This is a repository copy of *Rheumatoid arthritis treated with 6-months of first-line biologic or biosimilar therapy: an updated systematic review and network meta-analysis.*

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/142387/

Version: Accepted Version

**Article:**
Simpson, E.L. orcid.org/0000-0001-7353-5979, Ren, S. orcid.org/0000-0003-3568-7124, Hock, E.S. orcid.org/0000-0002-8617-8875 et al. (8 more authors) (2019) Rheumatoid arthritis treated with 6-months of first-line biologic or biosimilar therapy: an updated systematic review and network meta-analysis. International Journal of Technology Assessment in Health Care. ISSN 0266-4623

https://doi.org/10.1017/S0266462318003628

This article has been published in a revised form in International Journal of Technology Assessment in Health Care [http://doi.org/10.1017/S0266462318003628]. This version is free to view and download for private research and study only. Not for re-distribution, re-sale or use in derivative works. © Cambridge University Press

**Reuse**
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Rheumatoid arthritis treated with six-months of first-line biologic or biosimilar therapy: an updated systematic review and network meta-analysis

Authors: Emma L. Simpson¹, Shijie Ren¹, Emma S. Hock¹, John W. Stevens¹, Aymeric Binard², Yves-Marie Pers³, Rachel Archer¹, Suzy Paisley, Matthew D. Stevenson¹, Chloé Herpin⁴, Salah Ghabri⁴

Author affiliations: ¹ School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK; ² Department of Rheumatology, La Cavale Blanche University Hospital, 29200 Brest, France; ³ Clinical Immunology and Osteoarticular Diseases Therapeutic Unit, Lapeyronie University Hospital, Montpellier, France; ⁴ Department of Economic and Public Health Evaluation, French National Authority for Health (Haute Autorité de Santé, HAS)

ABSTRACT

Objective:

To estimate the effectiveness of first-line biologic disease modifying drugs (boDMARDs), and their approved biosimilars (bsDMARDs), compared to conventional (csDMARD) treatment, in terms of ACR (American College of Rheumatology) and EULAR (European League against Rheumatism) responses.

Methods:

Systematic literature search, on eight databases to January 2017, sought ACR and EULAR data from randomised controlled trials (RCTs) of boDMARDs / bsDMARDs (in combination with csDMARDs, or monotherapy). Two adult populations: methotrexate (MTX)-naïve patients with severe active RA; and csDMARD-experienced patients with moderate-to-severe active RA. Network meta-analyses (NMA) were conducted using a Bayesian Markov chain Monte Carlo simulation using a random effects model with a probit link function for ordered categorical.

Results:
Forty-six RCTs met the eligibility criteria. In the MTX-naive severe active RA population, no biosimilar trials meeting the inclusion criteria were identified. MTX plus methylprednisolone (MP) was most likely to achieve the best ACR response. There was insufficient evidence that combination boDMARDs was superior to intensive (two or more) csDMARDs. In the csDMARD-experienced, moderate-to-severe RA population, the greatest effects for ACR responses were associated with tocilizumab (TCZ) monotherapy, and combination therapy (plus MTX) with bsDMARD etanercept (ETN) SB4, boDMARD ETN and TCZ. These treatments also had the greatest effects on EULAR responses. No clear differences were found between the boDMARDs and their bsDMARDs.

**Conclusion:**

In MTX-naive patients, there was insufficient evidence that combination boDMARDs was superior to two or more csDMARDs. In csDMARD-experienced patients, boDMARDs and bsDMARDs were comparable and all combination boDMARDs / bsDMARDs were superior to single csDMARD.

**Keywords:** rheumatoid arthritis, biological agents, biosimilar, systematic review, network meta-analysis

**Acknowledgement**

The authors would like to thank the anonymous reviewers whose constructive comments have allowed to improve the quality of the manuscript. The authors would particularly like to thank members of the HAS Rheumatoid Arthritis scientific group for their useful comments during the 2nd HAS RA meeting group: Aymeric Binard, Morgane Beck, François Bocquet, Franck Maunoury, Hans-Martin Spath, Yves-Marie Pers and Sandrine Rollot. The authors would also like to thank David Scott and Fowzia Ibrahim for providing data from TACIT trial; Jackie Nam for providing data on the IDEA trial; David Scott and Adam Young for their help with trial selection in the original review; Gwenael Le Teuff for his comments on the methodology and the results of the network meta-analysis; Jaime Caro for his review of the statistical analysis plan of the network meta-analysis and his comments during the 2nd HAS RA meeting group.
Financial support

This work was supported in part by the French National Authority for health (Haute Autorité de Santé, HAS), and in part provided by the Health Technology Assessment (HTA) programme of the National Institute for Health Research (NIHR) on behalf of the National Institute for Health and Care Excellence (NICE) (project number 11/74/01). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the HTA programme, NIHR, National Health Service, UK Department of Health, or Haute Autorité de Santé.

Conflict of interest

The authors declare that they have no competing interest
The introduction of novel targeted therapies such as biologic original disease-modifying anti-rheumatic drugs (boDMARDs) has expanded the arsenal of available drugs for rheumatoid arthritis (RA). These are usually prescribed on failure of conventional synthetic disease-modifying anti-rheumatic drugs (csDMARDs). Biosimilars (bsDMARDs) of boDMARDs have recently been approved by the European Medicines Agency (EMA). The introduction of bsDMARDs onto the market may have the potential to provide a cheaper alternative to boDMARDs, provided they have similar effectiveness.

There are a wide range of treatment options offered, and few head-to-head randomised controlled trials (RCTs) between boDMARDs and bsDMARDs. A network meta-analysis allows a synthesis of all available evidence.

In 2017, the French national authority for health (Haute Autorité de Santé, HAS) initiated an economic evaluation of biological treatments for RA in csDMARD-experienced and MTX-naïve populations. The rationale of that evaluation emphasised two issues:

- boDMARDs place a substantial financial burden on healthcare systems and individual patients. The overall costs of boDMARDs should take into account the benefit of reducing the disease impact, however research is required to fully assess differences in cost-effectiveness between the currently available treatments
- A wide range of treatment options is offered, and there are few head-to-head randomised controlled trials (RCTs) between boDMARDs and bsDMARDs. A network meta-analysis can provide useful comparative evidence in order to define the best treatment strategies and a simultaneous comparison between treatments.

Unlike other reviews of biologics in RA, the current review included licensed bsDMARDs and boDMARDs compared with csDMARD therapy, and considered MTX-naive and csDMARD-experienced populations separately. Our study focused on the effectiveness criteria and aimed to estimate the short-term comparative effectiveness, on American College of Rheumatology (ACR) and European League Against Rheumatism (EULAR) responses, of first-line boDMARDs, and their EMA approved biosimilars. The choice of these outcomes was validated by the HAS clinical experts of the RA economic evaluation. Analysis of the safety outcomes were
beyond the scope of this study. This will be proposed in an independent chapter of the HAS on-going cost effectiveness report

MATERIALS AND METHODS

The review was conducted in accordance with the general principles recommended in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. (1) The review was an update of work by Stevenson et al 2016. (2)

Data sources

The following electronic bibliographic databases were searched: Medline and Medline in Process (via Ovid SP); Embase (via Ovid SP); Cochrane Database of Systematic Reviews (via Wiley); Cochrane Central Register of Controlled Trials (via Wiley); Health Technology Assessment Database (via Wiley); Database of Abstracts of Reviews of Effect (via Wiley); CINAHL (via EBSCOhost); and Toxline (via ProQuest). Original searches were performed on the databases from inception until May 2013 (2) and the searches were updated on 23rd January 2017. The exact search strategies are available from the authors. Electronic database searches were supplemented with searching of bibliographies of included trials, and information provided by trial authors.

Trial selection

Study design was restricted to RCTs. Two populations of adult RA patients were included: a MTX-naïve population with severe active RA (DAS28≥5.1); and adults with moderate to severe active RA (DAS28≥3.2) previously treated with, and inadequately responded to, csDMARDs (a csDMARD-experienced population). csDMARDs included methotrexate, sulfasalazine, leflunomide, with methotrexate being most commonly used. The following boDMARDs and bsDMARDs were included, as first-line biologic treatment, prescribed in accordance with EMA licensed indications: abatacept (ABT); adalimumab (ADA); certolizumab pegol (CTZ);
etanercept (ETN) and its biosimilar SB4 (Benepali); golimumab (GOL); infliximab (IFX) and its biosimilars CT-P13 (Inflectra or Remsima) and SB2 (Flixabi); and tocilizumab (TCZ). Following the date of our searches, biosimilars of adalimumab (Amgevita, Cyltezo, Imraldi, Solymbic) have been approved, however as these were not licensed at the time of searches they were not included. boDMARDs and bsDMARDs could be delivered either as monotherapy (as allowed by current French licensed indications), or in combination with csDMARDs (biologic plus csDMARD combination therapy is illustrated by + in Tables and Figures).

Included comparators were boDMARDs or bsDMARDs compared with each other, single csDMARD treatment, or intensive csDMARDs (two or more csDMARDs). Additionally, for the MTX-naïve population, management strategies involving further conventional DMARDs (e.g. SSZ, LEF), NSAIDs and corticosteroids. The outcomes sought were ACR responses or EULAR responses at follow-up between 22 weeks and 30 weeks. RCTs with early escape were only included if they reported a non-responder imputation. Two reviewers independently selected trials based on the review inclusion criteria, with any discrepancy resolved by a third reviewer.

Data collection and assessment of bias

One reviewer extracted data, and these data were checked by a second reviewer. Study arms where intervention treatments were administered in line with licensed indications were extracted. Where studies had treatment arms with unlicensed doses, these were not extracted. Two reviewers independently performed quality assessment based on the 2011 Cochrane risk of bias tool criteria 1.0. (3) Data extraction and quality assessment forms developed in Stevenson et al (2) were used. Any discrepancy was resolved by a third reviewer.

Network meta-analysis

Network meta-analyses (NMAs) of ACR and EULAR data at 22-30 weeks follow-up were conducted using Bayesian Markov chain Monte Carlo (MCMC) simulation. ACR and EULAR data were analysed as ordered categorical data with mutually exclusive categories: ACR has four ordered categories (no response, ACR20, ACR50 and ACR70) and EULAR has three categories (no response, moderate response and good response). Data were analysed using a probit link function(4) a random effects model to allow for heterogeneity in treatment effects across studies and assuming an homogeneous variance. Inconsistency between direct and
indirect evidence was assessed using a node-splitting method. The reference treatment was defined as any single csDMARD (mostly MTX). This choice was approved by the clinical experts involved in the HAS economic evaluation of boDMARDs strategies. Meta-regression was performed to explore whether duration of disease was a treatment effect modifier. Absolute goodness-of-fit of the model was assessed using residual deviance. The analyses assumed that biosimilar treatments were not the same as their parent treatment. All the analyses were performed in OpenBUGS using the R package R2OpenBUGS. For each analysis, the first 180,000 iterations were discarded to allow for the number of iterations required for convergence to the target distribution, and 20,000 further iterations were used to estimate parameters. Convergence to the target distribution were checked using Gelman Rubin diagnostics. The most effective treatment was determined by the probability of being ranked as the best treatment, which considered the size of the treatment effect and its associated uncertainty.
RESULTS

Included trials

The electronic database search identified 34,621 records and bibliography searching identified an additional 18 records. Following title and abstract sifting, 128 studies were assessed for eligibility, and 82 studies excluded (see Supplementary Figure S1 PRISMA flow diagram(1) for reasons for exclusion). Of the 17 trials excluded for not reporting ACR or EULAR responses within 22-30 weeks follow-up: 12 were trials with randomised phases shorter than 22 weeks; in one RCT participants in either arm could receive additional treatment at clinician’s discretion after three months; two had unlicensed comparators; one reported EULAR Boolean 28 at 1 year follow-up, but did not report ACR or EULAR responses; and one had the primary endpoint of safety and measured efficacy as the secondary endpoints of DAS44 and the health assessment questionnaire.

There were 46 trials meeting the review inclusion criteria, comprising ten RCTs(7-16) with a MTX-naïve population, and 36 RCTs(17-52) with a csDMARD-experienced population. Included trials are shown in Table 1. There was a balance across trials in terms of population characteristics and trial quality. There was some variation between trials in disease duration (Table 1), with the MTX-naive population ranging from 5-166 weeks, and csDMARD-experienced ranging from 94-676 weeks. However, meta-regression suggested that disease duration was not a treatment effect modifier for ACR response, with the estimated coefficient of the disease duration being close to zero. The limited number of RCTs did not allow us to perform meta-regression including disease duration for EULAR response in the two selected populations. We found no evidence of selective outcome reporting in the included trials. There was a low risk of bias in terms of blinding and analyses (Supplementary Figure S2). The majority of RCTs were blinded (74%), and reported analyses with either intent-to-treat or modified intent-to-treat (that is, all randomised patients who received at least one dose of trial drug were included in the analyses, 87% of RCTs). There was a higher risk of bias regarding randomisation, with unclear reporting of sequence generation and allocation concealment (54% and 52% of RCTs respectively).

Network meta-analyses

ACR response data were provided by ten trials of the MTX-naïve population, and thirty-four trials of the csDMARD-experienced population (Table 1). EULAR response data were provided by two trials of the MTX-
naïve population and nineteen trials of the csDMARD-experienced population (Table 1). Network diagrams are shown in Supplementary Figure S3. Results are presented using medians and 95% credible intervals (CrI) from posterior distributions, treatment rankings. The probabilities of being the best treatment were calculated for each analysis. The models fitted the data well with the total residual deviance close to the total number of data points.

**ACR responses**

Figure 1 shows the NMA for ACR responses. Results shown are the effect of each treatment relative to csDMARD on the probit scale, with negative values representing positive treatment effects (i.e. a smaller proportion of patients in the lower ACR categories).

In the MTX-naïve population, all treatments, except ADA monotherapy, were associated with beneficial median treatment effects relative to csDMARD with the greatest effect being associated with MTX plus MP (effect on probit scale -0.79, 95% CrI: -1.44 to -0.20) and IFX+(-0.74, 95% CrI: -1.14, -0.40). However, the treatment effects were superior against csDMARDs only for ADA+, ETN+, IFX+, intensive cDMARDs (two or more csDMARDs) and MTX plus MP at a conventional 5% level (Figure 1). Intensive cDMARDs and boDMARDs had similar responses. The difference between intensive csDMARDs and boDMARDs was not significant (data not shown). MTX plus MP was most likely to be the most effective intervention (median rank 1; probability of being the best 0.51). The estimated between-trial standard deviation was 0.10 (95% CrI: 0.02 to 0.34).

In the csDMARD-experienced population, all treatments, except placebo, were associated with beneficial median treatment effect relative to csDMARD with the greatest effects being associated with bsDMARD ETN (SB4)+ (-1.12, 95% CrI: -1.70 to -0.55), boDMARD ETN+ (-0.03, 95% CrI: -1.34 to -0.73), TCZ+ (-1.08, 95% CrI: -1.43 to -0.73) and TCZ monotherapy (-1.08, 95% CrI: -1.39 to -0.76)). The treatment effects were superior compared to csDMARD for all interventions except for ADA (borderline non-statistically significant) at a conventional 5% level (Figure 1). The treatment that was most likely to be the best was the bsDMARD of ETN (SB4) combination therapy (median rank 2; probability of being the best 0.37). The estimated between-trial standard deviation was 0.22 (95% CrI: 0.12 to 0.35).

**EULAR responses**
Figure 2 shows the NMA for EULAR responses. Results shown are the effect of each intervention relative to csDMARDs on the probit scale, with negative values representing positive treatment.

Only two trials provided EULAR response data in the MTX-naïve population. The results showed that ADA+ and GOL+ were associated with beneficial treatment effects relative to csDMARD with the greatest effect being associated with ADA+ (-0.64, 95% CrI: -1.15 to -0.15). However, the treatment effect was superior only for ADA+ at a conventional 5% level (Figure 2). ADA+ was most likely to be the most effective intervention of the three interventions with EULAR response data (ADA+, GOL+ and csDMARD) (median rank 1; probability of being the best 0.84). The estimated between-trial standard deviation was 0.14 (95% CrI: 0.03 to 0.48). However, this was based on only two trials.

In the csDMARD-experienced population, all treatments, except placebo, were associated with beneficial median treatment effects relative to csDMARD with the greatest effect being associated with TCZ+(-1.56, 95% CrI: -2.21 to -1.01), TCZ monotherapy (-1.47, 95% CrI: -2.15 to -0.89), ETN+ (-1.34, 95% CrI: -2.55 to -0.14) and bsDMARD ETN (SB4)+ (-1.36, 95% CrI: -2.78 to 0.05). The treatment effects were superior only for ETN+, GOL+, IFX+, CTZ+, TCZ (with and without MTX) at a conventional 5% level (Figure 2). The effects of combination therapies of bsDMARDs were comparable to boDMARDs (data not shown, available from authors on request). TCZ+ was the treatment that was most likely to be the most effective intervention (median rank 2; probability of being the best 0.33) for EULAR responses. The estimated between-trial standard deviation was 0.34 (95% CrI: 0.15 to 0.53).

DISCUSSION

In this review of ten RCTs, we found that in the MTX-naïve population, MTX plus MP, or intensive csDMARDs (that is, two or more csDMARDs), were comparable to boDMARD treatment for ACR responses. Thirty-six RCTs contributed data to NMAs of the csDMARD-experienced population. For both ACR and EULAR responses in this population, the greatest effects were associated with combination therapy (with MTX) of bsDMARD ETN, boDMARD ETN and TCZ, as well as TCZ monotherapy. The effects of combination therapies of bsDMARDs were comparable to boDMARDs (data not shown, available from authors on request). The ongoing results of the French economic evaluation of biologic treatments sequences for moderately-to-severely active RA should
confirm whether bsDMARDs strategies instead of boDMARDs for a csDMARD-experienced patients are cost-effective option. However, concerning the MTX-naïve patients, the above results that showed comparable effect between boDMARDs and intensive cDMARDs at six months seem in line with the recent RA French guidelines which do not recommend to reimburse the use of biological treatments for this indication (except when MTX is contraindicated).

One of the strengths of our study was to adopt strict definitions of outcomes and populations that are both consistent with the therapeutic strategies of RA (as they are defined by EULAR and ACR) and the existing indications in the RA French management.

Our approach was different from other NMAs performed by HTA institutions (e.g. CADTH, 2018; (53) ICER, 2017) (54) or Cochrane of biologics (e.g. Singh et al, 2016; (55) Singh et al, 2017 (56) Hazlewood et al, 2016) (57) who included many outcomes (e.g. DAS28, remission, radiographic progression) but did not use EULAR criterion. Moreover, the analysed population were not often the same as the CADTH (2018) (53) and ICER (2017) (54) who focused only on clinical effectiveness for moderately-to-severely active RA for patients who had an inadequate response to prior csDMARDs.

Despite many differences in terms of exclusion criteria, inclusion of targeted synthetic DMARD (e.g. baricitinib and tofacitinib) and the number of analysed studies, our findings on ACR criterion (e.g. ACR50) comparing boDMARDs showed similar findings for moderately-to-severely active RA as the above NMAs: most of boDMARDs (and bsDMARDs) compared with cDMARDs (i.e. MTX) showed a clinical benefit but they did not often allow to detect a significant difference when compared with each other.

Concerning MTX-naïve patients, our results were partially similar with those of Singh et al (2017) (56) that showed that biologic with MTX were associated with statistically benefits in terms of achievement of ACR50: In our analysis, boDMARDs (IFX+MTX, ETN+MTX, ADA+MTX) as well as intensive cDMARDs and MTX+ MP were more effective than cDMARDs.

The review had limitations. The searches were conducted in 2017. There was limited evidence on EULAR response in MTX-naïve patients, with no bsDMARD RCTs meeting the inclusion criteria of the review, and no
data for MTX plus MP or intensive csDMARDs for EULAR response. The limited number of RCTs did not allow us to perform meta-regression including disease duration for EULAR response. Only three RCTs of bsDMARDs meeting the inclusion criteria were identified. As bsDMARDs were compared only to their boDMARD, their inclusion in the NMA did not affect the results for other inventions since they did not form a closed loop in the network. Ideally, evidence synthesis based on remission or low disease activity would be used as these are established treatment targets and routinely used for monitoring patients in European clinical practice. The EMA guidelines consider that remission should be the primary endpoint in clinical trials, and can be either defined according to EULAR criteria (DAS28<2.6), or in accordance with the more strict EULAR–ACR criteria (Boolean or Index-based).

Nevertheless, there are few data in the RCTs on this criterion to make relevant comparisons, and so we compared the treatments on the ACR criteria, and the EULAR response, instead as these scores represent a relative change from baseline.

Our review included trials of bsDMARDs, while the majority of previous reviews have included only anti-TNF biologics. Our review differed from other reviews of biologics in RA, in that i) it was limited to first-line biologics; ii) it considered MTX-naive and csDMARD-experienced trials separately; iii) it was limited to trials reporting outcomes at 22-30 weeks follow-up; iv) it considered ACR and EULAR as ordered categorical data, whereas the reviews by Hazlewood et al. (2016) (57) Singh et al, 2016, (55), Singh et al, 2017 (56) and CADTH (2018) (53) treated these outcomes as binary which ignores the natural ordering and correlation between categories. Treating patient responses as mutually exclusive categories enables a simultaneous analysis of the data, including studies that do not provide information about some categories, and a single estimate of treatment effect.

The finding that intensive csDMARDs were comparable to boDMARDs for MTX-naive patients (data not shown) also agreed with a previous review. (57) Previous meta-analyses underscored the dearth of direct evidence of effectiveness difference between biological agents. (59, 60) Our findings suggested that TCZ monotherapy was most favourable for csDMARD-experienced RA patients. This result was coherent with the conclusions of a previous review(61) where TCZ was either of comparable or superior efficacy to other boDMARDs. However, this finding might be explained by the fact that, for boDMARDs with a significant effect toward inhibition of acute phase reactants (APR), such as TCZ or the Janus kinase inhibitors, DAS28 may overestimate clinical
response due to the high weight of APR components in the DAS28 formula. (62) Similarly CRP level is a component of ACR response. It is commented that there may be silent residual inflammation in the joints even though CRP is low.

In clinical practice, decision making in patients with RA is not the same as in clinical trials. The choice of treatment is the result of a complex decision process that must take into account disease activity, and physician and patient characteristics. For triple intensive csDMARD therapy, despite the efficacy, the question of treatment adherence and persistence remain.

CONCLUSION

Our findings provide data for the short term effectiveness of boDMARDs and bsDMARDs for use in the current HAS economic evaluation of DMARD strategies for RA. Adverse events were not addressed, but were in HAS’s economic decision model.

For MTX-naïve patients with severe active RA, MTX plus MP or intensive csDMARDs, at six months, were comparable to boDMARDs, with all these treatments being superior to a single csDMARD. For csDMARD-experienced patients with moderate to severe active RA, bsDMARDs were comparable to their boDMARD. Combination therapy with all boDMARDs and bsDMARDs were superior to csDMARD treatment, with bsDMARD ETN, boDMARD ETN, and TCZ likely to be the most effective.

REFERENCES


38. Matsubara T, Inoue H, Iwashashi M, Yamazaki A, Takeuchi T. A multi-center, double-dummy, double-blind study of subcutaneous (SC) abatacept (ABA) compared with intravenous (IV) ABA in Japanese

http://ard.bmj.com/content/71/Suppl_3/197.1.full.pdf


