Open Access Full Text Article

ORIGINAL RESEARCH

Systematic review and network meta-analysis of combination and monotherapy treatments in disease-modifying antirheumatic drug-experienced patients with rheumatoid arthritis: analysis of American College of Rheumatology criteria scores 20, 50, and 70

Michelle E Orme[1]
Katherine S MacGilchrist[2]
Stephen Mitchell[2]
Dean Spurden[3]
Alex Bird[3]

[1]Icera Consulting, Swindon, Wiltshire, UK; [2]Systematic Review Department, Abacus International, Bicester, Oxfordshire, UK; [3]Pfizer UK Limited, Tadworth, Surrey, UK

Correspondence: Katherine MacGilchrist
Abacus International, 6 Talisman Business Centre, Talisman Road, Bicester, Oxfordshire, OX26 6HR, UK
Tel +44 1869 355 642
Fax +44 1869 323 248
Email katherine.macgilchrist@abacusint.com

Background: Biologic disease-modifying antirheumatic drugs (bDMARDs) extend the treatment choices for rheumatoid arthritis patients with suboptimal response or intolerance to conventional DMARDs. The objective of this systematic review and meta-analysis was to compare the relative efficacy of EU-licensed bDMARD combination therapy or monotherapy for patients intolerant of or contraindicated to continued methotrexate.

Methods: Comprehensive, structured literature searches were conducted in Medline, Embase, and the Cochrane Library, as well as hand-searching of conference proceedings and reference lists. Phase II or III randomized controlled trials reporting American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 between 12 and 30 weeks' follow-up and enrolling adult patients meeting ACR classification criteria for rheumatoid arthritis previously treated with and with an inadequate response to conventional DMARDs were eligible. To estimate the relative efficacy of treatments whilst preserving the randomized comparisons within each trial, a Bayesian network meta-analysis was conducted in WinBUGS using fixed and random-effects, logit-link models fitted to the binomial ACR 20/50/70 trial data.

Results: The systematic review identified 10,625 citations, and after a review of 2450 full-text papers, there were 29 and 14 eligible studies for the combination and monotherapy meta-analyses, respectively. In the combination analysis, all licensed bDMARD combinations had significantly higher odds of ACR 20/50/70 compared to DMARDs alone, except for the rituximab comparison, which did not reach significance for the ACR 70 outcome (based on the 95% credible interval). The etanercept combination was significantly better than the tumor necrosis factor-α inhibitors adalimumab and infliximab in improving ACR 20/50/70 outcomes, with no significant differences between the etanercept combination and certolizumab pegol or tocilizumab. Licensed-dose etanercept, adalimumab, and tocilizumab monotherapy were significantly better than placebo in improving ACR 20/50/70 outcomes. Sensitivity analysis indicated that including studies outside the target population could affect the results.

Conclusion: Licensed bDMARDs are efficacious in patients with an inadequate response to conventional therapy, but tumor necrosis factor-α inhibitor combination therapies are not equally effective.

Keywords: bDMARD, rheumatoid arthritis, etanercept, systematic review, network meta-analysis, comparative effectiveness

Introduction

Rheumatoid arthritis (RA) is a chronic autoimmune inflammatory disease characterized by inflammation of the synovial lining of joints, tendons, and periarticular structures,[1] which affects approximately 0.8% of the UK population.[2] If untreated, RA leads to joint destruction, functional limitation and severe disability, and has a significant impact on health-related quality of life.[3–5] Therefore, RA imposes a significant economic burden on health-care systems and society in general.[6] Although the causes of RA are still obscure, research has shown that proinflammatory cytokines, such as tumor necrosis factor-α (TNF-α) and interleukin (IL)-6 or IL-1 play key roles in its pathogenesis.[7]

Conventional disease-modifying antirheumatic drugs (cDMARDs) are generally offered as first-line treatments (most commonly methotrexate [MTX] alone, or, for active disease, in combination with another DMARD). Biologic DMARDs (bDMARDs) offer a valuable treatment alternative, being recommended for patients with suboptimal response or intolerance to cDMARDs or where continued cDMARD therapy is contraindicated.[8,9]

A number of bDMARDs have been licensed for such use in the EU. TNF-α inhibitors include etanercept, adalimumab, infliximab, certolizumab pegol, and golimumab. In combination with MTX, the TNF-α inhibitors are each indicated for the treatment of moderate to severe active RA in adults when the response to DMARDs, including MTX, has been inadequate. In addition, adalimumab, etanercept, and certolizumab pegol are licensed as monotherapy in those patients intolerant of MTX or for whom continued MTX is inappropriate.

The costimulatory inhibitor abatacept and the anti-IL-6 therapy tocilizumab, in combination with MTX, are licensed for moderate to severe active RA in adults responding inadequately to previous therapy with one or more cDMARDs including MTX or a TNF-α inhibitor. Tocilizumab is also licensed as monotherapy in patients intolerant of MTX or for whom continued MTX is inappropriate. In addition, the anti-B-cell therapy rituximab, in combination with MTX, is licensed in adult patients with severe active RA with inadequate response or intolerance to other DMARDs including one or more TNF-α inhibitors.

The objective of this systematic review was to compare the clinical efficacy of EU licensed-dose bDMARD combinations for the treatment of adult RA patients after failure on one or more DMARDs, where efficacy was measured using American College of Rheumatology (ACR) response end points from randomized controlled trials (RCTs). A network meta-analysis (NMA) was performed to pool RCT evidence for bDMARDs via common control treatments (eg, MTX control), to provide estimates of relative treatment effects. The rationale for this approach was that there are few trials comparing bDMARDs head-to-head. Therefore, NMA can support inferences to the target RA population, as all the available evidence from relevant RCTs are used in the analysis.

As bDMARD monotherapies are used in a different part of the treatment pathway, ie, in a population intolerant of MTX or for whom continued MTX is inappropriate, a separate analysis of bDMARD monotherapies was performed.

Methods

The methods used for the review and meta-analysis of combination therapy are the same as for monotherapy, except where otherwise stated.

Systematic review
Study eligibility criteria

A protocol was written to define all aspects of the systematic review prior to commencement. The inclusion criteria are shown in Table 1. As the data used in a meta-analysis should be from sufficiently similar studies and outcomes to make the results meaningful and to reduce the influence of confounding factors, included studies had to report sufficient data for the ACR 20, 50, or 70 response to treatment end point (defined as a 20%, 50%, or 70% improvement in tender and swollen joints and the same level of improvement in three of the five following variables: patient and physician global assessments of overall disease activity; patient evaluation of pain (pain health assessment questionnaire [HAQ][10]); a score of physical disability; and blood acute-phase reactants). End points needed to be measured between 12 and 30 weeks from baseline. Studies in which more than 15% of patients had had previous TNF-α inhibitor treatment were excluded, because this population was more extensively pretreated and considered likely to be less responsive than the TNF-α inhibitor-naïve population. Studies were not restricted by date of publication or publication status.

Data extraction

The data sources to identify published RCTs and ongoing (as yet unpublished) RCTs included:

- Electronic databases accessed via OVID and the Cochrane Library: Medline in process and Medline 1950 to present; Embase 1980 to 2010 week 23 and 1980 to 2011 week 14; the Cochrane Library (Cochrane Database of

Table 1 Summary of inclusion criteria for the systematic review and meta-analysis

Study design	Phase II or III prospective, randomized controlled trials with at least one treatment arm containing a bDMARD combination therapy or bDMARD monotherapy
Population	Adult patients (≥18 years) meeting the ACR classification criteria for rheumatoid arthritis
	Previously treated with MTX or other DMARD
	≤15% of patients previously treated with TNF-α inhibitors
Interventions	Any bDMARD licensed in the EU
	Studies needed to include at least one treatment arm of bDMARD in combination with a DMARD or as a monotherapy
	No restrictions to drug dose or formulation, mode of delivery or duration of treatment
Comparators	DMARD (combination analysis) or placebo (monotherapy analysis)
	Other comparators where needed to connect the network or preserve randomization
Outcomes	ACR 20/50/70 response rate to treatment (defined as a 20%, 50%, or 70% improvement in the ACR score)
	Outcome reported between 12 and 30 weeks of follow-up
Language of publication	Non-English full-text papers were excluded, but English abstracts of non-English full-text papers were included

Abbreviations: ACR, American College of Rheumatology; bDMARD, biological disease-modifying antirheumatic drug; DMARD, disease-modifying antirheumatic drug; MTX, methotrexate; TNF-α, tumour necrosis factor-alpha.

Systematic Reviews, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials, and Health Technology Assessment Database). The original search was conducted June 17, 2010. The same search was rerun April 15, 2011 to identify any studies entered into the databases from June 1, 2010.

- Hand-searching of reference lists of included RCTs.
- Hand-searching of RCTs included in previously conducted systematic reviews/meta-analyses[11,12] and the Cochrane reviews on bDMARDs.[13–19]
- Reviewing the Canadian Agency for Drugs and Technologies in Health therapeutic review.[20]
- The following conference proceedings were searched (2005–10): ACR, European League Against Rheumatism, and the British Society for Rheumatology.
- The US Institutes of Health online registry and results database of clinical trials[21] was searched to identify ongoing studies.

The structured database search strings were designed to identify RCTs or systematic reviews indexed on Medline, and these strings were then modified for performing searches of Embase and the Cochrane Library to account for differences in syntax and thesaurus headings. Searches included terms for free text and Medical Subject Heading (MeSH) terms to identify RCTs of RA patients taking DMARDs or bDMARDs.

One reviewer screened the title and abstract of studies identified against the eligibility criteria. Full-text papers were then assessed to ensure studies met the criteria or for those studies where eligibility could not be determined from the title/abstract. Any uncertainties as to eligibility were referred to a second reviewer and resolved by consensus. Data were extracted from eligible publications into a predefined data-extraction table by one reviewer and verified by a second.

The data items collected included patient (average age, percentage female, disease duration, baseline severity of RA, MTX- or other DMARD-exposure and TNF-α exposure), intervention (treatment(s) received, dosage and dose schedule), study (study blinding and country(ies), number of patients randomized, follow-up period, frequency of withdrawals), and outcome (ACR 20/50/70) level parameters.

Quality assessment

Risk of bias was assessed using criteria set out in the National Institute for Health and Clinical Excellence (NICE) guidelines manual.[22] For studies included in the meta-analysis, a formal assessment of publication bias was conducted via funnel plots with Egger's linear regression test of asymmetry.[23,24]

Meta-analysis methodology

For this meta-analysis, the study arms were pooled into treatment groups; we were interested in those study arms where the intervention was used in accordance with the EU license, since these are the treatments used in clinical practice. Therefore, the treatments of interest for therapy in DMARD-experienced patients are licensed bDMARD combinations plus common control arms used to connect the network (Table 2).

In a separate analysis we considered a population of patients who are intolerant of MTX or for whom MTX is contraindicated; the treatments of interest here are licensed bDMARD monotherapies plus common control arms used to connect the network (Table 2). As other cDMARDs may be used as monotherapy if MTX is contraindicated, sulfasalazine is also a treatment of interest in the monotherapy analysis. Other control arms were included in the evidence networks to preserve randomization as well as other unlicensed arms. The results for these unlicensed treatments have been omitted from this publication.

Table 2 Active licensed treatment arms of interest

Combination therapy	Abatacept 10 mg/kg/4 weeks + MTX
	Adalimumab 40 mg/2 weeks + MTX
	Certolizumab pegol 200 mg/2 weeks + MTX
	Etanercept 2 × 25 mg/week + MTX
	Golimumab 50 mg/4 weeks + MTX
	Infliximab 3 mg/kg/8 weeks + MTX
	Rituximab 2 × 1000 mg + MTX
	Tocilizumab 8 mg/kg/4 weeks + MTX
	Controls – DMARD monotherapy
	(= MTX or sulfasalazine)
Monotherapy	Adalimumab 40 mg/2 weeks
	Certolizumab pegol 200 mg/2 weeks
	Etanercept 2 × 25 mg/week
	Tocilizumab 8 mg/kg/4 weeks
	Control (= placebo or sulfasalazine)

Abbreviations: DMARD, disease-modifying antirheumatic drug; MTX, methotrexate.

The analyses of the ACR 20/50/70 outcomes were conducted on an intent-to-treat (ITT) basis, or modified ITT (number actually receiving treatment at baseline) if the number randomized to treatment is not reported. An ITT analysis requires imputing outcomes for the missing participants, although there is no overall consensus on how to do this;[25] for the ACR 20/50/70 outcomes, it is assumed that missing participants did not achieve the required improvement (ie, a worst-case scenario).

Direct and indirect meta-analysis

A fixed and random-effects meta-analysis was conducted in Stata IC version 11.2 using the Metan package SJ9_2: sbe24_3 (StataCorp, College Station, TX).[26,27] The random-effects analysis used the method of DerSimonian and Laird, with the estimate of heterogeneity taken from the Mantel–Haenszel model. The fixed-effect analysis used the Mantel–Haenszel method. For binomial data analysis, if a study contains a zero observation (eg, no patients achieved ACR 70), Stata adds 0.5 to each cell of the trial by default.

Indirect comparisons between treatment (A) and other treatments of interest (B) via a common comparator (C) were made using the Bucher method[28,29] and the pooled odds ratio (OR) produced from the direct meta-analysis.

Network meta-analysis

In an NMA, treatment effects are calculated for all treatments using all available evidence in one simultaneous analysis.[30–32] NMA methods build on the principles of indirect comparisons[28,29] and preserve the randomized comparisons within each trial.

The models were fitted to the data using Bayesian Markov chain Monte Carlo methods (specifically Gibbs sampling),

using WinBUGS software version 1.4.3.[33–35] WinBUGS code for NMA of dichotomous and standard Bayesian random-effects meta-analysis was adapted from code developed by the MRC Biostatistics Unit[36] and the NICE Decision Support Unit.[37]

The WinBUGS models were run for a minimum of 100,000 iterations to ensure model convergence. Subsequently, two chains of 20,000 were sampled from the posterior distributions. These samples were used to calculate the median/mean and where relevant the 95% credible interval (CrI), which is the interval from percentiles 2.5 to the 97.5. For treatment effects, medians are presented as the best estimate for the central value, since means may be overly influenced by outliers.

Analysis of baseline risk

To calculate the absolute probability of responding to each treatment, we first conducted a standard direct random-effects meta-analysis that pooled data on the log-odds (or mean difference from baseline) of responding to the reference control treatment. The reference treatment is chosen to be the control that has the most data available, ie, the DMARD control in this analysis. The (mean and standard deviation) pooled log-odds (or mean differences) of responding to the reference treatment were then used as priors in the main NMA to inform the calculation of the absolute efficacy of each treatment.[38]

Analysis of treatment effects

For dichotomous end points, such as ACR 20/50/70, the NMA calculates the ORs for all treatments compared with other treatments. The base case models were random-effects models; fixed-effect models were used as sensitivity analyses. Random-effects NMA differs from fixed-effect NMA in that it allows the true treatment effect (eg, OR between two treatments) to vary between studies due to heterogeneity. In these random-effects models, a uniform (uninformative) prior is used for the between-studies standard deviation (as per Hasselblad[39] and Gelman[40]).

For the ACR 70 outcome, some studies reported zero events in the DMARD control arm, requiring a continuity correction to be applied. A fixed value of 0.5 was added to the numerator (and 1 to the denominator) for all arms of the affected trial.[37,41] The fixed-value correction overcomes computational errors, but it biases study estimates towards no difference and overestimates variances. Biases will be more apparent in trials where the treatment arms are of unequal effect.

Covariate analysis

Covariate analyses were conducted to explore potentially confounding factors. We conducted a study-level covariate analysis to take into account the following differences in study protocols (DMARD-experienced analysis): (1) Length of follow-up: the model included a study level continuous variable to adjust for the time point at which the response was measured (in weeks). Xweeks is a covariate centered at mean follow-up across the included studies, such that the coefficient βweeks estimates the incremental difference in (log) treatment effect for each week above/below the average follow-up across studies. (2) Studies where MTX was administered at a low dose: the Japanese maximum dose of 8 mg/week was used as a cutoff (Xmtx = 1 if study population received MTX within the normal dose range [maximum dose more than 8 mg/week]; 0 otherwise). The coefficient βmtx estimates the incremental (log) treatment effect between low-dose concomitant MTX and normal-dose concomitant MTX.

We conducted an additional covariate analysis to take into account the following study-arm level differences in patient characteristics: (1) average age at baseline, and (2) average disease duration at baseline. This covariate model included these continuous variables to adjust for differences in patient age and disease duration (in years) across study arms. Xage and Xduration are covariates centered at mean age and disease duration, respectively, such that the coefficient βage and βduration estimate the incremental difference in the (log) treatment effect for each year above/below the average age or disease duration across study arms.

Subgroup analyses

The following additional analyses were conducted for combination therapy:

1. As base case, but remove studies that included MTX-naïve patients, ie, subgroup is the MTX-experienced population. The studies removed for this analysis were Combe et al[42] and Genovese et al.[43]

2. As base case, but remove the certolizumab trials RAPID 1 and RAPID 2,[44,45] on the basis of study design/pattern of withdrawals where many trial participants, particularly those randomized to the MTX control arm, were withdrawn early because of lack of efficacy and given rescue medication prior to the primary end point at week 24 being reached. Week 16 withdrawals in RAPID 1: 62.8% placebo, 21.1% certolizumab 200 mg, 17.4% certolizumab 400 mg. Week 16 withdrawals in RAPID 2: 81.1% placebo, 21.1% certolizumab 200 mg, 21.1% certolizumab 400 mg.

3. As base case, but remove open-label etanercept trials of Kameda et al[46] and van Riel et al[47] and the etanercept trial with the early escape design of Genovese et al.[48]

4. As base case, but remove studies that contain (up to 15% of) patients who are TNF-α inhibitor-experienced, in addition to being MTX-experienced. These studies were: Genovese et al,[43] Keystone et al,[44] Kremer et al,[49,50] Maini et al,[51] Smolen et al,[45,52] and Zhang et al.[53] The rationale is that these patients have had more previous treatments.

5. As base case, but include data from the Klareskog et al trial[1,54] at 24 weeks. This trial is not representative of the inadequate responder population required by the systematic review selection criteria, as some of the population enrolled were MTX-naïve. Some previous meta-analyses have included this study, however, so its influence is examined here by including it in a sensitivity analysis. Data for this analysis were 24-week data from the unpublished clinical study report.[55]

Sensitivity analysis for monotherapy was conducted as follows: as base case, but include data from the TEMPO trial (24-week data from the unpublished clinical study report[55]). Subset analysis was not conducted: removing studies in MTX-naïve or TNF-α-experienced populations from the base case would leave too few remaining studies.

Model fit

The mean residual deviances provided an estimate of how well the values predicted by the model fit the observed dataset.[56] For an adequate model fit, the sum of the residual deviances should be approximately equal to the total number of study arms in the observed dataset. In addition the deviance information criterion (DIC) output by WinBUGs[57] was recorded. The model with the lowest DIC is estimated to be the model that would best predict a replicate dataset of the same structure as that currently observed.

Consistency of NMA estimates

An informal assessment of consistency was performed by comparing the treatment effects estimated via the NMA against the pair-wise direct meta-analysis results and results of the indirect Bucher analysis to identify potential discrepancies between the results from the different methods.

Furthermore, the network diagrams were examined to determine the number of independent loops in the network of evidence for which inconsistency in the evidence could occur.[38] Disregarding loops that occur solely from a multi-arm trial (since within-trial treatment effects are not independent), the size of any inconsistency was determined for each independent

loop using the Bucher method[28,29] and the Z-test (or chi-squared test if one edge of the loop is shared with other loops) to determine if the inconsistency was statistically significant.

Results
Literature search results

A total of 10,616 potentially relevant records were identified, excluding duplicates from the original search, of which 8175 were excluded on screening the title and abstract. On application of the inclusion criteria to the 2441 full-text papers, a further 2415 were excluded. Nine additional studies were identified from the updated search. Thirty-seven publications were included; 23 assessing combination therapy only[43–45,48–50,52,53,58–72]

eight monotherapy only,[73–80] and six monotherapy in addition to combination-therapy arms[42,46,47,51,81,82] (Figure 1).

Systematic review results for combination studies

Of the 29 studies with at least one combination-therapy arm (Figure 2) three assessed abatacept,[49,66,69] five adalimumab,[59,62,64,65,71] two certolizumab pegol,[44,45] six etanercept,[42,46–48,67,70] three golimumab,[50,63,82] six infliximab (one of which had an abatacept arm also, providing the only head-to-head comparison),[53,58,60,68,69,72] two rituximab,[61,81] and three tocilizumab.[43,51,52] All studies utilized licensed doses, with the exception of one golimumab study.[50]
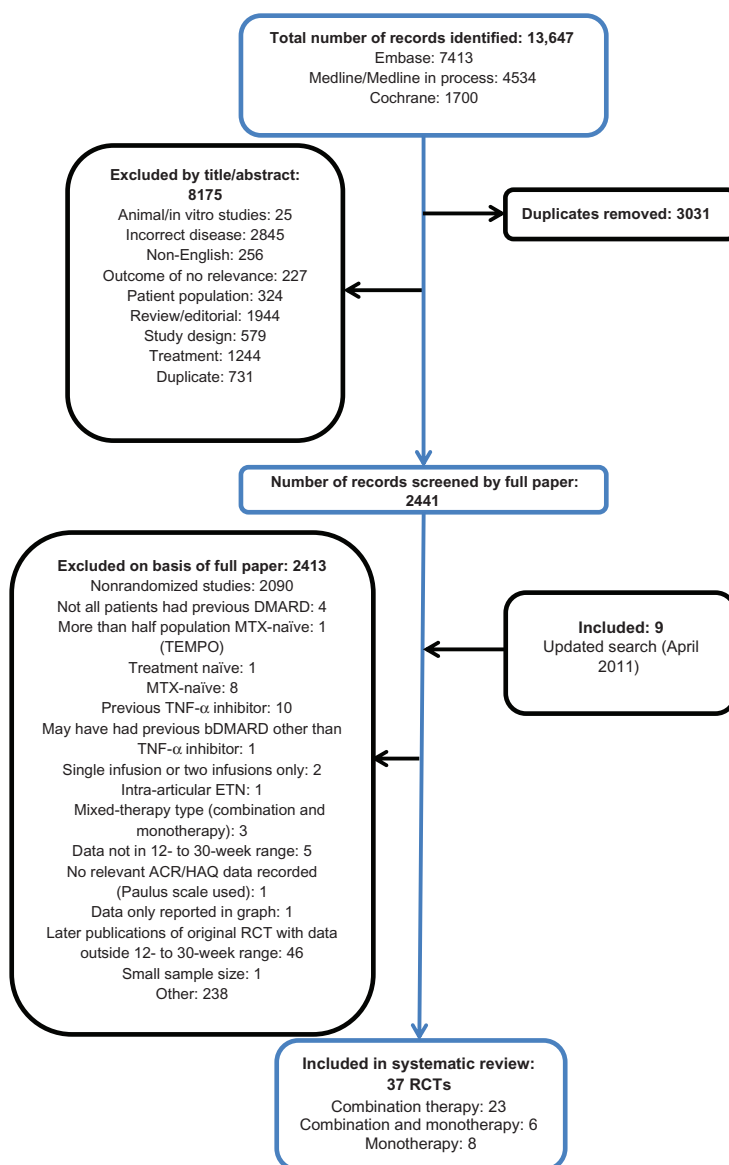
Figure 1 Flow diagram of included/excluded studies.
Abbreviations: ACR, American College of Rheumatology; bDMARD, biological disease-modifying anti-rheumatic drug; DMARD, disease-modifying anti-rheumatic drug; ETN, etanercept; HAQ, Health Assessment Questionnaire; MTX, Methotrexate; RCT, randomized controlled trial; TEMPO, Trial of Etanercept and Methotrexate with radiographic Patient Outcomes; TNF-α, tumour necrosis factor alpha.
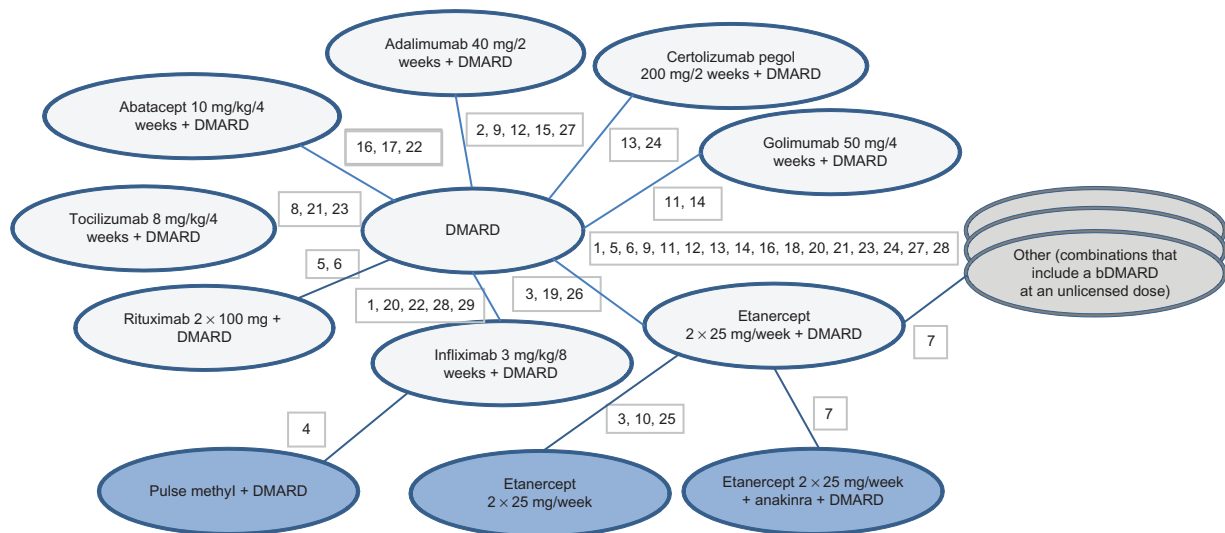
Figure 2 Network diagram for ACR20/50/70 outcomes for bDMARD combination therapies.

Notes: 1, Abe 2006; 2, Chen 2009; 3, Combe 2006; 4, Durez 2004; 5, Edwards 2004; 6, Emery 2010 (SERENE); 7, Genovese 2004; 8, Genovese 2008 (TOWARD); 9, Huang 2009; 10, Kameda 2010 (JESMR); 11, Kay 2008; 12, Keystone 2004 (DE019); 13, Keystone 2008 (RAPID 1); 14, Keystone 2009 (GO-FORWARD); 15, Kim 2007; 16, Kremer 2003; 17, Kremer 2006 (AIM); 18, Kremer 2010; 19, Lan 2004; 20, Maini 1999 (ATTRACT); 21, Maini 2006 (CHARISMA); 22, Schiff 2008 (ATTEST); 23, Smolen 2008 (OPTION); 24, Smolen 2009a (RAPID 2); 25, van Riel 2006 (ADORE); 26, Weinblatt 1999; 27, Weinblatt 2003 (ARMADA); 28, Westhovens 2006b (START); 29, Zhang 2006. DMARD 25 arms, 3039 patients; abatacept 10 mg/kg/4 weeks + DMARD 3 arms, 704 patients; adalimumab 40 mg/2 weeks + DMARD 5 arms, 495 patients; certolizumab pegol 200 mg/2 weeks + DMARD 2 arms, 639 patients; etanercept 2 × 25 mg/week + DMARD 6 arms, 500 patients; golimumab 50 mg/4 weeks + DMARD 2 arms, 124 patients; infliximab 3 mg/kg/8 weeks + DMARD 6 arms, 760 patients; rituximab 2 × 1000 mg + DMARD 2 arms, 212 patients; tocilizumab 8 mg/kg/4 weeks + DMARD 3 arms, 1058 patients.

Study and patient characteristics are summarized in Table 3. The majority of RCTs were double-blind, three being open-label.[46,47,60] In total the included studies randomized 11,490 patients. Included patients had active RA in spite of prior treatment with a DMARD. Not all studies reported baseline disease activity score (DAS). Of those that did, 13 involved populations with more severe RA: eleven trials had a mean or median baseline DAS28 of 5.9 or above,[44,45,47,51,52,59,61,66,69,81,82] and in two trials of either abatacept or infliximab, the authors noted the particularly severe or active nature of disease in the study population.[49,68] In two trials involving either etanercept or infliximab, the mean baseline DAS28 was between 5.0 and 5.2,[42,72] indicating that the population would have included some patients with severe RA and others with more moderate-severity disease. The definition of "active RA" was inconsistent across studies, with some requiring ≥ six tender joints and ≥six swollen joints, whilst others required ≥ twelve tender joints and ≥ten swollen joints. One study of infliximab in particular may have enrolled patients with less active RA, as its definition of active RA included having ≥ eight tender joints and only ≥ three swollen joints.[53] In most trials, the patient population was anti-TNFα inhibitor-naïve. Patients had a mean age of between 48 and 58 years and had on average suffered from RA for 5–10 years (around 9 months in Maini et al[51] and 13 years in Weinblatt et al[71]). These trials were, therefore,

broadly representative of the population of interest, namely, adult patients with moderate–severe active RA, previously treated with (and with insufficient response to) MTX or another DMARD, irrespective of disease duration.

The risk of bias, as assessed by NICE criteria, was considered low for the majority of included studies. For five studies, the risk of bias was unclear,[50,53,59,61,67] due to incomplete reporting. Only the study by van Riel et al[47] was considered to have a high risk of bias, as there was no concealment of treatment allocation (and several other parameters were unclear).

Data for the ACR 20/50/70 end points are presented in Table 4. The follow-up period was 24 weeks in 18 of the 29 trials,[42–46,48–50,52,61,62,64–66,70,71,81,82] ranging from 12 weeks[59,67] to 30.[68] Figures 3–5 show funnel plots for ACR 20/50/70, respectively, for all studies with DMARD control arm used in the combination-therapy meta-analysis. An asymmetrical funnel plot suggests publication bias or systematic difference between smaller and larger studies, and might therefore suggest that simple meta-analysis of the dataset was not appropriate. Funnel plots also highlight outlier studies, where the control-arm response is either particularly high (leading to an underestimate of the active treatment effect by comparison) or particularly low (leading to an overestimate of the active treatment's relative effect).[23,24] For ACR 20, there is a good, symmetrical spread of control responses either side

Table 3 Characteristics of included combination therapy studies

Study	Disease severity	Treatment and dose	Treatment group for analysis	Number of patients randomized	Mean age, years	Mean disease duration, years	Withdrawals, %
Abe et al[58]	Active disease despite MTX	INF 3 mg/kg + MTX	**INF 3 mg/kg/8 weeks + DMARD (MTX)**	49	55.2	9.1	2
		INF 10 mg/kg + MTX	INF 10 mg/kg/8 weeks + DMARD (MTX)	51	56.8	7.1	7.8
		MTX	**DMARD (MTX)**	47	55.1	7.5	10.6
Chen et al[59]	Active disease despite MTX DAS28 6.54 (median) Range 5.63–7.37	ADA 40 mg every other week + MTX	**ADA 40 mg/2 weeks + DMARD (MTX)**	35	53 (median)	6.2 (median)	8.6 (AEs leading to discontinuation)
		MTX	**DMARD (MTX)**	12	53 (median)	8.3 (median)	0 (AEs leading to discontinuation)
Combe et al[42]	Active RA despite SUL Mean baseline DAS28 5.1, 5.0, 5.2	ETN 25 mg twice weekly	ETN 2 × 25 mg/week	103	51.3	7.1	37 at 2 years
		SUL 2, 2.5, or 3 g daily	**DMARD (SUL)**	50	53.3	5.6	68 at 2 years
		ETN 25 mg twice weekly + SUL 2–3 g daily	**ETN 2 × 25 mg/week + DMARD (SUL)**	101	50.6	6.5	24 at 2 years
Durez et al[60]	RA > 1 year Severe, active disease despite MTX	IFN, 3 mg/kg IV infusion weeks 0, 2, and 6 + MTX	**INF 3 mg/kg/8 weeks + DMARD (MTX)**	12	48	10	NR
		MP, 1 g, single infusion week 0 + MTX	MP 1 mg + DMARD (MTX)	15	56	12	
Edwards et al[81]	Active disease despite MTX DAS mean 6.8 or 6.9	RTX 1000 mg IV day 1 and 15	RTX 2 × 1000 mg	40	54	9	5 at week 24 90 at year 2
		RTX 1000 mg iv day 1 and 15, cyclophosphamide 750 mg IV day 3 and 17	RTX 2 × 1000 mg + CYC	41	53	10	7.3 at week 24 78 at year 2
		RTX 1000 mg IV day 1 and 15 with MTX	**RTX 2 × 1000 mg + DMARD (MTX)**	40	54	12	2.5 at week 24 55 at year 2
		MTX	**DMARD (MTX)**	40	54	11	82.5 at week 24 85 at year 2
Emery et al[61]	Active disease despite MTX Mean DAS28 ESR 6.4–6.54 across groups	RTX 2 × 500 mg IV infusion day 1 and 15 + MTX	RTX 2 × 500 mg + DMARD (MTX)	168	51.9	7.1	3.0 at week 24
		RTX 2 × 1000 mg IV infusion day 1 and 15 + MTX	**RTX 2 × 1000 mg + DMARD (MTX)**	170	51.3	6.6	2.4 at week 24
		MTX	**DMARD (MTX)**	172	52.2	7.5	7.6 at week 24
Genovese et al[48]	Active RA despite MTX	ETN 25 mg twice weekly + MTX (mean) 16.1 mg/week	**ETN 2 × 25 mg/week + DMARD (MTX)**	80	54.4	9.7	7
		ETN 25 mg twice weekly + ANA 100 mg/day + MTX (mean) 15.7 mg/week	ETN 2 × 25 mg/week + ANA + DMARD (MTX)	81	55.7	10.6	20
		ETN 25 mg once weekly + ANA 100 mg/day + MTX (mean) 16.2 mg/week	ETN 1 × 25 mg/week + ANA + DMARD (MTX)	81	53.8	9.5	22

Study	Patient population	Intervention	Comparison group	N	Age	Disease duration	Withdrawals
Genovese et al[43]	Active, moderate to severe RA, despite conventional antirheumatic therapy DAS 28 6.6–6.7.	TOC 8 mg/kg IV every 4 weeks + cDMARD	**TOC 8 mg/kg/4 weeks + DMARD (any traditional)**	805	53	9.8	7 (withdrawn from initial treatment) 2 (rescue therapy)
		cDMARD	**DMARD (any traditional)**	415	54	9.8	10 (withdrawn from initial treatment) 11 (rescue therapy)
Huang et al[62]	Active disease despite MTX	ADA 40 mg every other week + MTX	**ADA 40 mg/2 weeks + DMARD (MTX)**	121	NR	NR	NR
		ADA 80 mg every other week + MTX	ADA 80 mg/2 weeks + DMA RD (MTX)	121	NR	NR	NR
		MTX	**DMARD (MTX)**	60	NR	NR	NR
Kameda et al[46]	Active RA despite MTX	ETN 25 mg twice weekly (MTX discontinued)	ETN 2 × 25 mg/week	74	58.1	10.6	16
		ETN 25 mg twice weekly + MTX 6–8 mg/week	**ETN 2 × 25 mg/week + DMARD (MTX)**	77	56.5	8.1	5
Kay et al[63]	Active disease despite MTX	GOL 100 mg SC every 4 weeks + MTX (≥10 mg/week)	GOL 100 mg/4 weeks + DMARD (MTX)	34	57.5	6.3	12.4 at week 16
		GOL 100 mg SC every 2 weeks + MTX (≥10 mg/week)	GOL 100 mg/2 weeks + DMARD (MTX)	34	53.5	9	
		GOL 50 mg SC every 4 weeks + MTX (≥10 mg/week)	**GOL 50 mg/4 weeks + DMARD (MTX)**	35	57	8.2	
		GOL 50 mg SC every 2 weeks + MTX (≥10 mg/week)	GOL 50 mg/2 weeks + DMARD (MTX)	34	48	8.2	
		MTX (≥10 mg/week)	**DMARD (MTX)**	35	52	5.6	17.1 at week 16
Keystone et al[44]	Active disease despite MTX DAS28 using ESR at baseline median range 6.9–7.0	CZP 200 mg + MTX (mean) 13.6 mg/week	**CZP 200 mg/2 weeks + DMARD (MTX)**	393	51.4	6.1	35.1 at week 52
		CZP 400 mg + MTX (mean) 13.6 mg/week	CZP 400 mg/2 weeks + DMARD (MTX)	390	52.4	6.2	29.7 at week 52
		MTX (mean) 13.4 mg/week	**DMARD (MTX)**	199	52.2	6.2	78.4 at week 52
Keystone et al[64]	Moderate–severe active disease despite MTX	ADA 40 mg every other week + MTX (mean) 16.7 mg/week	**ADA 40 mg/2 weeks + DMARD (MTX)**	207	56.1	11	22 at week 52
		ADA 20 mg every other week + MTX (mean) 16.3 mg/week	ADA 20 mg/2 weeks + DMARD (MTX)	212	57.3	11	
		MTX (mean) 16.7 mg/week	**DMARD (MTX)**	200	56.1	10.9	30
Keystone et al[82]	Active disease despite MTX Median DAS28 using ESR at baseline Range 5.9–6.1	GOL, 100 mg, injection every 4 weeks	GOL 100 mg/4 weeks	133	51 (median)	5.9 (median)	NR
		GOL 100 mg injection every 4 weeks + MTX	GOL 100 mg/4 weeks + DMARD (MTX)	89	50 (median)	6.7 (median)	
		GOL 50 mg injection every 4 weeks + MTX	**GOL 50 mg/4 weeks + DMARD (MTX)**	89	52 (median)	4.5 (median)	
		MTX	**DMARD (MTX)**	133	52 (median)	6.5 (median)	

Table 3 (*Continued*)

Study	Disease severity	Treatment and dose	Treatment group for analysis	Number of patients randomized	Mean age, years	Mean disease duration, years	Withdrawals, %
Kim et al[65]	NR	ADA 40 mg every other week + MTX (mean) 16.6 mg/week	**ADA 40 mg/2 weeks + DMARD (MTX)**	65	48.5	6.8	21.5 at week 24
		MTX (mean) 16.3 mg/week	**DMARD (MTX)**	63	49.8	6.9	36.5 at week 24
Kremer et al[50]	Moderate–severe active disease despite MTX for at least 3 months	GOL 2 mg/kg every 12 weeks	GOL 2 mg/kg	128	49.9	7.4	80–90 at week 48
		GOL 4 mg/kg every 12 weeks	GOL 4 mg/kg	129	48.4	8.4	
		GOL 2 mk/kg every 12 weeks + MTX	GOL 2 mg/kg + DMARD (MTX)	129	49.7	8.1	
		GOL 4 mg/kg every 12 weeks + MTX	GOL 4 mg/kg + DMARD (MTX)	128	49.6	9.4	
		MTX	**DMARD (MTX)**	129	50.2	7.4	
Kremer et al[49]	Active disease despite DMARD A high degree of baseline activity based on number of tender and swollen joints	ABA 2 mg/kg, day 1, 15, and 30, and then monthly + MTX 10–30 mg/week	ABA 2 mg/kg/4 weeks + DMARD (MTX)	105	54.4	9.6	22
		ABA 10 mg/kg, day 1, 15 and 30 and then monthly + MTX 10–30 mg/week	**ABA 10 mg/kg/4 weeks + DMARD (MTX)**	115	55.8	9.7	14
		MTX 10–30 mg/week	**DMARD (MTX)**	119	54.7	8.9	34
Kremer et al[66]	Active RA despite MTX DAS28 6.4 for both groups	ABA ~10 mg/kg on day 1, 15, 29, and monthly + MTX ≥10 mg/week	**ABA 10 mg/kg/4 weeks + DMARD (MTX)**	433	51.5	8.5	11
		MTX ≥ 10 mg/week	**DMARD (MTX)**	219	50.4	8.9	26
Lan et al[67]	Active disease despite previous MTX	ETN 25 mg twice weekly + MTX 12.5–20 mg/week	**ETN 2 × 25 mg/week + DMARD (MTX)**	29	47.6	>1	7
		MTX 12.5–20 mg/week	**DMARD (MTX)**	29	50.8	7	7
Maini et al[68]	Active disease despite MTX A considerable level of disease activity	INF 3 mg/kg every 8 weeks + MTX	INF 3 mg/kg/8 weeks + DMARD (MTX)	86	56	8.4	9–18
		INF 3 mg/kg every 4 weeks + MTX	INF 3 mg/kg/4 weeks + DMARD (MTX)	86	51	7.2	
		INF 10 mg/kg every 8 weeks + MTX	INF 10 mg/kg/8 weeks + DMARD (MTX)	87	55	9	
		INF 10 mg/kg every 4 weeks + MTX	INF 10 mg/kg/4 weeks + DMARD (MTX)	81	52	8.7	
		MTX	**DMARD (MTX)**	88	51	8.9	36
Maini et al[51]	Active disease despite MTX DAS28 ESR range 6.34–6.75	TOC, 2 mg/kg every 4 weeks	TOC 2 mg/kg every 4 weeks	53	52.2	9.19 mo	22.6
		TOC, 4 mg/kg every 4 weeks	TOC 4 mg/kg every 4 weeks	54	49.3	9.79 mo	20
		TOC, 8 mg/kg every 4 weeks	TOC 8 mg/kg every 4 weeks	52	50.1	9.21 mo	15
		TOC, 2 mg/kg every 4 weeks + MTX	TOC 2 mg/kg/4 weeks + DMARD (MTX)	52	49.2	9.33 mo	11.5
		TOC, 4 mg/kg every 4 weeks + MTX	TOC 4 mg/kg/4 weeks + DMARD (MTX)	49	50.2	7.82 mo	14
		TOC, 8 mg/kg every 4 weeks + MTX	**TOC 8 mg/kg/4 weeks + DMARD (MTX)**	50	50.1	10.62 mo	14
		MTX	**DMARD (MTX)**	49	50.9	11.24 mo	18

Study	Patient population	Treatment	Treatment	N			
Schiff et al[69]	Active disease despite MTX DAS28 6.8–6.9	ABA ~10 mg/kg on day 1, 15, 29, and monthly + MTX	**ABA 10 mg/kg/4 weeks + DMARD (MTX)**	156	49	7.9	5.8
		INF 3 mg/kg every 8 weeks + MTX	**INF 3 mg/kg/8 weeks + DMARD (MTX)**	165	49.1	7.3	7.9
Smolen et al[52]	Moderate–severe active RA despite prior MTX DAS 6.8	MTX	DMARD (MTX)	110	49.4	8.4	2.7
		TOC 4 mg/kg IV every 4 weeks + MTX 10–25 mg/week (mean) 14.7 mg	TOC 4 mg/kg/4 weeks + DMARD (MTX)	214	51.4	7.4	13
		TOC 8 mg/kg IV every 4 weeks + MTX 10–25 mg/week (mean) 14.5 mg	**TOC 8 mg/kg/4 weeks + DMARD (MTX)**	205	50.8	7.5	6.8
Smolen et al[45]	Active disease despite prior MTX DAS28 ESR 6.8–6.85	MTX (mean) 14.8 mg/week	DMARD (MTX)	204	50.6	7.8	7.4
		CZP 200 mg SC + MTX (mean) 12.5 mg/week	**CZP 200 mg/2 weeks + DMARD (MTX)**	246	52.2	6.1	29.3 at week 24
		CZP 400 mg SC + MTX (mean) 12.6 mg/week	CZP 400 mg/2 weeks + DMARD (MTX)	246	51.9	6.5	26.4 at week 24
van Riel et al[47]	Active RA despite MTX Mean baseline DAS28 6.2–6.3	MTX (mean) 12.2 mg/week	DMARD (MTX)	127	51.5	5.6	86.6 at week 24
		ETN 25 mg twice weekly	ETN 2 × 25 mg/week	160	53	10	10.7
		ETN 25 mg twice weekly + MTX ≥ 12.5 mg/week	**ETN 2 × 25 mg/week + DMARD (MTX)**	155	54	9.8	8.4
Weinblatt et al[71]	Active disease despite prior MTX	ADA 20 mg every other week + MTX (mean) 16.9 mg/week	ADA 20 mg/2 weeks + DMARD (MTX)	69	53.5	13.1	NR
		ADA 40 mg every other week + MTX (mean) 16.4 mg/week	**ADA 40 mg/2 weeks + DMARD (MTX)**	67	57.2	12.2	NR
		ADA 80 mg every other week + MTX (mean) 17.2 mg/week	ADA 80 mg/2 weeks + DMARD (MTX)	73	55.5	12.8	NR
Weinblatt et al[70]	Active disease despite MTX	ETN 25 mg twice weekly + MTX (mean) 19 mg/week	**ETN 2 × 25 mg/week + DMARD (MTX)**	62	56.0	11.1	NR
		MTX (mean) 18 mg/week	DMARD (MTX)	59	48	13	3
Westhovens et al[72]	Moderate–severe active disease despite MTX Mean DAS28 5.1	INF 3 mg/kg + MTX	**INF 3 mg/kg/8 weeks + DMARD (MTX)**	360	53	7.8 (median)	7.2 at week 22
		INF 10 mg/kg + MTX	INF 10 mg/kg/8 weeks + DMARD (MTX)	361	52	6.3 (median)	8.9 at week 22
Zhang[53]	Active RA despite MTX (definition of active RA was ≥3 swollen joints and ≥8 joints)	Placebo + MTX	**DMARD (MTX)**	363	52	8.4 (median)	6.3 at week 22
		INF, 3 mg/kg IV infusion weeks 0, 2, 6, and 14 + MTX	**INF 3 mg/kg/8 weeks + DMARD (MTX)**	87	47.9	85.6 mo	10.3
		Placebo + MTX	**DMARD (MTX)**	86	48.9	96.0 mo	17.4

Notes: Treatments in bold are treatments of interest (licensed doses); DMARD is the reference treatment.
Abbreviations: ABA, Abatacept; ADA, Adalimumab; ANA, anakinra; CZP, certolizumab pegol; DMARD, disease-modifying antirheumatic drug; ETN, etanercept; GOL, golimumab; INF, infliximab; mo, months; MP, methyl prednisolone; MTX, methotrexate; NR, not reported; RA, rheumatoid arthritis; RTX, rituximab; SUL, sulfasalazine; TOC, tocilizumab.

Table 4 ACR20/50/70 data used in the combination therapy NMA

Study	Population for sensitivity analysis	Treatment group for analysis	Follow-up (weeks)	Number of patients	ACR 20 n (%)	ACR 50 n (%)	ACR 70 n (%)
Abe et al[58]	MTX exp; MTX low dose (max 8 mg/wk)	**DMARD (MTX)**	14	47	11 (23.4%)	4 (8.5%)	0 (0.0%)
		INF 3 mg/kg/8 weeks + DMARD (MTX)	14	49	30 (61.2%)	15 (30.6%)	5 (10.2%)
		INF 10 mg/kg/8 weeks + DMARD (MTX)	14	51	27 (52.9%)	18 (35.3%)	8 (15.7%)
Chen et al[59]	MTX exp	**ADA 40 mg/2 weeks + DMARD (MTX)**	12	35	19 (54.3%)	12 (34.3%)	5 (14.3%)
		DMARD (MTX)	12	12	4 (33.3%)	2 (16.7%)	0 (0.0%)
Combe et al[42]	DMARD exp	**DMARD (SUL)**	24	50	14 (28.0%)	7 (14.0%)	1 (2.0%)
		ETN 2 × 25 mg/week	24	103	76 (73.8%)	48 (46.6%)	22 (21.4%)
		ETN 2 × 25 mg/week + DMARD (SUL)	24	100	74 (74.0%)	52 (52.0%)	25 (25.0%)
Durez et al[60]	MTX exp	PM 1 mg + DMARD (MTX)	14	15	1 (6.7%)	0 (0.0%)	0 (0.0%)
		INF 3 mg/kg/8 weeks + DMARD (MTX)	14	13	9 (69.2%)	6 (46.2%)	0 (0.0%)
Edwards et al[81]	MTX exp	**DMARD (MTX)**	24	40	15 (37.5%)	5 (12.5%)	2 (5.0%)
		RTX 2 × 1000 mg	24	40	26 (65.0%)	13 (32.5%)	6 (15.0%)
		RTX 2 × 1000 mg + CYC	24	41	31 (75.6%)	17 (41.5%)	6 (14.6%)
		RTX 2 × 1000 mg + DMARD (MTX)	24	40	29 (72.5%)	17 (42.5%)	9 (22.5%)
Emery et al[61]	MTX exp	**DMARD (MTX)**	24	172	40 (23.3%)	16 (9.3%)	9 (5.2%)
		RTX 2 × 500 mg + DMARD (MTX)	24	168	92 (54.8%)	44 (26.2%)	15 (8.9%)
		RTX 2 × 1000 mg + DMARD (MTX)	24	172	87 (50.6%)	45 (26.2%)	17 (9.9%)
Genovese et al[48]	MTX exp	**ETN 2 × 25 mg/week + DMARD (MTX)**	24	80	54 (67.5%)	33 (41.3%)	17 (21.3%)
		ETN 1 × 25 mg/week + ANA + DMARD (MTX)	24	81	41 (50.6%)	32 (39.5%)	19 (23.5%)
		ETN 2 × 25 mg/week + ANA + DMARD (MTX)	24	81	50 (61.7%)	25 (30.9%)	11 (13.6%)
Genovese et al[43]	DMARD exp; <15% TNF-α exp	**TOC 8 mg/kg/4 weeks + DMARD**	24	803	457 (56.9%)	282 (35.1%)	154 (19.2%)
		DMARD	24	413	91 (22.0%)	33 (8.0%)	11 (2.7%)
Huang et al[62]	MTX exp	**ADA 40 mg/2 weeks + DMARD (MTX)**	24	121	88 (72.7%)	49 (40.5%)	21 (17.4%)
		ADA 80 mg/2 weeks + DMARD (MTX)	24	121	86 (71.1%)	48 (39.7%)	21 (17.4%)
		DMARD (MTX)	24	60	41 (68.3%)	26 (43.3%)	11 (18.3%)
Kameda et al[46]	MTX exp; MTX low dose (max 8 mg/wk)	ETN 2 × 25 mg/week	24	74	47 (63.5%)	35 (47.3%)	19 (25.7%)
		ETN 2 × 25 mg/week + DMARD (MTX)	24	77	70 (90.9%)	50 (64.9%)	30 (39.0%)
Kay et al[63]	MTX exp	**DMARD (MTX)**	16	35	13 (37.1%)	2 (5.7%)	0 (0.0%)
		GOL 50 mg/4 weeks + DMARD (MTX)	16	35	21 (60.0%)	13 (37.1%)	3 (8.6%)
		GOL 50 mg/2 weeks + DMARD (MTX)	16	34	17 (50.0%)	8 (23.5%)	5 (14.7%)
		GOL 100 mg/4 weeks + DMARD (MTX)	16	34	19 (55.9%)	10 (29.4%)	6 (17.6%)
		GOL 100 mg/2 weeks + DMARD (MTX)	16	34	27 (79.4%)	11 (32.4%)	3 (8.8%)
Keystone et al[64]	MTX exp	**ADA 40 mg/2 weeks + DMARD (MTX)**	24	207	131 (63.3%)	81 (39.1%)	43 (20.8%)
		ADA 20 mg/2 weeks + DMARD (MTX)	24	212	129 (60.8%)	87 (41.0%)	37 (17.5%)
		DMARD (MTX)	24	200	59 (29.5%)	19 (9.5%)	5 (2.5%)
Keystone et al[44]	MTX exp; <15% TNF-α exp	**DMARD (MTX)**	24	199	27 (13.6%)	15 (7.5%)	6 (3.0%)
		CZP 200 mg/2 weeks + DMARD (MTX)	24	393	231 (58.8%)	146 (37.2%)	84 (21.4%)
		CZP 400 mg/2 weeks + DMARD (MTX)	24	390	237 (60.8%)	156 (40.0%)	80 (20.5%)

Study	Population	Treatment	Weeks	N	n (%)	n (%)
Keystone et al[82]	MTX exp	**DMARD (MTX)**	24	133	37 (27.8%)	18 (13.5%)
		GOL 100 mg/4 weeks	24	133	47 (35.3%)	26 (19.5%)
		GOL 50 mg/4 weeks + DMARD (MTX)	24	89	53 (59.6%)	33 (37.1%)
		GOL 100 mg/4 weeks + DMARD (MTX)	24	89	53 (59.6%)	29 (32.6%)
Kim et al[65]	MTX exp	**DMARD (MTX)**	24	63	23 (36.5%)	9 (14.3%)
		ADA 40 mg/2 weeks + DMARD (MTX)	24	65	40 (61.5%)	28 (43.1%)
Kremer et al[49]	MTX exp; <15% TNF-α exp	**DMARD (MTX)**	24	119	42 (35.3%)	14 (11.8%)
		ABA 2 mg/kg/4 weeks + DMARD (MTX)	24	105	44 (41.9%)	24 (22.9%)
Kremer et al[66]	MTX exp	**ABA 10 mg/kg/4 weeks + DMARD (MTX)**	24	115	69 (60.0%)	42 (36.5%)
		ABA 10 mg/kg/4 weeks + DMARD (MTX)	24	433	294 (67.9%)	173 (40.0%)
Kremer et al[50]	MTX exp; <15% TNF-α exp	**DMARD (MTX)**	24	219	87 (39.7%)	37 (16.9%)
		DMARD (MTX)	24	129	32 (24.8%)	12 (9.3%)
		GOL 2 mg/kg	24	128	29 (22.7%)	11 (8.6%)
		GOL 4 mg/kg	24	129	38 (29.5%)	15 (11.6%)
		GOL 2 mg/kg + DMARD (MTX)	24	129	48 (37.2%)	24 (18.6%)
		GOL 4 mg/kg + DMARD (MTX)	24	128	64 (50.0%)	32 (25.0%)
Lan et al[67]	MTX exp	**ETN 2 × 25 mg/week + DMARD (MTX)**	12	29	26 (89.7%)	19 (65.5%)
		DMARD (MTX)	12	29	10 (34.5%)	3 (10.3%)
Maini et al[68]	MTX exp	**DMARD (MTX)**	30	88	18 (20.5%)	4 (4.5%)
		INF 3 mg/kg/8 weeks + DMARD (MTX)	30	86	43 (50.0%)	23 (26.7%)
		INF 3 mg/kg/4 weeks + DMARD (MTX)	30	86	46 (53.5%)	25 (29.1%)
		INF 10 mg/kg/8 weeks + DMARD (MTX)	30	87	45 (51.7%)	27 (31.0%)
		INF 10 mg/kg/4 weeks + DMARD (MTX)	30	81	47 (58.0%)	21 (25.9%)
Maini et al[51]	MTX exp; <15% TNF-α exp	TOC 2 mg/kg/4 weeks	16	53	16 (30.2%)	3 (5.7%)
		TOC 4 mg/kg/4 weeks	16	54	33 (61.1%)	15 (27.8%)
		TOC 8 mg/kg/4 weeks	16	52	33 (63.5%)	21 (40.4%)
		TOC 2 mg/kg/4 weeks + DMARD (MTX)	16	52	33 (63.5%)	17 (32.7%)
		TOC 4 mg/kg/4 weeks + DMARD (MTX)	16	49	31 (63.3%)	18 (36.7%)
		TOC 8 mg/kg/4 weeks + DMARD (MTX)	16	50	37 (74.0%)	27 (54.0%)
		DMARD (MTX)	16	49	20 (40.8%)	14 (28.6%)
Schiff et al[69]	MTX exp	**ABA 10 mg/kg/4 weeks + DMARD (MTX)**	28	156	104 (66.7%)	63 (40.4%)
		DMARD (MTX)	28	110	46 (41.8%)	22 (20.0%)
		INF 3 mg/kg/8 weeks + DMARD (MTX)	28	165	98 (59.4%)	61 (37.0%)
Smolen et al[52]	MTX exp; <15% TNF-α exp	TOC 4 mg/kg/4 weeks + DMARD (MTX)	24	214	103 (48.1%)	66 (30.8%)
		TOC 8 mg/kg/4 weeks + DMARD (MTX)	24	205	121 (59.0%)	90 (43.9%)
Smolen et al[45]	MTX exp; <15% TNF-α exp	**DMARD (MTX)**	24	204	53 (26.0%)	22 (10.8%)
		DMARD (MTX)	24	127	11 (8.7%)	4 (3.1%)
		CZP 200 mg/2 weeks + DMARD (MTX)	24	246	141 (57.3%)	80 (32.5%)
		CZP 400 mg/2 weeks + DMARD (MTX)	24	246	142 (57.7%)	81 (32.9%)
van Riel et al[47]	MTX exp	ETN 2 × 25 mg/week	16	160	114 (71.3%)	67 (41.9%)
		ETN 2 × 25 mg/week + DMARD (MTX)	16	155	104 (67.1%)	62 (40.0%)
Weinblatt et al[70]	MTX exp	**DMARD (MTX)**	24	30	8 (26.7%)	1 (3.3%)
		ETN 2 × 25 mg/week + DMARD (MTX)	24	59	42 (71.2%)	23 (39.0%)

(Continued)

Table 4 (*Continued*)

Study	Population for sensitivity analysis	Treatment group for analysis	Follow-up (weeks)	Number of patients	ACR 20 n (%)	ACR 50 n (%)	ACR 70 n (%)
Weinblatt et al[71]	MTX exp	**DMARD (MTX)**	24	62	9 (14.5%)	5 (8.1%)	3 (4.8%)
		ADA 20 mg/2 weeks + DMARD (MTX)	24	69	33 (47.8%)	22 (31.9%)	7 (10.1%)
		ADA 40 mg/2 weeks + DMARD (MTX)	24	67	45 (67.2%)	37 (55.2%)	18 (26.9%)
		ADA 80 mg/2 weeks + DMARD (MTX)	24	73	48 (65.8%)	31 (42.5%)	14 (19.2%)
Westhovens et al[72]	MTX exp	**DMARD (MTX)**	22	363	87 (24.0%)	33 (9.1%)	16 (4.4%)
		INF 3 mg/kg/8 weeks + DMARD (MTX)	22	360	199 (55.3%)	110 (30.6%)	48 (13.3%)
		INF 10 mg/kg/8 weeks + DMARD (MTX)	22	361	205 (56.8%)	119 (33.0%)	54 (15.0%)
Zhang et al[53]	MTX exp; <15% TNF-α exp	**INF 3 mg/kg/8 weeks + DMARD (MTX)**	14	87	66 (75.9%)	38 (43.7%)	20 (23.0%)
		DMARD (MTX)	14	86	42 (48.8%)	22 (25.6%)	12 (14.0%)
TEMPO (sensitivity analysis only; data from CSR)[55]	Mixed naïve and experienced population	DMARD (MTX)	24	228	168 (73.7%)	93 (40.8%)	35 (15.4%)
		ETN 2 × 25 mg/week + DMARD (MTX)	24	231	189 (81.8%)	137 (59.3%)	83 (35.9%)

Notes: Treatments in bold are treatments of interest (licensed doses); DMARD is the reference treatment.
Abbreviations: ABA, abatacept; ADA, adalimumab; ANA, anakinra; CSR, corporate social responsibility; CYC, cyclophosphate; CZP, certolizumab pegol; DMARD, disease-modifying anti-rheumatic drug; ETN, etanercept; exp, experienced; GOL, golimumab; INF, infliximab; MTX, methotrexate; PM, pulse methylprednisolone; RTX, rituximab; SUL, sulfasalazine; TOC, tocilizumab.

of the mean response (Figure 3). RAPID 1,[44] RAPID 2,[45] TOWARD,[43] and ARMARDA[71] may underestimate the log-odds of ACR 20 response in the control arm, and hence overestimate the treatment effects (Figure 3). Conversely, AIM,[66] ATTEST,[69] Huang et al,[62] and Zhang et al[53] may overestimate the log-odds of ACR 20 response in the control arm, and hence underestimate the treatment effects (Figure 3). For ACR 50, there is a reasonable spread of control responses either side of the mean response (Figure 4). RAPID 1,[44] RAPID 2,[45] TOWARD,[43] and ATTRACT[68] may overestimate treatment effects, and CHARISMA,[51] ATTEST,[69] Huang et al,[62] and Zhang et al[53] underestimate them (Figure 4). For ACR 70, the spread of control responses is asymmetrical in the direction of lower-than-expected responses (Figure 5). OPTION[52] and TOWARD[43] may overestimate treatment effects, whereas CHARISMA,[51] Huang et al,[62] and Zhang et al[53] may underestimate them (Figure 5).

Meta-analysis results for combination-therapy analysis

The results from the NMA are shown in Table 5 (comparison versus DMARD control), with comparisons of etanercept versus other licensed bDMARDs in Table 6.

The random-effects model did not show a significant difference in ACR 70 for rituximab 2 × 1000 mg + DMARD compared to DMARD alone. Otherwise, all licensed bDMARD combinations have significantly higher odds of ACR 20/50/70 compared to DMARDs alone (based on the 95% CrI, Table 5).

The etanercept combination was significantly better than the other TNF-α inhibitors, adalimumab, and infliximab in improving ACR 20/50/70 outcomes (based on the 95% CrI, Table 6). The etanercept combination was also significantly better than abatacept in improving ACR 20/50/70 outcomes, significantly better than golimumab in improving ACR 20 and rituximab in improving ACR 70 (based on the 95% CrI, Table 6). There were no significant differences between the etanercept combination and certolizumab pegol or tocilizumab.

Regarding model selection, there were sufficient studies for random-effects models to be used. The base case NMA models displayed good convergence, and for all outcomes the random-effects model had the best fit based on lowest DIC and mean residual deviance (the sum of the residual deviances divided by the total number of study arms in the observed data set) (Table 7). For ACR 70 data, a continuity correction was applied in order to account for several instances of zero events in the control arms for this outcome. Between study heterogeneity, as shown by the standard deviation in
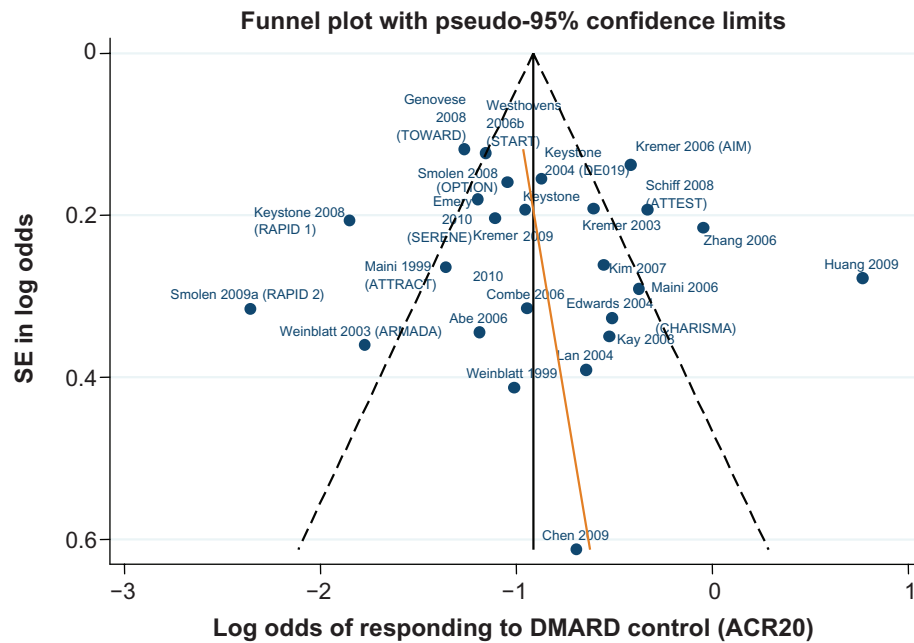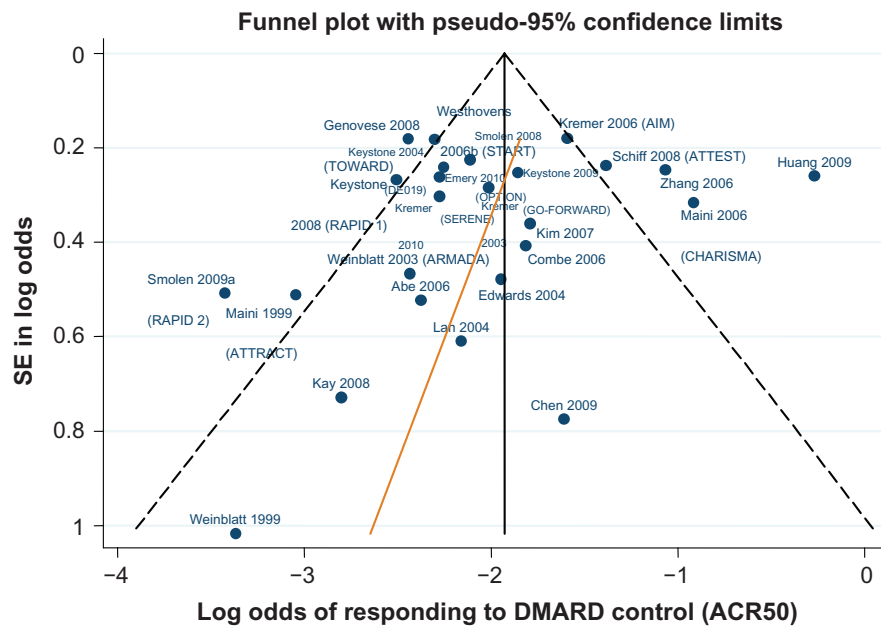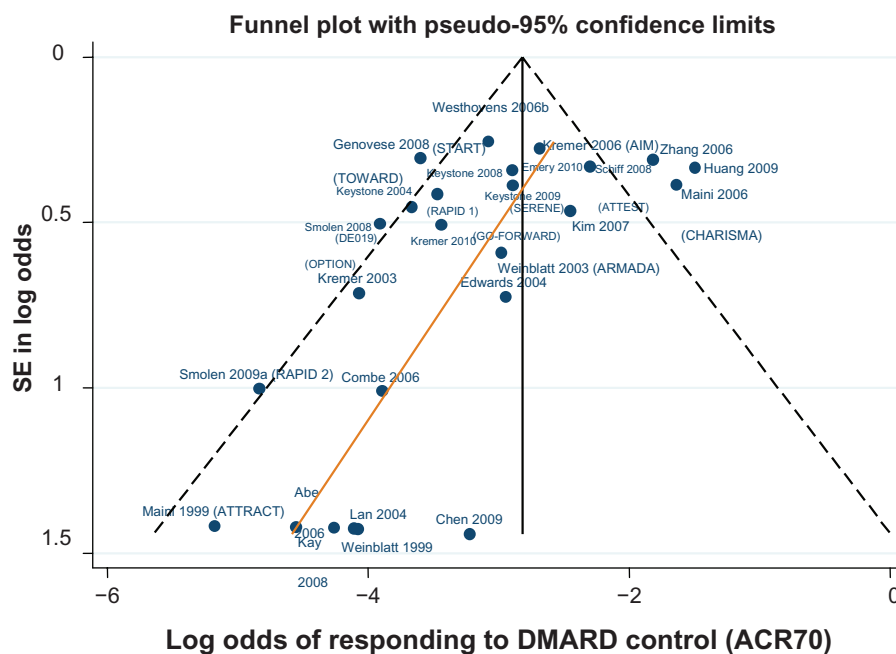
Funnel plot with pseudo-95% confidence limits

Figure 3 Funnel plot comparing the log odds of response across combination study control arms: log odds of DMARD control achieving ACR20.

treatment effects between studies (Table 7) was quite high among studies in the network (ACR 20 standard deviation [SD] on a logarithmic scale = 0.31, ACR 50 SD = 0.40, and ACR 70 SD = 0.50). This suggests that the predicted difference on a natural scale between a study's OR estimate and our NMA estimate may vary (between upper and lower limits) by 3.44 for ACR 20, 4.84 for ACR 50, and 7.23 for ACR 70. There is, therefore, greater uncertainty around the ACR 70 results than around ACR 50 or ACR 20.

The NMA results compare well with the direct head-to-head analysis (Table S1, Table 5) and with the Bucher indirect comparisons (Table S2, Table 6), though no formal test of consistency could be conducted, due to there being no independent loops of evidence. The NMA has a wider CrI compared to direct estimates from head-to-head trials: the lower bounds are similar, but the NMA estimates a much higher upper bound. Similarly, there is more uncertainty (in favor of etanercept) in the NMA estimates of etanercept

Funnel plot with pseudo-95% confidence limits

Figure 4 Funnel plot comparing the log odds of response across combination study control arms: log odds of DMARD control achieving ACR50.

Figure 5

Figure 5 Funnel plot comparing the log odds of response across combination study control arms: log odds of DMARD control achieving ACR70.

versus the other licensed combinations compared to the estimates obtained from the Bucher indirect comparison.

Table 8 shows the results of the study-level covariate analysis, which estimates the treatment effects taking into account the impact of low-dose MTX (maximum dose less than 8 mg/week) and length of follow-up for reporting the ACR outcomes. A low dose of background MTX did not have a statistically significant impact on ACR 20/50/70, and length of follow-up did not have a statistically significant impact on ACR 20 or ACR 50. The βmtx coefficient was statistically significant in the analysis of ACR 70 outcomes (based on the 95% CrI, Table 8), such that a longer length of follow-up was associated with higher odds of ACR 70 response. However, this single significant result should be viewed with caution, since the criteria for significance (type I error, the probability of rejecting the null hypothesis when it is true – < 5%) does not take into account multiple significance testing, ie, no correction for multiple testing was applied, and no reduction in the criteria for significance (to 1%, for example) was made to keep the type I error to a minimum, and as such this result could have occurred by chance.

In an additional covariate analysis of patient characteristics (at study-arm level), longer disease duration was associated with higher odds of ACR 50 and higher patient age with higher odds of ACR 70. Otherwise, age and disease duration were not statistically significant (Table S3). As the base case and covariate odds ratios for each treatment are not largely different, our conclusion regarding the differential effectiveness of etanercept vs adalimumab or infliximab remains unaltered.

The results of the pr-defined sensitivity analyses are shown in Table S4. Removing the RAPID 1/2 or TNF-α-exposed trials had very little impact on the treatment-effect estimates. Removing the etanercept studies had very little impact on the treatment-effect estimates for etanercept for ACR 50 and 70 outcomes, but lowered the treatment-effect estimates for etanercept and certolizumab for ACR 20. The inclusion of the TEMPO study lowered the treatment-effect estimates for etanercept for ACR 20/50/70.

Systematic review results for monotherapy studies

Fourteen studies qualified for inclusion in the analysis of bDMARD monotherapy (Figure 6): two studies with a licensed-dose adalimumab arm,[74,78] plus one additional study with nonlicensed adalimumab arms,[79] five trials including licensed-dose etanercept,[42,46,47,73,75] plus one additional study including nonlicensed etanercept,[80] and three studies including licensed tocilizumab.[51,76,77] There were two additional studies that included only nonlicensed rituximab and golimumab arms.[81,82] There were no studies included in the review that assessed a certolizumab monotherapy arm: the FAST4 WARD trial[83] was excluded on the basis that patients may have had a prior bDMARD (other than TNF-α).

Study characteristics and patient characteristics are summarized in Table 9. All studies were double-blind, with the exception of two that were open-label.[46,47] The range of

Table 5 American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 network meta-analysis base case results for combination treatments in DMARD-experienced patients: licensed biologic DMARD combinations versus DMARD alone

	Fixed effect			Random effects		
	OR v DMARD (95% CrI)	% of patients with event (95% CrI)	Probability of best, %	OR v DMARD (95% CrI)	% of patients with event (95% CrI)	Probability of best, %
ACR 20						
DMARD	–	29.3% (24.9%, 34%)	0.0%	–	29.3% (25%, 34%)	0.0%
ABA 10 mg/kg/4 weeks + DMARD	3.288 (2.597, 4.18)‡	57.6% (49.5%, 65.3%)	0.0%	3.255 (2.056, 5.159)‡	57.4% (44.8%, 69.2%)	0.0%
ADA 40 mg/2 weeks + DMARD	3.514 (2.65, 4.684)‡	59.3% (50.5%, 67.6%)	0.0%	3.439 (2.187, 5.303)‡	58.7% (46.3%, 69.8%)	0.0%
CZP 200 mg/2 weeks + DMARD	10.69 (7.383, 15.69)‡	81.6% (74.2%, 87.4%)	70.7%	11.06 (6.055, 21.06)‡	82.1% (70.7%, 90%)	64.2%
ETN 2 × 25 mg/week + DMARD	8.917 (5.25, 15.53)‡	78.7% (67.5%, 87%)	29.4%	9.341 (4.845, 19.29)‡	79.5% (66%, 89.2%)	35.1%
GOL 50 mg/4 weeks + DMARD	3.508 (2.18, 5.735)‡	59.2% (46.2%, 71.3%)	0.0%	3.387 (1.604, 6.863)‡	58.4% (39.4%, 74.6%)	0.2%
INF 3 mg/kg/8 weeks + DMARD	3.364 (2.71, 4.175)‡	58.2% (50.5%, 65.4%)	0.0%	3.347 (2.271, 4.983)‡	58.1% (47.1%, 68.5%)	0.0%
RTX 2 × 1000 mg + DMARD	3.592 (2.378, 5.47)‡	59.8% (48.3%, 70.5%)	0.0%	3.716 (1.915, 7.418)‡	60.6% (43.5%, 76.1%)	0.3%
TOC 8 mg/kg/4 weeks + DMARD	4.497 (3.619, 5.61)‡	65% (57.8%, 71.7%)	0.0%	4.399 (2.704, 7.125)‡	64.6% (51.7%, 75.5%)	0.2%
ACR 50						
DMARD	–	12% (9.4%, 15.1%)	0.0%	–	12% (9.4%, 15.1%)	0.0%
ABA 10 mg/kg/4 weeks + DMARD	3.57 (2.718, 4.727)‡	32.7% (24.9%, 41.9%)	0.0%	3.633 (2.093, 6.341)‡	33.1% (21.1%, 47.8%)	0.0%
ADA 40 mg/2 weeks + DMARD	3.899 (2.831, 5.446)‡	34.7% (25.9%, 44.9%)	0.0%	3.87 (2.303, 6.598)‡	34.5% (22.6%, 49%)	0.1%
CZP 200 mg/2 weeks + DMARD	9.013 (5.636, 14.94)‡	55.1% (41.7%, 68.5%)	37.2%	9.773 (4.604, 22.65)‡	57.1% (37.5%, 76.3%)	37.6%
ETN 2 × 25 mg/week + DMARD	10.22 (5.28, 22.3)‡	58.3% (40.5%, 76%)	61.0%	11.15 (4.947, 27.95)‡	60.4% (39.2%, 79.8%)	55.6%
GOL 50 mg/4 weeks + DMARD	4.532 (2.521, 8.237)‡	38.2% (24.5%, 54.3%)	0.8%	4.917 (2.051, 12.34)‡	40.1% (21.2%, 63.6%)	3.6%
INF 3 mg/kg/8 weeks + DMARD	3.538 (2.724, 4.612)‡	32.5% (24.9%, 41.2%)	0.0%	3.602 (2.246, 5.924)‡	32.9% (22.3%, 46.3%)	0.1%
RTX 2 × 1000 mg + DMARD	3.851 (2.28, 6.819)‡	34.3% (22.5%, 49.6%)	0.2%	4.103 (1.821, 9.73)‡	35.9% (19.2%, 57.8%)	1.4%
TOC 8 mg/kg/4 weeks + DMARD	5.838 (4.425, 7.773)‡	44.3% (35%, 54%)	0.8%	5.401 (2.911, 9.561)‡	42.3% (27.4%, 58%)	1.6%
ACR 70†						
DMARD	–	4.7% (3.3%, 6.5%)	0.0%	–	4.7% (3.4%, 6.5%)	0.0%
ABA 10 mg/kg/4 weeks + DMARD	3.524 (2.435, 5.213)‡	14.7% (9.4%, 22.6%)	0.0%	3.954 (1.974, 8.8)‡	16.3% (8.2%, 31.6%)	0.2%
ADA 40 mg/2 weeks + DMARD	3.963 (2.513, 6.5)‡	16.3% (9.8%, 26.2%)	0.0%	3.868 (1.91, 7.83)‡	16% (7.9%, 29.2%)	0.1%
CZP 200 mg/2 weeks + DMARD	11.43 (5.472, 27.86)‡	35.9% (19.8%, 59%)	25.5%	13.18 (4.489, 43.5)‡	39.2% (17.2%, 69.2%)	28.5%
ETN 2 × 25 mg/week + DMARD	18.99 (5.098, 130.6)‡	48.3% (19.4%, 86.6%)	69.5%	20.69 (4.921, 158.6)‡	50.6% (18.8%, 89%)	64.7%
GOL 50 mg/4 weeks + DMARD	4.73 (2.069, 12.2)‡	18.9% (8.6%, 39%)	1.3%	4.988 (1.401, 18.28)‡	19.6% (6.2%, 48.4%)	2.7%
INF 3 mg/kg/8 weeks + DMARD	3.549 (2.479, 5.133)‡	14.8% (9.5%, 22.4%)	0.0%	3.694 (2.021, 7.307)‡	15.4% (8.3%, 28%)	0.1%
RTX 2 × 1000 mg + DMARD	2.399 (1.168, 5.272)‡	10.6% (5%, 21.8%)	0.0%	2.644 (0.909, 8.387)	11.5% (4%, 30%)	0.2%
TOC 8 mg/kg/4 weeks + DMARD	7.991 (5.211, 12.79)‡	28.2% (18.4%, 41.2%)	3.7%	7.656 (3.442, 16.5)‡	27.2% (13.7%, 46.6%)	3.6%

Notes: †ACR 70 data with continuity correction; ‡licensed combination had significantly higher odds (based on the 95% CrI) compared to DMARD.
Abbreviations: ABA, abatacept; ADA, adalimumab; ANA, anakinra; CrI, credible interval (Bayesian probability interval); CZP, certolizumab pegol; DMARD, disease-modifying antirheumatic drug (MTX or SUL); ETN, etanercept; exp, experienced; GOL, golimumab; INF, infliximab; MTX, methotrexate; OR, odds ratio; RTX, rituximab; SUL, sulfasalazine; TOC, tocilizumab.

Table 6 American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 network meta-analysis base case results for combination treatments in DMARD-experienced patients: licensed ETN combination versus other licensed biologic DMARD combination

Treatment	Control	Fixed effects OR v control (95% CrI)	Random effects OR v control (95% CrI)
ACR 20			
ETN 2 × 25 mg/week + DMARD	ABA 10 mg/kg/4 weeks + DMARD	2.715 (1.521, 4.956)‡	2.858 (1.306, 6.815)‡
ETN 2 × 25 mg/week + DMARD	ADA 40 mg/2 weeks + DMARD	2.53 (1.405, 4.742)‡	2.72 (1.235, 6.357)‡
ETN 2 × 25 mg/week + DMARD	CZP 200 mg/2 weeks + DMARD	0.836 (0.437, 1.613)	0.846 (0.341, 2.173)
ETN 2 × 25 mg/week + DMARD	GOL 50 mg/4 weeks + DMARD	2.546 (1.235, 5.249)‡	2.759 (1.066, 7.88)‡
ETN 2 × 25 mg/week + DMARD	INF 3 mg/kg/8 weeks + DMARD	2.651 (1.509, 4.791)‡	2.786 (1.299, 6.301)‡
ETN 2 × 25 mg/week + DMARD	RTX 2 × 1000 mg + DMARD	2.48 (1.278, 4.958)‡	2.521 (0.966, 6.711)
ETN 2 × 25 mg/week + DMARD	TOC 8 mg/kg/4 weeks + DMARD	1.987 (1.115, 3.602)‡	2.121 (0.959, 5.107)
ACR 50			
ETN 2 × 25 mg/week + DMARD	ABA 10 mg/kg/4 weeks + DMARD	2.871 (1.395, 6.523)‡	3.07 (1.161, 8.969)‡
ETN 2 × 25 mg/week + DMARD	ADA 40 mg/2 weeks + DMARD	2.625 (1.249, 6.101)‡	2.882 (1.082, 8.347)‡
ETN 2 × 25 mg/week + DMARD	CZP 200 mg/2 weeks + DMARD	1.144 (0.492, 2.847)‡	1.143 (0.358, 3.715)
ETN 2 × 25 mg/week + DMARD	GOL 50 mg/4 weeks + DMARD	2.264 (0.924, 5.999)‡	2.277 (0.672, 7.943)
ETN 2 × 25 mg/week + DMARD	INF 3 mg/kg/8 weeks + DMARD	2.896 (1.426, 6.583)‡	3.098 (1.186, 8.671)‡
ETN 2 × 25 mg/week + DMARD	RTX 2 × 1000 mg + DMARD	2.662 (1.109, 6.817)‡	2.714 (0.826, 9.174)
ETN 2 × 25 mg/week + DMARD	TOC 8 mg/kg/4 weeks + DMARD	1.759 (0.849, 4.018)	2.068 (0.766, 6.284)
ACR 70†			
ETN 2 × 25 mg/week + DMARD	ABA 10 mg/kg/4 weeks + DMARD	5.405 (1.348, 39.22)‡	5.278 (1.016, 46.3)‡
ETN 2 × 25 mg/week + DMARD	ADA 40 mg/2 weeks + DMARD	4.826 (1.171, 34.53)‡	5.45 (1.07, 45.914)‡
ETN 2 × 25 mg/week + DMARD	CZP 200 mg/2 weeks + DMARD	1.661 (0.329, 13.06)	1.636 (0.244, 14.84)
ETN 2 × 25 mg/week + DMARD	GOL 50 mg/4 weeks + DMARD	4.055 (0.796, 31.279)	4.312 (0.604, 48.757)
ETN 2 × 25 mg/week + DMARD	INF 3 mg/kg/8 weeks + DMARD	5.395 (1.358, 38.16)‡	5.642 (1.126, 48.13)‡
ETN 2 × 25 mg/week + DMARD	RTX 2 × 1000 mg + DMARD	7.924 (1.686, 59.453)‡	8.058 (1.225, 78.37)‡
ETN 2 × 25 mg/week + DMARD	TOC 8 mg/kg/4 weeks + DMARD	2.385 (0.593, 16.28)	2.766 (0.535, 25.2)

Notes: †ACR 70 data with continuity correction; ‡licensed ETN combination has significantly higher odds of ACR outcome compared to other licensed biologic DMARD combination (based on the 95% CrI).
Abbreviations: ABA, abatacept; ADA, adalimumab; ANA, anakinra; CrI, credible interval (Bayesian probability interval); CZP, certolizumab pegol; DMARD, disease-modifying antirheumatic drug (MTX or SUL); ETN, etanercept; exp, experienced; GOL, golimumab; INF, infliximab; MTX, methotrexate; OR, odds ratio; RTX, rituximab; SUL, sulfasalazine; TOC, tocilizumab.

baseline disease severity, as measured by DAS28 score, was from moderate–severe (DAS28 5.0–5.2[42]) to very severe (DAS28 7.0–7.1[78,79]). Seven studies enrolled anti-TNF-α-naïve patients.[42,46,47,73–75,81] In one study, 14% of patients had prior exposure to etanercept or infliximab, but not in the 12 weeks prior to enrolment.[51] In another,[78] there had been no biologic treatment permitted in the 6 months prior to enrolment. In two,[77,84] the status was not reported (so for these, an assumption of patients being anti-TNF-α-naïve was made). The mean age ranged from 51[75] to 57[74] years. The percentage of female participants in any treatment arm varied from 61%[80] to 90%.[84] Mean disease duration ranged from 8.4 years[74] to 13 years.[75] The risk of bias was highest in the open-label studies.[46,47] Data for ACR 20/50/70 for the monotherapy analysis are presented in Table 10.

The patients enrolled in the adalimumab studies were broadly similar in terms of disease duration, but one of the adalimumab trials[78] may have involved patients who had had some prior biologic exposure (though not in the 6 months prior to enrolment), and may therefore have enrolled a group less likely to respond.

Tocilizumab studies had on average a shorter disease duration (8.3[77] and 8.5 years[84]) compared to the etanercept and adalimumab monotherapy studies. Figures 7–9 show funnel plots for ACR 20/50/70, respectively, for all studies with placebo control arm used in the monotherapy meta-analysis.

Meta-analysis results for monotherapy analysis

The results from the NMA are shown in Table 11 (comparison versus placebo control), with comparisons of etanercept versus other licensed bDMARDs shown in Table 12. Licensed-dose etanercept, adalimumab, and tocilizumab monotherapy were significantly better than placebo in improving ACR 20/50/70 outcomes (based on the 95% CrI, Table 11). Etanercept monotherapy was significantly better than sulfasalazine in improving ACR 20/50/70 outcomes (based on the 95% CrI, Table 12).

Table 7 Comparison of model fit for base case combination therapy network meta-analysis models

Analysis	DIC	Average residual deviance[††]	SD in treatment effects
ACR 20 combination DMARD-experienced (fixed effect)	615.154	1.306	NA
ACR 20 combination DMARD-experienced (random effects)	607.278	1.071	0.3151
ACR 50 combination DMARD experienced (fixed effect)	591.925	1.302	NA
ACR 50 combination DMARD-experienced (random effects)	579.904	1.015	0.4022
ACR 70 combination DMARD-experienced (fixed effect with cc)	530.861	1.272	NA
ACR 70 combination DMARD-experienced (random effects with cc)	521.232	1.028	0.5047
ACR 70 combination DMARD-experienced (fixed effect no cc)	DNC		
ACR 70 combination DMARD-experienced (random effects no cc)	DNC		

Note: [††]Sum of the residual deviance divided by total number of arms.
Abbreviations: NA, not applicable to fixed-effect models; DNC, did not converge; cc, continuity correction; DIC, deviance information criterion; SD, standard deviation.

Table 8 Results from combination therapy network meta-analysis study-level covariate analysis for American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 end point

Base case: random effects	OR v DMARD (95% CrI)[‡]		
	ACR 20	**ACR 50**	**ACR 70**[†]
ABA 10 mg/kg/4 weeks + DMARD	3.255 (2.056, 5.159)[‡]	3.633 (2.093, 6.341)[‡]	3.954 (1.974, 8.8)[‡]
ADA 40 mg/2 weeks + DMARD	3.439 (2.187, 5.303)[‡]	3.87 (2.303, 6.598)[‡]	3.868 (1.91, 7.83)[‡]
CZP 200 mg/2 weeks + DMARD	11.06 (6.055, 21.06)[‡]	9.773 (4.604, 22.65)[‡]	13.18 (4.489, 43.5)[‡]
ETN 2 × 25 mg/week + DMARD	9.341 (4.845, 19.29)[‡]	11.15 (4.947, 27.95)[‡]	20.69 (4.921, 158.6)[‡]
GOL 50 mg/4 weeks + DMARD	3.387 (1.604, 6.863)[‡]	4.917 (2.051, 12.34)[‡]	4.988 (1.401, 18.28)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.347 (2.271, 4.983)[‡]	3.602 (2.246, 5.924)[‡]	3.694 (2.021, 7.307)[‡]
RTX 2 × 1000 mg + DMARD	3.716 (1.915, 7.418)[‡]	4.103 (1.821, 9.73)[‡]	2.644 (0.909, 8.387)
TOC 8 mg/kg/4 weeks + DMARD	4.399 (2.704, 7.125)[‡]	5.401 (2.911, 9.561)[‡]	7.656 (3.442, 16.5)[‡]
Covariate analysis			
ABA 10 mg/kg/4 weeks + DMARD	3.314 (1.849, 5.878)[‡]	3.121 (1.656, 5.854)[‡]	2.994 (1.433, 6.802)[‡]
ADA 40 mg/2 weeks + DMARD	3.418 (2.051, 5.675)[‡]	3.862 (2.249, 6.904)[‡]	3.934 (2.035, 8.078)[‡]
CZP 200 mg/2 weeks + DMARD	11.18 (5.464, 23.63)[‡]	8.867 (3.968, 21.15)[‡]	10.5 (3.789, 37.14)[‡]
ETN 2 × 25 mg/week + DMARD	10.19 (3.733, 31.23)[‡]	24.67 (6.58, 109.5)[‡]	32.6 (4.276, 1399)[‡]
GOL 50 mg/4 weeks + DMARD	3.31 (1.436, 7.456)[‡]	6.091 (2.408, 16.59)[‡]	7.872 (2.3, 29.78)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.371 (2.117, 5.416)[‡]	3.496 (2.131, 5.891)[‡]	3.758 (2.044, 7.593)[‡]
RTX 2 × 1000 mg + DMARD	3.809 (1.759, 8.296)[‡]	3.734 (1.556, 9.128)[‡]	2.208 (0.793, 7.004)
TOC 8 mg/kg/4 weeks + DMARD	4.428 (2.483, 7.934)[‡]	5.335 (2.821, 9.869)[‡]	7.921 (3.773, 17.58)[‡]
Coefficients (on log scale)[‡]	**Median (95% CrI)**		
βmtx: standard MTX dose	−0.037 (−1.092, 1.002)	−0.563 (−1.855, 0.631)	−1.775 (−4.292, 0.13)
βweeks: weeks of follow-up	−0.004 (−0.055, 0.046)	0.039 (−0.018, 0.098)	0.086 (0.012, 0.163)[+]
Covariate analysis: adjusted odds ratios^	**OR v DMARD (95% CrI)**		
ABA 10 mg/kg/4 weeks + DMARD	3.265 (1.879, 5.621)[‡]	3.396 (1.871, 6.122)[‡]	3.487 (1.77, 7.447)[‡]
ADA 40 mg/2 weeks + DMARD	3.374 (2.002, 5.65)[‡]	4.203 (2.421, 7.558)[‡]	4.58 (2.342, 9.481)[‡]
CZP 200 mg/2 weeks + DMARD	11.02 (5.441, 23.06)[‡]	9.647 (4.354, 22.61)[‡]	12.2 (4.501, 43.03)[‡]
ETN 2 × 25 mg/week + DMARD	10.04 (3.618, 31.17)[‡]	26.9 (6.909, 122.9)[‡]	38.2 (4.743, 1636)[‡]
GOL 50 mg/4 weeks + DMARD	3.268 (1.357, 7.651)[‡]	6.612 (2.524, 18.99)[‡]	9.151 (2.527, 36.8)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.323 (2.086, 5.376)[‡]	3.8 (2.282, 6.521)[‡]	4.361 (2.335, 9.025)[‡]
RTX 2 × 1000 mg + DMARD	3.753 (1.761, 8.115)[‡]	4.064 (1.711, 9.807)[‡]	2.563 (0.949, 8.036)[‡]
TOC 8 mg/kg/4 weeks + DMARD	4.363 (2.441, 7.89)[‡]	5.797 (3.035, 10.88)[‡]	9.23 (4.318, 20.9)[‡]

Notes: ^Results adjusted to 24 weeks of follow-up and standard dose of MTX based on coefficients βweeks and βmtx; [‡]licensed combination had significantly higher odds (based on the 95% CrI) compared to DMARD alone; [†]for the ACR 70 network meta-analysis, a continuity correction (0.5) was applied to the data; the coefficients were not statistically significant except for [+] (for ACR 70 outcome, longer length of follow-up was associated with higher odds of ACR 70).
Abbreviations: ABA, abatacept; ADA, adalimumab; ANA, anakinra; CrI, credible interval (Bayesian probability interval); CZP, certolizumab pegol; DMARD, disease-modifying antirheumatic drug (MTX or SUL); ETN, etanercept; exp, experienced; GOL, golimumab; INF, infliximab; MTX, methotrexate; OR, odds ratio; RTX, rituximab; SUL, sulfasalazine; TOC, tocilizumab.
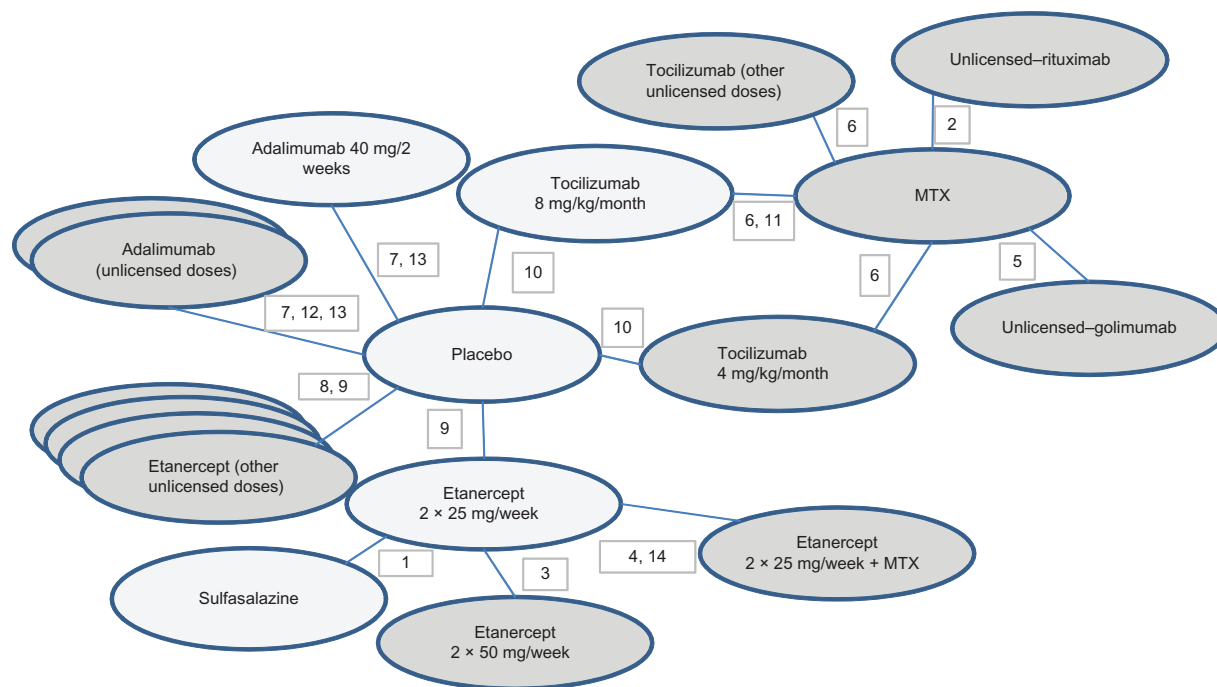
Figure 6 Network diagram for ACR20/50/70 outcomes for bDMARD monotherapy.
Notes: 1, Combe 2006; 2, Edwards 2004; 3, Johnsen 2006; 4, Kameda 2010 (JESMR); 5, Keystone 2009 (GO-FORWARD); 6, Maini 2006 (CHARISMA); 7, Miyasaka 2008 (Change); 8, Moreland 1997; 9, Moreland 1999; 10, Nishimoto 2004 (STREAM); 11, Nishimoto 2009 (SATORI); 12, van de Putte 2003; 13, van de Putte 2004; 14, van Riel 2006 (ADORE). Placebo 6 arms, 444 patients; MTX 4 arms, 488 patients; etanercept 2 × 25 mg/week, 5 arms, 441 patients; tocilizumab 8 mg/kg/4 weeks, 3 arms, 168 patients; adalimumab 40 mg/2 weeks, 2 arms, 204 patients; sulfasalazine 1 arm, 50 patients.

As expected, the NMA had wider confidence intervals compared to direct estimates from head-to-head trials. In general, the NMA models displayed fair convergence, though some of the ACR 70 models did not converge. The random-effects model had the best fit (Table 13). A continuity correction was applied to the ACR 70 data where zero events occurred in the control arms, and between-study heterogeneity estimates were high (ACR 20 SD on a logarithmic scale = 0.81, ACR 50 SD = 0.55, and ACR 70 SD = 0.76). This suggests that the predicted difference on a natural scale between a study's OR estimate and our NMA estimate may vary (between upper and lower limits) by 24.64 for ACR 20, 8.8 for ACR 50, and 19.8 for ACR 70 (Table 13). As a result of between-study heterogeneity, therefore, there is more uncertainty associated with the ACR 20 and ACR 70 treatment-effect estimates for monotherapy, compared to the ACR 50 outcome.

The NMA results compare well with direct head-to-head analysis (Table S5, Table 11) and with Bucher indirect comparisons (Table S6, Table 12). Examination of the monotherapy evidence network shows that there was one independent loop for which inconsistency of the direct and indirect evidence can be assessed (Figure S1). The analysis indicates that the inconsistency on this loop is not significant ($P > 0.05$ for ACR 20/50/70; Tables S7–S9).

A covariate analysis was not conducted, as there were too few monotherapy studies to make such an analysis robust. The results of the predefined sensitivity analyses are shown in Table S10. The inclusion of the TEMPO study lowers the treatment-effect estimates for etanercept monotherapy.

Discussion

bDMARDs, in combination with a conventional DMARD, have been shown to be efficacious in patients who have had an inadequate response to prior DMARD therapy, thus representing an important addition to the RA treatment algorithm for patients and their health-care provider. Based on the clinical data identified in a systematic review, we conducted NMAs obtaining pooled estimates of relative treatment effects, allowing pair-wise comparisons and ranking of licensed bDMARD therapies. We also conducted a separate analysis of bDMARD monotherapy treatments, which are licensed for use in patients who cannot tolerate MTX or for whom MTX is contraindicated. Our results show that all licensed bDMARD combinations have significantly higher odds (based on the 95% CrI) for ACR 20/50/70 compared to MTX or DMARD monotherapy, ACR 70 results for RTX being the only exception.

Table 9 Characteristics of included monotherapy studies

Study	Disease severity	Treatment and dose	Treatment group for analysis	Number of patients randomized	Mean age, years	Mean disease duration, years	Withdrawals, %
Johnsen et al[73]	Active RA despite DMARD DAS28 CRP 6.0–6.2	ETN 25 mg twice weekly	**ETN 2 × 25 mg/week**	26	50.5 (median)	12.5 (median)	12
		ETN 50 mg twice weekly	ETN 2 × 50 mg/week	51	55.0 (median)	15.0 (median)	16
Miyasaka[74]	Active disease despite prior treatment with ≥1 DMARD	Placebo every other week	**PLA**	87	53.4	8.4	59
		ADA 20 mg every other week	ADA 20 mg/2 weeks	87	54.8	10.0	38
		ADA 40 mg every other week	**ADA 40 mg/2 weeks**	91	56.9	9.9	35
		ADA 80 mg every other week	ADA 80 mg/2 weeks	87	54.3	9.5	25
Moreland et al[80]	Active disease despite 1–4 conventional DMARDs	Placebo	**PLA**	44	55	77% had duration	48
		ETN 0.25 mg/m²	ETN 0.25 mg/m²	46	54	>5 years	39
		ETN 2 mg/m²	ETN 2 mg/m²	46	52		22
		ETN 16 mg/m²	ETN 16 mg/m²	44	52		7
Moreland et al[75]	Active RA (advanced) despite DMARD (>90% MTX exposed)	Placebo	**PLA**	80	51	12	33
		ETN 10 mg sc twice weekly	ETN 2 × 10 mg/week	76	53	13	32
		ETN 25 mg sc twice weekly	**ETN 2 × 25 mg/week**	78	53	11	24
Nishimoto et al[77]	Active, relatively severe disease despite ≥1 DMARD	Placebo	**PLA**	53	Median 53.0	Median 8.4	47
		TOC 4 mg/kg every 4 weeks	TOC 4 mg/kg/4 weeks	54	Median 53.5	Median 7.3	4
		TOC 8 mg/kg every 4 weeks	**TOC 8 mg/kg/4 weeks**	55	Median 56.0	Median 8.3	7
Nishimoto et al[76]	Active disease despite prior MTX for ≥8 weeks DAS28 6.1	MTX 8 mg/week	MTX	66	50.8	8.7	48
		TOC 8 mg/kg iv every 4 weeks	**TOC 8 mg/kg/4 weeks**	61	52.6	8.5	11
van de Putte et al[79]	Active RA despite ≥1 DMARD DAS28 7.0–7.1	Placebo	**PLA**	70	50.2	9.4	33
		ADA 20 mg weekly	ADA 20 mg/weekly	72	53.7	10.4	14
		ADA 40 mg weekly	ADA 40 mg/weekly	70	52.6	10.0	17
		ADA 80 mg weekly	ADA 80 mg/weekly	72	53.2	10.1	8
van de Putte et al[78]	Active long-standing severe RA despite ≥1 DMARD DAS 7.07	Placebo	PLA	110	53.5	11.6	56.4
		ADA 20 mg every other week	ADA 20 mg/2 weeks	106	53.1	9.3	35.8
		ADA 20 mg weekly	ADA 20 mg/weekly	112	54.4	11.3	29.5
		ADA 40 mg every other week	**ADA 40 mg/2 weeks**	113	52.7	10.6	28.3
		ADA 40 mg weekly	ADA 40 mg/weekly	103	51.8	11.9	14.6

Notes: Combe et al,[42] Edwards et al,[81] Kameda et al,[46] Keystone et al,[82] Maini et al,[51] and van Riel et al[47] – characteristics as per Table 3; treatments in bold are treatments of interest (licensed doses); PLA is the reference treatment.
Abbreviations: ADA, Adalimumab; DAS, Disease Activity Score; DMARD, disease-modifying anti-rheumatic drugs; ETN, etanercept; MTX, methotrexate; PLA, placebo; RA, rheumatoid arthritis; TOC, Tocilizumab.

Table 10 American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 data used in the monotherapy network meta-analysis

Study	Population for sensitivity analysis	Treatment group for analysis	Follow-up (weeks)	Number of patients	ACR 20 n (%)	ACR 50 n (%)	ACR 70 n (%)
Combe et al[42]	DMARD exp	**SUL**	24	50	14 (28.0%)	7 (14.0%)	1 (2.0%)
		ETN 2 × 25 mg/week	24	103	76 (73.8%)	48 (46.6%)	22 (21.4%)
Edwards et al[81]	MTX exp	MTX	24	40	15 (37.5%)	5 (12.5%)	2 (5.0%)
		RTX 2 × 1000 mg	24	40	26 (65.0%)	13 (32.5%)	6 (15.0%)
Johnsen et al[73]	DMARD exp	**ETN 2 × 25 mg/week**	24	26	17 (65.4%)	10 (38.5%)	4 (15.4%)
		ETN 2 × 50 mg/week	24	51	30 (58.8%)	19 (37.3%)	8 (15.7%)
Kameda et al[46]	MTX exp	**ETN 2 × 25 mg/week**	24	74	47 (63.5%)	35 (47.3%)	19 (25.7%)
		ETN 2 × 25 mg/week + MTX	24	77	70 (90.9%)	50 (64.9%)	30 (39.0%)
Keystone et al[82]	MTX exp	MTX	24	133	37 (27.8%)	18 (13.5%)	7 (5.3%)
		GOL 100 mg/4 weeks	24	133	47 (35.3%)	26 (19.5%)	15 (11.3%)
Maini et al[51]	MTX exp	TOC 2 mg/kg/4 weeks	16	53	16 (30.2%)	3 (5.7%)	1 (1.9%)
		TOC 4 mg/kg/4 weeks	16	54	33 (61.1%)	15 (27.8%)	3 (5.6%)
		TOC 8 mg/kg/4 weeks	16	52	33 (63.5%)	21 (40.4%)	8 (15.4%)
		MTX	16	49	20 (40.8%)	14 (28.6%)	8 (16.3%)
Miyasaka[74]	DMARD exp	**PLA**	24	87	12 (13.8%)	5 (5.7%)	1 (1.1%)
		ADA 20 mg/2 weeks	24	87	25 (28.7%)	14 (16.1%)	9 (10.3%)
		ADA 40 mg/2 weeks	24	91	40 (44.0%)	22 (24.2%)	11 (12.1%)
		ADA 80 mg/2 weeks	24	87	44 (50.6%)	28 (32.2%)	13 (14.9%)
Moreland et al[80]	DMARD exp	**PLA**	12	44	6 (13.6%)	3 (6.8%)	NR
		ETN 0.25 mg/m²	12	46	15 (32.6%)	4 (8.7%)	NR
		ETN 2 mg/m²	12	46	21 (45.7%)	10 (21.7%)	NR
		ETN 16 mg/m²	12	44	33 (75.0%)	25 (56.8%)	NR
Moreland et al[75]	DMARD exp	**PLA**	24	80	9 (11.3%)	4 (5.0%)	1 (1.3%)
		ETN 2 × 10 mg/week	24	76	39 (51.3%)	18 (23.7%)	7 (9.2%)
		ETN 2 × 25 mg/week	24	78	46 (59.0%)	31 (39.7%)	12 (15.4%)
Nishimoto et al[77]	MTX exp	**PLA**	12	53	6 (11.3%)	1 (1.9%)	0 (0.0%)
		TOC 4 mg/kg/4 weeks	12	54	31 (57.4%)	14 (25.9%)	11 (20.4%)
		TOC 8 mg/kg/4 weeks	12	55	43 (78.2%)	22 (40.0%)	9 (16.4%)
Nishimoto et al[76]	MTX exp	MTX	24	66	17 (25.8%)	7 (10.6%)	4 (6.1%)
		TOC 8 mg/kg/4 weeks	24	61	49 (80.3%)	30 (49.2%)	18 (29.5%)
van de Putte et al[79]	DMARD exp	**PLA**	12	70	7 (10.0%)	1 (1.4%)	0 (0.0%)
		ADA 20 mg/week	12	72	36 (50.0%)	17 (23.6%)	8 (11.1%)
		ADA 40 mg/week	12	70	40 (57.1%)	19 (27.1%)	7 (10.0%)
		ADA 80 mg/week	12	72	39 (54.2%)	14 (19.4%)	6 (8.3%)
van de Putte et al[78]	DMARD exp	**PLA**	26	110	21 (19.1%)	9 (8.2%)	2 (1.8%)
		ADA 20 mg/2 weeks	26	106	38 (35.8%)	20 (18.9%)	9 (8.5%)
		ADA 20 mg/week	26	112	44 (39.3%)	23 (20.5%)	11 (9.8%)
		ADA 40 mg/2 weeks	26	113	52 (46.0%)	25 (22.1%)	14 (12.4%)
		ADA 40 mg/week	26	103	55 (53.4%)	36 (35.0%)	19 (18.4%)
van Riel et al[47]	MTX exp	**ETN 2 × 25 mg/week**	16	160	114 (71.3%)	67 (41.9%)	28 (17.5%)
		ETN 2 × 25 mg/week + MTX	16	155	104 (67.1%)	62 (40.0%)	29 (18.7%)
TEMPO (sensitivity analysis only; data from CSR)[55]	Mixed naïve and experienced population	MTX	24	228	168 (73.7%)	93 (40.8%)	35 (15.4%)
		ETN 2 × 25 mg/week	24	231	159 (68.8%)	90 (39.0%)	38 (16.5%)

Notes: Treatments in bold are treatments of interest (licensed doses); PLA is the reference treatment.
Abbreviations: ADA, adalimumab; CSR, corporate social responsibility; DMARD, disease-modifying antirheumatic drug; ETN, etanercept; exp, experienced; GOL, golimumab; INF, infliximab; MTX, methotrexate; NR, not reported; PLA, placebo; RTX, rituximab; SUL, sulfasalazine; TOC, tocilizumab.

For DMARD experienced patients, our results also show that the etanercept combination is significantly better than the adalimumab and infliximab combinations and comparable to the certolizumab combination in improving ACR 20/50/70 outcomes (based on the 95% CrI). Therefore, previous meta-analyses that pooled TNF-α inhibitors into a single group may have underestimated the efficacy of etanercept.[85,86]

The internal validity of any NMA is dependent upon three key considerations: RCT identification, individual RCT quality, and the degree of confounding bias because of similarity or consistency assumptions not being met.
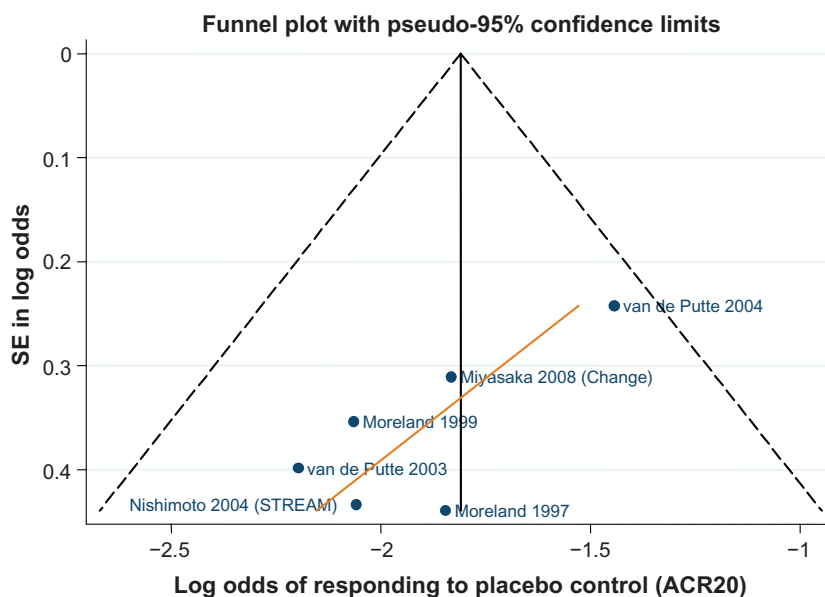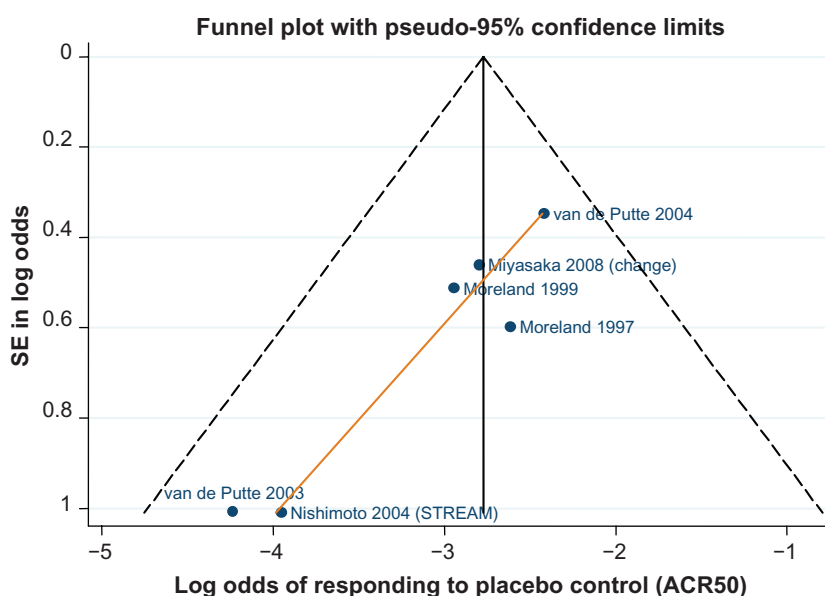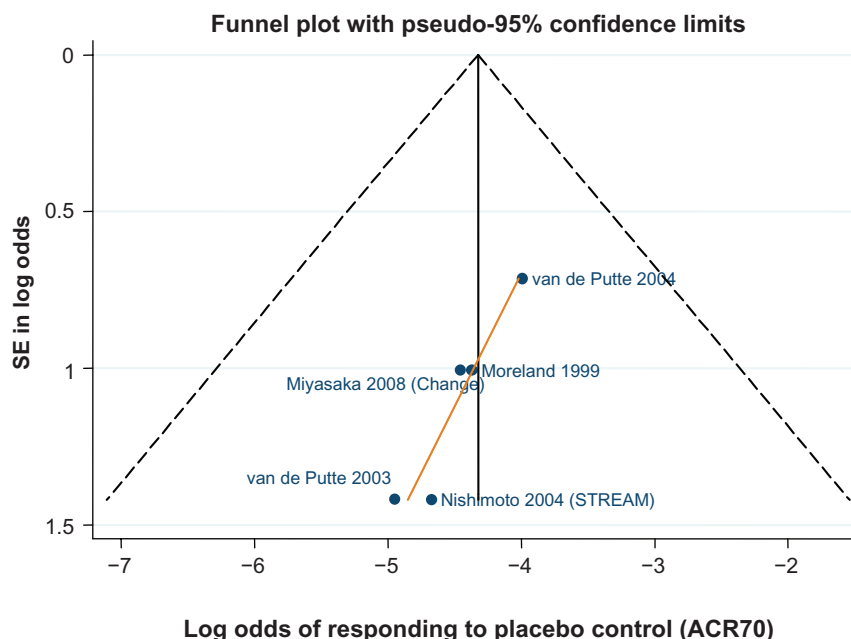
Funnel plot with pseudo-95% confidence limits

Figure 7 Funnel plot comparing the log odds of response across monotherapy study control arms: log odds of placebo control achieving ACR20.

Regarding the first of these, an extensive systematic review was conducted to ensure identification of all relevant RCTs. The extent of publication bias was assessed, the slope of the colored lines in the funnel plots (Figures 3–5, combination NMA; Figures 7–9, monotherapy NMA) indicating a small degree of publication bias. The network of RCTs was fairly balanced for most treatments. In the combination analysis, there was some network asymmetry, however; a greater weight of evidence was available for tocilizumab (three trials and 1058 patients) and a smaller such weight for golimumab (two arms and 124 patients).

Regarding the second consideration, quality assessment of individual RCTs did identify some open-label or early escape design studies that may have been more prone to bias, but the effect of including these in the base case was assessed – by sensitivity analyses – which showed that including these studies did not bias the treatment-effect estimates in favor of etanercept.

Regarding the third consideration, meta-analysis has the underlying assumption that trials and outcomes are sufficiently similar to allow data to be pooled, and the consistency assumption relies on there being no imbalance in modifiers

Funnel plot with pseudo-95% confidence limits

Figure 8 Funnel plot comparing the log odds of response across monotherapy study control arms: log odds of placebo control achieving ACR50.

Funnel plot with pseudo-95% confidence limits

Figure 9 Funnel plot comparing the log odds of response across monotherapy study control arms: log odds of placebo control achieving ACR70.

of relative treatment effects across studies. In our NMA, the similarity assumption was supported by the eligibility criteria applied for study selection, and the adjustment of the results by way of covariate analyses for the potential effect modifiers, low dosing of MTX, length of follow-up, age, and disease duration. This covariate adjustment aimed

to reduce the impact of any bias due to similarity and/or consistency violations. Low dosing of MTX did not have a statistically significant impact on ACR 20/50/70, nor did length of follow-up for ACR 20/50. Longer disease duration was associated with higher odds of ACR 50 and higher age with higher odds of ACR 70. Adjusting for age and

Table 11 American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 network meta-analysis base case results for monotherapy treatments in DMARD-experienced patients: licensed DMARD monotherapy versus placebo

	Fixed effect			Random effects		
	OR v PLA (95% CrI)	% of patients with event (95% CrI)	Probability of best, %	OR v PLA (95% CrI)	% of patients with event (95% CrI)	Probability of best, %
ACR 20						
PLA	–	14.1% (11.1%, 17.7%)	0.0%	–	14% (11.1%, 17.7%)	0.0%
ADA 40 mg/2 weeks	4.947 (3.163, 7.77)[‡]	44.8% (32.4%, 57.9%)	7.1%	5.125 (1.417, 18.62)[‡]	45.5% (18.5%, 75.7%)	6.9%
ETN 2 × 25 mg/week	11.85 (5.371, 29.52)[‡]	66% (45.6%, 83.4%)	28.0%	12 (1.733, 90.94)[‡]	66.3% (21.8%, 93.8%)	23.6%
SUL	1.598 (0.522, 5.172)	20.8% (7.6%, 46.4%)	0.0%	1.608 (0.105, 27.48)	20.7% (1.7%, 82%)	0.4%
TOC 8 mg/kg/4 weeks	26.17 (10, 76.19)[‡]	81.1% (61%, 92.8%)	64.9%	26.25 (3.883, 190.8)[‡]	81.1% (38.6%, 97%)	69.2%
ACR 50						
PLA	–	5.9% (3.9%, 8.7%)	0.0%	–	5.9% (3.9%, 8.7%)	0.0%
ADA 40 mg/2 weeks	4.818 (2.616, 9.344)[‡]	23.2% (12.5%, 39.8%)	7.1%	5.117 (1.819, 16.11)[‡]	24.3% (9.4%, 51.5%)	6.9%
ETN 2 × 25 mg/week	13.83 (4.745, 54.01)[‡]	46.4% (21.3%, 78.1%)	28.0%	13.46 (2.631, 80.29)[‡]	45.6% (13.3%, 84.1%)	23.6%
SUL	2.476 (0.58, 11.84)	13.4% (3.3%, 43.7%)	0.0%	2.379 (0.241, 25.64)	13% (1.4%, 62.7%)	0.4%
TOC 8 mg/kg/4 weeks	46.94 (7.572, 915.5)[‡]	74.7% (31.2%, 98.3%)	64.9%	55.08 (6.204, 1740)[‡]	77.5% (26.9%, 99.1%)	69.2%
ACR 70[†]						
PLA	–	1.3% (0.6%, 3.1%)	0.0%	–	1.3% (0.5%, 3.1%)	0.0%
ADA 40 mg/2 weeks	11.42 (3.866, 44.01)[‡]	13.2% (3.6%, 42.8%)	7.1%	11.71 (2.441, 77.5)[‡]	13.5% (2.5%, 55%)	6.9%
ETN 2 × 25 mg/week	19.49 (3.199, 788.9)[‡]	20.9% (3.3%, 92%)	28.0%	20.83 (1.56, 1740)[‡]	21.8% (1.7%, 96%)	23.6%
SUL	1.042 (0.025, 57.47)	1.4% (0%, 44.8%)	0.0%	1.14 (0.012, 178.8)	1.5% (0%, 71.5%)	0.4%
TOC 8 mg/kg/4 weeks	55.54 (5.138, 6469)[‡]	43% (5.4%, 98.9%)	64.9%	96.27 (3.992, 38820)[‡]	56.1% (4.5%, 99.8%)	69.2%

Notes: [†]ACR 70 data with continuity correction; [‡]licensed monotherapy has significantly higher odds of ACR outcome compared to PLA (based on the 95% CrI).
Abbreviations: ADA, adalimumab; CrI, credible interval (Bayesian probability interval); DNC, did not converge; DMARD, disease-modifying antirheumatic drug; ETN, etanercept; exp, experienced; MTX, methotrexate; OR, odds ratio; PLA, placebo; SUL, sulfasalazine; TOC, tocilizumab.

Table 12 American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 network meta-analysis base case results for monotherapy treatments in DMARD-experienced patients: licensed ETN monotherapy versus other licensed DMARD monotherapy

Treatment	Control	Fixed effects OR v control (95% CrI)	Random effects OR v control (95% CrI)
ACR 20			
ETN 2 × 25 mg/week	ADA 40 mg/2 weeks	2.396 (0.957, 6.623)	2.333 (0.238, 25.484)
ETN 2 × 25 mg/week	SUL	7.452 (3.522, 16.502)[‡]	7.524 (1.059, 52.521)[‡]
ETN 2 × 25 mg/week	TOC 8 mg/kg/4 weeks	0.454 (0.121, 1.702)	0.451 (0.03, 7.062)
ACR 50			
ETN 2 × 25 mg/week	ADA 40 mg/2 weeks	2.866 (0.824, 12.549)	2.641 (0.342, 20.704)
ETN 2 × 25 mg/week	SUL	5.593 (2.408, 14.723)[‡]	5.621 (1.196, 28.531)[‡]
ETN 2 × 25 mg/week	TOC 8 mg/kg/4 weeks	0.296 (0.012, 2.726)	0.244 (0.005, 4.144)
ACR 70[†]			
ETN 2 × 25 mg/week	ADA 40 mg/2 weeks	1.73 (0.197, 82.237)	1.743 (0.066, 195.274)
ETN 2 × 25 mg/week	SUL	19.117 (3.202, 526.039)[‡]	18.744 (1.47, 686.342)[‡]
ETN 2 × 25 mg/week	TOC 8 mg/kg/4 weeks	0.367 (0.002, 24.795)	0.201 (0, 58.072)

Notes: [†]ACR 70 data with continuity correction; [‡]licensed ETN has significantly higher odds of ACR outcome compared to other licensed DMARD (based on the 95% CrI). **Abbreviations:** ADA, adalimumab; CrI, credible interval (Bayesian probability interval); DNC, did not converge; DMARD, disease-modifying antirheumatic drug; ETN, etanercept; exp, experienced; OR, odds ratio; SUL, sulfasalazine; TOC, tocilizumab.

disease duration did not alter our conclusion. We further examined the influence of exposure to prior anti-TNF-α therapies and of incorporating the TEMPO trial, a trial that included some MTX-naïve patients or patients that were not MTX-inadequate responders, by sensitivity analyses: overall, removing subsets of trials had very little impact on treatment-effect estimates, but meta-analysis that included the TEMPO trial[85,87] underestimated the efficacy of etanercept in the DMARD-experienced/inadequate-response population because of the high MTX control arm response rate in patients previously untreated with MTX, ie, these patients were still able to benefit from MTX.

There remained heterogeneity among the studies in our NMA. The patient characteristic that differed across studies but that was not assessed as a covariate was the number of prior DMARD treatments. This is one area, therefore, where the similarity assumption may be challenged, and should be considered for covariate adjustment in future research.

One tocilizumab study, in particular – CHARISMA[51] – enrolled patients with a mean duration of disease of only 9.2 months.[51] This does not appear to have influenced the treatment estimates here, however. The CHARISMA study is small compared to the other tocilizumab studies – OPTION[52] and TOWARD[43] – so will have less weight in the meta-analysis. The random-effects direct meta-analysis of tocilizumab versus DMARD did not indicate any significant heterogeneity in effect on ACR 20 between OPTION, CHARISMA, and TOWARD (ACR 20 $I^2 = 0\%$, $P = 0.86$; ACR 50 $I^2 = 30.6\%$, $P = 0.24$; ACR 70 $I^2 = 59.7\%$, $P = 0.08$). In the direct meta-analysis of tocilizumab versus DMARD for ACR 50 and ACR 70, CHARISMA had a lower ACR 50 and ACR 70 treatment effect (relative to DMARD) compared to OPTION and TOWARD. This was somewhat counterintuitive, as one would expect that patients with shorter disease duration (fewer previous lines of treatment) would have better response to treatment than would patients with longer

Table 13 Comparison of model fit for monotherapy network meta-analysis models

Analysis	DIC	Average residual deviance[††]	SD in treatment effects
ACR 20 monotherapy DMARD-experienced (fixed effects)	280.685	1.531	NA
ACR 20 monotherapy DMARD-experienced (random effects)	265.259	1.012	0.817
ACR 50 monotherapy DMARD-experienced (fixed effects)	255.059	1.224	NA
ACR 50 monotherapy DMARD-experienced (random effects)	252.167	1.042	0.554
ACR 70 monotherapy DMARD-experienced (fixed effects with cc)	208.357	1.244	NA
ACR 70 monotherapy DMARD-experienced (random effects with cc)	205.674	1.045	0.761
ACR 70 monotherapy DMARD-experienced (random effects, no cc)	DNC		
ACR 70 monotherapy DMARD-experienced (fixed effects, no cc)	DNC		

Note: [††]Sum of the residual deviance divided by total number of arms. **Abbreviations:** ACR 20/50/70, American College of Rheumatology (ACR) criteria scores of 20, 50, and 70; DMARD, disease-modifying antirheumatic drug; NA, not applicable to fixed-effect models; DNC, did not converge; cc, continuity correction; DIC, deviance information criterion; SD, standard deviation.

disease duration (who have had more previous treatments). The different ACR 50 and ACR 70 relative effects observed in CHARISMA, therefore, may be due to factors other than disease duration, and we conclude that the short duration of disease in the CHARISMA study population did not impact on the treatment-effect estimates.

Differences in placebo/MTX responses across trials were assessed by way of funnel plots, identifying some studies within the network that under- or overestimated the response to placebo/MTX, meaning that they would over- or underestimate, respectively, the treatment effect. From review of the funnel plots, it can be deduced that the overall treatment effects on ACR 20 and ACR 50 may be overestimated for certolizumab pegol.[44,45] The low response to MTX observed in the certolizumab pegol RAPID 1[44] and RAPID 2[45] trials (see Figures 3 and 4) may be explained by the early escape trial design, whereby patients who had failed to respond at weeks 12–14 were withdrawn from treatment at week 16 and classified as nonresponders. More than half of the patients were withdrawn from the MTX control arms, whereas a lower percentage of certolizumab combination-arm patients were withdrawn. This imbalance in withdrawals may have had an impact on the treatment effects measured by these studies: week 16 withdrawals in RAPID 1 – 62.8% placebo, 21.1% certolizumab 200 mg, 17.4% certolizumab 400 mg; week 16 withdrawals in RAPID2 – 81.1% placebo, 21.1% certolizumab 200 mg, 21.1% certolizumab 400 mg. The ITT primary outcome at 24 weeks suggested a greater treatment effect for CZP compared to placebo than was the case before early escape.

Treatment effects may be overestimated for tocilizumab on ACR 20 and ACR 50[43] and on ACR 70.[43,52] Infliximab treatment effects may be underestimated for ACR 20 and ACR 70.[53,69] For ACR 50, two studies[53,69] underestimated the treatment effect, and one study[68] provided an overestimate. For adalimumab, the effect may be an underestimate for ACR 50 and 70.[62] For ACR 20, one study may have underestimated the adalimumab effect[62] and another overestimated it.[71] The treatment effect of abatacept on ACR 20 may be an underestimate.[66] There were no etanercept studies that were outliers on the funnel plots, suggesting that the treatment effects for etanercept were within the bounds of what might be expected.

The assumption of consistency between the direct and indirect evidence could not be assessed formally in the combination analyses, as there were no independent loops of evidence in the network: for ACR outcomes, there was only one study[69] that compared one licensed-treatment combination (infliximab) directly to another (abatacept) head-to-head. The combination network was comprised solely of indirect comparisons via MTX/DMARD. However, the results of direct meta-analyses and of the indirect Bucher were compared to base case results from the NMA to gauge consistency. For example, etanercept vs DMARD direct (data from etanercept vs DMARD trials only) was compared to etanercept vs DMARD as estimated by the NMA, and etanercept vs other bDMARDs indirect (as no head-to-head data) was compared to etanercept vs other bDMARDs as estimated by the NMA. The NMA had a wider confidence interval compared to the direct estimates from head-to-head trials: when comparing confidence intervals, the lower bounds were similar but the NMA estimated a much higher upper bound. Similarly, there is more uncertainty (in favor of etanercept) in the NMA estimates of etanercept versus the other licensed combinations compared to the estimates obtained from the Bucher indirect comparison. In the monotherapy analyses, one loop of evidence was present in the network, enabling a formal test of the consistency assumption. This test indicated that the direct and indirect treatment-effect estimates were not statistically significantly different, indicating that the consistency assumption held.

The relative treatment-effect estimates observed in our NMA were not influenced by any prior distribution estimates, as noninformative priors were used, meaning that prior to the data being applied, any result was taken to be equally likely, and that the posterior results were driven by the data. Model selection in our NMA was based on the best model fit, as indicated by the lowest DIC and average residual deviance values.

Regarding the heterogeneity observed among studies in the network, it may be argued that this might present a challenge to the similarity assumption. It does, however, better support the external validity of these NMA results: this variation in patient populations is more likely to reflect real-world practice.

Our outcome measure was ACR response, a good short-term measure of disease response that is widely reported in clinical trials of RA. Other measures, such as the HAQ, might be more relevant for longer-term progression measurement. However, HAQ is not so broadly reported and is not as sensitive in measuring short-term changes in RA symptoms.

Our data relate to the population of adult patients with active, moderate–severe RA who have failed on or had an inadequate response to MTX or other conventional DMARDs. In these patients, the treatments evaluated were effective. In relation to other NMAs, our data illustrate that evidence

from the different TNF-α inhibitors should not be combined together in meta-analyses, because efficacy differs between drugs in this class. Etanercept is a fusion protein including a soluble fragment of human p75-soluble TNF receptor and human immunoglobulin G_1, whereas adalimumab and infliximab are MAbs directed against TNF. Differences in the kinetics and mode of action between etanercept and the MAbs have been reported,[88] and these differences may provide a plausible biological rationale for the differences in treatment-effect estimates observed in our NMA. Differences in the findings of published NMAs of biological DMARDs in RA have been reviewed and attributed to methodological shortcomings and inconsistencies.[89] Our NMA was performed incorporating many of the recommended criteria,[89] for a high-quality NMA including clear statement of the population (DMARD-MTX-inadequate responders, as distinct from TNF-α-inadequate responders, or DMARD- or MTX-naïve populations), analyzing monotherapy and combination therapy in separate networks (thereby avoiding lumping of mono- and combination therapy without controlling for concomitant DMARD use), exploring heterogeneity and effect modification by covariate analyses, and examining the influence of particular trials or sets of trials by sensitivity analyses.

Our data do not address treatment effects in the population of patients who have failed TNF-α treatment, as this is a later stage in the treatment pathway. Likewise, further review work would be required to gain treatment-effect estimates for bDMARDs in a moderate-RA population, which implies introducing bDMARDs at an earlier stage. Future NMAs, whilst mindful of the risk of multiplicity, should consider covariate adjustment for the number of prior DMARD treatments, C-reactive protein, or baseline HAQ, if sufficient data are available.

Disclosure

References

1. Lee DM, Weinblatt ME. Rheumatoid arthritis. *Lancet*. 2001;358(9285):903–911.
2. Symmons D, Turner G, Webb R, et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology (Oxford)*. 2002;41(7):793–800.
3. Alishiri GH, Bayat N, Salimzadeh A, et al. Health-related quality of life and disease activity in rheumatoid arthritis. *J Res Med Sci*. 2011;16(7):897–903.
4. Smith HS, Smith AR, Seidner P. Painful rheumatoid arthritis. *Pain Physician*. 2011;14(5):E427–E458.
5. Strand V, Sharp V, Koenig AS, et al. Comparison of health-related quality of life in rheumatoid arthritis, psoriatic arthritis and psoriasis and effects of etanercept treatment. *Ann Rheum Dis*. 2012;71(7):1143–1150.
6. Boonen A, Mau W. The economic burden of disease: comparison between rheumatoid arthritis and ankylosing spondylitis. *Clin Exp Rheumatol*. 2009;27(4 Suppl 55):S112–S117.
7. Foster W, Carruthers D, Lip GY, Blann AD. Inflammatory cytokines, endothelial markers and adhesion molecules in rheumatoid arthritis: effect of intensive anti-inflammatory treatment. *J Thromb Thrombolysis*. 2010;29(4):437–442.
8. Thwaites C, Finney A. Rheumatoid arthritis. 2: Exploring treatment options to achieve early control and remission. *Nurs Times*. 2010;106(10):18–20.
9. National Institute for Health and Clinical Excellence. Rheumatoid arthritis (CG79): full guideline. London, UK: NICE; 2009. Available from: http://guidance.nice.org.uk/CG79/Guidance. Accessed October 30, 2012.
10. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum*. 1980;23(2):137–145.
11. Smolen JS, Landewe R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Ann Rheum Dis*. 2010;69(6):964–975.
12. Chen YF, Jobanputra P, Barton P, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. *Health Technol Assess*. 2006;10(42):iii–iv, xi–xiii, 1–229.
13. Maxwell LJ, Singh JA. Abatacept for rheumatoid arthritis: a Cochrane systematic review. *J Rheumatol*. 2010;37(2):234–245.
14. Navarro-Sarabia F, Ariza-Ariza R, Hernandez-Cruz B, Villanueva I. Adalimumab for treating rheumatoid arthritis. *J Rheumatol*. 2006;33(6):1075–1081.
15. Blumenauer B, Judd M, Cranney A, et al. Etanercept for the treatment of rheumatoid arthritis. *Cochrane Database Syst Rev*. 2003(4):CD004525.
16. Singh JA, Noorbaloochi S, Singh G. Golimumab for rheumatoid arthritis: a systematic review. *J Rheumatol*. 2010;37(6):1096–1104.
17. Blumenauer B, Judd M, Wells G, et al. Infliximab for the treatment of rheumatoid arthritis. *Cochrane Database Syst Rev*. 2002(3):CD003785.
18. Singh JA, Christensen R, Wells GA, et al. Biologics for rheumatoid arthritis: an overview of Cochrane reviews. *Sao Paulo Med J*. 2010;128(5):309–310.
19. Singh JA, Beg S, Lopez-Olivo MA. Tocilizumab for rheumatoid arthritis: a Cochrane systematic review. *J Rheumatol*. 2011;38(1):10–20.
20. Canadian Agency for Drugs and Technologies in Health. *Clinical and Economic Overview: Biological Response Modifier Agents for Adults with Rheumatoid Arthritis*. Ottawa, ON: CADTH; 2010. Available from: http://www.cadth.ca/media/pdf/TR_RA_Clinical_and_Economic_Overview_e.pdf. Accessed May 2, 2012.
21. ClinicalTrials.gov [homepage on the Internet]. US National Institutes of Health. Available from: http://clinicaltrials.gov. Accessed April 11, 2012.
22. National Institute for Health and Clinical Excellence. Appendix D: Methodology checklist – randomised controlled trials. London, UK: NICE; 2009. Available from: http://www.nice.org.uk/media/633/21/The_guidelines_manual_2009_-_Appendix_D_Methodology_checklist_-_randomised_controlled_trials.pdf. Accessed October 30, 2012.
23. Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR. Empirical assessment of effect of publication bias on meta-analyses. *BMJ*. 2000;320(7249):1574–1577.
24. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–634.

25. Higgins JPT, Deeks JJ, Altman DG, editors. Intention-to-treat issues for dichotomous data. In: *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: John Wiley & Sons; 2009.

26. Harris R, Bradburn M, Deeks JJ, Harbord R, Altman DG, Sterne J. Metan: fixed- and random-effects meta-analysis. *Stata J*. 2008;8(1):3–28.

27. [No authors listed]. Software updates (Metan). *Stata J*. 2008;8(4):594.

28. 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.

29. Glenny AM, Altman DG, Song F, et al. Indirect comparisons of competing interventions. *Health Technol Assess*. 2005;9(26):1–134, iii–iv.

30. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ*. 2005;331(7521):897–900.

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

32. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. 2004;23(20):3105–3124.

33. The BUGS project, MRC Biostatistics Unit. Bayesian Inference Using Gibbs Sampling (BUGS) project. WinBUGS 1.4.3. Available from: http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml. Accessed October 30, 2012.

34. Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health Care Evaluation*. Chichester, UK: John Wiley & Sons; 2004.

35. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res*. 2001;10(4):277–303.

36. The BUGS Project. Winbugs 1.4 beta version. New examples for WinBUGS. Available from: http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/examples.shtml. Accessed October 30, 2012.

37. Dias S, Welton N, Sutton A, Ades AE. NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analyisis of randomised controlled trials. London, UK: NICE; 2011. Available from: http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm

38. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU technical support document 5: evidence synthesis in the baseline natural history model. London, UK: NICE; 2011. Available from: http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm

39. Hasselblad V. Meta-analysis of multitreatment studies. *Med Decis Making*. 1998;18(1):37–43.

40. Gelman A. Prior distributions for variance parameters in heirarchical models. *Bayesian Anal*. 2006;1(3):515–533.

41. Higgins JPT, Deeks JJ, Altman DG, editors. Special topics in statistics: rare events. Studies with zero-cell counts. In: *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: John Wiley & Sons; 2009.

42. Combe B, Codreanu C, Fiocco U, et al. Etanercept and sulfasalazine, alone and combined, in patients with active rheumatoid arthritis despite receiving sulfasalazine: a double-blind comparison. *Ann Rheum Dis*. 2006;65(10):1357–1362.

43. Genovese MC, McKay JD, Nasonov EL, et al. Interleukin-6 receptor inhibition with tocilizumab reduces disease activity in rheumatoid arthritis with inadequate response to disease-modifying antirheumatic drugs: the tocilizumab in combination with traditional disease-modifying antirheumatic drug therapy study. *Arthritis Rheum*. 2008;58(10):2968–2980.

44. Keystone E, Heijde D, Mason D Jr, et al. Certolizumab pegol plus methotrexate is significantly more effective than placebo plus methotrexate in active rheumatoid arthritis: findings of a fifty-two-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. *Arthritis Rheum*. 2008;58(11):3319–3329.

45. Smolen J, Landewe RB, Mease P, et al. Efficacy and safety of certolizumab pegol plus methotrexate in active rheumatoid arthritis: the RAPID 2 study. A randomised controlled trial. *Ann Rheum Dis*. 2009;68(6):797–804.

46. Kameda H, Ueki Y, Saito K, et al. Etanercept (ETN) with methotrexate (MTX) is better than ETN monotherapy in patients with active rheumatoid arthritis despite MTX therapy: a randomized trial. *Mod Rheumatol*. 2010;20(6):531–538.

47. van Riel PL, Taggart AJ, Sany J, et al. Efficacy and safety of combination etanercept and methotrexate versus etanercept alone in patients with rheumatoid arthritis with an inadequate response to methotrexate: the ADORE study. *Ann Rheum Dis*. 2006;65(11):1478–1483.

48. Genovese MC, Cohen S, Moreland L, et al. Combination therapy with etanercept and anakinra in the treatment of patients with rheumatoid arthritis who have been treated unsuccessfully with methotrexate. *Arthritis Rheum*. 2004;50(5):1412–1419.

49. Kremer JM, Westhovens R, Leon M, et al. Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. *N Engl J Med*. 2003;349(20):1907–1915.

50. Kremer J, Ritchlin C, Mendelsohn A, et al. Golimumab, a new human anti-tumor necrosis factor alpha antibody, administered intravenously in patients with active rheumatoid arthritis: forty-eight-week efficacy and safety results of a phase III randomized, double-blind, placebo-controlled study. *Arthritis Rheum*. 2010;62(4):917–928.

51. Maini RN, Taylor PC, Szechinski J, et al. Double-blind randomized controlled clinical trial of the interleukin-6 receptor antagonist, tocilizumab, in European patients with rheumatoid arthritis who had an incomplete response to methotrexate. *Arthritis Rheum*. 2006;54(9):2817–2829.

52. Smolen JS, Beaulieu A, Rubbert-Roth A, et al. Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial. *Lancet*. 2008;371(9617):987–997.

53. Zhang FC, Hou Y, Huang F, et al. Infliximab versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a preliminary study from China. *Int J Rheum Dis*. 2006;9(2):127–130.

54. Klareskog L, van der Heijde D, de Jager JP, et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. *Lancet*. 2004;363(9410):675–681.

55. Pfizer. TEMPO CSR 36848. 24-week data, Table 9.4.2.2A. Data on file.

56. Cooper NJ, Sutton AJ, Lu G, Khunti K. Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Arch Intern Med*. 2006;166(12):1269–1275.

57. The BUGS Project. DIC: deviance information criterion. Available from: http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/dicpage.shtml. Accessed October 30, 2012.

58. Abe T, Takeuchi T, Miyasaka N, et al. A multicenter, double-blind, randomized, placebo controlled trial of infliximab combined with low dose methotrexate in Japanese patients with rheumatoid arthritis. *J Rheumatol*. 2006;33(1):37–44.

59. Chen DY, Chou SJ, Hsieh TY, et al. Randomized, double-blind, placebo-controlled, comparative study of human anti-TNF antibody adalimumab in combination with methotrexate and methotrexate alone in Taiwanese patients with active rheumatoid arthritis. *J Formos Med Assoc*. 2009;108(4):310–319.

60. Durez P, Nzeusseu Toukap A, Lauwerys BR, et al. A randomised comparative study of the short term clinical and biological effects of intravenous pulse methylprednisolone and infliximab in patients with active rheumatoid arthritis despite methotrexate treatment. *Ann Rheum Dis*. 2004;63(9):1069–1074.

61. Emery P, Deodhar A, Rigby WF, et al. Efficacy and safety of different doses and retreatment of rituximab: a randomised, placebo-controlled trial in patients who are biological naive with active rheumatoid arthritis and an inadequate response to methotrexate (Study Evaluating Rituximab's Efficacy in MTX iNadequate rEsponders (SERENE)). *Ann Rheum Dis*. 2010;69(9):1629–1635.

62. Huang F, Zhang FC, Bao CD, et al. Adalimumab plus methotrexate for the treatment of rheumatoid arthritis: a multi-center randomized, double-blind, placebo-controlled clinical study. *Zhonghua Nei Ke Za Zhi*. 2009;48(11):916–921. Chinese.

63. Kay J, Matteson EL, Dasgupta B, et al. Golimumab in patients with active rheumatoid arthritis despite treatment with methotrexate: a randomized, double-blind, placebo-controlled, dose-ranging study. *Arthritis Rheum*. 2008;58(4):964–975.

64. Keystone EC, Schiff MH, Kremer JM, et al. Once-weekly administration of 50 mg etanercept in patients with active rheumatoid arthritis: results of a multicenter, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum*. 2004;50(2):353–363.

65. Kim HY, Lee SK, Song YW, et al. A randomized, double-blind, placebo-controlled, phase III study of the human anti-tumor necrosis factor antibody adalimumab administered as subcutaneous injections in Korean rheumatoid arthritis patients treated with methotrexate. *Int J Rheum Dis*. 2007;10(1):9–16.

66. Kremer JM, Genant HK, Moreland LW, et al. Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: a randomized trial. *Ann Intern Med*. 2006;144(12):865–876.

67. Lan JL, Chou SJ, Chen DY, Chen YH, Hsieh TY, Young M Jr. A comparative study of etanercept plus methotrexate and methotrexate alone in Taiwanese patients with active rheumatoid arthritis: a 12-week, double-blind, randomized, placebo-controlled study. *J Formos Med Assoc*. 2004;103(8):618–623.

68. Maini R, St Clair EW, Breedveld F, et al. Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. ATTRACT Study Group. *Lancet*. 1999;354(9194):1932–1939.

69. Schiff M, Keiserman M, Codding C, et al. Efficacy and safety of abatacept or infliximab vs placebo in ATTEST: a phase III, multi-centre, randomised, double-blind, placebo-controlled study in patients with rheumatoid arthritis and an inadequate response to methotrexate. *Ann Rheum Dis*. 2008;67(8):1096–1103.

70. Weinblatt ME, Kremer JM, Bankhurst AD, et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med*. 1999;340(4):253–259.

71. Weinblatt ME, Keystone EC, Furst DE, et al. Adalimumab, a fully human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. *Arthritis Rheum*. 2003;48(1):35–45.

72. Westhovens R, Yocum D, Han J, et al. The safety of infliximab, combined with background treatments, among patients with rheumatoid arthritis and various comorbidities: a large, randomized, placebo-controlled trial. *Arthritis Rheum*. 2006;54(4):1075–1086.

73. Johnsen AK, Schiff MH, Mease PJ, et al. Comparison of 2 doses of etanercept (50 vs 100 mg) in active rheumatoid arthritis: a randomized double blind study. *J Rheumatol*. 2006;33(4):659–664.

74. Miyasaka N. Clinical investigation in highly disease-affected rheumatoid arthritis patients in Japan with adalimumab applying standard and general evaluation: the CHANGE study. *Mod Rheumatol*. 2008;18(3):252–262.

75. Moreland LW, Schiff MH, Baumgartner SW, et al. Etanercept therapy in rheumatoid arthritis. A randomized, controlled trial. *Ann Intern Med*. 1999;130(6):478–486.

76. Nishimoto N, Miyasaka N, Yamamoto K, et al. Study of active controlled tocilizumab monotherapy for rheumatoid arthritis patients with an inadequate response to methotrexate (SATORI): significant reduction in disease activity and serum vascular endothelial growth factor by IL-6 receptor inhibition therapy. *Mod Rheumatol*. 2009;19(1):12–19.

77. Nishimoto N, Yoshizaki K, Miyasaka N, et al. Treatment of rheumatoid arthritis with humanized anti-interleukin-6 receptor antibody: a multi-center, double-blind, placebo-controlled trial. *Arthritis Rheum*. 2004;50(6):1761–1769.

78. van de Putte LB, Atkins C, Malaise M, et al. Efficacy and safety of adalimumab as monotherapy in patients with rheumatoid arthritis for whom previous disease modifying antirheumatic drug treatment has failed. *Ann Rheum Dis*. 2004;63(5):508–516.

79. van de Putte LBA, Rau R, Breedveld FC, et al. Efficacy and safety of the fully human anti-tumour necrosis factor a monoclonal antibody adalimumab (D2E7) in DMARD refractory patients with rheumatoid arthritis: a 12 week, phase II study. *Ann Rheum Dis*. 2003;62(12):1168–1177.

80. Moreland LW, Baumgartner SW, Schiff MH, et al. Treatment of rheumatoid arthritis with a recombinant human tumor necrosis factor receptor (p75)-Fc fusion protein. *N Engl J Med*. 1997;337(3):141–147.

81. Edwards JC, Szczepanski L, Szechinski J, et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. *N Engl J Med*. 2004;350(25):2572–2581.

82. Keystone EC, Genovese MC, Klareskog L, et al. Golimumab, a human antibody to tumour necrosis factor {alpha} given by monthly subcutaneous injections, in active rheumatoid arthritis despite methotrexate therapy: the GO-FORWARD Study. *Ann Rheum Dis*. 2009;68(6):789–796.

83. Fleischmann R, Vencovsky J, van Vollenhoven RF, 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*. 2009;68(6):805–811.

84. Nishimoto N, Miyasaka N, Yamamoto K, Kawai S, Takeuchi T, Azuma J. Long-term safety and efficacy of tocilizumab, an anti-IL-6 receptor monoclonal antibody, in monotherapy, in patients with rheumatoid arthritis (the STREAM study): evidence of safety and efficacy in a 5-year extension study. *Ann Rheum Dis*. 2009;68(10):1580–1584.

85. Bergman GJD, Hochberg MC, Boers M, Wintfeld N, Kielhorn A, Jansen JP. Indirect comparison of tocilizumab and other biologic agents in patients with rheumatoid arthritis and inadequate response to disease-modifying antirheumatic drugs. *Semin Arthritis Rheum*. 2010;39(6):425–441.

86. National Institute for Health and Clinical Excellence. Rheumatoid arthritis – adalimumab, etanercept and infliximab (TA130). London, UK: NICE; 2007. Available from: http://guidance.nice.org.uk/TA130. Accessed October 30, 2012.

87. National Institute for Health and Clinical Excellence. Rheumatoid arthritis (after the failure of previous anti-rheumatic drugs) – golimumab (TA225). London, UK: NICE; 2011. Available from: http://guidance.nice.org.uk/TA225. Accessed October 30, 2012.

88. Tracey D, Klareskog L, Sasso EH, Salfeld JG, Tak PP. Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. *Pharmacol Ther*. 2008;117(2):244–279.

89. Thorlund K, Druyts E, Aviña-Zubieta JA, Wu P, Mills E. Why the findings of published multiple treatment comparison meta-analyses of biologic treatments for rheumatoid arthritis are different: an overview of recurrent methodological shortcomings. *Ann Rheum Dis*. Epub October 20, 2012.

Supplementary materials

Table S1 Direct meta-analysis of American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 outcomes: combination therapy

Treatment	Fixed effect		Random effects	
	OR v DMARD (95% CI)	*P*-value	OR v DMARD (95% CI)	*P*-value
ACR 20				
ABA 10 mg/kg/4 weeks + DMARD	2.996 (2.338, 3.839)	<0.001	2.997 (2.339, 3.841)	<0.001
ADA 40 mg/2 weeks + DMARD	3.433 (2.589, 4.552)	<0.001	3.332 (1.667, 6.661)	0.001
CZP 200 mg/2 weeks + DMARD	10.583 (7.284, 15.377)	<0.001	10.57 (6.991, 15.983)	<0.001
ETN 2 × 25 mg/week + DMARD	8.198 (4.724, 14.224)	<0.001	8.092 (4.645, 14.095)	<0.001
GOL 50 mg/4 weeks + DMARD	3.432 (2.105, 5.596)	<0.001	3.437 (2.107, 5.608)	<0.001
INF 3 mg/kg/8 weeks + DMARD	3.406 (2.726, 4.257)	<0.001	3.359 (2.488, 4.536)	<0.001
RTX 2 × 1000 mg + DMARD	3.553 (2.345, 5.384)	<0.001	3.554 (2.345, 5.387)	<0.001
TOC 8 mg/kg/4 weeks + DMARD	4.48 (3.595, 5.582)	<0.001	4.472 (3.589, 5.572)	<0.001
ACR 50				
ABA 10 mg/kg/4 weeks + DMARD	3.281 (2.444, 4.406)	<0.001	3.278 (2.441, 4.403)	<0.001
ADA 40 mg/2 weeks + DMARD	3.622 (2.62, 5.007)	<0.001	3.881 (1.407, 10.702)	0.009
CZP 200 mg/2 weeks + DMARD	8.927 (5.453, 14.614)	<0.001	9.077 (4.706, 17.51)	<0.001
ETN 2 × 25 mg/week + DMARD	9.779 (4.822, 19.832)	<0.001	9.409 (4.635, 19.102)	<0.001
GOL 50 mg/4 weeks + DMARD	4.493 (2.468, 8.18)	<0.001	4.577 (2.145, 9.764)	<0.001
INF 3 mg/kg/8 weeks + DMARD	3.525 (2.671, 4.652)	<0.001	3.455 (2.295, 5.202)	<0.001
RTX 2 × 1000 mg + DMARD	3.791 (2.208, 6.509)	<0.001	3.792 (2.207, 6.513)	<0.001
TOC 8 mg/kg/4 weeks + DMARD	5.841 (4.373, 7.802)	<0.001	5.57 (3.844, 8.069)	<0.001
ACR 70				
ABA 10 mg/kg/4 weeks + DMARD	3.76 (2.42, 5.83)	<0.001	3.77 (2.05, 6.93)	<0.001
ADA 40 mg/2 weeks + DMARD	3.66 (2.32, 5.79)	<0.001	3.78 (1.28, 11.15)	0.016
CZP 200 mg/2 weeks + DMARD	11 (5.06, 23.94)	<0.001	10.18 (4.67, 22.22)	<0.001
ETN 2 × 25 mg/week + DMARD	15.58 (3.67, 66.2)	<0.001	15.64 (3.7, 66.18)	<0.001
GOL 50 mg/4 weeks + DMARD	4.85 (2.01, 11.67)	<0.001	4.77 (1.98, 11.5)	<0.001
INF 3 mg/kg/8 weeks + DMARD	3.18 (2.17, 4.66)	<0.001	3 (2.03, 4.44)	<0.001
RTX 2 × 1000 mg + DMARD	2.33 (1.11, 4.88)	0.025	2.52 (0.87, 7.31)	0.089
TOC 8 mg/kg/4 weeks + DMARD	8.13 (5.13, 12.88)	<0.001	7.27 (3.33, 15.89)	<0.001

Note: *P* < 0.05 is statistically significant.
Abbreviations: ABA, abatacept; ADA, adalimumab; ANA, anakinra; CI, confidence interval; CZP, certolizumab pegol; DMARD, disease-modifying antirheumatic drug (MTX or SUL); ETN, etanercept; exp, experienced; GOL, golimumab; INF, infliximab; MTX, methotrexate; OR, odds ratio; RTX, rituximab; SUL, sulfasalazine; TOC, tocilizumab.

Table S2 Bucher indirect meta-analysis of American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 outcomes: combination therapy

Treatment	Control	Fixed effect		Random effects	
		OR v control (95% CI)	*P*-value	OR v control (95% CI)	*P*-value
ACR 20					
ETN 2 × 25 mg/week + DMARD	ABA 10 mg/kg/4 weeks + DMARD	2.736 (1.495, 5.008)	0.001	2.7 (1.47, 4.959)	0.001
ETN 2 × 25 mg/week + DMARD	ADA 40 mg/2 weeks + DMARD	2.388 (1.286, 4.436)	0.006	2.429 (1, 5.899)	0.050
ETN 2 × 25 mg/week + DMARD	CZP 200 mg/2 weeks + DMARD	0.775 (0.398, 1.508)	0.452	0.766 (0.383, 1.53)	0.449
ETN 2 × 25 mg/week + DMARD	GOL 50 mg/4 weeks + DMARD	2.389 (1.143, 4.99)	0.021	2.354 (1.123, 4.935)	0.023
ETN 2 × 25 mg/week + DMARD	INF 3 mg/kg/8 weeks + DMARD	2.407 (1.328, 4.362)	0.004	2.409 (1.282, 4.528)	0.006
ETN 2 × 25 mg/week + DMARD	RTX 2 × 1000 mg + DMARD	2.307 (1.157, 4.602)	0.018	2.277 (1.138, 4.556)	0.020
ETN 2 × 25 mg/week + DMARD	TOC 8 mg/kg/4 weeks + DMARD	1.83 (1.011, 3.313)	0.046	1.809 (0.996, 3.287)	0.052
ACR 50					
ETN 2 × 25 mg/week + DMARD	ABA 10 mg/kg/4 weeks + DMARD	2.98 (1.386, 6.411)	0.005	2.87 (1.333, 6.18)	0.007
ETN 2 × 25 mg/week + DMARD	ADA 40 mg/2 weeks + DMARD	2.7 (1.241, 5.876)	0.012	2.424 (0.704, 8.355)	0.161
ETN 2 × 25 mg/week + DMARD	CZP 200 mg/2 weeks + DMARD	1.095 (0.463, 2.594)	0.836	1.037 (0.395, 2.723)	0.942
ETN 2 × 25 mg/week + DMARD	GOL 50 mg/4 weeks + DMARD	2.176 (0.862, 5.498)	0.100	2.056 (0.729, 5.8)	0.173
ETN 2 × 25 mg/week + DMARD	INF 3 mg/kg/8 weeks + DMARD	2.774 (1.298, 5.929)	0.008	2.723 (1.202, 6.169)	0.016
ETN 2 × 25 mg/week + DMARD	RTX 2 × 1000 mg + DMARD	2.58 (1.059, 6.282)	0.037	2.481 (1.018, 6.05)	0.046
ETN 2 × 25 mg/week + DMARD	TOC 8 mg/kg/4 weeks + DMARD	1.674 (0.78, 3.594)	0.186	1.689 (0.76, 3.757)	0.199
ACR 70					
ETN 2 × 25 mg/week + DMARD	ABA 10 mg/kg/4 weeks + DMARD	4.144 (0.914, 18.784)	0.065	4.149 (0.867, 19.84)	0.075
ETN 2 × 25 mg/week + DMARD	ADA 40 mg/2 weeks + DMARD	4.257 (0.935, 19.385)	0.061	4.138 (0.682, 25.104)	0.123
ETN 2 × 25 mg/week + DMARD	CZP 200 mg/2 weeks + DMARD	1.416 (0.274, 7.31)	0.678	1.536 (0.298, 7.909)	0.608
ETN 2 × 25 mg/week + DMARD	GOL 50 mg/4 weeks + DMARD	3.212 (0.591, 17.462)	0.177	3.279 (0.606, 17.743)	0.168
ETN 2 × 25 mg/week + DMARD	INF 3 mg/kg/8 weeks + DMARD	4.899 (1.098, 21.858)	0.037	5.213 (1.171, 23.213)	0.030
ETN 2 × 25 mg/week + DMARD	RTX 2 × 1000 mg + DMARD	6.687 (1.317, 33.953)	0.022	6.206 (1.035, 37.223)	0.046
ETN 2 × 25 mg/week + DMARD	TOC 8 mg/kg/4 weeks + DMARD	1.916 (0.42, 8.739)	0.401	2.151 (0.418, 11.083)	0.360

Note: *P* < 0.05 is statistically significant.
Abbreviations: ABA, abatacept; ADA, adalimumab; ANA, anakinra; CI, confidence interval; CZP, certolizumab pegol; DMARD, disease-modifying antirheumatic drug (MTX or SUL); ETN, etanercept; exp, experienced; GOL, golimumab; INF, infliximab; MTX, methotrexate; OR, odds ratio; RTX, rituximab; SUL, sulfasalazine; TOC, tocilizumab.

Table S3 Results from combination therapy NMA study arm level (patient characteristics) covariate analysis for American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 end point

	OR v DMARD (95% CrI)[‡]		
	ACR 20	**ACR 50**	**ACR 70**[†]
Base case: random effects			
ABA 10 mg/kg/4 weeks + DMARD	3.255 (2.056, 5.159)[‡]	3.633 (2.093, 6.341)[‡]	3.954 (1.974, 8.8)[‡]
ADA 40 mg/2 weeks + DMARD	3.439 (2.187, 5.303)[‡]	3.87 (2.303, 6.598)[‡]	3.868 (1.91, 7.83)[‡]
CZP 200 mg/2 weeks + DMARD	11.06 (6.055, 21.06)[‡]	9.773 (4.604, 22.65)[‡]	13.18 (4.489, 43.5)[‡]
ETN 2 × 25 mg/week + DMARD	9.341 (4.845, 19.29)[‡]	11.15 (4.947, 27.95)[‡]	20.69 (4.921, 158.6)[‡]
GOL 50 mg/4 weeks + DMARD	3.387 (1.604, 6.863)[‡]	4.917 (2.051, 12.34)[‡]	4.988 (1.401, 18.28)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.347 (2.271, 4.983)[‡]	3.602 (2.246, 5.924)[‡]	3.694 (2.021, 7.307)[‡]
RTX 2 × 1000 mg + DMARD	3.716 (1.915, 7.418)[‡]	4.103 (1.821, 9.73)[‡]	2.644 (0.909, 8.387)
TOC 8 mg/kg/4 weeks + DMARD	4.399 (2.704, 7.125)[‡]	5.401 (2.911, 9.561)[‡]	7.656 (3.442, 16.5)[‡]
Covariate analysis[^]			
ABA 10 mg/kg/4 weeks + DMARD	3.242 (2.181, 4.765)[‡]	3.626 (2.643, 5.051)[‡]	4.187 (2.627, 7.196)[‡]
ADA 40 mg/2 weeks + DMARD	4.151 (2.662, 6.658)[‡]	5.221 (3.281, 8.649)[‡]	4.774 (2.312, 10.8)[‡]
CZP 200 mg/2 weeks + DMARD	11.34 (6.8, 20.36)[‡]	10.8 (6.096, 22.74)[‡]	15.17 (6.535, 41.99)[‡]
ETN 2 × 25 mg/week + DMARD	8.183 (4.108, 17.72)[‡]	9.377 (4.661, 23.88)[‡]	23.06 (5.343, 150.1)[‡]
GOL 50 mg/4 weeks + DMARD	3.674 (1.802, 6.893)[‡]	5.491 (2.821, 10.48)[‡]	6.67 (2.42, 18.11)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.334 (2.418, 4.616)[‡]	3.634 (2.692, 4.763)[‡]	4.019 (2.708, 6.159)[‡]
RTX 2 × 1000 mg + DMARD	3.634 (2.072, 6.462)[‡]	3.913 (2.341, 7.101)[‡]	2.42 (1.149, 5.519)[‡]
TOC 8 mg/kg/4 weeks + DMARD	4.466 (2.991, 6.703)[‡]	5.927 (4.274, 8.365)[‡]	9.456 (5.74, 16.06)[‡]
Coefficients (on log scale)[‡]	**Median (95% CrI)**		
βage: Age at baseline (average for study arm)	0.016 (−0.05, 0.079)	0.051 (−0.01, 0.115)	0.134 (0.037, 0.229)[†]
βduration: Duration of disease (average for study arm)	0.025 (−0.053, 0.104)	0.074 (0.005, 0.149)[†]	0.099 (−0.002, 0.218)

Notes: [^]Results centered at study-arm mean (mean age across study arms = 52.43 years, mean disease duration = 7.86 years); [‡]licensed combination had significantly higher odds (based on the 95% CrI) compared to DMARD alone; [†]for the ACR 70 network meta-analysis, a continuity correction (0.5) was applied to the data; the coefficients were not statistically significant except for + (longer disease duration was associated with higher odds of ACR 50; higher age was associated with higher odds of ACR 70).
Abbreviations: ABA, abatacept; ADA, adalimumab; ANA, anakinra; CrI, credible interval (Bayesian probability interval); CZP, certolizumab pegol; DMARD, disease-modifying antirheumatic drug (MTX or SUL); ETN, etanercept; exp, experienced; GOL, golimumab; INF, infliximab; MTX, methotrexate; OR, odds ratio; RTX, rituximab; SUL, sulfasalazine; TOC, tocilizumab.

Table S4 Combination therapy network meta-analysis sensitivity analysis results for American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 end point

	OR v DMARD (95% CrI)[‡]		
	ACR 20	ACR 50	ACR 70[†]
Base case: random effects			
ABA 10 mg/kg/4 weeks + DMARD	3.255 (2.056, 5.159)[‡]	3.633 (2.093, 6.341)[‡]	3.954 (1.974, 8.8)[‡]
ADA 40 mg/2 weeks + DMARD	3.439 (2.187, 5.303)[‡]	3.87 (2.303, 6.598)[‡]	3.868 (1.91, 7.83)[‡]
CZP 200 mg/2 weeks + DMARD	11.06 (6.055, 21.06)[‡]	9.773 (4.604, 22.65)[‡]	13.18 (4.489, 43.5)[‡]
ETN 2 × 25 mg/week + DMARD	9.341 (4.845, 19.29)[‡]	11.15 (4.947, 27.95)[‡]	20.69 (4.921, 158.6)[‡]
GOL 50 mg/4 weeks + DMARD	3.387 (1.604, 6.863)[‡]	4.917 (2.051, 12.34)[‡]	4.988 (1.401, 18.28)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.347 (2.271, 4.983)[‡]	3.602 (2.246, 5.924)[‡]	3.694 (2.021, 7.307)[‡]
RTX 2 × 1000 mg + DMARD	3.716 (1.915, 7.418)[‡]	4.103 (1.821, 9.73)[‡]	2.644 (0.909, 8.387)
TOC 8 mg/kg/4 weeks + DMARD	4.399 (2.704, 7.125)[‡]	5.401 (2.911, 9.561)[‡]	7.656 (3.442, 16.5)[‡]
As base case, but excluding MTX-naïve studies (MTX-experienced population)			
ABA 10 mg/kg/4 weeks + DMARD	3.31 (1.83, 5.52)[‡]	3.86 (1.91, 7.09)[‡]	4.91 (1.87, 11.1)[‡]
ADA 40 mg/2 weeks + DMARD	3.65 (2.04, 6.12)[‡]	4.27 (2.17, 7.75)[‡]	4.17 (1.67, 8.92)[‡]
CZP 200 mg/2 weeks + DMARD	12 (5.46, 23.67)[‡]	10.99 (4.08, 24.9)[‡]	17.01 (3.97, 52.62)[‡]
ETN 2 × 25 mg/week + DMARD	12.08 (3.83, 30.26)[‡]	28.5 (5.23, 104.6)[‡]	111 (3.04, 702.1)[‡]
GOL 50 mg/4 weeks + DMARD	3.67 (1.46, 7.69)[‡]	5.91 (1.96, 13.91)[‡]	6.62 (1.29, 20.6)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.87 (2.24, 6.3)[‡]	4.02 (2.08, 7.16)[‡]	4.06 (1.63, 9.1)[‡]
RTX 2 × 1000 mg + DMARD	4.07 (1.74, 8.21)[‡]	4.67 (1.63, 10.76)[‡]	3.71 (0.88, 10.88)
TOC 8 mg/kg/4 weeks + DMARD	4.48 (1.97, 8.9)[‡]	5.25 (1.92, 11.62)[‡]	8.18 (2.14, 22.14)[‡]
As base case, but excluding RAPID 1 and RAPID 2			
ABA 10 mg/kg/4 weeks + DMARD	3.239 (1.951, 5.364)[‡]	3.608 (2.024, 6.503)[‡]	4.055 (1.914, 9.659)[‡]
ADA 40 mg/2 weeks + DMARD	3.418 (2.111, 5.53)[‡]	3.884 (2.252, 6.853)[‡]	3.885 (1.871, 8.409)[‡]
CZP 200 mg/2 weeks + DMARD	NA	NA	NA
ETN 2 × 25 mg/week + DMARD	9.477 (4.771, 19.78)[‡]	11.3 (4.84, 29.37)[‡]	21.2 (4.751, 157.8)[‡]
GOL 50 mg/4 weeks + DMARD	3.364 (1.563, 7.297)[‡]	4.979 (2.064, 12.92)[‡]	4.869 (1.38, 18.76)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.343 (2.187, 5.148)[‡]	3.628 (2.203, 6.106)[‡]	3.767 (1.967, 7.976)[‡]
RTX 2 × 1000 mg + DMARD	3.732 (1.824, 7.873)[‡]	4.103 (1.792, 10.01)[‡]	2.685 (0.854, 9.246)
TOC 8 mg/kg/4 weeks + DMARD	4.389 (2.539, 7.569)[‡]	5.371 (2.841, 9.792)[‡]	7.659 (3.288, 17.58)[‡]
As base case, but excluding low-quality ETN studies			
ABA 10 mg/kg/4 weeks + DMARD	3.271 (2.309, 4.614)[‡]	3.614 (2.057, 6.371)[‡]	4.063 (1.873, 9.751)[‡]
ADA 40 mg/2 weeks + DMARD	3.439 (2.375, 4.934)[‡]	3.876 (2.289, 6.699)[‡]	3.918 (1.837, 8.591)[‡]
CZP 200 mg/2 weeks + DMARD	10.93 (6.786, 18.51)[‡]	9.78 (4.502, 22.36)[‡]	13.37 (4.062, 54.14)[‡]
ETN 2 × 25 mg/week + DMARD	8.53 (4.776, 16.04)[‡]	11.55 (4.85, 29.68)[‡]	20.83 (4.695, 182.9)[‡]
GOL 50 mg/4 weeks + DMARD	3.505 (1.878, 6.219)[‡]	4.953 (2.037, 12.79)[‡]	4.954 (1.387, 19.9)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.345 (2.485, 4.552)[‡]	3.614 (2.257, 5.956)[‡]	3.737 (1.912, 7.952)[‡]
RTX 2 × 1000 mg + DMARD	3.701 (2.139, 6.261)[‡]	4.046 (1.78, 9.64)[‡]	2.748 (0.838, 9.316)
TOC 8 mg/kg/4 weeks + DMARD	4.411 (3.055, 6.369)[‡]	5.376 (2.918, 9.762)[‡]	7.541 (3.207, 18.02)[‡]
As base case, but excluding studies in TNF-α-exposed patients			
DMARD	–	–	–
ABA 10 mg/kg/4 weeks + DMARD	3.479 (1.598, 7.638)[‡]	3.543 (1.59, 7.82)[‡]	3.345 (1.472, 8.01)[‡]
ADA 40 mg/2 weeks + DMARD	3.418 (1.906, 6.213)[‡]	3.908 (2.148, 7.375)[‡]	3.814 (1.915, 7.752)[‡]
CZP 200 mg/2 weeks + DMARD	NA	NA	NA
ETN 2 × 25 mg/week + DMARD	9.755 (4.387, 22.63)[‡]	11.32 (4.564, 31.03)[‡]	19.55 (4.814, 147.4)[‡]
GOL 50 mg/4 weeks + DMARD	3.315 (1.292, 8.238)[‡]	5.2 (1.905, 14.96)[‡]	4.722 (1.449, 18.3)[‡]
INF 3 mg/kg/8 weeks + DMARD	3.457 (1.926, 6.357)[‡]	4.08 (2.169, 7.935)[‡]	4.224 (2.137, 9.672)[‡]
RTX 2 × 1000 mg + DMARD	3.813 (1.576, 9.494)[‡]	4.145 (1.593, 11.29)[‡]	2.671 (0.934, 8.725)
TOC 8 mg/kg/4 weeks + DMARD	NA	NA	NA
As base case, but including TEMPO			
ABA 10 mg/kg/4 weeks + DMARD	3.249 (1.767, 6.024)[‡]	3.629 (1.916, 6.932)[‡]	4.11 (1.92, 9.757)[‡]
ADA 40 mg/2 weeks + DMARD	3.402 (1.947, 5.956)[‡]	3.896 (2.159, 7.215)[‡]	3.853 (1.836, 8.394)[‡]
CZP 200 mg/2 weeks + DMARD	11.24 (5.128, 25.42)[‡]	9.821 (4.05, 25.4)[‡]	13.26 (4.306, 48.4)[‡]
ETN 2 × 25 mg/week + DMARD	5.137 (2.816, 10.08)[‡]	5.482 (2.891, 11.75)[‡]	6.448 (2.832, 19.11)[‡]
GOL 50 mg/4 weeks + DMARD	3.308 (1.373, 8.033)[‡]	5.153 (1.943, 14.73)[‡]	4.801 (1.369, 18.95)[‡]

(Continued)

Table S4 (*Continued*)

	OR v DMARD (95% CrI)‡		
	ACR 20	**ACR 50**	**ACR 70†**
INF 3 mg/kg/8 weeks + DMARD	3.403 (2.048, 5.69)‡	3.664 (2.149, 6.397)‡	3.769 (1.946, 7.951)‡
RTX 2 × 1000 mg + DMARD	3.813 (1.631, 8.947)‡	4.13 (1.648, 10.85)‡	2.688 (0.878, 9.141)
TOC 8 mg/kg/4 weeks + DMARD	4.392 (2.317, 8.34)‡	5.344 (2.648, 10.42)‡	7.655 (3.283, 17.96)‡

Notes: ‡Licensed combination had significantly higher odds (based on the 95% CrI) compared to DMARD alone; †for the ACR 70 network meta-analysis, a continuity correction (0.5) was applied to the data.
Abbreviations: ABA, abatacept; ADA, adalimumab; ANA, anakinra; CrI, credible interval (Bayesian probability interval); CZP, certolizumab pegol; DMARD, disease-modifying antirheumatic drug (MTX or SUL); ETN, etanercept; exp, experienced; GOL, golimumab; INF, infliximab; MTX, methotrexate; NA, not applicable; OR, odds ratio; RTX, rituximab; SUL, sulfasalazine; TNF, tumor necrosis factor; TOC, tocilizumab.

Table S5 Direct meta-analysis of American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 outcomes: licensed DMARD monotherapy versus placebo in DMARD-experienced patients

Treatment	Fixed effect		Random effects	
	OR v PLA (95% CI)	**P-value**	**OR v PLA (95% CI)**	**P-value**
ACR 20				
ADA 40 mg/2 weeks	5.280 (3.211, 8.680)	<0.001	5.298 (3.226, 8.701)	<0.001
ETN 2 × 25 mg/week	11.34 (4.958, 25.938)	<0.001	11.34 (4.958, 25.938)	<0.001
TOC 8 mg/kg/4 weeks	28.069 (9.689, 81.318)	<0.001	28.069 (9.689, 81.318)	<0.001
ACR 50				
ADA 40 mg/2 weeks	4.633 (2.398, 8.953)	<0.001	4.588 (2.374, 8.866)	<0.001
ETN 2 × 25 mg/week	12.532 (4.159, 37.76)	<0.001	12.532 (4.159, 37.76)	<0.001
TOC 8 mg/kg/4 weeks	34.667 (4.459, 269.54)	0.001	34.667 (4.459, 269.54)	0.001
ACR 70				
ADA 40 mg/2 weeks	10.861 (3.045, 38.736)	<0.001	10.126 (2.837, 36.145)	<0.001
ETN 2 × 25 mg/week	25.714 (3.215, 205.639)	0.002	25.714 (3.215, 205.639)	0.002
TOC 8 mg/kg/4 weeks	40.714 (2.276, 728.176)	0.012	40.714 (2.276, 728.176)	0.012

Notes: SUL not connected to placebo, so no direct results; $P < 0.05$ is statistically significant.
Abbreviations: ADA, adalimumab; CI, confidence interval; DMARD, disease-modifying antirheumatic drug; ETN, etanercept; OR, odds ratio; PLA, placebo; SUL, sulfasalazine; TOC, tocilizumab.

Table S6 Bucher indirect meta-analysis of American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 outcomes: licensed DMARD monotherapy in DMARD-experienced patients

Treatment	Control	Fixed effect		Random effects	
		OR v control (95% CI)	**P-value**	**OR v control (95% CI)**	**P-value**
ACR 20					
ETN 2 × 25 mg/week	ADA 40 mg/2 weeks	2.148 (0.818, 5.639)	0.121	2.14 (0.816, 5.616)	0.122
ETN 2 × 25 mg/week	TOC 8 mg/kg/4 weeks	0.404 (0.105, 1.555)	0.187	0.404 (0.105, 1.555)	0.187
ACR 50					
ETN 2 × 25 mg/week	ADA 40 mg/2 weeks	2.705 (0.749, 9.774)	0.129	2.731 (0.756, 9.872)	0.125
ETN 2 × 25 mg/week	TOC 8 mg/kg/4 weeks	0.361 (0.035, 3.71)	0.392	0.361 (0.035, 3.71)	0.392
ACR 70					
ETN 2 × 25 mg/week	ADA 40 mg/2 weeks	2.368 (0.207, 27.089)	0.488	2.539 (0.222, 29.066)	0.454
ETN 2 × 25 mg/week	TOC 8 mg/kg/4 weeks	0.632 (0.018, 22.108)	0.800	0.632 (0.018, 22.108)	0.800

Notes: SUL not connected to placebo, so no direct results; $P < 0.05$ is statistically significant.
Abbreviations: ADA, adalimumab; CI, confidence interval; DMARD, disease-modifying antirheumatic drug; ETN, etanercept; OR, odds ratio; SUL, sulfasalazine; TOC, tocilizumab.

Table S7 Direct and indirect meta-analysis of ACR50 on Tocilizumab 4 mg/kg/4 weeks—Tocilizumab 8 mg/kg/4 weeks loop

Meta-analysis of ACR50	Comparator	OR	Lower 95% CI	Upper 95% CI	P value
Direct					
Tocilizumab 4 mg/kg/4 weeks	MTX	0.962	0.407	2.271	0.929
Tocilizumab 4 mg/kg/4 weeks	Placebo	18.2	2.296	144.262	0.006
Tocilizumab 8 mg/kg/4 weeks	MTX	3.663	0.784	17.123	0.099
Tocilizumab 8 mg/kg/4 weeks	Placebo	34.667	4.459	269.54	0.001
Indirect via MTX					
Tocilizumab 8 mg/kg/4 weeks	Tocilizumab 4 mg/kg/4 weeks	3.808	0.652	22.251	0.138
Indirect via placebo					
Tocilizumab 8 mg/kg/4 weeks	Tocilizumab 4 mg/kg/4 weeks	1.905	0.103	35.109	0.665
Inconsistency on loop		1.999	0.066	60.328	0.690

Abbreviations: OR, odds ratio; CI, confidence interval; MTX, methotrexate.

Table S8 Direct and indirect meta-analysis of ACR20 on Tocilizumab 4 mg/kg/4 weeks—Tocilizumab 8 mg/kg/4 weeks loop

Meta-analysis of ACR20	Comparator	OR	Lower 95% CI	Upper 95% CI	P value
Direct					
Tocilizumab 4 mg/kg/4 weeks	MTX	2.279	1.034	5.02	0.041
Tocilizumab 4 mg/kg/4 weeks	Placebo	10.558	3.859	28.885	0.000
Tocilizumab 8 mg/kg/4 weeks	MTX	5.417	1.195	24.546	0.028
Tocilizumab 8 mg/kg/4 weeks	Placebo	28.069	9.689	81.318	0.000
Indirect via DMARD					
Tocilizumab 8 mg/kg/4 weeks	Tocilizumab 4 mg/kg/4 weeks	2.377	0.432	13.084	0.320
Indirect via placebo					
Tocilizumab 8 mg/kg/4 weeks	Tocilizumab 4 mg/kg/4 weeks	2.659	0.615	11.498	0.191
Inconsistency on loop		0.894	0.094	8.465	0.922

Abbreviations: OR, odds ratio; CI, confidence interval; MTX, methotrexate; DMARD, disease modifying anti-rheumatic drug.

Table S9. Direct and indirect meta-analysis of ACR70 on Tocilizumab 4 mg/kg/4 weeks—Tocilizumab 8 mg/kg/4 weeks loop

Meta-analysis of ACR70	Comparator	OR	Lower 95% CI	Upper 95% CI	P value
Direct					
Tocilizumab 4 mg/kg/4 weeks	MTX	0.281	0.068	1.161	0.079
Tocilizumab 4 mg/kg/4 weeks	Placebo	40.932	2.328	719.842	0.011
Tocilizumab 8 mg/kg/4 weeks	MTX	4.668	0.304	71.629	0.269
Tocilizumab 8 mg/kg/4 weeks	Placebo	40.714	2.276	728.176	0.012
Indirect via DMARD					
Tocilizumab 8 mg/kg/4 weeks	Tocilizumab 4 mg/kg/4 weeks	16.612	0.765	360.725	0.074
Indirect via placebo					
Tocilizumab 8 mg/kg/4 weeks	Tocilizumab 4 mg/kg/4 weeks	0.995	0.017	58.049	0.998
Inconsistency on loop		16.701	0.102	2739.729	0.279

Abbreviations: OR, odds ratio; CI, confidence interval; MTX, methotrexate; DMARD, disease modifying anti-rheumatic drug.

Table S10 Monotherapy network meta-analysis sensitivity analysis results for American College of Rheumatology (ACR) criteria scores of 20, 50, and 70 end point

| | OR v PLA (95% CrI) | | |
	ACR 20	ACR 50	ACR 70[†]
Base case: random effects			
ADA 40 mg/2 weeks	5.125 (1.417, 18.62)[‡]	5.117 (1.819, 16.11)[‡]	11.71 (2.441, 77.5)[‡]
ETN 2 × 25 mg/week	12 (1.733, 90.94)[‡]	13.46 (2.631, 80.29)[‡]	20.83 (1.56, 1740)[‡]
SUL	1.608 (0.105, 27.48)	2.379 (0.241, 25.64)	1.14 (0.012, 178.8)
TOC 8 mg/kg/4 weeks	26.25 (3.883, 190.8)[‡]	55.08 (6.204, 1740)[‡]	96.27 (3.992, 38820)[‡]
As base case, but including TEMPO			
ADA 40 mg/2 weeks	5.076 (1.509, 17.53)[‡]	5.127 (1.902, 15.24)[‡]	11.38 (2.67, 74.92)[‡]
ETN 2 × 25 mg/week	8.209 (1.763, 38.24)[‡]	12.85 (3.355, 62.92)[‡]	21.17 (2.73, 316.1)[‡]
SUL	1.089 (0.101, 11.94)	2.242 (0.292, 20.22)	1.002 (0.018, 41.62)
TOC 8 mg/kg/4 weeks	39.39 (7.95, 197.4)[‡]	48.87 (10.22, 310.7)[‡]	40.62 (4.862, 821.2)[‡]

Notes: [‡]Licensed treatments have significantly higher odds (based on the 95% CrI) compared to PLA; [†]for the ACR 70 network meta-analysis, a continuity correction (0.5) was applied to the data.
Abbreviations: ADA, adalimumab; CrI, credible interval (Bayesian probability interval); ETN, etanercept; exp, experienced; OR, odds ratio; PLA, placebo; SUL, sulfasalazine; TOC, tocilizumab.
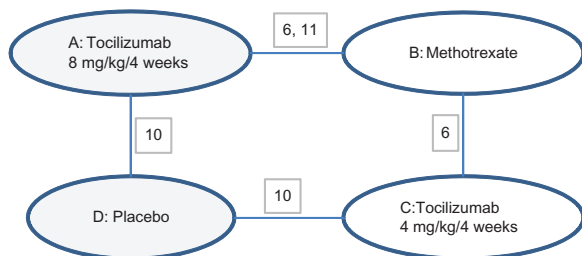

Figure S1 Part of the monotherapy evidence network containing the tocilizumab 4 mg/kg/4 weeks–tocilizumab 8 mg/kg/4 weeks loop.
Notes: 6, Maini 2006 (CHARISMA); 10, Nishimoto 2004 (STREAM); 11, Nishimoto 2009 (SATORI).