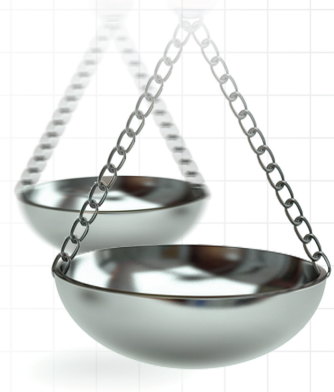
Letter in Reply



Journal of **Comparative Effectiveness Research**

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

Ian N Bruce*[,1], Sarowar Golam[2], Jason Steenkamp[3], Pearl Wang[3], Evelyn Worthington[3], Barnabas Desta[4], Konstantina Psachoulia[5], Wilma Erhardt[2] & Raj Tummala[5]

[1]Division of Musculoskeletal & Dermatological Sciences, Centre for Epidemiology Versus Arthritis, The University of Manchester & NIHR Manchester Biomedical Research Centre, Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK

[2]Global Market Access & Pricing, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

[3]EVERSANA™, Burlington, ON, L7N 3H8, Canada

[4]BioPharmaceuticals Business Unit, AstraZeneca, Gaithersburg, MD 20878, USA

[5]Respiratory, Inflammation & Autoimmunity, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, MD 20878, USA

*Author for correspondence: ian.bruce@manchester.ac.uk

First draft submitted: 25 October 2022; Accepted for publication: 4 November 2022; Published online: 14 December 2022

Keywords: comparative effectiveness research • indirect treatment comparison • randomized controlled trials • simulated treatment comparison • systemic lupus erythematosus

We appreciate the insightful comments from the author(s) of the Letter to the Editor regarding our recent publication, 'Indirect treatment comparison of anifrolumab efficacy versus belimumab in adults with systemic lupus erythematosus' [1]. At the outset we would say that it is hugely encouraging that we now have more than one approved medication for the management of moderate to severe systemic lupus erythematosus (SLE), a situation which should be of benefit to our patients. We welcome the opportunity to address some of the key points highlighted by the Letter to the Editor.

The first point (Concern 1) deals with the clinical profile of patients included in the TULIP studies as compared with the BLISS studies. The author(s) considered TULIP to be less representative of the characteristic SLE population, highlighting the drop in effective sample size during matching-adjusted indirect comparison (MAIC) as supporting evidence. Although, the author(s) are correct to identify there are differences in the mean patient baseline characteristics between these trials, necessitating the use of population-adjusted indirect treatment comparison (ITC) methods to reduce potential bias, it is wrong to state that TULIP is unrepresentative of the target SLE population. Both the TULIP and BLISS trials enrolled patients with moderate to severe SLE despite standard therapy, which ultimately is reflected in comparable target populations approved by the US FDA and European regulators. In both the TULIP [2,3] and BLISS [4–6] trials, patient eligibility required active disease despite stable therapy as evidenced by a SLE Disease Activity Index (SLEDAI) of $\geq$6 [2–5] (or for BLISS-SC, $\geq$8 [6]) and a positive autoantibody. A key difference in the TULIP trials was the stated need for a clinical SLEDAI $\geq$4 and evidence of at least one British Isles Lupus Assessment Group (BILAG) 2004 A or 2 BILAG 2004 B scores confirming the presence of active clinical disease. The TULIP trials did therefore recruit patients with more clinical disease, albeit still within the moderate to severe spectrum of active SLE (again, this is reflected by the approved indication [7,8]). Overall; however, we contend that these are broadly similar trials with comparable inclusion criteria. The gold standard of evidence for a comparison between drugs would be a head-to-head clinical trial comparing anifrolumab to belimumab. In the absence of such a trial, population-adjusted ITCs can give us a reasonable estimate that we can interpret with care, bearing in mind methodological limitations.

In Concern 2, the author(s) bring up several points regarding the ITC methodologies: study selection, BILAG index version, selection of effect modifiers, and the selection and appropriateness of the methodologies used.

First, the author(s) consider whether additional belimumab studies should have been included. Both MAIC and simulated treatment comparison (STC) are fundamentally pairwise analyses that require pooling of studies to inform each treatment node. It is not appropriate to pool every study conducted for a treatment of interest, but rather to select the most comparable studies of those available such that the comparison has the highest chance of fulfilling the assumption of exchangeability. Long-term extensions, post marketing efficacy studies and studies restricted to a geographical region rather than in a global patient population did not substantially meet the comparability standard to pool with phase III, randomized, double-blind, placebo-controlled, global clinical studies like TULIP and BLISS. The author(s) particularly highlight BLISS-SC, and we agree that this study holds particular relevance. It is precisely why we conducted a sensitivity analysis comparing TULIP-1 and TULIP-2 to BLISS-SC alone, as it was our opinion that it is not methodologically appropriate to pool BLISS-SC with BLISS-52 and BLISS-76. Aside from the obvious difference in administration route between belimumab i.v. and s.c., the eligibility criteria of BLISS-SC also differed from the other BLISS studies (and from TULIP) in that the population required a higher SLEDAI score at entry (SLEDAI $\geq$8 instead of $\geq$6).

The authors also highlighted the difference in BILAG index versions between the BLISS and TULIP studies. We accept that this is a reasonable limitation to identify. However, we do not believe that this precludes a comparison of the trials. The two BILAG index versions are broadly comparable as they both share the same 'intention-to-treat' principle at the core of the scoring systems [9]. This consistent intention-to-treat principle means that any differences between them are unlikely to substantially change the conclusions of our study.

Additionally in Concern 2, the author(s) question the choice of STC as the primary analysis and suggest both MAICs and STCs may perform poorly. We support the use of STC as the primary analysis given recent research [10]. However, there are different schools of thought and to be as transparent and broad as possible, we included both the MAIC and STC results in the body of the publication. MAICs and STCs are methodologies accepted by key Health Technology Assessment agencies such as The National Institute for Heath and Care Excellence (NICE), and we conducted our analyses based on the NICE Decision Support Unit technical support documents [11]. No ITC will ever be completely free from bias, whether population adjusted or otherwise and neither will naively comparing two separate placebo-controlled trials by their absolute values alone. Ultimately, conducting population adjusted ITCs was the most robust approach possible given the data available in order to aid clinicians in decision making for their patients.

In Concern 3, the author(s) highlight some potential areas of confusion regarding between-trial heterogeneity in prespecified results between TULIP-1 and TULIP-2. Regarding the odds ratio for SLE Responder Index-4 (SRI-4), the SRI-4 results from TULIP-1 that were used in this ITC were the same values agreed to by the FDA and reported in the US Prescribing Information [8]. This aligned with the prespecified outcome definition used in TULIP-2 as well as the BLISS trials, and how SRI-4 is reported in the product labels for anifrolumab [7,8].

The author(s) also note that only four of the baseline characteristics were substantially different between BLISS and TULIP, whereas we adjusted for 12. There is always a balance to be struck between adjusting for every characteristic (i.e., tighter matching but lower sample size) and only the few most critical characteristics (i.e., may still miss numerically small but clinically important differences, but preserves sample size). If the characteristics are not substantially different, it does not add any bias to adjust for them and it is, in fact, a more conservative approach since confidence intervals become wider with the smaller and more well-matched sample size, making statistical significance harder to achieve.

Finally in Concern 3, the author(s) assert that it is counterintuitive that anifrolumab showed a larger relative treatment effect when adjusting toward a lower disease-activity population to match BLISS. Belimumab has previously showed higher treatment effect in more severely active patients (for example the 'high disease activity population' with a SELENA SLEDAI of 10 or more according to the EU product information) [12]. In contrast, anifrolumab has a more consistent effect across a wider range of disease activity, with positive treatment differences in BILAG-based Composite Lupus Assessment (BICLA) response at week 52 for anifrolumab versus placebo in patients with 'low disease activity' (SLEDAI-2K <10: 54.2% vs 37.2%) and 'high disease activity' (SLEDAI-2K $\geq$10: 44.7% vs 28.2%) [13]. This is reflected in the broader EMA label for anifrolumab compared with belimumab with regards to disease severity [7,14].

In the absence of head-to-head trials, we agree it is imperative that ITCs are as transparent and robust as possible. Like any ITC, the analyses conducted in our publication have limitations and we attempted to address these in our discussion and are happy to clarify further in this letter. We do however disagree with the authors that the conclusions are 'unhelpful to appropriate decision making' on account of those limitations. We would underscore

that both STC and MAIC are widely accepted methodologies for population adjusted analyses, as described by major HTA bodies such as NICE. The limitations of our analyses, the data and the ITC methodologies themselves have been described in our publication and are a critical component in interpreting the results. Overall, these limitations do not prevent a meaningful comparison across the clinical trials in question. As we enter a new era of having more choice of therapies for our patients living with SLE, such comparative analyses will provide another level of support to help clinicians, patients and payers in their decision making.

Financial & competing interests disclosure

This work was supported by AstraZeneca. I Bruce has received grant support from Genzyme/Sanofi, GSK, Roche and UCB; consulting fees from AstraZeneca, Lilly, GSK, Merck Serono, UCB and ILTOO; and was a speaker for AstraZeneca, GSK and UCB. S Golam, B Desta, K Psachoulia, W Erhardt and R Tummala are employees of AstraZeneca. J Steenkamp, P Wang and E Worthington are employees of EVERSANA™, which was contracted by AstraZeneca to work on this study. 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.

Medical writing and editorial support were provided by R Franklin of JK Associates Inc., part of Fishawack Health, and were funded by AstraZeneca.

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. Furie RA, Morand EF, Bruce IN, Manzi S, Kalunian KC, Vital EM. 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).
3. 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).
4. 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. (Munch.)* 63(12), 3918–3930 (2011).
5. Navarra SV, Guzman 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).
6. 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).
7. European Medicines Agency. Saphnelo™ Summary of Product Information. www.ema.europa.eu/en/medicines/human/EPAR/saphnelo
8. SAPHNELO™ Prescribing Information. (2021). www.accessdata.fda.gov/drugsatfda_docs/label/2021/761123s000lbl.pdf
9. Isenberg DA, Rahman A, Allen E *et al.* BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. *Rheumatology (Oxford)* 44(7), 902–906 (2005).
10. 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).
11. National Institute of Health and Care Excellence. Population-adjusted indirect comparisons (MAIC and STC). https://www.sheffield.ac.uk/nice-dsu/tsds/population-adjusted
12. 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).
13. Vital EM, Merrill JT, Morand EF *et al.* Anifrolumab efficacy and safety by type I interferon gene signature and clinical subgroups in patients with SLE: *post hoc* analysis of pooled data from two phase III trials. *Ann. Rheum. Dis.* 81(7), 951–961 (2022).
14. European Medicines Agency. BENLYSTA® Summary of Product Information. www.ema.europa.eu/en/documents/product-information/benlysta-epar-product-information_en.pdf