

# Comparative effectiveness of biologics for the management of rheumatoid arthritis: systematic review and network meta-analysis

Rafael Alfonso-Cristancho<sup>1</sup> · Nigel Armstrong<sup>2</sup> · Ramesh Arjunji<sup>1</sup> · Rob Riemsma<sup>2</sup> · Gill Worthy<sup>2</sup> · Rita Ganguly<sup>1</sup> · Jos Kleijnen<sup>2,3</sup>

Received: 6 April 2016 / Revised: 7 September 2016 / Accepted: 23 September 2016 / Published online: 10 October 2016  
© International League of Associations for Rheumatology (ILAR) 2016

**Abstract** Our aim was to establish the comparative effectiveness of rheumatoid arthritis (RA) biologics, using a systematic review and network meta-analysis. The systematic review used randomized controlled trials (RCTs) in adults with RA who failed treatment with conventional disease-modifying agents for rheumatoid disease (cDMARDs). We compared the effectiveness of abatacept, adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, and rituximab to tocilizumab, a recent biologic with a different mechanism of action (anti-IL-6 receptor). A network meta-analysis (NMA) included the indirect and direct evidence previously selected. In total, 207 articles were included describing 68 RCTs. The NMA showed that tocilizumab monotherapy was superior to standard care (ACR20, OR 13.27, 95 % CrI [3.958, 43.98]; ACR50, 17.45 [10.18, 31.24]; ACR70, 37.77 [7.226, 216.3]; EULAR, 10.42 [1.963, 54.8]); and methotrexate (MTX; ACR50, OR 5.44 [4.142, 7.238]; ACR70, 7.364 [1.4, 30.83]; EULAR, 4.226 [1.184, 15.58]) at 26 weeks. Similarly, the combination of tocilizumab + MTX was significantly better than standard care/placebo and MTX alone for ACR20, ACR50, ACR70, and EULAR at 26 weeks (OR 18.63 [5.32, 66.81]; 24.27 [14.5, 41.91]; 46.13 [10.08, 277]; 14.23 [2.493, 84.02]; 4.169 [2.267, 7.871]; 5.44 [4.142, 7.238]; 8.731 [4.203, 19.29]; 7.306 [4.393, 13.04], respectively). At 52 weeks, compared to

MTX alone, tocilizumab + MTX was significantly better for ACR20 and ACR50 response. Few significant differences were found between tocilizumab (alone or in combination) and any other biologics. Results must be considered in context with the limitations of the available evidence. This NMA suggests that tocilizumab was superior to cDMARDs and as effective as other biologics for RA.

**Keywords** Anti-IL-6 receptor · Biologics · Disease-modifying agents for rheumatoid disease · Rheumatoid arthritis · Tocilizumab

## Abbreviations

ABA	Abatacept
ACR	American College of Rheumatology
ADA	Adalimumab
CE	Cost-effectiveness
CEA	Cost-effectiveness analysis
CRP	C-reactive protein
CZP	Certolizumab pegol
DMARDs	Disease-modifying antirheumatic drugs
ESR	Erythrocyte sedimentation rate
ETA	Etanercept
EULAR	European League Against Rheumatism
GOL	Golimumab
GSK	GlaxoSmithKline
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
INF	Infliximab
MTX	Methotrexate
OR	Odds Ratio
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses

**Electronic supplementary material** The online version of this article (doi:10.1007/s10067-016-3435-2) contains supplementary material, which is available to authorized users.

✉ Rafael Alfonso-Cristancho  
rafael.x.alfonso@gsk.com

<sup>1</sup> GlaxoSmithKline, Collegeville, PA, USA

<sup>2</sup> Kleijnen Systematic Reviews Ltd, York, UK

<sup>3</sup> School for Public Health and Primary Care (CAPHRI), Maastricht University, Maastricht, The Netherlands

PRO	Patient relevant/reported outcomes
QALY	Quality-adjusted life year
QoL	Quality of life
RA	Rheumatoid arthritis
RCT	Randomized controlled trial
RF	Rheumatoid factor
RIT	Rituximab
RR	Relative risk
SE	Standard error
SMD	Standard mean difference
TNF- $\alpha$	Tumor necrosis factor- $\alpha$
TCZ	Tocilizumab

## Introduction

Rheumatoid arthritis (RA) is a chronic systemic inflammatory disease that causes cartilage damage, bone erosions, and eventually joint deformity. These impairments are associated with limitations in daily activities, work productivity, and quality of life. Other tissues and organs, including the heart and lungs, may also be affected causing additional health issues. Approximately 1 % of the population is diagnosed with RA worldwide but can vary by country [1–5]. The diagnosis of RA increases after the fourth decade of life and is three times more likely in women than men. Regional differences in the prevalence of RA have been described [3, 5]. In 2005, the prevalence rate of RA in the USA was estimated to be 0.6 % of the adult population [3], compared to Australia and New Zealand where the prevalence rates ranged from 2.1 to 3.5 %, in a similar period. In a recent systematic review, authors reported median annual incidence rates for the total population in south Europe of 16.5 (range 9 to 24) cases per 100,000, compared to 29 (range 24 to 36) for north European countries, and 38 (range 31 to 45) cases per 100,000, for North America [5].

Regardless of the geography, RA is associated with a significant burden to healthcare systems [6] and society [7]. The total costs of RA in the UK, including indirect costs and work-related disability, have been estimated at between £3.8 and £4.75 billion per year, before 2000 [8]. In the USA, by 2003, the estimate of the total cost of arthritis and other rheumatic conditions was approximately \$128 billion, equivalent to 1.2 % of the 2003 US gross domestic product [9]. More recent publications [10–14] reported the annual mean total cost per patient to range from \$4700 to \$24,920 US dollars. Annual mean indirect costs in these publications had wider ranges but are difficult to compare due to differences in methodology. These estimates are constantly increasing due to aging of the population and availability of new health technologies for the management of RA, which often come at higher costs to the healthcare system.

Recent advances in the management of RA focus on biologic disease-modifying antirheumatic drugs (DMARDs). The first set of biologic agents to be approved for RA was

the TNF antagonists: etanercept, infliximab, and adalimumab. Subsequently, two other TNF antagonists were also approved certolizumab pegol and golimumab. Other biologic DMARDs with different mechanisms of action to the TNF antagonists have been also approved for the management of RA: rituximab, a genetically engineered chimeric anti-CD20 monoclonal antibody that depletes B cells; abatacept, a soluble human fusion protein that selectively modulates T-cell costimulation; anakinra, an interleukin (IL)-1 antagonist; and tocilizumab (TCZ), a humanized monoclonal antibody against the IL-6 receptor inhibitor (IL-6R). Despite the number of available therapies, there continues to be an unmet need for newer agents for bio-naive and TNF refractory patients.

Current evidence-based guidelines recommend that patients must fail conventional DMARDs (cDMARDs) before treatment with biologics, as monotherapy or in combination with cDMARDs, to improve outcomes, such as radiological damage, symptom control, function, and health-related quality of life (HRQoL) in patients with moderate to severe RA [15, 16]. In 2012, a comprehensive review from the Agency of Healthcare Quality (AHRQ) in the USA, compared the therapeutic effects of the different drugs for RA [17], but it was not possible to draw specific conclusions about which therapy is best for patients with RA. Considering the multiple alternatives and different mechanism of action between biologic DMARDs and in the scarcity of head-to-head data, it is important to assess comparative effectiveness of biologics in RA to inform decision makers, physicians, and patients about the best alternative for each patient. Therefore, to address this need, we conducted a systematic review and a network meta-analysis of treatment strategies that incorporate biologic and cDMARD.

## Materials and methods

### Methods of study selection, quality assessment, and appraisal

A systematic review was performed to identify randomized controlled trials of approved biologics. Generally, the reviews followed the Centre for Reviews and Dissemination guidance for undertaking reviews in healthcare [18] and the Cochrane Collaboration Handbook [20]. A range of databases were searched up to January, 2014, as well as searches of trial registries, and the references of identified research and review articles. All relevant studies regardless of language or publication status (published, unpublished, in press or in progress) were considered eligible for this systematic review. Objectively derived search filters for randomized controlled trials were used. No date or language limits were applied; only studies in humans were sought in multiple datasets using specific keywords (Appendix). Screening of titles and abstracts was done by two reviewers independently. The full texts of

potentially relevant studies were obtained and assessed for inclusion by one reviewer with checking of a 10 % random sample by a second reviewer. Disagreements were resolved through discussion or referral to a third reviewer where necessary. Quality assessment was carried out independently by two reviewers using the Cochrane Collaboration quality assessment checklist [19, 20]. Consensus was used to resolve any disagreements. Data extraction sheets were designed and piloted using Microsoft Excel 2007. Data extraction was carried out by one reviewer, and then checked by a second reviewer. Any disagreements were resolved by consensus. A flow diagram of the numbers of studies included and excluded at each stage was provided following guidance in the PRISMA statement (Fig. 1).

**Quantitative analysis and meta-analysis methods**

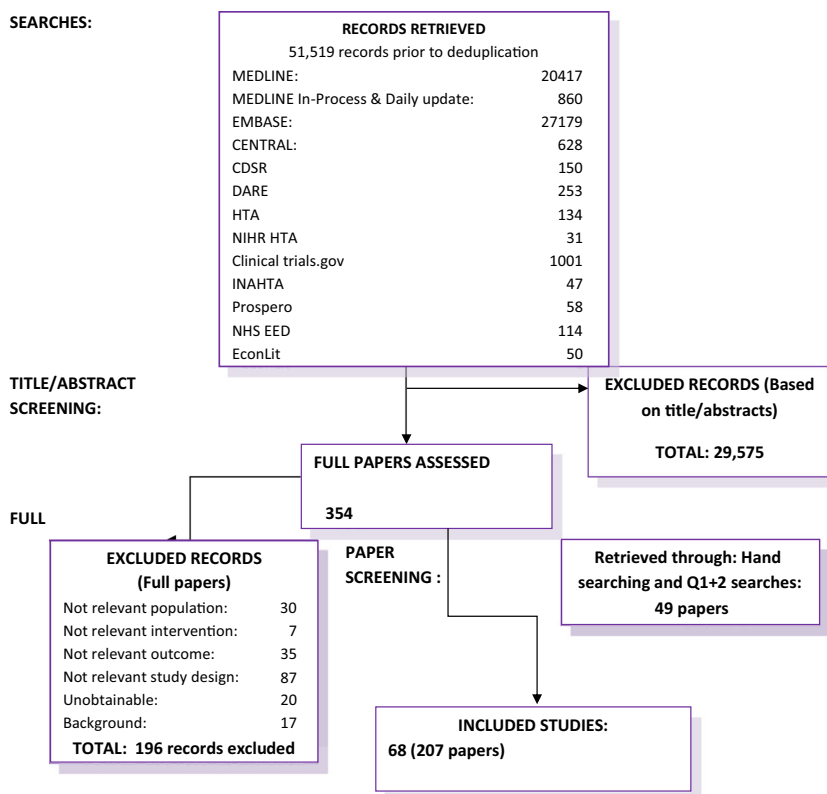
For the few cases when “head-to-head” comparisons of biologic treatments were available, the quantitative analysis was conducted in line with the Cochrane Handbook [20]. However, for the majority of other, non-head-to-head, trials, mixed treatment comparison meta-analysis was conducted using Bayesian methods [21]. The comparison treatment was extracted based on the definition provided by the authors in each RCT. Most of these studies used specific comparators (i.e., placebo or MTX), or used the physician’s criteria with a broader set of treatments in the control arms (i.e., DMARD or Standard Care). The main outcomes considered included:

ACR 20, ACR 50, ACR 70, and EULAR score (moderate or good). Data were collected for outcomes reported at 26 weeks ( $\pm 2$  weeks) and 52 weeks ( $\pm 2$  weeks). Dichotomous outcomes (e.g., number of patients responding according to ACR 20 criteria) were reported as odds ratios (ORs) with 95 % confidence intervals (CI). Pooled effect sizes and 95 % CIs were calculated for the direct head-to-head comparisons using fixed or random effects models, as appropriate, where trials were considered to be clinically and statistically homogeneous. Analyses used Cochrane Review Manager version 5.1 (RevMan 5.1). Publication bias was assessed where there were sufficient numbers of trials.

The most recently approved biologic agent for the management of RA, TCZ, was used as the reference biologic treatment for the analysis. In the absence of trials directly comparing the biologics, indirect treatment comparisons and NMA were performed. Comparisons were performed at three different levels of aggregation: (1) all doses for bDMARDs and cDMARDs were considered as separate interventions, and results were presented for the recommended doses of each DMARD (Appendix 3, Table 2). For cDMARDs, only MTX at the recommended dose was included in this analysis; (2) all doses for cDMARDs were combined as treated as the same intervention; and (3) all doses were combined for all interventions. Doses below the recommended dose for each of the biological DMARD were excluded.

If it was possible to form a connected network of trial evidence (where all treatments could be linked directly or indirectly) then a

**Fig. 1** Summary of study flow



Bayesian NMA was performed using WinBUGs version 1.4 (<http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml>). Vague priors [normal (0, 0.0001)] were used for estimating the trial baselines, treatment differences, and the random effects standard deviation [uniform (0,2)]. A burn-in period of 10,000 simulations was used, followed by a further 20,000 simulations, which were used to obtain parameter estimates. Model fit was investigated using the deviance information criterion (DIC) and residual deviance and, depending on the amount of available evidence, the fit of fixed and random effects models were compared. The results from the most appropriate model are presented as OR with 95 % credible intervals (CrI).

The assumptions of homogeneity, similarity, and consistency as described by Song et al. [22] were assessed, as without these the NMA results may not be valid. Clinical homogeneity was assessed by reviewing the baseline characteristics of the trial populations, including baseline severity of RA, age, concomitant treatments, and comorbidities. Since most biologics were not directly compared, consistency could not be assessed with the available evidence.

## Results

### Systematic review

A total of 158 manuscripts were identified and included through the specific search strategy (Appendix 1, search strategy and results). In addition, 49 papers were retrieved through extended manual searches. In total, 207 publications were included for the assessment of RCTs of clinical effectiveness of biologics in RA. These 207 papers described 68 individual trials in total (Appendix 2, references of studies included, and Appendix 3, Table 1).

All trials included adult patients with moderate to severe active RA, who had an inadequate response to DMARDs. Most of these studies ( $n = 32$ ) recruited patients worldwide and were between 24 to 26 weeks in duration ( $n = 40$ ). Twelve trials were performed only in North America, 11 trials only in Japan, and 3 only in Korea. Five trials were performed only in Europe, one only in China, and two trials in South and North America. For two trials, it was not reported where they were performed. Eighteen trials were 52 weeks in duration, and eight trials followed patients for 104 weeks, but most presented data at 24 weeks. The definition of active RA differed considerably between trials. Most trials accepted the use of rescue medication (24 trials), although this was not reported in 39 trials, while five trials explicitly did not allow the use of rescue medication. In most trials, patients had failed on previous cDMARDs (60 trials), but in five trials, patients had failed on previous TNF treatment, and in three trials, patients had failed on both (Appendix 3, Table 1).

Overall, the methodological quality of the 68 trials was inconsistent. Despite our efforts, a number of the quality assessment criteria could not be assessed and were described as “unclear” due to the inadequate reporting of methods in the trial publications. Poor reporting made the assessment of the randomization method particularly difficult (60/68 trials were “unclear”), and allocation concealment (49/68 trials were “unclear”). Given the difficulties in assessing the methodological quality of a number of the trials, the reliability of the data is uncertain. Despite this uncertainty, 38/68 trials had low risk of bias for selective outcome reporting. In terms of individual trial, the methodological quality varied and some had particular methodological problems of concern, which may affect the robustness of their results. Of note, 8/68 trials had described the randomization method and 17/68 trials described allocation concealment adequately. In 20 trials patients, care providers and outcome assessors were clearly described as blinded but only 11 trials used a true ITT analysis. Assessor blinding or independent verification of outcomes is particularly important for subjective outcomes such as treatment response (ACR and EULAR). Prior knowledge of the treatment being received may lead intentionally or unintentionally to the biased assessment of outcomes in favor of the study drug of interest.

### Network meta-analysis

The 68 trials (Appendix, Table 1) were used to form evidence networks for the analysis of outcome. A total of 24 evidence networks were created and included in the analyses (four main outcomes: ACR 20, ACR 50, ACR 70, and EULAR; two time points: 26 and 52 weeks; and three levels for pooling different doses of interventions).

The results for the main effectiveness outcomes at 26 weeks from the NMA using separate doses for all bDMARDs and only using MTX for cDMARDs at the regular dose (level 1) are presented in Table 1 for all comparisons with TCZ alone, and in Table 2 for all comparisons with TCZ in combination with MTX. These results showed that TCZ alone was superior than standard care/placebo in achieving ACR 20, ACR 50, ACR 70, and EULAR response at 26 weeks of follow-up (OR 13.27, 95 % CrI (3.958, 43.98); 17.45 (10.18, 31.24); 37.77 (7.226, 216.3); 10.42 (1.963, 54.8), respectively). TCZ alone was also significantly better than MTX alone for achieving ACR 50, ACR 70, and EULAR response at 26 weeks of follow-up (OR 5.44, 95 % CrI (4.142, 7.238), 7.364 (1.4, 30.83), and 4.226 (1.184, 15.58), respectively). All the results were estimated using random effects models as these provided a better fit to the data compared to the fixed effect models (lower DIC values). The level 1 analysis for the trials with 52-week results contained fewer trials, and so fixed effect results were presented for ACR 20 and 50 at 52 weeks as these models provided a better fit to the data.

**Table 1** Primary endpoint comparisons with tocilizumab alone (ACR odds ratio vs TCZ-IV8, median [95 % CrI]) at 26 and 52 weeks

Comparator	26 (±2) weeks		52 (±4) weeks	
	ACR 20	ACR 50	ACR 70	ACR 20
ABA-IV10 + MTX	0.997 (0.270, 3.908)	1.155 (0.683, 1.942)	1.803 (0.249, 8.327)	0.88 (0.503, 1.526)
ADA-40bi	2.591 (0.890, 7.544)	2.896 (1.897, 4.418)	2.671 (0.680, 12.51)	NA
ADA-40bi + MTX	0.716 (0.199, 2.501)	0.905 (0.513, 1.583)	1.335 (0.218, 6.807)	1.859 (0.841, 4.141)
CZP-200	0.5917 (0.096, 3.65)	0.628 (0.342, 1.126)	0.6799 (0.0266, 10.7)	NA
CZP-200 + MTX	0.323 (0.061, 1.704)	NA	0.80 (0.088, 5.528)	0.451 (0.231, 0.864)
ETA-50w	2.394 (0.667, 8.34)	4.916 (2.996, 8.167)	4.912 (0.819, 23.24)	3.48 (1.842, 6.628)
ETA-50w + MTX	1.107 (0.246, 4.558)	4.281 (2.667, 6.977)	1.534 (0.163, 6.682)	2.019 (1.045, 3.895)
GOL-SC50 + MTX	0.762 (0.143, 4.293)	1.026 (0.463, 2.246)	1.515 (0.171, 11.44)	NA
INF-3/8 + MTX	1.445 (0.275, 7.549)	1.427 (0.7621, 2.666)	1.897 (0.241, 14.32)	1.176 (0.207, 6.709)
MTX	2.968 (0.954, 9.361)	5.44 (4.142, 7.238)	7.364 (1.4, 30.83)	NA
RIT-1000 + MTX	0.7685 (0.210, 2.79)	0.8774 (0.512, 1.526)	1.503 (0.238, 7.291)	1.529 (0.815, 2.891)
Standard care	13.27 (3.958, 43.98)	17.45 (10.18, 31.24)	37.77 (7.226, 216.3)	3.567 (2.212, 5.761)
Model (DIC)	Random (385.4, 26 studies)	Random (385.5, 27 studies)	Random (324.1, 26 studies)	0.4578 (0.142, 1.375)
				NA
				Fixed (170.9, 11 studies)
				Fixed (150.9, 10 studies)

No ACR 70 and EULAR responses at 52 weeks; no data for INF 3/8, ABA-IV 10, and GOL-SC 50. Odds ratios >1 are in favor of the intervention in the title of each table. Significant results [95 % credible interval not containing 1] are italicized

ABA abatacept, ADA adalimumab, ETA etanercept, INF infliximab, CZP certolizumab pegol, GOL golimumab, RIT rituximab, TCZ tocilizumab, DMARD disease-modifying antirheumatic drug, standard care usual care without additional intervention or placebo

**Table 2** Primary endpoint comparisons with tocilizumab + MTX (ACR odds ratio vs TCZ-IV8 + MTX, median (95 % CrI)) at 26 and 52 weeks

Comparator	26 (±2) weeks		52 (±2) weeks	
	ACR 20	ACR 50	ACR 70	ACR 20
ABA-IV10	NA	NA	NA	NA
ABA-IV10 + MTX	1.404 (0.565, 3.698)	1.607 (1.06, 2.426)	2.12 (0.643, 6.219)	0.953 (0.626, 1.436)
ADA-40bi	3.646 (1.026, 12.95)	4.027 (2.447, 6.607)	3.196 (0.711, 19.77)	1.331 (0.805, 2.241)
ADA-40bi + MTX	1.01 (0.434, 2.262)	1.256 (0.779, 2.003)	1.599 (0.587, 4.436)	NA
CZP-200	0.833 (0.133, 5.311)	0.8746 (0.503, 1.467)	0.832 (0.035, 12.27)	2.005 (0.986, 4.127)
CZP-200 + MTX	0.454 (0.119, 1.761)	NA	0.9641 (0.177, 4.637)	NA
ETA-50w	3.364 (1.061, 10.63)	6.819 (4.515, 10.4)	5.844 (1.545, 22.14)	0.489 (0.279, 0.83)
ETA-50w + MTX	1.556 (0.491, 4.537)	5.961 (4.115, 8.75)	1.809 (0.397, 5.515)	1.745 (0.37, 1.438)
GOL-SC50 + MTX	1.073 (0.273, 4.475)	1.425 (0.6907, 2.908)	1.813 (0.3441, 9.299)	3.77 (2.224, 6.344)
INF-3/8 + MTX	2.025 (0.519, 7.89)	1.986 (1.144, 3.413)	2.254 (0.5353, 11.6)	2.185 (1.27, 3.735)
MTX	4.169 (2.267, 7.871)	5.44 (4.142, 7.238)	8.731 (4.203, 19.29)	NA
RIT-1000 + MTX	1.086 (0.45, 2.63)	1.224 (0.774, 1.916)	1.796 (0.59, 5.259)	NA
Standard care	18.63 (5.32, 66.81)	24.27 (14.5, 41.91)	46.13 (10.08, 277)	NA
TCZ-IV8	1.408 (0.499, 3.984)	1.389 (1.004, 1.925)	1.185 (0.313, 5.965)	0.477 (0.078, 2.021)
Model (DIC)	Random (385.4, 26 studies)	Random (385.5, 27 studies)	Random (324.1, 26 studies)	0.81 (0.576, 1.129)
				Fixed (170.9, 11 studies)
				Fixed (150.9, 10 studies)
				Fixed (52.2, 3 studies)

Odds ratios >1 are in favor of the intervention in the title of each table. Significant results are italicized

In the level 1 analysis, the combination of TCZ and MTX was significantly better than standard care/placebo and MTX alone for ACR 20, ACR 50, ACR 70, and EULAR response outcomes at 26 weeks follow-up (OR 18.63 (95 % CrI 5.32, 66.81), 24.27 (95 % CrI 14.5, 41.91), 46.13 (95 % CrI 10.08, 277), 14.23 (95 % CrI 2.493, 84.02), 4.169 (95 % CrI 2.267, 7.871), 5.44 (95 % CrI 4.142, 7.238), 8.731 (95 % CrI 4.203, 19.29), 7.306 (95 % CrI 4.393, 13.04), respectively). At 52 weeks, when compared to MTX alone, TCZ + MTX were significantly better for ACR 20 and ACR 50 response. However, no comparison was possible with standard care/placebo at 52 weeks due to lack of data. TCZ + MTX was significantly better than etanercept alone for ACR 20, ACR 50, and ACR 70 responses at 26 weeks and ACR 20 and ACR 50 responses at week 52. Compared to adalimumab monotherapy, TCZ + MTX was significantly better for ACR 20 and ACR 50 response at 26 weeks. TCZ + MTX was significantly better in ACR responses for the following comparisons: abatacept + MTX for ACR 50 at 26 weeks; adalimumab + MTX for ACR 50 at 52 weeks; etanercept + MTX for ACR 50 at 26 weeks and ACR 20 at 52 weeks; infliximab + MTX for ACR 50 at 26 weeks, and TCZ alone for ACR 50 at 26 weeks. Only one comparison favored the comparator: CZP-200 + MTX, which was for ACR 20 at 52 weeks follow-up (OR 0.489, 95 % CrI (0.279, 0.83)). Comparisons with abatacept, adalimumab alone, certolizumab alone, golimumab alone or in combination, and infliximab alone could not be assessed due to lack of data.

For EULAR responses, TCZ + MTX was significantly better than adalimumab + MTX at 26 weeks and infliximab + MTX or abatacept + MTX at 52 weeks.

The results for the level 2 NMA—using separate doses for all bDMARDs and all doses for cDMARDs (Tables 3 and 4) and level 3—using combined doses for all bDMARDs and cDMARDs (Tables 5 and 6) resulted in slightly different results due to the inclusion of additional trials with different doses for cDMARDs and bDMARDs (licensed doses and higher doses). However, overall, the results were consistent. In the level 3 analyses, TCZ alone and TCZ + DMARD were significantly better to DMARD alone and standard of care for all outcomes assessed, with one exception, the EULAR score response at 26 weeks was no longer significant in favor of TCZ alone and TCZ + DMARD when compared with standard care/placebo. If there were significant differences in comparison with other bDMARDs, they were in favor of TCZ alone and TCZ + DMARD, but in different magnitudes as previously described.

### Discussion

In this systematic review, 68 trials were included to assess the clinical effectiveness of biologic DMARDs in RA. Overall, the methodological quality of the included trials was

**Table 3** Primary endpoint comparisons with tocilizumab alone (ACR odds ratio vs TCZ-IV8, median [95 % CrI]) at 26 and 52 weeks

Comparator	26 (±2) weeks			52 (±4) weeks		
	ACR 20	ACR 50	ACR 70	EULAR	ACR 20	ACR 50
ABA-IV10 + DMARD	2.427 (1.259, 4.786)	1.214 (0.265, 5.716)	1.672 (0.717, 3.77)	NA	1.232 (0.801, 1.898)	2.175 (1.253, 3.831)
ADA-40bi	NA	2.779 (0.613, 12.64)	2.437 (1.174, 5.208)	2.931 (0.936, 9.106)	NA	NA
ADA-40bi + DMARD	1.92 (0.881, 4.155)	1.154 (0.263, 5.32)	1.567 (0.66, 3.563)	1.613 (0.342, 7.809)	2.549 (1.211, 5.225)	6.304 (2.727, 14.07)
CZP-200	0.7326 (0.150, 3.428)	0.725 (0.111, 4.939)	1.035 (0.361, 2.729)	0.829 (0.106, 6.116)	NA	NA
CZP-200 + DMARD	0.816 (0.348, 1.936)	1.015 (0.149, 6.992)	0.242 (0.027, 1.186)	0.381 (0.088, 1.62)	0.626 (0.349, 1.089)	1.207 (0.574, 2.47)
ETA-50w	3.759 (1.667, 8.391)	2.332 (0.518, 11.17)	3.727 (1.505, 8.526)	NA	2.768 (1.786, 4.3)	3.11 (1.757, 5.576)
ETA-50w + DMARD	1.543 (0.623, 3.717)	2.123 (0.455, 9.628)	1.886 (0.776, 4.367)	NA	2.526 (1.623, 3.951)	3.733 (2.053, 6.785)
GOL-SC50 + DMARD	1.792 (0.72, 4.494)	1.081 (0.198, 6.176)	1.586 (0.593, 4.057)	1.176 (0.207, 6.709)	NA	NA
INF-3/8 + DMARD	3.211 (0.888, 12.01)	1.848 (0.186, 19.89)	1.822 (0.585, 6.018)	NA	2.278 (1.438, 3.614)	2.563 (1.348, 4.814)
DMARD	6.52 (3.581, 12.15)	4.5 (1.276, 16.34)	7.845 (3.98, 15.7)	4.226 (1.184, 15.58)	4.877 (3.469, 6.891)	8.364 (5.241, 13.26)
RIT-1000 + DMARD	1.668 (0.717, 3.903)	0.9759 (0.194, 4.985)	1.677 (0.654, 4.088)	0.953 (0.233, 3.912)	0.629 (0.202, 1.739)	0.728 (0.091, 3.344)
Standard care	16.17 (6.82, 38.72)	13.73 (2.976, 64.99)	33.68 (12.46, 90.57)	10.42 (1.963, 54.8)	NA	NA
Model (DIC)	Random (626.4, 43 studies)	Random (618.8, 43 studies)	Random (531.9, 45 studies)	Random (258.8, 18 studies)	Fixed (264.7, 16 studies)	Fixed (215.5, 13 studies)

Odds ratios >1 are in favor of the intervention in the title of each table. Significant results are italicized

**Table 4** Primary endpoint comparisons with tocilizumab + DMARD (ACR odds ratio vs TCZ-IV8 + DMARD, median [95 % CrI]) at 26 and 52 weeks

Comparator	52 (±2) weeks						
	ACR 20	ACR 50	ACR 70	EULAR	ACR 20	ACR 50	EULAR
ABA-IV10 + DMARD	1.569 (0.87, 2.883)	1.08 (0.344, 3.308)	1.828 (0.98, 3.436)	NA	1.109 (0.754, 1.618)	1.761 (1.126, 2.769)	4.583 (2.534, 8.226)
ADA-40bi	NA	2.464 (0.49, 12.34)	2.664 (1.121, 6.902)	4.095 (0.792, 20.35)	NA	NA	NA
ADA-40bi + DMARD	1.24 (0.65, 2.329)	1.018 (0.369, 2.909)	1.703 (0.923, 3.184)	2.245 (0.813, 6.581)	2.288 (1.114, 4.63)	5.112 (2.353, 10.65)	NA
CZP-200	0.472 (0.102, 2.119)	0.646 (0.118, 3.506)	1.14 (0.4422, 2.656)	1.151 (0.11, 11.56)	NA	NA	NA
CZP-200 + DMARD	0.528 (0.247, 1.157)	0.902 (0.181, 4.469)	0.268 (0.033, 1.175)	0.532 (0.228, 1.28)	0.564 (0.326, 0.942)	0.978 (0.501, 1.837)	1.545 (0.795, 2.904)
ETA-50w	2.43 (1.193, 4.881)	2.07 (0.585, 7.257)	4.079 (1.97, 8.154)	NA	2.275 (1.525, 3.358)	3.021 (1.854, 4.988)	NA
ETA-50w + DMARD	0.9998 (0.439, 2.142)	1.885 (0.604, 5.628)	2.073 (1.031, 4.001)	NA	NA	NA	NA
GOL-SC50 + DMARD	1.158 (0.505, 2.599)	0.963 (0.243, 3.807)	1.728 (0.784, 3.871)	1.634 (0.436, 6.043)	NA	2.077 (1.19, 3.519)	5.617 (3.514, 9.085)
INF-3/8 + DMARD	2.075 (0.611, 7.091)	1.648 (0.209, 13.47)	1.994 (0.745, 5.801)	NA	4.387 (3.343, 5.808)	6.752 (4.871, 9.366)	5.477 (4.06, 7.447)
DMARD	4.216 (2.718, 6.589)	3.995 (1.947, 8.198)	8.62 (5.777, 13.39)	5.893 (3.507, 10.37)	0.566 (0.184, 1.543)	0.591 (0.076, 2.576)	NA
DMARD	1.075 (0.513, 2.246)	0.865 (0.248, 2.922)	1.836 (0.871, 3.785)	1.331 (0.603, 3.039)	NA	NA	NA
Standard care	10.47 (4.803, 22.76)	12.18 (2.746, 53.4)	37.1 (14.44, 100.4)	14.69 (1.919, 110.3)	NA	NA	NA
TCZTCZ-IV8	0.647 (0.330, 1.248)	0.887 (0.235, 3.246)	1.092 (0.598, 2.11)	1.398 (0.438, 4.502)	0.9 (0.66, 1.223)	0.8094 (0.575, 1.132)	NA
Model (DIC)	Random (626.4, 43 studies)	Random (618.8, 43 studies)	Random (531.9, 45 studies)	Random (258.8, 18 studies)	Fixed (264.7, 16 studies)	Fixed (215.5, 13 studies)	Fixed (72.7, 4 studies)

Odds ratios >1 are in favor of the intervention in the title of each table. Significant results are italicized

**Table 5** Primary endpoint comparisons with tocilizumab alone (ACR odds ratio vs TCZ, median [95 % CrI]) at 26 and 52 weeks

Comparator	26 (±2) weeks			52 (±4) weeks		
	ACR 20	ACR 50	ACR 70	EULAR	ACR 20	ACR 50
ABA + DMARD	1.759 (0.744, 4.413)	0.795 (0.205, 3.014)	1.54 (0.651, 3.564)	NA	1.801 (1.201, 2.711)	2.109 (1.205, 3.75)
ADA	2.898 (1.247, 6.85)	2.213 (0.627, 7.776)	2.268 (1.08, 4.837)	1.539 (0.386, 6.25)	NA	NA
ADA + DMARD	1.273 (0.547, 3.004)	0.692 (0.196, 2.516)	1.325 (0.601, 2.961)	3.533 (0.564, 23.53)	4.491 (2.916, 6.904)	2.064 (1.084, 3.969)
CZP	0.589 (0.138, 2.517)	0.806 (0.177, 3.692)	0.3313 (0.012, 3.452)	0.215 (0.023, 2.228)	NA	NA
CZP + DMARD	0.655 (0.27, 1.611)	0.4352 (0.094, 2.016)	0.681 (0.238, 1.767)	0.759 (0.137, 4.405)	1.613 (1.154, 2.246)	1.214 (0.602, 2.504)
ETA	3.162 (1.344, 7.57)	1.68 (0.434, 6.489)	3.403 (1.445, 7.697)	NA	2.875 (1.917, 4.317)	3.094 (1.72, 5.541)
ETA + DMARD	1.252 (0.459, 3.262)	1.272 (0.334, 4.562)	1.761 (0.758, 3.844)	NA	3.441 (2.271, 5.201)	3.721 (2.079, 6.782)
GOL + DMARD	1.372 (0.545, 3.53)	1.229 (0.319, 4.872)	1.683 (0.661, 4.258)	NA	NA	NA
INF + DMARD	2.499 (0.667, 9.615)	1.186 (0.151, 9.547)	1.668 (0.543, 5.565)	2.471 (0.287, 21.6)	2.745 (1.782, 4.213)	2.325 (1.251, 4.359)
DMARD	5.148 (2.461, 11.1)	2.914 (0.995, 8.79)	7.319 (3.847, 14.47)	9.775 (2.246, 44.71)	7.676 (5.642, 10.5)	8.293 (5.249, 13.39)
RIT + DMARD	1.323 (0.527, 3.3)	0.632 (0.155, 2.534)	1.571 (0.6155, 3.922)	2.574 (0.454, 15.16)	0.996 (0.32, 2.809)	0.903 (0.147, 3.704)
Standard care	15.73 (6.619, 38.03)	7.825 (2.284, 27.53)	30.15 (10.89, 88.37)	2.736 (0.565, 14.55)	NA	NA
Model (DIC)	Random (719.5, 49 studies)	Random (849.1, 44 studies)	Random (586.4, 49 studies)	Random (271.0, 18 studies)	Fixed (420.5, 17 studies)	Fixed (249.3, 15 studies)

Odds ratios >1 are in favor of the intervention in the title of each table. Significant results are italicized

**Table 6** Primary endpoint comparisons with tocilizumab + DMARD (ACR odds ratio vs TCZ + DMARD, median [95 % CrI]) at 26 and 52 weeks

Comparator	26 (±2) weeks			52 (±2) weeks			
	ACR 20	ACR 50	ACR 70	EULAR	ACR 20	ACR 50	EULAR
ABA + DMARD	1.547 (0.843, 2.909)	1.217 (0.451, 3.255)	1.557 (0.814, 2.912)	NA	1.242 (0.857, 1.802)	1.71 (1.09, 2.684)	4.583 (2.534, 8.226)
ADA	2.551 (1.035, 6.173)	3.362 (0.983, 11.63)	2.279 (0.932, 5.734)	1.099 (0.216, 5.783)	NA	NA	NA
ADA + DMARD	1.118 (0.64, 1.948)	1.058 (0.451, 2.48)	1.326 (0.744, 2.371)	2.522 (0.696, 9.647)	3.099 (2.083, 4.589)	1.669 (0.962, 2.912)	NA
CZP	0.515 (0.137, 1.922)	1.234 (0.333, 4.508)	0.332 (0.011, 3.649)	0.154 (0.017, 1.445)	NA	NA	NA
CZP + DMARD	0.573 (0.298, 1.125)	0.668 (0.199, 2.243)	0.682 (0.29, 1.476)	0.541 (0.167, 1.744)	1.113 (0.791, 1.554)	0.988 (0.524, 1.845)	NA
ETA	2.768 (1.414, 5.379)	2.583 (0.895, 7.366)	3.434 (1.72, 6.479)	NA	1.992 (1.373, 2.87)	2.493 (1.556, 4.055)	1.545 (0.795, 2.904)
ETA + DMARD	1.094 (0.5, 2.312)	1.939 (0.733, 4.959)	1.769 (0.935, 3.143)	NA	2.373 (1.64, 3.462)	3.011 (1.873, 4.984)	NA
GOL + DMARD	1.2 (0.608, 2.408)	1.878 (0.667, 5.371)	1.695 (0.803, 3.485)	1.754 (0.310, 10.28)	NA	NA	NA
INF	NA	NA	NA	NA	NA	NA	NA
INF + DMARD	2.18 (0.671, 7.049)	1.818 (0.287, 11.53)	1.678 (0.635, 4.792)	NA	1.889 (1.267, 2.807)	1.886 (1.11, 3.192)	5.617 (3.514, 9.085)
DMARD	4.519 (3.034, 6.808)	4.472 (2.524, 7.923)	7.355 (5.263, 10.66)	6.947 (3.381, 14.46)	5.296 (4.081, 6.897)	6.708 (4.913, 9.391)	5.477 (4.06, 7.447)
RIT + DMARD	1.159 (0.598, 2.246)	0.969 (0.33, 2.829)	1.588 (0.762, 3.224)	1.832 (0.559, 6.092)	0.687 (0.224, 1.902)	0.732 (0.1233, 2.879)	NA
Standard care	13.79 (6.106, 31.19)	11.87 (4.227, 34.92)	30.23 (10.97, 89.38)	1.954 (0.463, 8.795)	NA	NA	NA
TCZ	0.878 (0.405, 1.844)	1.526 (0.499, 4.642)	1.004 (0.536, 1.9)	0.709 (0.179, 2.869)	0.69 (0.51, 0.933)	0.81 (0.578, 1.13)	NA
Model (DIC)	Random (719.5, 49 studies)	Random (849.1, 44 studies)	Random (586.4, 49 studies)	Random (271.0, 18 studies)	Fixed (420.5, 17 studies)	Fixed (249.3, 15 studies)	Fixed (72.6, 5 studies)

Odds ratios > 1 are in favor of the intervention in the title of each table. Significant results are italicized

inconsistent. There was considerable clinical heterogeneity within the systemic review, particularly with regard to the patient populations and dosing schedules used. Differences in clinical and demographic characteristics, as well as previous exposure, failure with previous treatments (i.e., DMARD IR or TNF IR), and differences in the definition of standard care as comparator, in each trial population may influence the analysis and the results of the NMA. Some of these characteristics were measured and reported in some but not all the trials systematically. These are important limitations, and the impact of those differences cannot be fully quantified in the current analysis and must be considered when putting in context these results. In addition, as most NMA, the availability of new evidence will require additional analysis to incorporate that information and provide new estimates about the effectiveness of biologic treatment in RA. Safety endpoints were not included in this analysis, in part due to higher heterogeneity and differences across trials about the definitions and reporting of adverse events. Therefore, results of these analyses should be interpreted in the context of those limitations. Another limitation of this study is the use of TCZ, the only IL-6 inhibitor receptor currently available, as the reference for comparison as the most recently approved biologic in this indication. During the study period, only TCZ published RCTs were available and therefore limited the generalizability of the study findings to the IL-6 class. With the availability of new biologics, including new therapies, better generalizability of the results may be obtained by including those studies.

The NMA was done at three levels, allowing only recommended doses for all drugs at level 1, combining all cDMARDs in level 2, and combining all cDMARDs and doses for bDMARDs (except subrecommended doses) in level 3. Overall, the results of these analyses were similar for all three levels, providing consistency to the results. Results of the network analyses showed that TCZ alone and TCZ + MTX are superior to DMARDs alone and standard care on most outcomes assessed, where enough information was available to perform the analysis. Few significant differences were found between TCZ alone and any of the biologicals alone or in combination with MTX. Only one comparison favored the comparator. Certolizumab + MTX showed significantly better response for ACR 20 at 52 weeks of follow-up compared to TCZ alone in the level 1 analysis. This difference was not found in any of the other analyses performed.

Other NMA for biologics in RA have been published focusing on different evidence networks and endpoints. Recently, Tvete et al. [23] conducted a multiple treatment comparison analysis using ACR 50 as the dependent variable and dose level and disease duration as the independent variable for assessing the comparable relative effect between nine biologics (adalimumab, certolizumab, etanercept, golimumab, infliximab, anakinra, abatacept, rituximab, and tocilizumab) and placebo or DMARD. In contrast to our analysis, Tvete et al. aggregated



ACR 50 response across all periods (12 to 54 weeks) and did not consider other response variables (ACR 20, ACR 70, EULAR response). This analysis, based on 54 publications, embraced all treatment and comparators arms over all publications. The authors found the drug effect to be dependent on dose level, but not on disease duration, and the impact of a high versus low dose level was the same for all drugs. Similar to our analysis, differences in patient characteristics between trials was not fully incorporated, and the authors concluded that all biologic agents were more effective than placebo.

## Conclusions

The systematic review of RCTs of biologics in adults with RA who have failed treatment with conventional disease-modifying agents for rheumatoid disease (cDMARD) showed inconsistencies and varying degrees of methodological quality. Network meta-analysis allows combining information from different clinical trials and performing comparisons within the context of the available evidence. The resulting assessment of clinical effectiveness, using the network meta-analyses, showed that TCZ alone and TCZ + MTX were superior to cDMARDs alone and standard care on all outcomes assessed and as effective as other biologics in RA. Few significant differences were found between TCZ alone and any of the biologicals alone or in combination with MTX. As new evidence and agents becomes available, additional analysis must be performed to improve the quality of these comparisons.

**Acknowledgments** This study/analysis was funded by GlaxoSmithKline. The authors would like to acknowledge Jessica Panish for her support and contribution to this manuscript. Copy-edit assistance was provided by Alanna Franchetti of MedErgy and was funded by GSK.

## Compliance with ethical standards

**Conflicts of interest** • R Alfonso, R Arjunji, and R Ganguly are employees of and own stocks in GSK.

• N Armstrong, R Riemsma, G Worthy, and J Kleijnen are employees of KSR.

• KSR received research funding to conduct the systematic review and the analysis for this research from GSK.

All authors had full control of all primary data and agree to allow the journal to review their data if requested. The manuscript does not report a clinical study or contain patient data.

## References

- Badley E, DesMeules M (2003) Introduction. In: Badley E, DesMeules M, eds. *Arthritis in Canada: an ongoing challenge*. Ottawa, Canada: Health Canada. <http://www.lupuscanada.org/english/resources/pdfs/Arthritis-in-Canada-Health-Canada.pdf>. Accessed 13 September 2013.
- Symmons D, Turner G, Webb R et al (2002) The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology (Oxford)* 41:793–800
- Helmick CG, Felson DT, Lawrence RC et al (2008) Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. *Arthritis Rheum* 58:15–25
- Wong R, Davis AM, Badley E, Grewal R, Mohammed M (2010) Prevalence of arthritis and rheumatic diseases around the world: a growing burden and implications for health care needs. In: *Models in Care in Arthritis, Bone & Joint Disease (MOCA)* <http://www.modelsofcare.ca/pdf/10-02.pdf>. Accessed 13 September 2013.
- Alamanos Y, Voulgari PV, Drosos AA (2006) Incidence and prevalence of rheumatoid arthritis, based on the 1987 American College of Rheumatology criteria: a systematic review. *Semin Arthritis Rheum* 36:182–188
- Filipovic I, Walker D, Forster F, Curry AS (2011) Quantifying the economic burden of productivity loss in rheumatoid arthritis. *Rheumatology (Oxford)* 50:1083–1090
- Sokka T, Kautiainen H, Pincus T et al (2010) Work disability remains a major problem in rheumatoid arthritis in the 2000s: data from 32 countries in the QUEST-RA study. *Arthritis Res Ther* 12:R42
- Pugner KM, Scott DI, Holmes JW, Hieke K (2000) The costs of rheumatoid arthritis: an international long-term view. *Semin Arthritis Rheum* 29:305–320
- Centers for Disease Control and Prevention (2011) *Arthritis-related statistics*. Atlanta, GA, USA: CDC. [http://www.cdc.gov/arthritis/data\\_statistics/arthritis\\_related\\_stats.htm](http://www.cdc.gov/arthritis/data_statistics/arthritis_related_stats.htm). Accessed 13 September 2013.
- Birnbaum HG, Pike C, Banerjee R, Waldman T, Cifaldi M (2012) Changes in utilization and costs for patients with rheumatoid arthritis, 1997 to 2006. *Pharmacoeconomics* 30:323–336
- Kawatkar AA, Jacobsen SJ, Levy GD, Medhekar SS, Venkatasubramanian KV, Herrinton LJ (2012) Direct medical expenditure associated with rheumatoid arthritis in a nationally representative sample from the medical expenditure panel survey. *Arthritis Care and Research* 64:1649–1656
- Kleinman NL, Cifaldi MA, Smeeding JE, Shaw JW, Brook RA (2013) Annual incremental health benefit costs and absenteeism among employees with and without rheumatoid arthritis. *Journal of Occupational & Environmental Medicine* 55:240–244
- Simons WR, Rosenblatt LC, Trivedi DN (2012) The economic consequences of rheumatoid arthritis: analysis of Medical Expenditure Panel Survey 2004, 2005, and 2006 data. *Journal of Occupational & Environmental Medicine* 54:48–55
- Poole CD, Currie CJ, Holden S, Singh A (2012) Impact of disease severity on the total, health-related, financial costs of treatment of people with rheumatoid arthritis in the United Kingdom. Presented at: 17th Annual International Meeting of the International Society for Pharmacoeconomics and Outcomes Research, ISPOR; 2–6 June 2012; Washington, DC, USA. *Value Health* 15:A36–A37
- National Collaborating Centre for Chronic Conditions (2009) *Rheumatoid arthritis: national clinical guideline for management and treatment in adults* [Internet]. London, UK: Royal College of Physicians. <http://www.nice.org.uk/nicemedia/live/12131/43326/43326.pdf>. Accessed 13 September 2013.
- Singh JA, Saag KG, Bridges SL et al (2016) 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. *Arthritis Rheumatol* 68:1–26
- Donahue KE, Jonas DE, Hansen RA et al (2012) Drug therapy for rheumatoid arthritis in adults: an update. Rockville, MD, USA: Agency for Healthcare Research and Quality. <http://www.ncbi.nlm.nih.gov/books/NBK97388/>. Accessed 13 September 2013.
- Centre for Reviews and Dissemination (2009) *Systematic reviews: CRD's guidance for undertaking reviews in health care*. York, UK: University of York. <http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm>. Accessed 1 April 2014.

19. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD et al (2011) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Br Med J* 343:d5928
20. Higgins JPT, Green S, eds- (2011) *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration. <http://www.cochrane-handbook.org/>. Accessed 1 April 2014.
21. Hoaglin D, Hawkins N, Jansen J, Scott D, Itzler R, Cappelleri J et al (2011) Conducting indirect treatment comparison and network meta-analysis studies: report of the ISPOR task force on indirect treatment comparisons: part 2. *Value Health* 14:429–437
22. Song F, Altman DG, Glenny A-M, Deeks JJ (2003) Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *Br Med J* 326:472
23. Tvette IF, Natvig B, Gasemyr J, Meland N, Roine M, Klemp M (2015) Comparing effects of biologic agents in treating patients with rheumatoid arthritis: a multiple treatment comparison regression analysis. *PLoS One* 10:e0137258