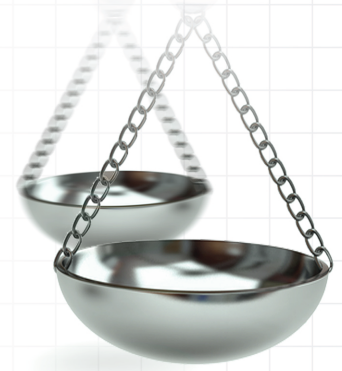




Letter to the Editor: indirect treatment comparison of anifrolumab efficacy versus belimumab in adults with systemic lupus erythematosus

Nick Ballew^{*1} , Aneela Mian², Roger Abramino Levy³ & Matt Bradley⁴

¹ GSK, Value Evidence & Outcomes, Collegeville, PA 19426, USA

² GSK, Global Medical Affairs, Brentford, Middlesex, UK

³ GSK, Specialty Care, Global Medical Affairs, Collegeville, PA 19426, USA

⁴ GSK, Value Evidence & Outcomes, Brentford, Middlesex, UK

*Author for correspondence: Tel.: +1 610 917 7000; nick.g.ballew@gsk.com

First draft submitted: 10 June 2022; Accepted for publication: 17 August 2022; Published online: 14 December 2022

Keywords: clinical trials • comparative effectiveness research • evidence-based medicine • health technology assessment • immunology • pharmaceuticals • randomized controlled trials • rheumatology

We read with interest the publication titled “Indirect treatment comparison of anifrolumab efficacy versus belimumab in adults with systemic lupus erythematosus” by Bruce *et al.* [1]. Using population-adjusted indirect comparison methods, the authors report that in anifrolumab-treated patients with moderate-to-severe systemic lupus erythematosus (SLE), the odds of response (reduction of ≤ 4 points in SLE Disease Activity Index 2000 [SLEDAI-2K] and SLE Responder Index-4 [SRI-4]) is over twofold higher than that observed in belimumab-treated patients (odds ratio [OR] varies slightly depending on the outcome and method) [1].

We believe that indirect treatment comparisons (ITCs) are a valuable tool to compare the efficacy and safety of alternative treatment options when direct data are unavailable. However, we have identified limitations of the approach used that have not been described fully and which impact the interpretation of the authors’ findings.

Our concerns are:

- The clinical profile of patients included in the TULIP studies is highly selective and under-represented in the broader, more characteristic SLE population recruited to the BLISS studies.
- The methodological approach selected is subject to bias and lacks transparency.
- The results are not in line with observations from the individual clinical trials.

Concern 1

An important property of an ITC is that the patient populations are comparable [2]. Where this is not the case, but there is sufficient overlap of the populations, an adjusted analysis can be performed [3]. The authors are aware of this issue, citing ‘insufficient overlap in patient characteristics’ to reject otherwise eligible belimumab studies from the analysis [1]. However, while attempting to ‘mitigate the impact of between trial heterogeneity’ [1], the authors have overlooked the critical issue of insufficient overlap of the narrower TULIP population with the broader BLISS population to allow a robust adjustment to be performed.

The differences between the SLE populations in the TULIP versus BLISS studies are important: almost all patients in the TULIP studies had severe disease (British Isles Lupus Assessment Group [BILAG] index 1A and/or $\leq 2B$), predominantly in the mucocutaneous and musculoskeletal domains [4,5]; in contrast, the BLISS studies included a broader SLE cohort with involvement beyond skin and joints [6–8]. This difference may be explained by the more restrictive inclusion criteria of the TULIP studies [4,5], resulting in an SLE population with more skin and joint involvement at the expense of patients with other serious organ involvement such as hematological and renal.

From the data presented in the paper [1], it is evident that the TULIP population is under-represented in the broader, more characteristic SLE population recruited to the BLISS studies. The most obvious illustration comes from the matching-adjusted indirect comparison (MAIC):

- In the MAIC, the effective sample size from the TULIP studies was reduced by approximately 90% from the actual sample size, indicating very poor overlap [1,3]. This points to under-representation of the TULIP population in the broader SLE population of the BLISS studies. With such a markedly reduced effective sample size, inferences depend heavily on a small number of individuals, which may lead to inflation of response, generating spurious results. As the MAIC weights are not provided, it is impossible to assess the extent of this bias.

Concern 2

It is unclear how the feasibility assessment criteria were applied to each study [1]. This is critical since the analysis is highly dependent on the choice of studies. To date, belimumab has been evaluated in ten randomized controlled trials (including five placebo-controlled phase III trials in SLE), in long-term extensions and multiple post-marketing efficacy and safety studies [9]. As of April 2021, >7200 patients with SLE had received belimumab in clinical studies [9]. Yet the breadth of belimumab evidence has been excluded from the analysis [1]. Of particular note is the exclusion of the phase III BLISS-SC study (as noted by the authors as bioequivalent) [10,11]; no explanation is offered why the primary analysis is restricted to the belimumab intravenous formulation. BLISS-SC was included in the supplementary analysis [11] and, notwithstanding our overall concerns, these results change the paper's conclusions since there was no significant difference between treatments across any outcomes. It is concerning that there is no comment on this, given the conclusion from the primary analysis.

There is no discussion in the paper that the BLISS and TULIP studies used different BILAG index versions [4,5,7,8]. This impacts not only the availability of effect modifiers common to both trials but also, potentially, the comparison of outcomes incorporating BILAG (SRI-4 response and BILAG flare at week 52).

The selection of effect modifiers is critical to the validity of the simulated treatment comparison (STC) [3]. The authors identified 14 effect modifiers, reducing to 12 to enable the MAIC [1]. However, the selection is unclear, supported by an unreferenced literature review of the predictors of SLE treatment outcomes. The authors note that prognostic variables are not required in the STC/MAIC and reflect on the challenges of differentiating between prognostic variables and effect modifiers in their discussion. This makes the absence of reporting of the 'extensive sensitivity analyses. . . to explore the impact of adjusting for different subsets of TEMs' troubling [1].

The authors assert their preference for STC methods based on recent research suggesting superior performance in eliminating bias when compared with MAIC. However, there are other simulation studies published recently that detail scenarios (similar to that investigated here; conducting the adjustment in presence of poor population overlap, comparing a conditional to marginal effect for the STC, and a small effective sample size in the MAIC) where, even in the presence of effect modifiers, both STC and MAIC can produce results with substantial bias and, in some cases, produce results that are less accurate than a standard ITC [12,13].

It has been demonstrated that MAIC analyses perform poorly when populations have limited overlap, generating results that, '... may even increase bias compared with standard indirect comparisons' [12–14]. The authors acknowledge this concern and hence base the primary analysis on the STC [1]. Given this known bias, we find it curious that the MAIC results in the paper are highly consistent with the STC results. We believe rather than validating the robustness of the STC, as claimed by the authors, the MAIC results actually demonstrate that the STC is equally subject to bias arising from limited population overlap.

Concern 3

The pooled unadjusted OR for anifrolumab versus placebo for SRI-4 (Table 3; and likely used in the analysis) appears to be an overestimate [1]. The 1.63 OR is higher than what would be expected based on a pooling of the prespecified results in the primary publications (1.33) or if the revised *post hoc* definition for TULIP-1 is considered (1.56).

The authors adjust for between trial heterogeneity with regard to the TULIP and BLISS studies but appear to ignore heterogeneity between the two TULIP studies [4,5,7,8]: SRI-4 response favored placebo in TULIP-1 but anifrolumab in TULIP-2. This has implications for the plausibility of the STC/MAIC results and leads to overly narrow 95% CIs. Pooling the study results of TULIP-1 and TULIP-2 gives an OR of 1.33 (CI 0.99–1.79 fixed

effects/CI 0.54–3.25 random effects). This is non-significant, in clear contrast to the statistically significant benefit for belimumab on SRI-4 response in the individual and pooled BLISS studies.

To perform the STC, the authors adjust for 12 baseline characteristics considered to be effect modifiers despite that only four from their selection are substantially different between BLISS and TULIP. Arguably increased disease activity (BILAG Index A and B, SLEDAI score) would be expected to result in a larger treatment effect for both belimumab and anifrolumab [15,16]. Thus, it is counterintuitive that this analysis generates a larger anifrolumab treatment effect in a lower disease activity population.

Summary

In summary, this ITC fails to meet fundamental requirements for a robust population adjusted analysis: that there is sufficient overlap that adjustments can be made [2,3]. BLISS-SC is a relevant study, and the rationale for its exclusion is not clearly articulated. We have highlighted other examples where the methodology may have introduced bias and lacks transparency. The cumulative effect of these issues leads to conclusions that are not in line with observations based on the results of the individual clinical studies. Without appropriate discussion of these limitations, we consider the findings of this paper are unhelpful to appropriate decision-making in the choice of SLE treatments.

Author contributions

All authors contributed equally to the development of this letter.

Financial & competing interests disclosure

The production of this letter to the editor was funded by GSK. All authors are employees of GSK. N Ballew, RA Levy and M Bradley hold stocks and shares in GSK. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Editorial and submission support was provided by H Taylor, Fishawack Indicia Ltd., UK, part of Fishawack Health and was funded by GSK.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

References

1. Bruce IN, Golam S, Steenkamp J *et al.* Indirect treatment comparison of anifrolumab efficacy versus belimumab in adults with systemic lupus erythematosus. *J. Comp. Eff. Res.* 11(10), 765–777 (2022).
2. 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* 33(5), 607–617 (2013).
3. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. *Med. Decis. Making* 38(2), 200–211 (2018).
4. Morand EF, Furie R, Tanaka Y *et al.* Trial of anifrolumab in active systemic lupus erythematosus. *N. Engl. J. Med.* 382(3), 211–221 (2020).
5. Furie RA, Morand EF, Bruce IN *et al.* Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase III trial. *Lancet Rheumatol.* 1(4), e208–e219 (2019).
6. Manzi S, Sanchez-Guerrero J, Merrill JT *et al.* Effects of belimumab, a B lymphocyte stimulator-specific inhibitor, on disease activity across multiple organ domains in patients with systemic lupus erythematosus: combined results from two Phase III trials. *Ann. Rheum. Dis.* 71(11), 1833–1838 (2012).
7. Furie R, Petri M, Zamani O *et al.* A Phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. *Arthritis Rheum.* 63(12), 3918–3930 (2011).
8. Navarra SV, Guzmán RM, Gallacher AE *et al.* Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase III trial. *Lancet* 377(9767), 721–731 (2011).
9. Levy RA, Gonzalez-Rivera T, Khamashta M *et al.* 10 Years of belimumab experience: what have we learnt? *Lupus* 30(11), 1705–1721 (2021).
10. European Medicines Agency. Belimumab summary of product characteristics (2021). www.ema.europa.eu/en/documents/overview/benlysta-epar-medicine-overview_en.pdf

11. Stohl W, Schwarting A, Okada M *et al.* Efficacy and safety of subcutaneous belimumab in systemic lupus erythematosus: a fifty-two-week randomized, double-blind, placebo-controlled study. *Arthritis Rheumatol.* 69(5), 1016–1027 (2017).
12. Remiro-Azocar A, Heath A, Baio G. Methods for population adjustment with limited access to individual patient data: a review and simulation study. *Res. Synth. Methods* 12(6), 750–775 (2021).
13. Remiro-Azocar A, Heath A, Baio G. Parametric G-computation for compatible indirect treatment comparisons with limited individual patient data. *Res. Synth. Methods.* 13(6), 716–744 (2022).
14. Phillippo DM, Dias S, Ades AE, Welton NJ. Assessing the performance of population adjustment methods for anchored indirect comparisons: a simulation study. *Stat. Med.* 39(30), 4885–4911 (2020).
15. Van Vollenhoven RF, Petri MA, Cervera R *et al.* Belimumab in the treatment of systemic lupus erythematosus: high disease activity predictors of response. *Ann. Rheum. Dis.* 71(8), 1343–1349 (2012).
16. European Medicines Agency. Saphnelo (anifrolumab) - medicine overview (2022). www.ema.europa.eu/en/documents/overview/saphnelo-epar-medicine-overview_en.pdf