.25 to 0.35)	<0.001	0.29	(0.22 to 0.39)	<0.001
Withdrawals due to any cause	CZP + MTX	MTX + placebo	965	Risk ratio	78	0.40	(0.36 to 0.45)	<0.001	0.39	(0.3 to 0.52)	<0.001
Any adverse event	CZP + MTX	MTX + placebo	965	Risk ratio	49	1.22	(1.09 to 1.35)	<0.001	1.20	(1.02 to 1.43)	0.03
Any serious adverse event	CZP + MTX	MTX + placebo	965	Risk ratio	0	2.13	(1.23 to 3.7)	0.01	2.13	(1.23 to 3.7)	0.01
Malignancies (all)	CZP + MTX	MTX + placebo	965	Risk ratio	18	2.04	(0.43 to 10)	0.37	1.67	(0.26 to 11.11)	0.59
Injection reactions	CZP + MTX	MTX + placebo	965	Risk ratio	0	6.67	(0.87 to 50)	0.07	5.88	(0.78 to 50)	0.09

Table 11 Summary of absolute risk reduction meta-analysis results (as in the submission)

Outcome	Treatment of interest	Control treatment	N	Effect measure	Statistical heterogeneity (I^2)	Fixed effects			Random effects		
						Effect size	CI	p value	Effect size	CI	p value
ACR20 response at 6 months	CZP + MTX	MTX + placebo	965	Risk difference	0	0.47	(0.41 to 0.52)	<0.001	0.47	(0.42 to 0.52)	<0.001
ACR50 response at 6 months	CZP + MTX	MTX + placebo	965	Risk difference	0	0.30	(0.25 to 0.34)	<0.001	0.30	(0.25 to 0.34)	<0.001
ACR70 response at 6 months	CZP + MTX	MTX + placebo	965	Risk difference	0	0.17	(0.14 to 0.21)	<0.001	0.17	(0.13 to 0.2)	<0.001
Mortality	CZP + MTX	MTX + placebo	965	Risk difference	0	0.00	(-0.01 to 0.01)	0.53	0.00	(-0.01 to 0.01)	0.51
Withdrawal due to lack of efficacy at 3 months	CZP + MTX	MTX + placebo	965	Risk difference	89	-0.49	(-0.54 to -0.43)	<0.001	-0.51	(-0.68 to -0.33)	<0.001
Withdrawals due to any cause	CZP + MTX	MTX + placebo	965	Risk difference	84	-0.49	(-0.54 to -0.43)	<0.001	-0.50	(-0.64 to -0.36)	<0.001
Any adverse event	CZP + MTX	MTX + placebo	965	Risk difference	69	0.12	(0.06 to 0.18)	<0.001	0.11	(-0.01 to 0.23)	0.07
Any serious adverse event	CZP + MTX	MTX + placebo	965	Risk difference	0	0.05	(0.02 to 0.08)	<0.001	0.05	(0.02 to 0.08)	<0.001
Malignancies (all)	CZP + MTX	MTX + placebo	965	Risk difference	48	0.01	(-0.01 to 0.02)	0.3	0.00	(-0.01 to 0.02)	0.57
Injection reactions	CZP + MTX	MTX + placebo	965	Risk difference	0	0.02	(0.01 to 0.03)	<0.001	0.02	(0.01 to 0.03)	<0.001
Van der Heijde mTSS change from baseline at 6 months	CZP + MTX	MTX + placebo	965	Mean difference	0	-1.06	(-1.56 to -0.57)	<0.001	-1.06	(-1.56 to -0.57)	<0.001

Table 12 Results from indirect comparison of ACR20 data at 6 months (odds ratios) for combination treatments (as in the submission)

	MTX	CZP + MTX	Adalimumab + MTX	Etanercept + MTX	Infliximab + MTX	Rituximab + MTX	Tocilizumab + MTX
MTX	1.00	0.09 0.06 to 0.14 P<0.01 I^2 = 14%	0.21 0.1 to 0.41 P<0.01 I^2 = 71%	0.15 0.05 to 0.39 P<0.01 I^2 = 0%	0.34 0.18 to 0.65 P<0.01 I^2 = 79%	0.23 0.09 to 0.59 P<0.01 I^2 = 0%	0.26 0.17 to 0.39 P<0.01 I^2 = 100%
CZP + MTX	10.57 6.99 to 15.98 P<0.01 I^2 = 14%	1.00	2.17 0.96 to 4.87 P=0.06	1.56 0.53 to 4.53 P=0.42	3.64 1.7 to 7.8 P<0.01	2.41 0.86 to 6.74 P=0.1	2.70 1.5 to 4.85 P<0.01
Adalimumab + MTX	4.88 2.43 to 9.78 P<0.01 I^2 = 71%	0.46 0.21 to 1.04 P=0.06	1.00	0.72 0.21 to 2.4 P=0.59	1.68 0.65 to 4.32 P=0.28	1.11 0.34 to 3.59 P=0.86	1.24 0.55 to 2.8 P=0.6
Etanercept + MTX	6.79 2.53 to 18.21 P<0.01 I^2 = 0%	0.64 0.22 to 1.87 P=0.42	1.39 0.42 to 4.66 P=0.59	1.00	2.34 0.72 to 7.58 P=0.16	1.55 0.39 to 6.06 P=0.53	1.73 0.59 to 5.05 P=0.31
Infliximab + MTX	2.90 1.53 to 5.51 P<0.01 I^2 = 79%	0.27 0.13 to 0.59 P<0.01	0.60 0.23 to 1.53 P=0.28	0.43 0.13 to 1.39 P=0.16	1.00	0.66 0.21 to 2.07 P=0.48	0.74 0.34 to 1.59 P=0.44
Rituximab + MTX	4.39 1.71 to 11.3 P<0.01 I^2 = 0%	0.42 0.15 to 1.17 P=0.1	0.90 0.28 to 2.91 P=0.86	0.65 0.17 to 2.53 P=0.53	1.51 0.48 to 4.74 P=0.48	1.00	1.12 0.4 to 3.15 P=0.83
Tocilizumab + MTX	3.92 2.58 to 5.95 P<0.01 I^2 = 100%	0.37 0.21 to 0.67 P<0.01	0.80 0.36 to 1.81 P=0.6	0.58 0.2 to 1.68 P=0.31	1.35 0.63 to 2.9 P=0.44	0.89 0.32 to 2.51 P=0.83	1.00

Table 13 Results from indirect comparison of ACR50 data at 6 months (odds ratios) for combination treatments (as in the submission)

	MTX	CZP + MTX	Adalimumab + MTX	Etanercept + MTX	Infliximab + MTX	Rituximab + MTX	Tocilizumab + MTX
MTX	1.00	0.11 0.06 to 0.21 P<0.01 I^2 = 31%	0.15 0.09 to 0.25 P<0.01 I^2 = 30%	0.05 0.01 to 0.42 P=0.01 I^2 = 0%	0.30 0.16 to 0.56 P<0.01 I^2 = 67%	0.19 0.06 to 0.6 P<0.01 I^2 = 0%	0.15 0.09 to 0.26 P<0.01 I^2 = 0%
CZP+ MTX	9.08 4.71 to 17.51 P<0.01 I^2 = 31%	1.00	1.35 0.58 to 3.15 P=0.49	0.49 0.06 to 4.26 P=0.52	2.74 1.12 to 6.74 P=0.03	1.75 0.48 to 6.47 P=0.4	1.40 0.61 to 3.24 P=0.43
Adalimumab + MTX	6.72 3.93 to 11.49 P<0.01 I^2 = 30%	0.74 0.32 to 1.73 P=0.49	1.00	0.36 0.04 to 3.05 P=0.35	2.03 0.9 to 4.59 P=0.09	1.30 0.37 to 4.53 P=0.68	1.04 0.49 to 2.19 P=0.92
Etanercept + MTX	18.53 2.36 to 145.51 P=0.01 I^2 = 0%	2.04 0.23 to 17.76 P=0.52	2.76 0.33 to 23.18 P=0.35	1.00	5.60 0.65 to 48.11 P=0.12	3.58 0.34 to 37.52 P=0.29	2.86 0.34 to 23.98 P=0.33
Infliximab + MTX	3.31 1.79 to 6.11 P<0.01 I^2 = 67%	0.36 0.15 to 0.9 P=0.03	0.49 0.22 to 1.11 P=0.09	0.18 0.02 to 1.53 P=0.12	1.00	0.64 0.18 to 2.31 P=0.49	0.51 0.23 to 1.14 P=0.1
Rituximab + MTX	5.17 1.68 to 15.98 P<0.01 I^2 = 0%	0.57 0.15 to 2.1 P=0.4	0.77 0.22 to 2.68 P=0.68	0.28 0.03 to 2.93 P=0.29	1.56 0.43 to 5.65 P=0.49	1.00	0.80 0.23 to 2.77 P=0.72
Tocilizumab + MTX	6.47 3.84 to 10.9 P<0.01 I^2 = 0%	0.71 0.31 to 1.65 P=0.43	0.96 0.46 to 2.03 P=0.92	0.35 0.04 to 2.93 P=0.33	1.96 0.87 to 4.38 P=0.1	1.25 0.36 to 4.33 P=0.72	1.00

Table 14 Results from indirect comparison of ACR70 data at 6 months (odds ratios) for combination treatments (as in the submission)

	MTX	CZP + MTX	Adalimumab + MTX	Etanercept + MTX	Infliximab + MTX	Rituximab + MTX	Tocilizumab + MTX
MTX	1.00	0.10 0.05 to 0.21 P<0.01 I^2 = 0%	0.16 0.08 to 0.32 P<0.01 I^2 = 23%	0.09 0 to 1.55 P=0.1 I^2 = 100%	0.30 0.19 to 0.48 P<0.01 I^2 = 0%	0.18 0.04 to 0.9 P=0.04 I^2 = 100%	0.07 0.03 to 0.2 P<0.01 I^2 = 0%
CZP + MTX	10.18 4.67 to 22.22 P<0.01 I^2 = 0%	1.00	1.61 0.56 to 4.64 P=0.38	0.89 0.04 to 17.52 P=0.94	3.10 1.25 to 7.67 P=0.01	1.85 0.31 to 10.99 P=0.5	0.72 0.2 to 2.67 P=0.63
Adalimumab + MTX	6.34 3.09 to 12.98 P<0.01 I^2 = 23%	0.62 0.22 to 1.79 P=0.38	1.00	0.55 0.03 to 10.73 P=0.69	1.93 0.82 to 4.52 P=0.13	1.15 0.2 to 6.66 P=0.88	0.45 0.13 to 1.6 P=0.22
Etanercept + MTX	11.48 0.64 to 204.24 P=0.1 I^2 = 100%	1.13 0.06 to 22.25 P=0.94	1.81 0.09 to 35.2 P=0.69	1.00	3.50 0.19 to 64.52 P=0.4	2.08 0.08 to 56.16 P=0.66	0.82 0.04 to 17.45 P=0.9
Infliximab + MTX	3.28 2.07 to 5.2 P<0.01 I^2 = 0%	0.32 0.13 to 0.8 P=0.01	0.52 0.22 to 1.21 P=0.13	0.29 0.02 to 5.28 P=0.4	1.00	0.60 0.11 to 3.16 P=0.54	0.23 0.07 to 0.73 P=0.01
Rituximab + MTX	5.52 1.11 to 27.43 P=0.04 I^2 = 100%	0.54 0.09 to 3.22 P=0.5	0.87 0.15 to 5.04 P=0.88	0.48 0.02 to 12.98 P=0.66	1.68 0.32 to 8.91 P=0.54	1.00	0.39 0.06 to 2.66 P=0.34
Tocilizumab + MTX	14.06 4.95 to 39.93 P<0.01 I^2 = 0%	1.38 0.38 to 5.08 P=0.63	2.22 0.63 to 7.87 P=0.22	1.23 0.06 to 26.2 P=0.9	4.28 1.37 to 13.4 P=0.01	2.55 0.38 to 17.28 P=0.34	1.00

Table 15 Results from indirect comparison of ACR20 data at 3 months (odds ratios) for combination treatments (as in the submission)

	MTX	CZP + MTX	Adalimumab + MTX	Etanercept + MTX	Infliximab + MTX	Tocilizumab + MTX
MTX	1.00	0.11 0.08 to 0.16 P<0.01 I^2 = 8%	0.11 0.05 to 0.25 P<0.01 I^2 = 100%	0.14 0.03 to 0.56 P=0.01 I^2 = 64%	0.19 0.08 to 0.47 P<0.01 I^2 = 0%	0.25 0.11 to 0.58 P<0.01 I^2 = 100%
CZP + MTX	8.99 6.29 to 12.86 P<0.01 I^2 = 8%	1.00	1.02 0.42 to 2.45 P=0.97	1.24 0.29 to 5.27 P=0.77	1.74 0.67 to 4.53 P=0.24	2.24 0.89 to 5.65 P=0.09
Adalimumab + MTX	8.86 3.96 to 19.81 P<0.01 I^2 = 100%	0.99 0.41 to 2.38 P=0.97	1.00	1.23 0.24 to 6.16 P=0.8	1.71 0.52 to 5.68 P=0.38	2.21 0.68 to 7.12 P=0.19
Etanercept + MTX	7.23 1.78 to 29.29 P=0.01 I^2 = 64%	0.80 0.19 to 3.41 P=0.77	0.82 0.16 to 4.1 P=0.8	1.00	1.40 0.27 to 7.33 P=0.17	1.80 0.35 to 9.26 P=0.48
Infliximab + MTX	5.17 2.13 to 12.54 P<0.01 I^2 = 0%	0.57 0.22 to 1.5 P=0.24	0.58 0.18 to 1.93 P=0.38	0.72 0.14 to 3.75 P=0.69	1.00	1.29 0.38 to 4.4 P=0.69
Tocilizumab + MTX	4.02 1.71 to 9.42 P<0.01 I^2 = 100%	0.45 0.18 to 1.13 P=0.09	0.45 0.14 to 1.46 P=0.19	0.56 0.11 to 2.86 P=0.48	0.78 0.23 to 2.66 P=0.69	1.00

Table 16 Results from indirect comparison of ACR50 data at 3 months (odds ratios) for combination treatments (as in the submission)

	MTX	CZP + MTX	Etanercept + MTX	Infliximab + MTX	Tocilizumab + MTX
MTX	1.00	0.14 0.08 to 0.23 P<0.01 I^2 = 0%	0.05 0.01 to 0.18 P<0.01 I^2 = 0%	0.21 0.06 to 0.69 P=0.01 I^2 = 100%	0.35 0.15 to 0.82 P=0.01 I^2 = 0%
CZP + MTX	7.19 4.33 to 11.93 P<0.01 I^2 = 0%	1.00	0.36 0.09 to 1.4 P=0.14	1.52 0.42 to 5.53 P=0.53	2.54 0.96 to 6.76 P=0.06
Etanercept + MTX	20.14 5.66 to 71.68 P<0.01 I^2 = 0%	2.80 0.71 to 10.99 P=0.14	1.00	4.25 0.74 to 24.21 P=0.1	7.13 1.56 to 32.59 P=0.01
Infliximab + MTX	4.74 1.44 to 15.61 P=0.01 I^2 = 100%	0.66 0.18 to 2.41 P=0.53	0.24 0.04 to 1.34 P=0.1	1.00	1.68 0.39 to 7.19 P=0.49
Tocilizumab + MTX	2.83 1.22 to 6.52 P=0.01 I^2 = 0%	0.39 0.15 to 1.04 P=0.06	0.14 0.03 to 0.64 P=0.01	0.60 0.14 to 2.55 P=0.49	1.00

Table 17 Results from indirect comparisons of ACR70 data at 3 months for combination treatments (as in the submission)

	MTX	CZP + MTX	Etanercept + MTX	Infliximab + MTX	Tocilizumab + MTX
MTX	1.00	0.10 0 to 0.03 P=0.39 I^2 = 0%	0.09 0.08 to 0.21 P<0.01 I^2 = 0%	0.30 0 to 0.32 P<0.01 I^2 = 0%	0.20 0.19 to 1.55 P<0.01 I^2 = 100%
CZP + MTX	10.18 0.79 to 0.73 P=7.67 I^2 = 0%	1.00	0.89 0.56 to 17.52 P<0.01	3.10 0.04 to 4.64 P<0.01	2.04 1.25 to 17.52 P=0.38
Etanercept + MTX	11.48 0.89 to 0.12 P=64.52 I^2 = 0%	1.13 0.64 to 0.04 P=44.07	1.00	3.50 0 to 35.2 P=0.94	2.30 0.19 to 0 P=0.69
Infliximab + MTX	3.28 0.25 to 0.29 P<0.01 I^2 = 0%	0.32 2.07 to 0.1 P=1.47	0.29 0.22 to 0.8 P<0.01	1.00	0.66 0 to 5.28 P=0.13
Tocilizumab + MTX	5.00 0.39 to 0 P=3.41 I^2 = 100%	0.49 2.57 to 0.13 P<0.01	0.44 0.3 to 1.37 P<0.01	1.52 0.02 to 2.1 P=0.17	1.00

Table 18 Results from indirect comparison of ACR20 data at 6 months (odds ratios) for mono-therapy (as in the submission)

	Placebo	CZP	Adalimumab	Etanercept
Placebo	1.00	0.12 0.06 to 0.25 P<0.01 I^2 = 0%	0.24 0.15 to 0.38 P<0.01 I^2 = 0%	0.09 0.04 to 0.21 P<0.01 I^2 = 0%
CZP	8.42 3.98 to 17.81 P<0.01 I^2 = 0%	1.00	1.99 0.82 to 4.81 P=0.13	0.78 0.26 to 2.37 P=0.66
Adalimumab	4.24 2.65 to 6.77 P<0.01 I^2 = 0%	0.50 0.21 to 1.22 P=0.13	1.00	0.39 0.15 to 1.01 P=0.05
Etanercept	10.81 4.76 to 24.53 P<0.01 I^2 = 0%	1.28 0.42 to 3.90 P=0.66	2.55 0.99 to 6.56 P=0.05	1.00

Table 19 Results from indirect comparison of ACR50 data at 6 months (odds ratios) for mono-therapy (as in the submission)

	Placebo	CZP	Adalimumab	Etanercept
Placebo	1.00	0.13 0.04 to 0.39 P<0.01 I^2 = 100%	0.26 0.14 to 0.49 P<0.01 I^2 = 0%	0.08 0.03 to 0.25 P<0.01 I^2 = 0%
CZP	7.63 2.56 to 22.77 P<0.01 I^2 = 100%	1.00	1.98 0.56 to 7.00 P=0.29	0.62 0.13 to 2.94 P=0.55
Adalimumab	3.86 2.04 to 7.30 P<0.01 I^2 = 0%	0.51 0.14 to 1.79 P=0.29	1.00	0.32 0.09 to 1.12 P=0.08
Etanercept	12.25 4.08 to 36.78 P<0.01 I^2 = 0%	1.60 0.34 to 7.57 P=0.55	3.17 0.88 to 11.30 P=0.08	1.00

Table 20 Results from indirect comparison of ACR70 data at 6 months (odds ratios) for mono-therapy (as in the submission)

	Placebo	CZP	Adalimumab	Etanercept
Placebo	1.00	0.07 0.004 to 1.33 P=0.08 I^2 = 0%	0.11 0.03 to 0.38 P<0.01 I^2 = 0%	0.07 0.01 to 0.55 P=0.01 I^2 = 0%
CZP	13.49 0.75 to 242.5 P=0.08 I^2 = 0%	1.00	1.52 0.07 to 34.91 P=0.79	0.95 0.03 to 32.97 P=0.98
Adalimumab	8.89 2.63 to 30.03 P<0.01 I^2 = 0%	0.66 0.03 to 15.14 P=0.79	1.00	0.62 0.06 to 6.85 P=0.70
Etanercept	14.26 1.81 to 112.4 P=0.01 I^2 = 0%	1.06 0.03 to 36.83 P=0.98	1.61 0.15 to 17.65 P=0.70	1.00

Table 21 Results from indirect comparison of ACR20 data at 3 months (odds ratios) for mono-therapy (as in the submission)

	Placebo	CZP	Adalimumab	Etanercept
Placebo	1.00	0.14 0.07 to 0.28 P<0.01 I^2 = 100%	0.13 0.06 to 0.30 P<0.01 I^2 = 48%	0.19 0.10 to 0.38 P<0.01 I^2 = 100%
CZP	7.00 3.51 to 13.9 P<0.01 I^2 = 100%	1.00	0.92 0.32 to 2.69 P=0.88	1.33 0.50 to 3.52 P=0.56
Adalimumab	7.57 3.34 to 17.2 P<0.01 I^2 = 48%	1.08 0.37 to 3.16 P=0.88	1.00	1.44 0.50 to 4.19 P=0.50
Etanercept	5.25 2.65 to 10.4 P<0.01 I^2 = 100%	0.75 0.28 to 1.98 P=0.56	0.69 0.24 to 2.02 P=0.50	1.00

Table 22 Results from indirect comparison of ACR50 data at 3 months (odds ratios) for mono-therapy (as in the submission)

	Placebo	CZP	Adalimumab	Etanercept
Placebo	1.00	0.04 0.01 to 0.30 P<0.01 I^2 = 0%	0.09 0.03 to 0.29 P<0.01 I^2 = 7%	0.12 0.05 to 0.31 P<0.01 I^2 = 0%
CZP	25.20 3.32 to 191.0 P<0.01 I^2 = 0%	1.00	2.37 0.23 to 24.15 P=0.47	3.01 0.32 to 28.08 P=0.33
Adalimumab	10.62 3.43 to 32.92 P<0.01 I^2 = 7%	0.42 0.04 to 4.29 P=0.47	1.00	1.27 0.29 to 5.53 P=0.75
Etanercept	8.38 3.27 to 21.51 P<0.01 I^2 = 0%	0.33 0.04 to 3.11 P=0.33	0.79 0.18 to 3.44 P=0.75	1.00

Table 23 Results from indirect comparison of ACR70 data at 3 months (odds ratios) for mono-therapy (as in the submission)

	Placebo	CZP	Adalimumab	Etanercept
Placebo	1.00	0.09 0.005 to 1.62 P=0.10 I^2 = 0%	0.06 0.01 to 0.31 P<0.01 I^2 = 0%	0.22 0.06 to 0.80 P=0.02 I^2 = 0%
CZP	11.31 0.62 to 207.1 P=0.10 I^2 = 0%	1.00	0.67 0.02 to 19.16 P=0.82	2.44 0.10 to 59.05 P=0.58
Adalimumab	16.87 3.18 to 89.54 P<0.01 I^2 = 0%	1.49 0.05 to 42.61 P=0.82	1.00	3.64 0.44 to 30.27 P=0.23
Etanercept	4.64 1.26 to 17.11 P=0.02 I^2 = 0%	0.41 0.02 to 9.93 P=0.58	0.27 0.03 to 2.29 P=0.23	1.00

WAS THE REVIEW UP-TO-DATE?

Yes. Searches conducted up to April 2009

GENERAL COMMENTS

The review incorporated the following elements:

a) *Clear statement of method*
b) *Comprehensive ascertainment of relevant literature*
c) *Minimal/explicit bias (for which adjustment can be made) introduced in the process of summarising the available literature*
d) *Appropriate meta-analysis*
e) *Other useful features (please state)*

d) alone of the above

On this basis the review can be classified as:

A. *Systematic review with a meta-analysis*
B. Systematic review with no meta-analysis (or with a generally inappropriate meta-analysis, the results of which should be ignored)
C. Comprehensive overview, with clearly stated method
D. Review with clearly stated method
E. General review
F. Other (please state)

A. of the above: the methods were however not reported in sufficient detail and there were some inconsistencies/unclarities (for example including the FAST4WARD study even though it was assessed as not investigating a licensed dose)

Appendix 8 Searches for Study C87014

- Centerwatch,
- ACR (abstracts 2006-2008; author search of back issues of Arthritis and Rheumatism to cover all other years of ACR Annual conference),
- EULAR abstracts 2002 -2008,
- British Society for Rheumatology,
- FDA,
- internet searches.

The following documents containing information about the C87014 trial were identified:

- ClinicalTrials.gov : full entry on the trial but no reference to further publications; http://clinicaltrials.gov/ct2/show/NCT00544154
- Centerwatch: full entry on the trial, but no reference to further publications; http://www.centerwatch.com/drug-information/fda-approvals/drug-details.aspx?DrugID=1022
- FDA: background document from FDA approval of Cimzia, mentions CZP study with 247 patients but provides no reference to further publication; http://www.accessdata.fda.gov/drugsatfda_docs/label/2009/125271s000lbl.pdf
- Internet searches:
  - press conference on trial C87014;

    http://www.accessmylibrary.com/coms2/summary_0286-923087_ITM
  - comment about trial C87014;

    http://findarticles.com/p/articles/mi_m0PDG/is_3_3/ai_n6056518/

Appendix 9 Modelled ACR response rates at 6 months by treatment arm

For economic modelling the proportion of patients achieving an ACR20 ACR50 and ACR70 response in trial control arms (MTX for combined therapy, and placebo for monotherapy) was aggregated and used to calculate the proportion of patients achieving these outcomes for CZP and each of the comparators. The most recent results from the manufacturer are shown below (taken from the manufacturer's response to request for clarification).

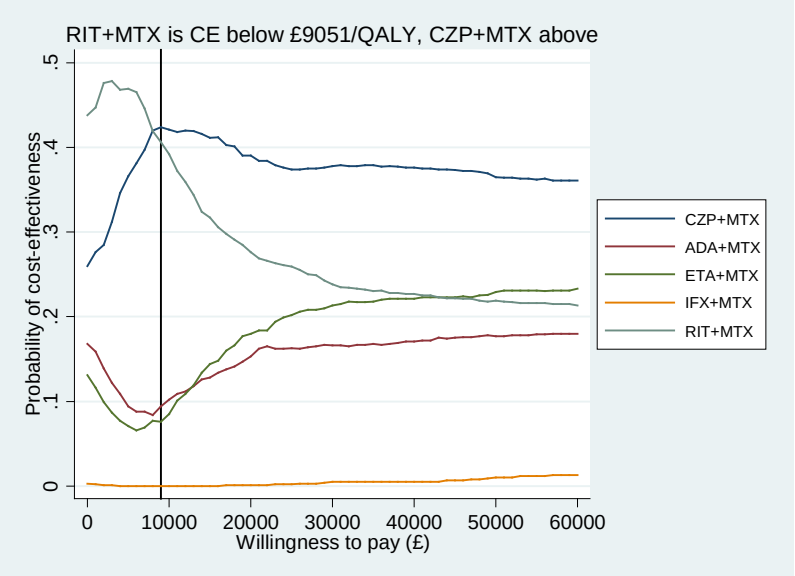
Risk and lower and upper 95% confidence intervals for all treatments at 6 months

Treatment	ACR 20			ACR 50			ACR 70		
	Risk	Low 95% CI	High 95% CI	Risk	Low 95% CI	High 95% CI	Risk	Low 95% CI	High 95% CI
Infliximab + MTX	48.2%	32.9%	63.8%	26.1%	16.1%	39.5%	11.3%	7.4%	16.7%
Rituximab + MTX	58.4%	35.4%	78.3%	35.6%	15.2%	63.1%	17.6%	4.1%	51.4%
Tocilizumab + MTX	55.7%	45.3%	65.6%	40.9%	29.1%	53.8%	35.2%	16.1%	60.7%
Adalimumab	41.3%	30.5%	52.9%	16.2%	9.3%	26.8%	11.2%	3.6%	29.9%
Adalimumab + MTX	61.0%	43.8%	75.8%	41.8%	29.6%	55.1%	19.7%	10.7%	33.4%
Etanercept	64.2%	44.1%	80.3%	38.1%	17.0%	64.9%	16.8%	2.5%	61.5%
Etanercept + MTX	68.5%	44.8%	85.4%	66.4%	20.1%	94.0%	30.7%	2.4%	88.8%
Placebo	13.1%	NA	NA	5.7%	NA	NA	1.0%	NA	NA
Certolizumab pegol	55.9%	37.5%	72.9%	31.4%	13.3%	57.7%	12.3%	0.8%	71.6%
Certolizumab pegol + MTX	77.2%	69.1%	83.7%	49.2%	33.5%	65.2%	28.2%	15.3%	46.2%

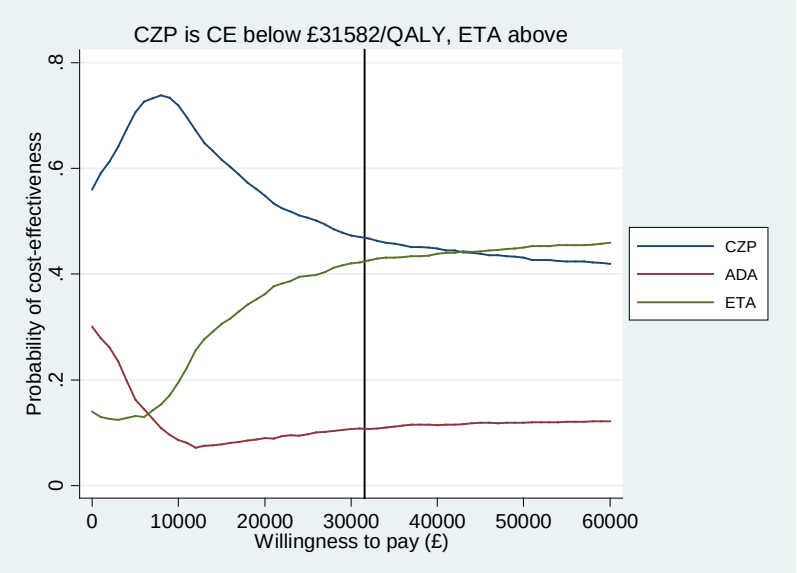
Appendix 10 Cost-effectiveness acceptability frontiers

From the clarification document:

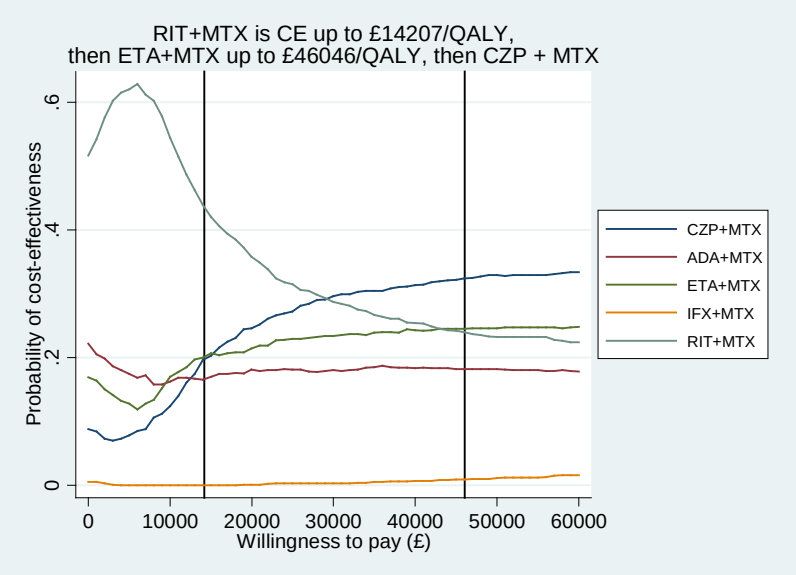
Cost-effectiveness acceptability frontier for scenario A2 (with PAS), for regimes in combination with MTX



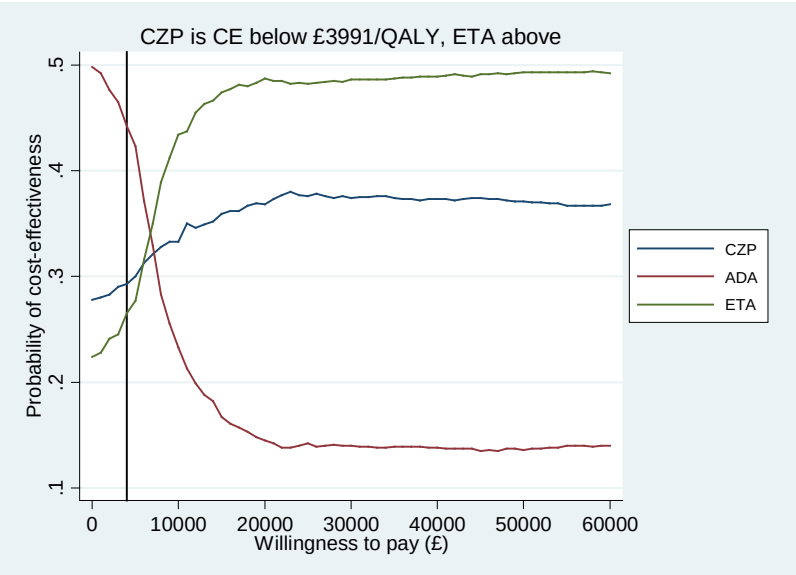
Cost-Effectiveness Acceptability Frontier for scenario A2 (with PAS), for mono-therapy regimens



Cost-Effectiveness Acceptability Frontier for scenario B2 (without PAS), for regimes in combination with MTX



Cost-Effectiveness Acceptability Frontier for scenario B2 (without PAS), for mono-therapy regimens



Appendix 11 Quality assessment using the ScHARR-TAG economic modelling checklist

Title

Certolizumab pegol (CIMZIA®) for the treatment of Rheumatoid Arthritis

A statement of the problem

Yes, a statement of the problem has been given.

A discussion of the need for modelling

No discussion of the need for modelling was included.

A description of the relevant factors and outcomes

Yes a description was provided. The relevant factors and outcomes were more precisely described in the clarification document.

A description of model including: type of model; time frame; perspective; and setting

Yes, a description was provided. A lifetime Markov model was used. The time frame for the analysis is 45 years. Patients enter the model after an inadequate response to MTX and at the start of treatment with CZP or at the start of treatment with a comparator. The first cycle lasts for 6 months, at the end of that cycle patients are assigned to one of four response groups based on assumed risks of responding to the relevant treatment, as measured on the ACR scale. Only patients who obtain an adequate response in the first timestep continue on the modelled initial therapy. There are no state transitions other than discontinuation of treatment or death.

A description of data sources, with description of respective strengths and weaknesses

The variable estimates used in the model were obtained from CZP studies (RAPID 1, RAPID 2 and FAST4WARD) and published literature. A description of those data was provided but strengths and weaknesses have not been discussed.

Key assumptions relating to model structure and data stated

Yes, key assumptions relating to the model structure and data used were stated, particularly precisely in the clarification document

Disease specific factors included within modelling (Items to be specified in conjunction with expert clinical input)

There were no disease specific factors included within modelling

Validation

There was no validation of the model results in the submission.

Results

Model results are reported in the submission in an appropriate format.

Sensitivity analysis results

Results of sensitivity analyses are reported in the submission.

9 References

1 NICE. Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis (Guidance TA130). 2007. Report No.:

2 Hyrich KL, Lunt M, Watson KD, Symmons DP, Silman AJ. Outcomes after switching from one anti-tumor necrosis factor alpha agent to a second anti-tumor necrosis factor alpha agent in patients with rheumatoid arthritis: results from a large UK national cohort study. *Arthritis Rheum* 2007; **56**(1):13-20.

3 Luqmani R, Hennell S, Estrach C, Basher D, Birrell F, Bosworth A, *et al*. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of rheumatoid arthritis (after the first 2 years). *Rheumatology (Oxford)* 2009; **48**(4):436-439.

4 Bongartz T, Sutton AJ, Sweeting MJ, Buchan I, Matteson EL, Montori V. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. *JAMA* 2006; **295**(19):2275-2285.

5 Shoor S. Review: anti-tumor necrosis factor antibody therapy for rheumatoid arthritis increases risk for serious infection and malignancy. *ACP J Club* 2006; **145**(3):65.

6 Raval A, khavan-Toyserkani G, Brinker A, Avigan M. Brief communication: characteristics of spontaneous cases of tuberculosis associated with infliximab. *Ann Intern Med* 2007; **147**(10):699-702.

7 Ramos-Casals M, Brito-Zeron P, Munoz S, Soria N, Galiana D, Bertolaccini L, *et al*. Autoimmune diseases induced by TNF-targeted therapies: analysis of 233 cases. *Medicine (Baltimore)* 2007; **86**(4):242-251.

8 Kwon HJ, Cote TR, Cuffe MS, Kramer JM, Braun MM. Case reports of heart failure after therapy with a tumor necrosis factor antagonist. *Ann Intern Med* 2003; **138**(10):807-811.

9 Socal MP, Vargas AP, Laporte EA, Costa AF, Picon PD. Progressive multifocal leuko-encephalopathy induced by monoclonal antibodies- a systematic review: preliminary data. *Poster presentation HTA conferance, Montreal 2008* 2008.

10 Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; **50**(6):683-691.

11 Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics* 2008; **26**(9):753-767.

12 Furst DEG. Open-label, pilot protocol of patients with rheumatoid arthritis who switch to infliximab after an incomplete response to etanercept: the opposite study.[see comment]. *Annals of the Rheumatic Diseases* 2007; **66**(7):893-899.

13 De Filippis L, Caliri A, Anghelone S. Improving outcome in tumour necrosis factor alpha blocking agents etanercept and infliximab in

patients with active rheumatoid arthritis. *Panminerva Med* 2006; **48**:129-135.

14 Kim H-Y. A randomized, double-blind, placebo-controlled, phase III study of the human anti-tumor necrosis factor antibody adalimumab administered as subcutaneous injections in Korean rheumatoid arthritis patients treated with methotrexate. *APLAR Journal of Rheumatology* 2007; **10**(1):9-16.

15 Keystone E, Kavanaugh A, Sharp J, Tannenbaum H, Hua Y, Teoh L, *et al*. Radiographic, Clinical, and Functional Outcomes of Treatment with Adalimumab (a human anti-tumour necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled 52-week trial. *Arthritis Rheum* 2004; **50**(5):1400-1411.

16 Fleischmann R, Vencovsky J, van VR, Borenstein D, Box J, Coteur G, *et al*. Efficacy and safety of certolizumab pegol monotherapy every 4 weeks in patients with rheumatoid arthritis failing previous disease-modifying antirheumatic therapy: the FAST4WARD study. *Ann Rheum Dis* 2008; **Epub**.

17 Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R. Indirect comparisons of competing interventions. *Health Technology Assessment* 2005; **9**(26):1-201.

18 Nixon RM, Bansback N, Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Stat Med* 2007; **26**(6):1237-1254.

19 Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, *et al*. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995; **38**(6):727-735.