Advances in Rheumatology

RESEARCH

Open Access

Compared efficacy of rituximab, abatacept, and tocilizumab in patients with rheumatoid arthritis refractory to methotrexate or TNF inhibitors agents: a systematic review and network meta-analysis

Alisson Pugliesi[1†], Amanda Borges de Oliveira[2†], Ana Beatrice Oliveira[3], Ricardo Xavier[4], Licia Maria Henrique da Mota[5,6], Manoel Barros Bertolo[1], Miguel Angel Gonzalez-Gay[7], Gustavo Citera[8] and Luiz Sergio Fernandes de Carvalho[3,9,10*]

Abstract

Background Our aim was to compare the efficacy of rituximab, tocilizumab, and abatacept in individuals with rheumatoid arthritis (RA) refractory to treatments with MTX or TNFi agents.

Methods We searched 6 databases until January 2023 for phase 2–4 RCTs evaluating patients with RA refractory to MTX or TNFi therapy treated with rituximab, abatacept, and tocilizumab (intervention arm) compared to controls. Study data were independently assessed by two investigators. The primary outcome was considered as achieving ACR70 response.

Results The meta-analysis included 19 RCTs, with 7,835 patients and a mean study duration of 1.2 years. Hazard ratios for achieving an ACR70 response at six months were not different among the bDMARDs, however, we found high heterogeneity. Three factors showing a critical imbalance among the bDMARD classes were identified: baseline HAQ score, study duration, and frequency of TNFi treatment in control arm. Multivariate meta-regression adjusted to these three factors were conducted for the relative risk (RR) for ACR70. Thus, heterogeneity was attenuated ($I2 = 24\%$) and the explanatory power of the model increased ($R2 = 85\%$). In this model, rituximab did not modify the chance of achieving an ACR70 response compared to abatacept (RR = 1.773, 95%CI 0.113–10.21, $p = 0.765$). In contrast, abatacept was associated with RR = 2.217 (95%CI 1.554–3.161, $p < 0.001$) for ACR70 compared to tocilizumab.

Conclusion We found high heterogeneity among studies comparing rituximab, abatacept, and tocilizumab. On multivariate metaregressions, if the conditions of the RCTs were similar, we estimate that abatacept could increase the chance of reaching an ACR70 response by 2.2-fold compared to tocilizumab.

[†]Alisson Pugliesi and Amanda Borges de Oliveira: Both authors had a primary role in this study and should figure with equal importance.

*Correspondence:
Luiz Sergio Fernandes de Carvalho
lsergio@clarityhealth.com.br
Full list of author information is available at the end of the article

Key messages

Abatacept could increase the chance of reaching an ACR70 response by 2.2-fold compared to tocilizumab.

Keywords Rheumatoid arthritis, Rituximab, Abatacept, Tocilizumab, bDMARD, American college of rheumatology, Network meta-analysis, Meta-regression

Introduction

Successive paradigm shifts have marked the treatment of rheumatoid arthritis (RA) in the past two decades. Following the changing environment, clinical discussion around the choice of immunobiologicals is increasingly complex. Traditionally used as a first-line treatment after methotrexate (MTX) failure, therapy with tumor necrosis factor inhibitors (TNFi) improves the quality of life of most individuals with RA. However, at least one third of individuals do not respond to these agents [1, 2].

Treatments with rituximab (RTX), abatacept (ABA), and tocilizumab (TCZ) are available options [3–5]. These three drugs have demonstrated good efficacy compared to placebo, and similar or greater efficacy vs adalimumab in head-to-head clinical trials, however, there are no head-to-head studies comparing efficacy between them [19, 31]. Such comparisons, unfortunately, will probably never be carried out. In these scenarios, network meta-analysis (NMA) could allow an approach to shed light on this uncertain data.

There are three network meta-analyses [6–8] available comparing drugs used in the treatment of patients with RA refractory to TNFi or MTX therapies. Despite pointing to a similar effectiveness among drugs, it is important to highlight that few data were available at the time of publication of these studies, with their low statistical power suggesting a risk of type II error greater than 30%. Other meta-analyses of randomized or observational studies [9, 10] presented a low quality and/or highly heterogeneous data. The essential problem with not having a conclusive study demonstrating an equivalence among drugs is that public policies based on cost minimization can be arbitrarily implemented.

Thus, the absence of head-to-head trials between RTX, ABA and TCZ and the low statistical power of previous NMA make a new assessment of published studies and a new study analyzing the combination of the efficacy of such drugs essential. It is also important to understand the existing heterogeneity among trials, finding ways to mitigate it to allow an adequate comparison among treatments.

In this sense, the aim of this study is to perform a systematic review seeking to evaluate randomized clinical trials (RCTs) on the efficacy of RTX, TCZ, and ABA in individuals with RA refractory to MTX or TNFi therapies.

Methods

Study design

This systematic review used a NMA to investigate efficacy of RTX, TCZ or ABA in patients with RA in RCTs. The review was registered in PROSPERO under number CRD42020167953. This study followed PRISMA statement [11].

Eligibility criteria

Inclusion criteria were restricted to RCTs phase 2–4 designed to evaluate treatment response in individuals with RA refractory to treatment with TNFi agents or MTX. Patients who did not achieve at least an ACR20 response were considered refractory. All RCTs should assess the efficacy of treatments and include a RTX, ABA, or TCZ intervention arm and a placebo arm treated with or without MTX and/or TNFi agents (control arm). Only RCTs that used the ACR response criterion (20, 50 and/or 70) and whose exposure to therapy (either intervention or control) was at least 24 weeks were considered eligible. Just articles in English and that were published in full were selected.

We excluded literature reviews, editorials, conference records, congress annals or abstracts, in vitro and animal model studies, qualitative studies, cross-sectional studies, case–control studies, systematic reviews, studies that evidenced alternative treatments using antibiotics, studies with an exposure time of less than 24 weeks, researches that considered patients with other types of autoimmune diseases, investigations of the use of TCZ, RTX, or ABA to treat other diseases, studies that included patients with comorbidities (pregnancy, breastfeeding, histoplasmosis, coccidioidomycosis, HIV, tuberculosis etc.), and analyses of individuals with hypersensitivity to RTX, ABA, or TCZ.

Search strategy

To identify all RCTs that assess the relationship between the efficacy of RTX, TCZ, and ABA in individuals with RA refractory to treatment with TNFi agents or MTX, searches for original articles were conducted in the MedLine (PubMed), Cochrane Library, Embase, Web of Science, Scopus, and Latin American

and Caribbean Literature in Health Sciences (LILACS) databases until January 12, 2023.

The search strategy used is detailed in the Additional file 1: Table S1. The term search was focused on titles, keywords, and abstracts.

Study selection and data extraction

The first study selection stage consisted of reading the titles, further encompassing abstracts and keywords. Titles and abstracts of articles identified by the search strategy were independently assessed by two authors of the present study. In the second stage, the investigators independently evaluated the full text of the articles and made their selections according to the pre-specified eligibility criteria. The Rayyan QCRI application (RayyanSystems Inc., MA,EUA) was used to manage duplicate files.

In the second stage, the investigators independently evaluated the full text of the articles and made their selections according to the pre-specified eligibility criteria. Primary and secondary outcomes were assessed. The primary outcome was an ACR70 response according to the American College of Rheumatology (ACR) criteria, i.e., at least 70% improvement in the number of swollen and painful joints and improvement in three of these five parameters: physician's global assessment of the disease, patient's global assessment of the disease, patient's assessment of pain, C-reactive protein or erythrocyte sedimentation rate, and Health Assessment Questionnaire (HAQ) score. The ACR score measures the rheumatic disease activity and is used to measure differences in relation to the baseline, showing a good relationship with the quality of life and functionality [12]. The ACR20, ACR50, and ACR70 responses refer to an improvement of 20%, 50%, and 70% in the ACR score, respectively.

The ACR70 response was chosen as the primary outcome due to its high specificity in relation to the functionality and quality of life improvement [13]. The secondary outcomes were the ACR50 and ACR20 response criteria compared among treatments to identify differences in the ACR response pattern.

Data extracted from studies were inserted, organized, and standardized in a Microsoft Excel® (Microsoft Corporation, One Microsoft Way Redmond, Washington, EUA) spreadsheet.

Methodological quality and risk of bias

The methodological quality and risk of bias evaluation was carried out independently by two investigators and disagreements were solved by consensus. The certainty of evidence for each study was assessed using the *Grading of Recommendations Assessment, Development and Evaluation* (GRADE) system [14]. The risk of bias was assessed using the Cochrane Collaboration tool, which considers six dimensions: random sequence generation, allocation concealment, blinding of participants and professionals, blinding of outcome assessment, incomplete outcome data, and selective outcome reporting [15].

Assessing consistency and heterogeneity

Aiming to explore the log Hazard Ratio(logHR) association to achieve an ACR70 response, a meta-regression analysis was performed, wherein the ABA treatment effects depended on the study follow-up, on the presence of TNFi therapies in the control arm, and on the mean baseline HAQ score [16]. Sensitivity analyses were performed by excluding the trials with the longest exposure duration (in patient-years) by drug class.

Statistical analysis

A random effects model was used for the Bayesian network meta-analysis (bNMA) [16]. Means and standard deviations were used to represent variables within the RCTs. All outcomes were expressed as hazard ratios (HRs) and corresponding 95% confidence intervals (CIs). The analysis was performed using the Markov chain Monte Carlo methods. Three chains were suitable, yielding 5,000 iterations (20,000 per chain) and giving rise to the subsequent distributions of the model parameters. The Brooks-Gelman-Rubin method [17] was used for the convergence diagnostics. The model goodness of fit was assessed through the residual deviance. The I^2 statistic was used to investigate the statistical heterogeneity and a comparison-adjusted funnel plot was used to identify any publication bias [18]. The 'node splitting' approach was adopted to measure the existing degree of inconsistency [18].

Statistical analyses were performed using the R Studio software, version 1.1.4 (R v.4.0.1 and R_Studio v.1.1.463, Auckland, NZ) [19].

Results

Using the afore mentioned search terms and platforms, we identified 600 citations (details in Additional file 1 section). After excluding duplicates, studies not including TCZ, ABA, or RTX therapies and for being observational studies, we found 19 trials to be included for qualitative synthesis and meta-analysis. The PRISMA flowchart used for selecting studies can be found in Additional file 1: Fig. S1.

For our primary outcome meta-analysis, we included 19 RCTs with 7,835 patients (9,402 patient-years) randomized to the intervention arm or to the control arm. The mean age of patients included in the present study was 52.3 years; 77.2% of these patients were women; the mean disease duration was 8.7 years; and the mean

follow-up was 1.2 years. Details of the basic trial characteristics are presented in Additional file 1: Table S2.

Among the 19 selected trials, 31.5% (six) included TCZ as intervention arm, 31.5% (six) included RTX as active arm, and 31.5% (six) included ABA as intervention arm. Only one study compared ABA with RTX. Additional file 1: Fig. S2 shows the network of direct comparisons among the treatment arms of included studies.

All studies had a low risk of bias according to the Cochrane Collaboration tool (Additional file 1: Table S3) and were of high quality according to the GRADE system (Additional file 1: Table S4). As shown in Additional file 1: Figure S3A-S3C, there was no significant publication bias in funnel charts and no significant small-study bias according to the Egger tests.

ACR70

In the meta-analysis of random effects for the chance of achieving an ACR70 response at six months, six studies [20–25] with ABA therapy were selected. In these studies, 437 (25.6%) patients in the intervention arm and 250 (19.8%) patients in the control arm achieved an ACR70 response at six months and the observed HR was 1.35 (95% CI 1.17–1.55 *p*: 0.013, $I^2 = 41\%$ *p* for heterogeneity = 0.013). For RTX, six studies were selected [26–31], where 350 (23.5%) patients in the intervention arm and 107 (12%) patients in the control arm achieved an ACR70 response at six months, indicating a HR of 2.43 (95% CI 1.99–2.96 $p < 0.001$, $I^2 = 64\%$ *p* for heterogeneity < 0.001). For TCZ, six studies were selected [32–37], with 294 (18.5%) patients in the intervention arm and 97 (11%) patients in the control arm achieving an ACR70 response at six months, with a HR of 1.53 (95% CI 1.24–1.89 $p = 0.002$, $I^2 = 76\%$ *p* for heterogeneity < 0.001) (Additional file 1: Fig. S4A).

As shown in Table 1 and, direct and indirect comparisons suggest a superiority of ABA, TCZ, or RTX vs. placebo, with TCZ showing a slightly lower magnitude of effect (HR 2.765, 95% CI 1.240–6.692 $p = 0.009$) compared to ABA (HR 3.423, 95% CI 1.422– 8.709 $p < 0.001$) and to RTX (HR 3.494, 95% CI 1.530–8.658 $p < 0.001$) in achieving ACR70 response at six months. There was no statistical difference between these drugs and the TNFi therapy, considering 4 studies of ABA vs TNFi (HR 1.043 (95% CI 0.422, 2.242), $p = 0.59$), 1 study of TCZ vs TNFi (HR 0.839 (95% CI 0.373, 3.804), $p = 0.45$) and 1 study of RTX vs TNFi (HR 1.059 (95% CI 0.339, 2.581), $p = 0.51$). For all analyses, there was a slight inconsistency between direct and indirect measures (Table 1, Indirect comparisons among bDMARDs), but they showed a substantial heterogeneity among RCTs (> 40%) (Additional file 1: Fig. S4A).

ACR50

For the chances of achieving an ACR50 response, we selected the same studies as selected for the ACR70 response in six months. For ABA, 730 (43%) patients in the intervention arm and 439 (34.7%) patients in the placebo arm were involved, with a HR of 1.28 (95% CI 1.17–1.41 $p \leq 0.001$, $I^2 = 54\%$ *p* for heterogeneity = 0.005). For

Table 1 Indirect comparisons among bDMARDs in active rheumatoid arthritis expressed in hazard ratios for achieving 70% of American College of Rheumatology response at six months with sensitive analyses

	ACR70 response rate, HR (95% CI), *p* value	
	All trials	**Excluding one trial from each class***
Abatacept versus Rituximab	1.017 (0.373, 2.845), $p = 0.84$	1.029 (0.324, 2.959), $p = 0.88$
Tocilizumab versus Rituximab	0.791 (0.254, 2.461), $p = 0.32$	0.811 (0.224, 2.618), $p = 0.40$
Tocilizumab versus Abatacept	0.806 (0.263, 2.505), $p = 0.37$	0.814 (0.237, 2.614), $p = 0.43$
Abatacept versus TNFi	1.043 (0.422, 2.242), $p = 0.59$	1.063 (0.402, 2.492), $p = 0.56$
Tocilizumab versus TNFi	0.839 (0.373, 3.804), $p = 0.45$	0.814 (0.331, 4.007), $p = 0.41$
Rituximab versus TNFi	1.059 (0.339, 2.581), $p = 0.51$	1.042 (0.305, 2.886), $p = 0.67$
Abatacept versus placebo	3.423 (1.422, 8.709), $p < 0.001$	3.438 (1.413, 8.762), $p < 0.001$
Tocilizumab versus placebo	2.765 (1.240, 6.692), $p = 0.009$	2.525 (1.096, 6.921), $p = 0.015$
Rituximab versus placebo	3.494 (1.530, 8.658), $p < 0.001$	3.509 (1.522, 8.676), $p < 0.001$

Random effects standard deviation for all trials 0.854 (0.537, 1.408), Inconsistency factor (distance between direct and indirect effects) are 0.021 (− 2.661, 2.607, $p = 0.988$), 0.032 (− 2.592, 2.557, $p = 0.934$), 0.038 (− 2.711, 2.557, $p = 0.984$), − 1.247 (− 3.628, 1.148 $p = 0.271$), − 1.028 (− 3.924, 1.018 p = 0.176), 0.150 (− 1.764, 2.109, $p = 0.879$), 0.714 (− 1.353, 2.751, $p = 0.468$), − 1.278 (− 3.576, 1.094, $p = 0.248$), 0.666 (− 1.304, 2.635, $p = 0.476$), respectively for the node-split comparisons Abatacept versus Rituximab, Tocilizumab versus Rituximab, Tocilizumab versus Abatacept, Abatacept versus TNFi, Tocilizumab versus TNFi, Rituximab versus TNFi, Abatacept versus placebo, Tocilizumab versus placebo, Rituximab vs placebo. For ACR70 network meta-analysis model fit statistics posterior mean of the residual deviance ($D_{res} = 40.1$), and deviance information criterion (DIC) = 77.1

TNFi tumor necrosis factor inhibitors; *HR* hazard ratio; *ACR70* 70% of American College of Rheumatology

* Sensitivity analysis: excluding trials with largest exposure (in patient-years) from each class

RTX, 586 (39.4%) patients in the intervention arm and 201 (22.6%) patients in the control arm were involved, with a HR of 1.94 (95% CI 1.70–2.20 $p \leq 0.001$, $I^2 = 67\%$ p for heterogeneity ≤ 0.001). For TCZ, 537 (34%) patients in the intervention arm and 178 (20.3%) patients in the control arm were involved, with a HR of 1.75 (95% CI 1.52–2.02 $p \leq 0.001$, $I^2 = 79\%$ p for heterogeneity ≤ 0.001) (Additional file 1: Fig. S4B).

Indirect comparisons among the intervention arms showed that ABA vs. RTX present a HR = 1.052 (95% CI 0.514–2.131, $p = 0.986$), TCZ vs. RTX present a HR = 0.833 (95% CI 0.378–1.835 $p = 0.93$), and TCZ vs. ABA present a HR = 0.879 (95% CI 0.393–1.961 $p = 0.91$) for achieving ACR50 in six months.

ACR20

Finally, for the chances of achieving an ACR20 response, 1,108 (65%) patients in the intervention arm and 644 (51%) patients in the placebo arm were included for ABA, with a HR of 1.63 (95% CI 1.53–1.75, $p \leq 0.001$, $I^2 = 84\%$, p for heterogeneity ≤ 0.001). For RTX, 880 (59.2%) patients in the intervention arm and 343 (38.5%) patients in the control arm were involved, with a HR of 1.75 (95% CI 1.61–1.91 p ≤ 0.001, $I^2 = 71\%$ p for heterogeneity ≤ 0.001). For TCZ, 843 (53.2%) patients in the intervention arm and 292 (33.4%) patients in the control arm were involved, with a HR of 1.90 (95% CI 1.73–2.10 $p \leq 0.001$, $I^2 = 85\%$ p for heterogeneity ≤ 0.001) (Additional file 1: Fig. S4C).

When indirectly compared to each other, ABA vs. RTX show a HR of 1.013 (95% CI 0.669–1.535 $p = 0.890$) and TCZ vs. ABA show a HR of 1.014 (95% CI 0.626–1.618 $p = 0.92$).

Explaining the heterogeneity among RCTs

The spider chart shown in Fig. 1 presents an important imbalance among studies in the frequency of TNFi treatment in the control arm. RCTs with TCZ less often included TNFi agents in the control arm: ABA, 50%; RTX, 43%; and TCZ, 17% ($p = 0.048$). Besides, the baseline HAQ score was significantly lower (ABA 1.68 ± 0.15, RTX 1.68 ± 0.23, and TCZ 1.50 ± 0.17; $p = 0.049$). The mean follow-up time (total study duration) among RCTs was significantly longer in trials with ABA (25 ± 18.75 months) compared to those with RTX (9.43 ± 3.21) and TCZ (6.00 ± 0) ($p = 0.015$). Baseline 28-joint Disease Activity Score (DAS28) also showed an imbalance among trials, but this variable was excluded from the meta-regressions for presenting a high collinearity with the baseline HAQ score (Additional file 1: Fig. S5A).

Meta-regressions

Meta-regressions were conducted to assess the impact of treatments on the ACR70 response in a manner adjusted to the baseline HAQ score, study follow-up time (total study duration), and frequency of TNFi treatment in the control arm (Additional file 1: Fig. S5B). These 3 explanatory variables were found after exploring all potential predictors of ACR70 response such as percentage of women, glucocorticoid use, patients' age, MTX dose, duration of RA, baseline HAQ, baseline DAS28, study follow-up time and frequency of TNFi treatment in the control arm.

These explanatory variables were chosen in a forward stepwise process, and all the 3 variables were associated with the ACR70 response in univariate regressions and were imbalanced among treatment arms. As shown in Table 2, each 0.1 point over the mean baseline HAQ in each study was associated with a RR of 2.043 (95%CI 1.032–14.44, p for difference = 0.028, $R^2 = 13\%$, $I^2 = 89\%$, p for heterogeneity ≤ 0.001) for achieving an ACR70 response (Additional file 1: Fig. S4B). Additionally, the presence of an TNFi therapy in the control arm of any RCT was associated with a RR of 0.316 (95% CI 0.134–0.746, p for difference = 0.009, $R^2 = 27\%$, $I^2 = 85\%$, p for heterogeneity ≤ 0.001) (Additional file 1: Fig. S4A) and each additional month of follow-up was associated with a RR of 0.980 (95% CI 0.961–0.996, p for difference = 0.048, $R^2 = 9\%$, $I^2 = 71\%$, p for heterogeneity ≤ 0.001) for achieving an ACR70 response.

In a multivariate model with the ACR70 response as an outcome and the covariates treatments in the intervention arm (RTX vs. ABA and TCZ vs. ABA), baseline HAQ score, follow-up time, and frequency of TNFi treatment in the control arm (Table 2), the heterogeneity was reduced ($I^2 = 24\%$, p for heterogeneity = 0.27) and the explanatory power of the model increased ($R^2 = 85\%$). In this model, ABA did not modify the chance of achieving an ACR70 response when compared to RTX (1.773, 95% CI 0.113–10.21, $p = 0.765$). In contrast, ABA was associated with a RR of 2.217 (95% CI 1.554–3.161, p for difference ≤ 0.001) when compared to TCZ for achieving an ACR70 response. It means that, if the conditions of the RCTs were similar, ABA could increase the chance of achieving an ACR70 response by 2.22-fold when compared to TCZ. In this model, the meta-regression of RTX vs TCZ was not included in Model 4 as this covariate showed high collinearity with other covariates (variance inflation factor of 9 with ABA vs TCZ).

Sensitivity analyses

Sensitivity analyses were performed considering the exclusion of the RCTs with the longest exposure

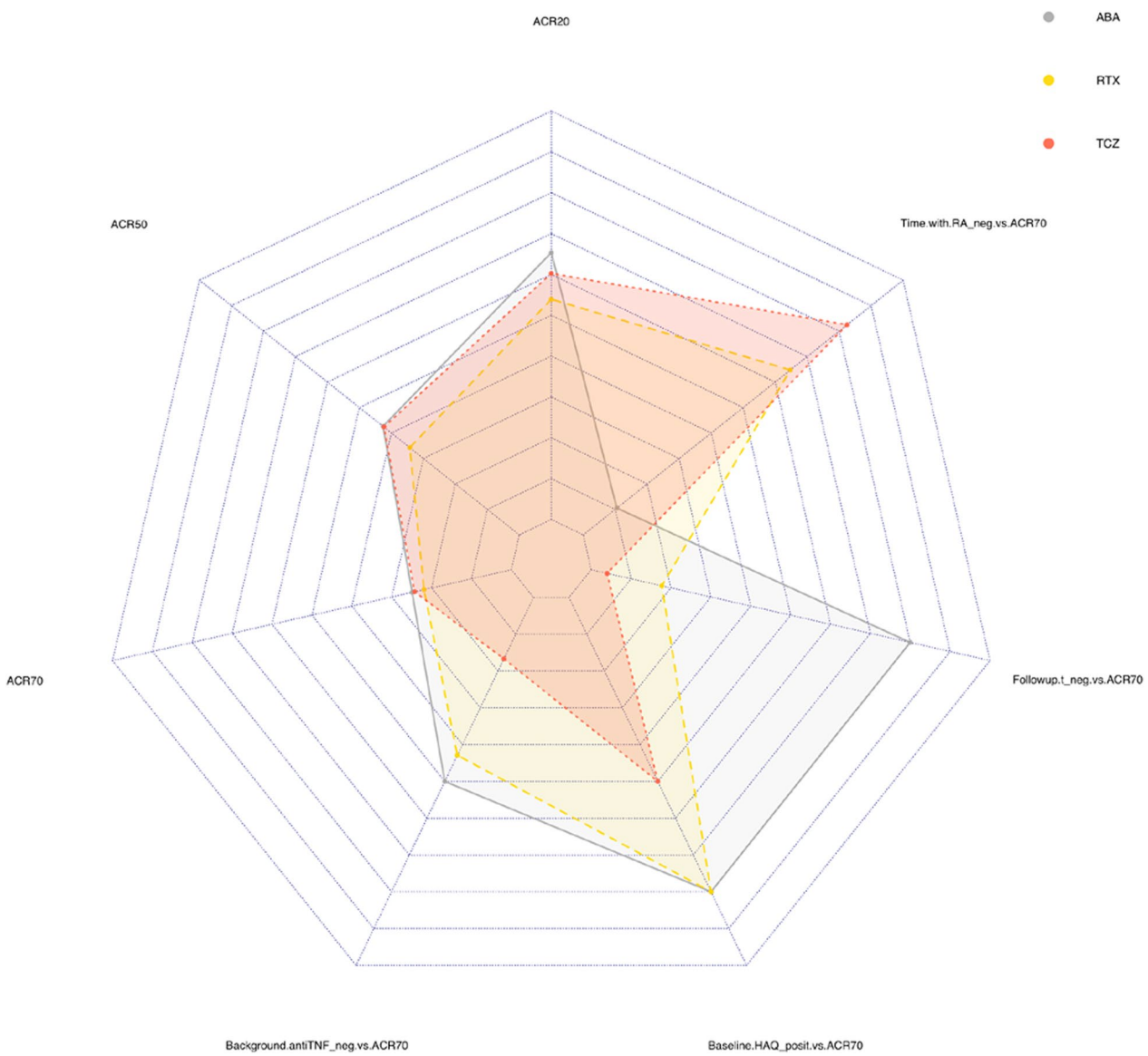

Fig. 1 The Spider chart describes the imbalance among RCTs characteristics included in the meta-analysis according to the non-TNFi treatment used in active arm (ABA, RTX or TCZ). The suffix "_posit.vs.ACR70" represent the independent variables that show positive association with ACR70 response, and the suffix "_neg.vs.ACR70" represent the independent variables that show negative association with ACR70 response. Example: the higher the follow-up time (suffix "_neg.vs. ACR70"), the lower the chance of achieving ACR70 response. Followup.t: follow-up time (total study duration); ACR70: 70% achieving ACR response; ACR50: 50% achieving ACR response; ACR20: 20% achieving ACR response. Background TNFi: percentage of individuals on tumor necrosis factor (TNF) inhibitors; HAQ: Health Assessment Questionnaire score; ABA: abatacept; RTX: rituximab; TCZ: tocilizumab; ACR: American College of Rheumatologists

durations (highest number of patient-years), one study being excluded for each class of non-TNFi biological drug [20, 25, 33]. As shown in Table 1, the exclusion of such studies did not modify the results.

Discussion

The present NMA showed no significant differences between tocilizumab, abatacept and rituximab to achieve ACR70 response in six months in patients with RA

Table 2 Meta-regression models for achieving 70% of American College of Rheumatology response at total of study duration (mean = 14 months) as dependent variable

	RR	95% CI		*p*
		Lower bound	Upper bound	
Model 1				
Mean baseline HAQ (each additional 0.1 point)	2.0433	1.0328	14.4460	0.028
Model 2				
Background TNFi in control arm (yes vs no)	0.3166	0.1345	07460	0.009
Model 3				
Follow-up time (each additional 1 month)	0.9714	0.9522	0.9920	0.007
Model 4				
Mean baseline HAQ (each additional 0.1 point)	2.0332	1.0141	14.6164	0.045
Background TNFi in control arm (yes vs. no)	0.2187	0.1341	0.3567	< 0.001
Follow-up time (each additional 1 month)	0.9763	0.9504	0.9899	0.016
ABA versus RTX	1.7736	0.1134	10.2165	0.625
ABA versus TCZ	2.2171	1.5541	3.1614	< 0.001

RR relative risk; *HAQ* Health Assessment Questionnaire score; *TNFi* tumor necrosis factor inhibitors; *ABA* abatacept; *RTX* rituximab; *TCZ* tocilizumab; *ACR70* 70% of American College of Rheumatology

refractory to methotrexate or TNFi agents, but these findings yielded high heterogeneity. Heterogeneity was mostly explained by 3 factors: baseline HAQ score, study duration, and frequency of TNFi treatment in control arm. To control these unbalanced characteristics, multivariate meta-regression adjusted to these 3 factors were performed and showed that ABA was associated with RR of 2.22 (95% CI 1.55–3.16, $p < 0.001$) for ACR70 compared to TCZ, with attenuated heterogeneity and high explanatory power.

In meta-regressions, each increment of 0.1 point in the baseline HAQ score was associated with a twofold greater chance of achieving an ACR70 response, while such chance was reduced by 68% with the use of an TNFi therapy in the control arm and by 2% with each additional follow-up month. These findings are data-driven rather than hypothesis-driven. However, we believe that RCTs that select individuals with better quality of life at baseline (despite uncontrolled RA) may have unmeasured factors (biological, psychological, social) that may increase their response of TCZ, ABA or RTX. In the case of TNFi users, it is more obvious to think that marginal gain with a second bDMARDs will be smaller than individuals not taking TNFi. Finally, the longer the exposure time to bDMARDs after 6 months may increase the chance to identify new individuals with disease flares or therapy failures.

As shown in the spider chart, ABA trials had a significantly longer duration compared to RTX and TCZ trials, while RCTs with ABA and RTX included TNFi agents in the control arm more frequently than the TCZ ones. Both these characteristics reduce the propensity of trials with ABA to drive an ACR70 response. Thus, in multivariate meta-regressions compensating for the effect of unbalanced characteristics, it is suggested that the use of ABA increases the chance of achieving an ACR70 response compared to the use of TCZ. However, it is important to understand that these are hypothesis-generating findings that need to be confirmed in a specific RCT.

The findings of our NMA comply with that of three previous meta-analyses[7–9]. Largely, comparisons among treatments were focused on the effectiveness of TNFi vs. non-TNFi agents. On one hand, two of these studies showed a statistical power (1 – beta) > 80% for comparisons between TNFi and non-TNFi agents. On the other hand, comparisons among the three non-TNFi agents showed a statistical power < 65%. The present study reached a post-hoc statistical power for non-inferiority of 86.3% for the RTX vs. ABA comparison and of 99.2% for the TCZ vs. ABA comparison on the ACR70 response.

However, the presented result conflicts with that of observational studies, such as the study by Gottenberg et al.[38], where authors compared the effectiveness and safety of RTX, ABA, and TCZ in a population of patients with RA refractory to treatment with TNFi agents. The primary outcome of such study consisted of drug retention without the occurrence of death from any cause, discontinuation of the drug studied, initiation of a new biologic or a combination of conventional disease-modifying antirheumatic drugs or increase in the dose of oral corticosteroids at more than 10 mg a day at two consecutive visits. Secondary endpoints included European League Against Rheumatism (EULAR) response

at months 6, 12, and 24. A good EULAR response was defined as a decrease in disease activity score in 28 joints— erythrocyte sedimentation rate (DAS28-ESR) more than 1.2 points and resulting score 3.2 or less. A moderate EULAR response was defined as a decrease in DAS28-ESR more than 0.6 points and resulting score 5.1 or less [38]. Among the main findings, RTX and TCZ had similar results, both being more effective in controlling the disease than ABA. Despite the evident reduction in the primary outcome and the better European League Against Rheumatism (EULAR) response with RTX and TCZ compared to ABA, there was no difference in terms of major outcomes, such as deaths, serious infections, cardiovascular events, and cancer. Besides the fact that it is an observational study that used massive data imputation [39], some considerations about this research may justify the differences found compared to our findings: High rates of loss to follow-up in patients on Tocilizumab; Different frequency of administration between RTX (at least six months) and ABA and TCZ (monthly); Use of response metrics (DAS28-ESR and EULAR response) that give great weight to laboratory markers of inflammatory activity and therefore favor TCZ.

The comparison among non-TNFi agents is essential for choosing the ideal anti-inflammatory strategy. Additionally, it contributes to cost-effectiveness analyses establishing which drug could bring the best balance between gains in quality of life and morbidity reduction in relation to the costs of therapies [40]. However, when therapies have a similar efficacy, it is worth minimizing the risk of false negative results [41]. Through the present study, it is possible to conclude that ABA, TCZ, and RTX have a similar efficacy when accounting solely the effects of co-treatments, but when accounting the full spectrum of unbalanced factors across RCT arms, our findings suggest that ABA increases the chance of achieving an ACR70 response in comparison to TCZ.

This study further exemplifies why it is essential to comprehensively understand the heterogeneity among studies. Though meta-regression results should be understood as an hypothesis-generating evidence, they enable testing a scenario where components of heterogeneity among RCTs become more balanced.

Limitations

Some important limitations must be emphasized. During the extraction of information from the studies, some data were not available for all trials, and the analysis of only English published trials can generate a systematic data collection bias. In general, unpublished studies and observational studies could have a different impact on the results; however, there is great heterogeneity in terms of selection and statistical treatment in observational studies, which could hinder the data analysis.

In contrast, there are two important issues to be considered as limitations to the interpretation of these results. Firstly, there is a high degree of heterogeneity among studies for all drugs. The high variability in the relative treatment effects threatens the external validity of the study evidence and limits its generalization [42]. Secondly, the NMA does not compensate for other factors unrelated to the drugs that make up the treatment arms [43]. When adjusted to the baseline quality of life, follow-up time, and presence of TNFi therapies in the control arm, meta-regressions showed that the use of ABA increased the chance of achieving an ACR70 response by 2.2-fold compared to the use of TCZ, with a low heterogeneity and a high explanatory power ($R^2 > 80\%$).

We have not evaluated safety outcomes. The argument for evaluating only efficacy outcome is that there was no signal a priori for safety concerns or critical differences in safety among TCZ, ABA and RTX that could be novel, considering prior meta-analyses. It is important to note that the observed frequency of severe adverse events in the included RCTs would not allow enough statistical power for safety outcomes.

Another limitation of our study was our inability to ascertain the reasons for TNFi agent or MTX failure and the number of prior TNFi therapy failures. Patients may have been treated with different doses or for different TNFi treatment durations before considering an IR. Another potential limitation is the number of comparisons made for this analysis, which may favor spurious associations. However, the statistical power achieved in this study suggests that chances of false negative results are extremely low.

Additionally, we did not evaluate the impact of different dosing or posology used across RCTs to improve the clustering. Yet, considering the broad inclusion criteria, which resulted in a study population that best resembles real patients, and the choice of viable interventions, in the case of clinical trials, the results of this study have external validity with applicability in daily practice.

Conclusion

In this study, a systematic review was carried out using a NMA to compare the efficacy of RTX, TCZ, and ABA in individuals with RA refractory to treatment with TNFi agents or MTX. The NMA showed no significant difference among the studied drugs in achieving an ACR70 response, therefore, they have a similar effectiveness, with low inconsistency, but high heterogeneity among studies. Based on the result of multivariate meta-regressions, by mathematically equalizing the conditions of the RCTs, we estimate that ABA could

increase the chance of reaching an ACR70 response by 2.2-fold compared to TCZ. With the present results, it is advisable to evaluate the introduction ABA or RTX before TCZ in refractory RA. However, novel head-to-head clinical trials are still needed to confirm these findings.

Abbreviations

ABA	Abatacept
ACR	American college of rheumatology
bNMA	Bayesian network meta-analysis
Cis	Confidence intervals
GRADE	Grading of recommendations assessment: development and evaluation
HAQ	Health assessment questionnaire
HRs	Hazard ratios
LILACS	Latin American and Caribbean literature in health sciences
MTX	Methotrexate
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PROSPERO	Prospective register of systematic reviews
RA	Rheumatoid arthritis
RCTs	Randomized clinical trials
RR	Relative risk
RTX	Rituximab
TCZ	Tocilizumab
TNF	Tumor necrosis factor

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42358-023-00298-z.

Additional file 1. Supplemental Results and Methods.

Acknowledgements
We acknowledge all support received by Brazilian Society of Rheumatology.

Author contributions
Concept and design: LSFC, ABO, AP. Acquisition of data: ABO, AP. Analysis and interpretations of data: LSFC, MAGG, MBB, RX, GC. Drafting of the manuscript: LSFC, ABO, AP. Critical revision of the paper for important intellectual content: AP, MAGG, MBB, RX, GC, LMHM. Statistical analysis: LSFC, ABO. Provision of study materials: MAGG, LMHM, MBB, RX, GC. Obtaining funding: LSFC, MAGG. Administrative, technical, or logistic support: MAGG, LMHM, MBB, RX, GC. Supervision: LSFC, MAGG.

Funding
This work was supported by grants 437413/2018–7 from the Brazilian National Research Council (CNPq) and 2019/09068–3 from São Paulo State Foundation for Research (FAPESP).

Availability of data and materials
Full systematic review data and codes in R for data processing and meta-analysis are available under request.

Declarations

Ethical approval and consent to participate
This systematic review and meta-analysis was registered in PROSPERO (CRD42020167953). Ethical Approval and consent to participate were not included in the protocol since all data.

Consent for publication
All authors reviewed the last version of this manuscript and.

Competing interests
AP: Received speakes fees from AbbVie, Pfizer, Roche, Lilly, Novartis, UCB, Boehringer Ingelheim, Janssen. Research Grants from BMS. ABO: None. ABO: None. RX: Speaker/consultant and research grants from: Abbvie, Pfizer, Jansen, Lilly, Novartis. LMHM: Has received personal or institutional support from AbbVie, Janssen, Pfizer and Roche; has delivered speeches at events related to this work and sponsored by AbbVie, Janssen, Pfizer, Roche and UCB. MBB: Has participated in clinical and/or experimental studies related this work and sponsored by Roche; has delivered speeches at events related to this work and sponsored by AbbVie and Pfizer. MAGG: has received research support from AbbVie, MSD, Jansen and Roche, and has received consultation fees or has participated in a company-sponsored speaker´s bureau from AbbVie, Pfizer, Roche, Sanofi, Celgene, Sobi and MSD. GC: received speaker fees from AbbVie, Pfizer, BMS, Roche, Lilly, and Novartis; LSFC: Speaker/Consultant fees from Abbott, American Heart Association, Roche, Amgen, NovoNordisk, Libbs. Research grants from AstraZeneca and Amgen.

Author details
[1]Hospital de Clínicas, Faculty of Medical Sciences, State University of Campinas (UNICAMP), Campinas, SP, Brazil. [2]Neurosurgery Unit of the Hospital de Base Do Distrito Federal Base, Brasília, DF, Brazil. [3]Health Sciences Research and Teaching Foundation of Brasília (ESCS/FEPECS), Brasília, DF, Brazil. [4]Rheumatology Service, Federal University of Rio Grande Do Sul, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil. [5]Hospital Universitário de Brasília, University of Brasília (HUB-UnB-EBSERH), Brasília, DF, Brazil. [6]Medical Sciences at Faculty of Medicine, University of Brasília (UnB), Brasília, DF, Brazil. [7]Division of Rheumatology, Hospital Universitário Marqués de Valdecilla, IDIVAL, Universidad de Cantabria, Santander, Spain. [8]Section of Rheumatology, Instituto de Rheabilitacion Psicofisica, Buenos Aires, Argentina. [9]Laboratory of Data for Quality of Care and Outcomes Research (LaDaQCOR), Catholic University of Brasília (UCB), QS 07, Lote 01, Taguatinga Sul-Taguatinga, Brasília, DF 71966-700, Brazil. [10]Medical Data Analyst, Data Lab, Clarity Healthcare Intelligence, Jundiai, SP, Brazil.

Received: 5 April 2022 Accepted: 10 April 2023
Published online: 06 July 2023

References

1. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. Lancet. 2010;376:1094–108.
2. Hyrich KL, Lunt M, Watson KD, Symmons DP, Silman AJ, British Society for Rheumatology Biologics Register. Outcomes after switching from one anti-tumor necrosis factor alpha agent to a second anti-tumor necrosis factor alpha agent in patients with rheumatoid arthritis: results from a large UK national cohort study. Arthritis Rheum. 2007;56:13–20.
3. Kavanaugh A, Rosengren S, Lee SJ, Hammaker D, Firestein GS, Kalunian K, et al. Assessment of rituximab's immunomodulatory synovial effects (ARISE trial). 1: clinical and synovial biomarker results. Ann Rheum Dis. 2008;67:402–8.
4. Moreland LW, Alten R, Van den Bosch F, Appelboom T, Leon M, Emery P, et al. Costimulatory blockade in patients with rheumatoid arthritis: a pilot, dose-finding, double-blind, placebo-controlled clinical trial evaluating CTLA-4Ig and LEA29Y eighty-five days after the first infusion. Arthritis Rheum. 2002;46:1470–9.
5. Keystone EC, Anisfeld A, Ogale S, Devenport JN, Curtis JR. Continued benefit of tocilizumab plus disease-modifying antirheumatic drug therapy in patients with rheumatoid arthritis and inadequate clinical responses by week 8 of treatment. J Rheumatol. 2014;41:216–26.
6. Salliot C, Finckh A, Katchamart W, Lu Y, Sun Y, Bombardier C, et al. Indirect comparisons of the efficacy of biological antirheumatic agents in rheumatoid arthritis in patients with an inadequate response to conventional disease-modifying antirheumatic drugs or to an anti-tumour necrosis factor agent: a meta-analysis. Ann Rheum Dis. 2011;70:266–71.
7. Bergman GJ, 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;39:425–41.
8. Schoels M, Aletaha D, Smolen JS, Wong JB. Comparative effectiveness and safety of biological treatment options after tumour necrosis factor

α inhibitor failure in rheumatoid arthritis: systematic review and indirect pairwise meta-analysis. Ann Rheum Dis. 2012;71:1303–8.
9. Stevenson M, Archer R, Tosh J, Simpson E, Everson-Hock E, Stevens J, et al. Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying antirheumatic drugs and after the failure of conventional disease-modifying antirheumatic drugs only: systematic review and economic evaluation. Health Technol Assess. 2016;20:1–610.
10. Singh JA, Wells GA, Christensen R, Tanjong Ghogomu E, Maxwell L, Macdonald JK, et al. Adverse effects of biologics: a network meta-analysis and Cochrane overview. Cochrane Database Syst Rev. 2011;2011:CD008794.
11. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71.
12. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd. etal 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis. 2010;69:1580–8.
13. Ward MM, Guthrie LC, Alba MI. Brief report: rheumatoid arthritis response criteria and patient-reported improvement in arthritis activity: is an American College of Rheumatology twenty percent response meaningful to patients? Arthritis Rheumatol. 2014;66:2339–43.
14. Brasil. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia. Diretrizes Metodológicas. Sistema GRADE – manual de graduação da qualidade da evidência e força de recomendação para tomada de decisão em saúde. Brasília: Ministério da Saúde; 2014. http://bvsms.saude.gov.br/bvs/publicacoes/diretrizes_metodologicas_sistema_grade.pdf
15. Higgins JPT, Altman DG, Sterne JAC. Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 (updated March 2011), The Cochrane Collaboration, Chichester (UK); 2011.
16. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making. 2013;33:607–17.
17. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. J Comput Graph Stat. 1998;7:34–55.
18. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med. 2010;29:932–44.
19. R Core Team (2022). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
20. Schiff M, Weinblatt ME, Valente R, van der Heijde D, Citera G, Elegbe A, et al. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: two-year efficacy and safety findings from AMPLE trial. Ann Rheum Dis. 2014;73:86–94.
21. Genovese MC, Schiff M, Luggen M, Le Bars M, Aranda R, Elegbe A, et al. Longterm safety and efficacy of abatacept through 5 years of treatment in patients with rheumatoid arthritis and an inadequate response to tumor necrosis factor inhibitor therapy. J Rheumatol. 2012;39:1546–54.
22. Schiff M, Keiserman M, Codding C, Songcharoen S, Berman A, Nayiager S, et al. Clinical response and tolerability to abatacept in patients with rheumatoid arthritis previously treated with infliximab or abatacept: open-label extension of the ATTEST Study. Ann Rheum Dis. 2011;70:2003–7.
23. Gaultney J, Benucci M, Iannazzo S, Nappi C, Sion K, Sabater FJ. Trial-based cost-effectiveness of abatacept for rheumatoid arthritis patients in Italy. Expert Rev Pharmacoecon Outcomes Res. 2016;16:409–17.
24. Genovese MC, Becker JC, Schiff M, Luggen M, Sherrer Y, Kremer J, et al. Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. N Engl J Med. 2005;353:1114–23.
25. Kremer JM, Genant HK, Moreland LW, Russell AS, Emery P, Abud-Mendoza C, et al. Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: a randomized trial. Ann Intern Med. 2006;144:865–76.
26. Emery P, Fleischmann R, Filipowicz-Sosnowska A, Schechtman J, Szczepanski L, Kavanaugh A, et al. The efficacy and safety of rituximab in patients with active rheumatoid arthritis despite methotrexate treatment: results of a phase IIB randomized, double-blind, placebo-controlled, dose-ranging trial. Arthritis Rheum. 2006;54:1390–400.
27. Porter D, van Melckebeke J, Dale J, Messow CM, McConnachie A, Walker A, et al. Tumour necrosis factor inhibition versus rituximab for patients with rheumatoid arthritis who require biological treatment (ORBIT): an open-label, randomised controlled, non-inferiority, trial. Lancet. 2016;388:239–47.
28. Keystone E, Burmester GR, Furie R, Loveless JE, Emery P, Kremer J, et al. Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. Arthritis Rheum. 2008;59:785–93.
29. Emery P, Deodhar A, Rigby WF, Isaacs JD, Combe B, Racewicz AJ, 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;69:1629–35.
30. Mease PJ, Cohen S, Gaylis NB, Chubick A, Kaell AT, Greenwald M, et al. Efficacy and safety of retreatment in patients with rheumatoid arthritis with previous inadequate response to tumor necrosis factor inhibitors: results from the SUNRISE trial. J Rheumatol. 2010;37:917–27.
31. Brown S, Everett CC, Naraghi K, Davies C, Dawkins B, Hulme C, McCabe C, Pavitt S, Emery P, Sharples L, Buch MH. Alternative tumour necrosis factor inhibitors (TNFi) or abatacept or rituximab following failure of initial TNFi in rheumatoid arthritis: the SWITCH RCT. Health Technol Assess. 2018;22:1–280.
32. Gabay C, Emery P, van Vollenhoven R, Dikranian A, Alten R, Pavelka K, et al. Tocilizumab monotherapy versus adalimumab monotherapy for treatment of rheumatoid arthritis (ADACTA): a randomised, double-blind, controlled phase 4 trial. Lancet. 2013;381:1541–50.
33. Bijlsma JWJ, Welsing PMJ, Woodworth TG, Middelink LM, Pethö-Schramm A, Bernasconi C, et al. Early rheumatoid arthritis treated with tocilizumab, methotrexate, or their combination (U-Act-Early): a multicentre, randomised, double-blind, double-dummy, strategy trial. Lancet. 2016;388:343–55.
34. Kivitz A, Olech E, Borofsky M, Zazueta BM, Navarro-Sarabia F, et al. Subcutaneous tocilizumab versus placebo in combination with disease-modifying antirheumatic drugs in patients with rheumatoid arthritis. Arthritis Care Res. 2014;66:1653–61.
35. Lindegaard HM, Johansen P, Gröndal G, Jensen EC, Juul L, Schlemmer AM, Agular B, Hansen I. Doubling the single-dose infusion rate of tocilizumab in rheumatoid arthritis is safe and efficacious. Scand J Rheumatol. 2016;45:262–6.
36. Smolen JS, Beaulieu A, Rubbert-Roth A, Ramos-Remus C, Rovensky J, Alecock E, Woodworth T, Alten R, OPTION Investigators. Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial. Lancet. 2008;371:987–97.
37. Emery P, Keystone E, Tony HP, Cantagrel A, van Vollenhoven R, Sanchez A, Alecock E, Lee J, Kremer J. IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial. Ann Rheum Dis. 2008;67:1516–23.
38. Gottenberg JE, Morel J, Perrodeau E, Bardin T, Combe B, Dougados M, et al. Comparative effectiveness of rituximab, abatacept, and tocilizumab in adults with rheumatoid arthritis and inadequate response to TNF inhibitors: prospective cohort study. BMJ. 2019;364:l67.
39. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol. 2017;17:162.
40. Fischhoff B. The realities of risk-cost-benefit analysis. Science. 2015;350:aaa6516.
41. Briggs AH, O'Brien BJ. The death of cost-minimization analysis? Health Econ. 2001;10:179–84.
42. Ades AE, Lu G, Higgins JP. The interpretation of random-effects meta-analysis in decision models. Med Decis Making. 2005;25:646–54.
43. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity–subgroups, meta-regression, bias, and bias-adjustment. Med Decis Making. 2013;33:618–40.

Publisher's Note